Nucleic Acids Research, 1999, Vol. 27, No. 5 **1323–1330**

Targeted gene repair directed by the chimeric RNA/DNA oligonucleotide in a mammalian cell-free extract

Allyson Cole-Strauss, Howard Gamper, William K. Holloman[1], Madeline Muñoz, Nikki Cheng and Eric B. Kmiec*

Department of Microbiology and Immunology, Jefferson Center for Biomedical Research, Thomas Jefferson University, 700 Butler Avenue, Doylestown, PA 18901, USA and [1]Department of Microbiology, Cornell University School of Medicine, 1300 York Avenue, New York, NY 10021, USA

Received November 18, 1998; Revised and Accepted January 8, 1999

ABSTRACT

Chimeric oligonucleotides consisting of RNA and DNA residues have been shown to catalyze site-directed genetic alteration in mammalian cells both *in vitro* and *in vivo*. Since the frequency of these events appears to be logs higher than the rates of gene targeting, a process involving homologous recombination, we developed a system to study the mechanisms of chimera-directed gene conversion. Using a mammalian cell-free extract and a genetic readout in *Escherichia coli*, we find that point mutations and single base deletions can be corrected at frequencies of ~0.1% and 0.005%, respectively. The reaction depends on an accurately designed chimera and the presence of functional hMSH2 protein. The results of genetic and biochemical studies reported herein suggest that the process of mismatch repair functions in site-directed gene correction.

INTRODUCTION

The success of genetic knock-outs in animals to create models of human diseases has matured the conception of gene targeting from experimental technique to therapeutic possibility. Yet significant barriers remain; prominent among these is the low frequency of site-specific targeting in mammalian cells. For every targeted event, $>10^5$ or 10^6 non-specific events occur (1) even though mammalian cells possess the enzymatic machinery to conduct homologous targeting (2,3). At the center of the targeting problem is the infrequent rate of homologous recombination (HR) in mammalian cells and it is likely that this frequency is heavily influenced by the level of enzymatic activity in each cell. Differences in cell type, cell cycle position and the level of certain expressed proteins may also regulate this rate (4,5). The last of these parameters led certain investigators to identify genes and proteins that might improve targeting efficiency because they promote homologous recombination reactions *in vitro*. In spite of the success in identifying some candidates (RAD51 and RAD52), no successful application to enhance the frequency of gene targeting has yet been made. In fact, the human Rad51 protein has been found to become phosphorylated after DNA damage (6), a modification that obliterates its recombinase ability.

Our work in the enzymology of recombinational repair (reviewed in 7) led us to conclude that specific gene targeting as a gene therapy is currently problematic. Hence, we sought to develop a new strategy, one in which mutated genes are repaired rather than replaced. Based on *in vitro* studies that revealed the importance of RNA in mediating DNA pairing events (8,9), we designed an oligonucleotide that could localize to a specific site and be effective in gene targeting. The molecule is a double-stranded chimera consisting of RNA and DNA residues, capped at both ends by sequences which fold into a hairpin (10). By virtue of its structure, the half-life in the cell and in serum is substantially longer than similar duplex molecules (11). The chimera also contains a single nucleotide that differs from the target sequence and, upon binding, forms a mismatched base with the specific targeted nucleotide. The resulting helical distortion is subsequently recognized by the cell's DNA repair machinery and the base pair corrected using the DNA sequence of the chimera as a template. The strategy of converting single nucleotides or introducing short base insertions or deletions differs greatly from all gene replacement methods currently in use. Gene conversion frequencies surpassing 10^{-1} have been achieved in a variety of cell types using episomal and genomic targets (11–18).

Although the successes in various mammalian cell lines and animal models have provided an impetus to consider therapeutic applications, we sought a more fundamental understanding of the mechanism of chimera-directed gene repair. To this end, a mammalian cell-free extract that can promote conversion *in vitro* was developed. The system relies on directed correction of mutant antibiotic resistance genes in a mammalian cell-free extract that can be quantitated after transformation of *Escherichia coli*, using genetic readout to score for positive correction events. In this report, we present an analysis of conversion of three different targets, mediated by the chimeric oligonucleotide, and provide preliminary results addressing the genetic basis of the mechanism.

*To whom correspondence should be addressed. Tel: +1 215 489 4903; Fax: +1 215 489 4922; Email: eric.kmiec@mail.tju.edu

MATERIALS AND METHODS

Plasmids

The tet^S plasmids were derived from pBR322 (Fig. 1). pT^{S}m153 contains a single base change (T→A) in the tetracycline (tet) gene at position 153. The mutation generates a stop codon at position 153 in the reading frame, rendering the plasmid sensitive to tetracycline, and creates a *Bfa*I restriction site. In addition, pT^{S}m153 has a silent mutation (A→G) at position 325 which generates a *Tse*I restriction site. pT^{S}Δ208 contains a single nucleotide deletion at position 208 which disrupts the reading frame of the tetracycline gene (see Fig. 4). This deletion also introduces a *Sph*I restriction site. The kan^S plasmid pK^{S}m4021 (Fig. 1) contains a single base change (T→G) at base 4021 relative to the parental plasmid pWE15 (Stratagene). All mutations were created using a standard overlapping PCR procedure. This transversion creates both a stop codon in the kanamycin (kan) gene and a diagnostic *Bfa*I site. All these plasmids contain a functional ampicillin gene.

Oligonucleotides

Chimeric oligonucleotides were synthesized using commercially available phosphoramidites on 1000 Å CPG supports. Following release and deprotection, crude product was purified to at least 85% homogeneity as judged by column chromatography. Oligonucleotide concentrations were determined spectrophometrically using a conversion factor of 33 µg/ml per $A_{260\ nm}$ unit.

Cells

HuH-7 cells were obtained from the laboratory of Dr Clifford Steer (University of Minnesota). They were cultured in DMEM, 10% FBS, 2 mM glutamine, 0.5% penicillin/streptomycin. LoVo cells were obtained from ATCC (CCL-229) and cultured in Ham's F12 with 10% FBS, 2 mM glutamine and 0.5% penicillin/streptomycin. Electromax DH10B cells were from Life Technologies and BMH71-18 cells were from Clontech. Pertinent genetic alterations include recA1 for DH10B and mutS::Tn*10* for BMH71-18 cells, respectively.

Cell-free extract

Cell-free extracts were prepared by a modification of Li and Kelly (19), using cultured HuH-7 cells. Briefly, ~2 × 10^8 cells were harvested from flasks by scraping. They were washed immediately in cold hypotonic buffer (20 mM HEPES, pH 7.5, 5 mM KCl, 1.5 mM $MgCl_2$, 1 mM DTT) with 250 mM sucrose. Cells were resuspended in cold hypotonic buffer without sucrose and after 15 min lysed with 25 strokes of a Dounce homogenizer using a tight fitting pestle. Lysed cells were incubated on ice for 60 min and centrifuged at 12 000 *g* for 15 min. The cytoplasmic fraction is enriched with nuclear proteins due to the extended co-incubation of the fractions following cell breakage. The supernatant was immediately aliquoted and frozen at –80°C. Protein concentration of the extract was determined by the Bradford assay.

In vitro reaction

Reaction mixtures of 100 µl consisted of 15–30 µg of cytoplasmic extract, 800–2100 ng of chimeric oligonucleotide and 1 µg of plasmid DNA in 20 mM Tris, pH 7.4, 15 mM $MgCl_2$, 0.4 mM DTT and 1.0 mM ATP. Samples were incubated at 37°C for 45 min and deproteinized by two phenol/chloroform (1:1) extractions. They were ethanol precipitated with 0.1 vol 3 M sodium acetate, pH 4.8, and 2 vol ethanol at –20°C for 2 h or overnight. Samples were centrifuged at 15 000 r.p.m. and 4°C for 30 min (Eppendorf 5417R centrifuge). Pellets were resuspended in 50 µl H_2O and stored at –20°C. Reaction mixtures for antibody studies were identical to those described above except that prior to assembling the reagents, 500 ng of αhMSH2 or IgG (Santa Cruz Biotechnology) were mixed with the extract at 4°C for 1–2 h. The other reaction components were then added and the reaction processed as described above.

Electroporation and plating of samples

Except where noted, plasmid (5 µl; 100 ng) was transfected into DH10B cells (20 µl) by electroporation (400 V, 300 µF, 4 kΩ) in a Cell-Porator apparatus (Life Technologies). Afterwards, the cells (20 µl) were transferred to 1 ml SOC in a 1.5 ml Eppendorf tube and incubated at 37°C for 1 h. Aliquots of this suspension were spread on tet (12 µg/ml) or kan (50 µg/ml) plates without dilution and on ampicillin (100 µg/ml) plates after a 10^3–10^4 dilution. Colonies were counted on an Accucount 1000 (Biologics) after overnight incubation at 37°C. Gene conversion frequency was expressed as the ratio of tet- or kan-resistant colonies per 10^5 ampicillin-resistant colonies. When higher tet or kan colony counts were desired, the cell suspension was transferred into 4 ml SOC with tet (final concentration 2.4 µg/ml) or kan (10 µg/ml) in a 14 ml Falcon snap-cap tube and shaken at 37°C for 3 h. Cells were pelleted and resuspended in 200 µl SOC prior to plating as usual. Individual tet^R and kan^R colonies were picked into 5 ml LB with tet or kan and grown overnight. The cell pellet from a 200 µl aliquot of each culture was resuspended in 30 µl H_2O, heated to 95°C for 10 min and 10 µl used for PCR amplification. The remaining culture was pelleted and frozen for DNA analyses by miniprep.

PCR amplification

In the tet system, PCR primers generated a 475 bp fragment. The 5′ and 3′ primer sequences were CAT/CGA/TAA/GCT/TTA/ATG/C and GCT/CAT/GAG/CCC/GAA/GTG/GC. The PCR conditions were 94°C/30 s, 59°C/30 s, 72°C/30 s for 30 cycles in a Perkin Elmer 2400 thermocycler. For restriction analysis of convertants at position 153, each sample was digested with *Bfa*I to verify gene correction and with *Tse*I to confirm the presence of the silent mutation. Restriction analysis of convertants at position 208 utilized digestion with *Sph*I. In the kan system, the 5′ and 3′ primers were CAG/GGG/ATC/AAG/ATC/TGA and CGA/AGT/CAC/TGT/TGC/AGC/TCG. The PCR conditions were 94°C/30 s, 55°C/30 s, 72°C/30 s for 35 cycles. Loss of a *Bfa*I restriction site indicated correction of the kan gene at position 4021. Restriction digests were electrophoresed at 120 V for 2 h in a 1.5% agarose gel containing TAE buffer (0.04 M Tris–acetate, 1 mM EDTA) and 0.5 µg/ml ethidium bromide. A subset of PCR fragments were subjected to automatic DNA sequencing in both directions following purification on a QIAquick spin column (Qiagen).

RESULTS

The repair assay

A diagram of the plasmids and chimeras used for measuring targeted gene repair is depicted in Figure 1. The assay is based on

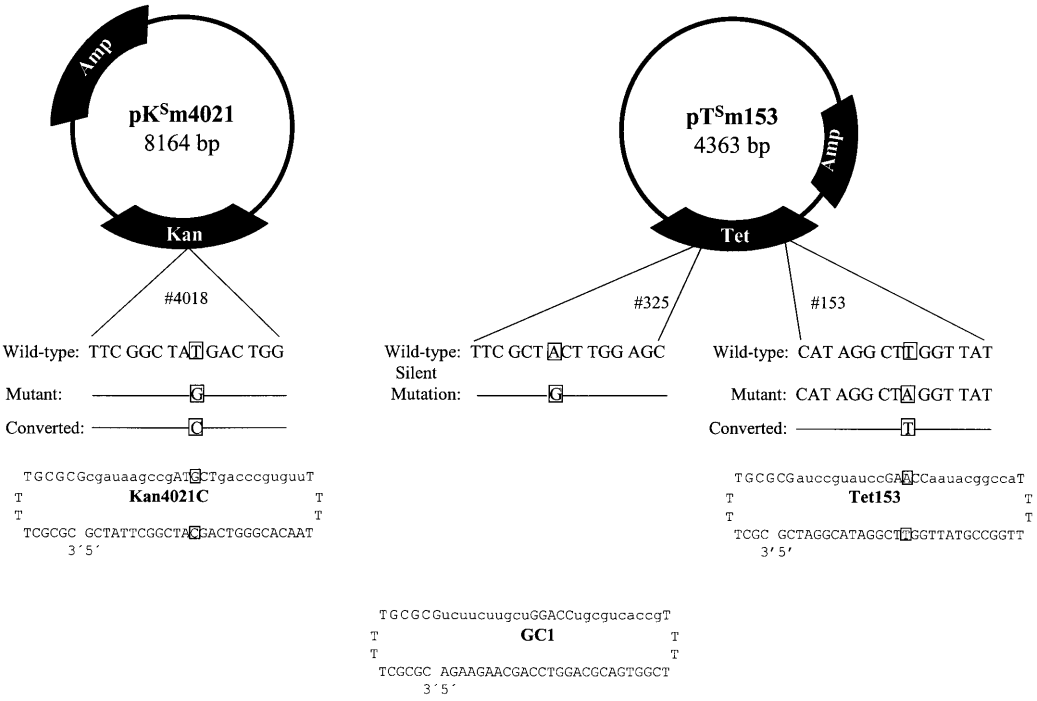

Figure 1. Kan^S and tet^S plasmids and correcting chimeras. Model systems for detecting targeted base pair substitution by chimeric oligonucleotides. Bases at the site of correction are boxed. Lower case letters in the chimeric oligonucleotides denote where the DNA backbone is replaced by a 2′-*O*-methyl backbone. GC1 is a non-specific chimera bearing no complementarity to either the tet^S or kan^S target sites.

the targeted conversion of genes encoding antibiotic resistance in *E.coli*. Plasmids pT^{S}m153 and pK^{S}m4021 contain separate and different point mutations within the coding region of genes conferring resistance to tet (pT^{S}m153) and kan (pK^{S}m4021). In the case of pT^{S}m153, which is derived from pBR322, the T residue at position 153 has been changed to A, thereby generating a stop codon and a diagnostic *Bfa*I site. In addition, the tet gene also contains a silent mutation at position 325 (A→G). This alteration in sequence has no effect on genetic readout but enables confirmation that converted plasmids did not arise from wild-type plasmid contamination. The site was constructed so that a *Tse*I cleavage site is available for control analyses. Plasmid pK^{S}m4021 is derived from pWE15 by changing a T residue at position 4021 to G. This change introduces a stop codon in the kan gene and an associated *Bfa*I restriction site.

Chimeric oligonucleotides used to repair the antibiotic resistance genes share the same design as chimeras previously reported to work in human cells. These molecules consist of a 30 bp long duplex bracketed by 4 base long hairpin loops. The nicked dumb-bell structure in conjunction with 2′-*O*-methyl backbone segments in the upper strand impart high thermal stability and enhanced nuclease resistance. Left of the nick is a 5 bp long G-C clamp inserted to prevent fraying. To the right of the nick is a 25 bp long targeting region which is homologous to the plasmid DNA with the exception of a single centralized base pair. For back mutation of the antibiotic resistance genes, the chimeras specify functional sequences and so differ from the targets at the site of the point mutation. Chimeric oligonucleotide Tet153 restores the wild-type T residue in the tet gene while the Kan4021C chimera directs a G→C conversion in the kan gene. The resultant TAC codon in the kan gene specifies Tyr just like the wild-type TAT codon, but also distinguishes corrected plasmids from contaminating wild-type plasmids. The chimeric oligonucleotide designated GC1 is a non-specific control molecule containing no homology to either plasmid.

The extract was produced by breaking cells by hypotonic lysis. The cell debris and subcellular organelles were separated by centrifugation at 12 000 *g* and the clarified cytoplasmic fraction was saved (see Materials and Methods for details). A 1 h holding period prior to centrifugation enriched the extract in nuclear proteins. The reaction mixture consisted of the cell-free extract, supercoiled plasmid and chimeric oligonucleotide in a buffer containing 20 mM Tris, pH 7.4, 15 mM $MgCl_2$, 0.4 mM DTT and 1 mM ATP. After incubation, the samples were extracted with phenol/chloroform and the DNA precipitated by ethanol. The DNA (predominately nicked circular) was then electroporated into strains of *E.coli* that are either deficient in RecA activity (DH10B) or MutS activity (BMH71-18). Unpublished studies have shown that both RecA and MutS proteins are required for chimera-mediated repair in *E.coli* (R.Metz, personal communication). Hence, by using mutant strains, it is unlikely that converted plasmids arise from repair activity in the bacterial cell. As a genetic readout, clones arising on plates containing kan or tet were picked and propagated as isolated colonies for RFLP analysis and sequencing. Colony counts were normalized for variable electroporation and plating efficiencies by determining the number of tet^R or kan^R colonies generated per 10^5 amp^R colonies for each reaction.

Table 1, top, illustrates results of chimera-directed genetic repair of the pT^{S}m153 plasmid in an extract from HuH-7 cells. In this experiment the amounts of plasmid and chimeric oligonucleotide

were held constant while the level of extract was increased. A dose-dependent increase in tet-resistant colonies was observed up to 12 µg extract, beyond which the activity plateaued. The non-specific chimera, GC1, produced no colonies. In the absence of the extract, no tet^R colonies were observed, reinforcing the conclusion that the repair of plasmid DNA occurs in the extract. Similar results were observed in the kan system (Table 1, bottom). A dose–response effect was observed and non-specific, or incomplete, reactions did not facilitate repair. The molecular ratio of chimera to plasmid used in these experiments (200:1 in the tet system and 350:1 in the kan system) was at the upper end of the linear dose–response curve for chimera (data not shown) and gave a maximal frequency of conversion of ~0.1%. The tet^R colonies arose in a strain which lacks functional RecA protein, while the kan^R colonies arose in a strain which lacks a functional MutS protein.

Table 1. Correction of mutant antibiotic resistance genes

	Tet Substitutory System Plasmid	Chimera	HuH-7 extract	Tet^R colonies/ 10^5 amp^R colonies
1	pT^Sm153	Tet153T	1.5 µg	2.0
2	↓	↓	6 µg	39
3			12 µg	93
4			24 µg	99
5			-	0
6	↓	GC1	24 µg	0

	Kan Substitutory System Plasmid	Chimera	HuH-7 extract	Kan^Rcolonies/ 10^5 AmpR colonies
1	pK^Sm4021	Kan 4021C	1.5 µg	1.3
2	↓	↓	6 µg	75
3			12 µg	330
4			24 µg	280
5			-	0.6
6	↓	GC1	24 µg	0

Each reaction contained 1 µg of plasmid and either 1.4 µg of Tet153T, 1.2 µg of Kan4021C or 1.4 µg of GC1. Genetic readout utilized DH10B (recA1) cells in the tet system and BMH71-18 (mutS::Tn*10*) cells in the kan system. Colony counts reflect an average of five independent experiments.

A second control experiment was performed in which reaction mixtures were deficient in one or more components and tet^R or kan^R colonies were analyzed again (Table 2). The complete reactions resulted in the appearance of 120 $tet^R/10^5$ amp^R colonies and 270 $kan^R/10^5$ amp^R colonies, respectively. Again, incomplete reactions were devoid of repair activity. Reaction number 7 was a composite of reactions 2 and 3 in which the plasmid (reaction 2) or the chimera (reaction 3) was incubated separately with the extract. For reaction 7, the products of these reactions were combined after phenol extraction and ethanol precipitation and were electroporated into BMH71-18 or DH10B *E.coli*. This control guards against the possibility that either molecule is modified in the extract, enabling a reconstruction of the active repair structure in the bacteria. No colonies have been observed under these conditions.

Conversion at the DNA level was monitored by RFLP analysis of plasmid DNA from representative colonies following PCR amplification. For the $tet^S \rightarrow tet^R$ system, plasmid pT^Sm153 contains two diagnostic restriction sites. First, the targeted site (residue 153) is part of a *Bfa*I recognition sequence. DNA from uncorrected plasmid yields three fragments when cut with *Bfa*I

Table 2. Gene correction requires cell-free extract and chimeric oligonucleotides

	Tet Substitutory System Plasmid	Chimera	HuH-7 extract	Tet^R colonies/ 10^5 Amp^R colonies
1	pT^Sm153	Tet153T	30 µg	100
2	↓	-	↓	0
3	-	Tet153T		
4	pT^Sm153	↓	-	
5	↓	-	-	
6	-	Tet153T	-	
7	#2 + #3			↓

	Kan Substitutory System Plasmid	Chimera	HuH-7 extract	Kan^R colonies/ 10^5 Amp^R colonies
1	pK^Sm4021	Kan4021C	30 µg	270
2	↓	-	↓	0
3	-	Kan4021C		
4	pK^Sm4021	↓	-	
5	↓	-	-	
6	-	Kan4021C	-	
7	#2 + #3			↓

The reaction contained when indicated 1 µg of plasmid and either 0.7 µg of Tet153T or 1.4 µg of Kan4021C. Plasmid DNA was electroporated into DH10B cells in the tet system and BMH-71-18 cells in the kan system. For details of 7, see text.

while the corrected plasmid produces two fragments. Second, a silent mutation at position 325 introduces a *Tse*I recognition site in DNA from both mutant and corrected plasmids. As a result, the amplified DNA yields three fragments when treated with *Tse*I.

As shown in Figure 2A, RFLP analyses confirmed gene conversion at the targeted site only. Lanes 2 and 3 represent control digestions of uncorrected plasmid by *Bfa*I and *Tse*I, respectively, while lanes 4–13 are paired digestions of plasmid DNA isolated from tet^R colonies. Although *Bfa*I digestion patterns change as predicted (lanes 4, 6, 8, 10 and 12), the *Tse*I site has remained intact (lanes 5, 7, 9, 11 and 13). As described above, the presence of this *Tse*I site in plasmids harboring the *Bfa*I change eliminates the possibility of plasmid contamination. Correction of the targeted nucleotides in plasmid pK^Sm4021 has been similarly confirmed by RFLP analysis (Fig. 2B) as the conversion eliminates a *Bfa*I restriction site.

DNA sequence analysis of bacterial colonies tet^R or kan^R was used to confirm that the plasmids harboring altered RFLP patterns contained a corrected sequence. In Figure 3A, a representative clone from the experiment described in Figure 2 is presented at the DNA sequence level. The coding strand sequence surrounding position 153 is presented in the top panel. The highlighted nucleotide, T, has been converted from the mutant nucleotide A. As an internal control to guard against possible wild-type plasmid contamination, the sequence at position 325 is presented in Figure 3A and the silent mutation G remains unchanged. Similar results were obtained when kan^R colonies were sequenced. As shown in Figure 3B, the G·C base pair was converted to C·G. In both cases, >10 clones displaying an antibiotic-resistant cutting pattern by RFLP were picked randomly and each sample had the sequence alterations described above.

Several systems have now been developed using chimeric molecules as described for conversion of a single base pair (11–18). To expand the range of applications of chimera-directed

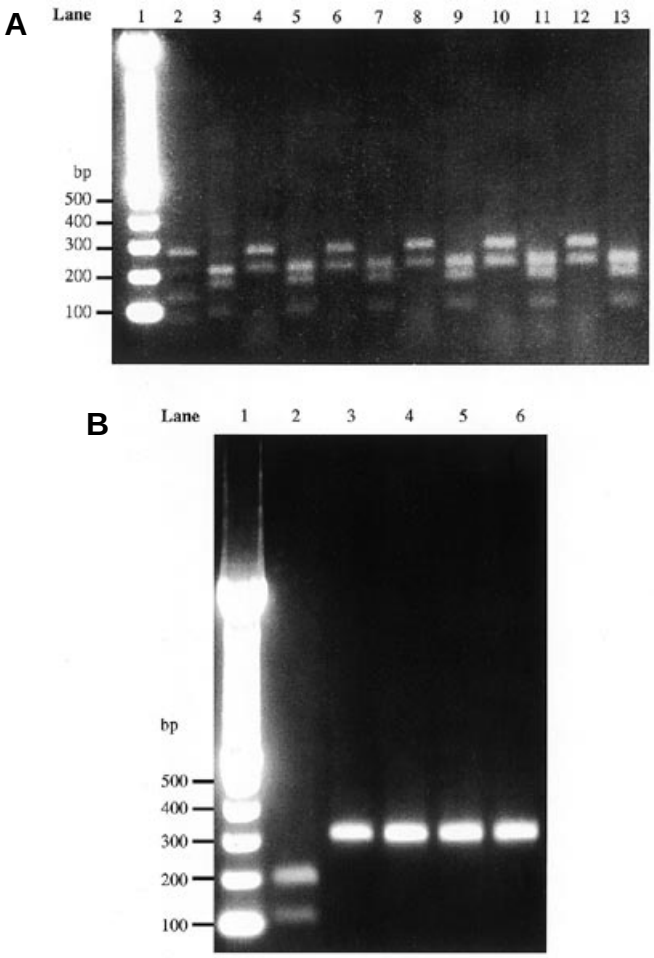

Figure 2. (**A**) Confirmation of tetracycline gene correction by RFLP analyses. Lane 1, 100 bp ladder. Lanes 2 and 3, 475 bp fragment PCR amplified from pT^{S}m153 plasmid and digested with *Bfa*I and *Tse*I, respectively. The tet^{S} *Bfa*I digestion pattern yields bands corresponding to 266, 130 and 79 bp. The *Tse*I silent mutation generates fragments of 205, 174 and 96 bp. Lanes 4 and 5, 475 bp fragment amplified from a corrected pT^{R}m153 plasmid digested with *Bfa*I and *Tse*I, respectively. The tet^{R} *Bfa*I digestion pattern yields fragments of 266 and 209 bp. The *Tse*I silent mutation is present and yields the characteristic fragments. Lanes 6–13, same as lanes 4 and 5. (**B**) Confirmation of kanamycin gene correction by RFLP analysis. Lane 1, 100 bp ladder; lane 2, 306 bp PCR fragment from pK^{S}m4021 plasmid digested with *Bfa*I generating fragments of 197 and 109 bp; lanes 3–6, 306 bp PCR fragments from corrected pK^{R}m4021 plasmid digested with *Bfa*I. Correction eliminates the cut site.

conversion, we attempted to repair a *deletion* in plasmid pT^{S}Δ208. This plasmid was engineered so that a one base deletion at position 208 (Fig. 4A) generates a termination codon and an *Sph*I restriction site. The chimeric oligonucleotide utilized here was similar in design to those of previous experiments except that it contained an extra residue in the central DNA region relative to the targeted plasmid sequence. It was predicted, therefore, to facilitate the insertion of a T at position 208, restoring an intact coding region. As seen in Table 3, an increasing number of tet-resistant clones were observed in a dose-dependent reaction. The reaction required all components being present in the mixture as before (data not shown) and sequence analysis confirmed that the targeted site of insertion was corrected in the tet^{R} clones (Fig. 4B). A T residue is placed between T207 and G209 in the coding region of the tet gene. No other alterations were observed in the bases immediately surrounding the targeted base (Fig. 4B) or the rest of the plasmid (data not shown). The frequency of insertion, however, is ~30-fold lower than direct nucleotide exchange, illustrated in Tables 1 and 2. In both cases, however, the same design of chimeric oligonucleotide catalyzes both point mutation and deletion mutation gene correction.

Table 3. Correction of a 1 bp deletion using chimeric oligonucleotides and cell-free extract

	Tet Deletion System Plasmid	Chimera	HuH-7 Extract	Tet^{R} colonies/ 10^{5} Amp^{R} colonies
1	pT^{S}Δ208	TetΔ208T	1 μg	0.0087
2	↓	↓	4 μg	0.62
3	↓	↓	8 μg	2.8
4	↓	↓	16 μg	2.4
5	↓	↓	-	0
6	↓	GC1	16 μg	0

Each reaction contained 1 μg of plasmid DNA and either 1.4 μg of TetΔ208T or GC1. DH10B cells were used for genetic analysis.

Cell-free extracts from HuH-7 cells were the source of the gene correction activity in this system. Since both single nucleotide change and targeted insertion occurred, we tested the hypothesis that mismatch repair was the pathway used in chimera-directed conversion. LoVo cells are defective in mismatch repair due to partial deletions in both copies of the hMSH2 gene (20). When cell-free extracts were prepared from LoVo cells, we observed only 13% of the activity compared with the HuH-7 extract in converting tet^{S}→tet^{R} plasmids (Table 4). Consistent with this extract lacking a critical protein and not containing a block to repair, addition of LoVo extract to HuH-7 extract did not inhibit HuH-7-mediated repair (Table 4). These results suggest that chimera-directed conversion proceeds through a step involving mismatch recognition. As an additional control, we tested whether anti-hMSH2 antibodies could inhibit activity. The inclusion of this antibody in the complete reaction mixture (plasmid, chimeric oligonucleotide and HuH-7 cell-free extract) reduced the activity by two thirds whereas normal rabbit IgG was only slightly inhibitory. Taken together, these data imply that hMSH2 is an important component of the chimera-directed gene conversion process.

DISCUSSION

The use of various types of oligonucleotides to ablate or correct dysfunctional gene expression (21) has recently gained support through studies in mammalian cell culture and animals. In some cases, correction of a defect is facilitated at the level of the RNA using antisense molecules (22) or ribozymes (23). Such strategies are important advances but do not accomplish inheritable changes. RNA/DNA chimeric oligonucleotides that act at the DNA level and direct base changes in the genome aim to provide permanent genetic correction (11–18). The frequency of conversion obtained by our group and others is several logs higher than previously reported for gene targeting protocols including those using single-stranded vectors (24) or short single-stranded or double-stranded DNA fragments (25). These other methodologies,

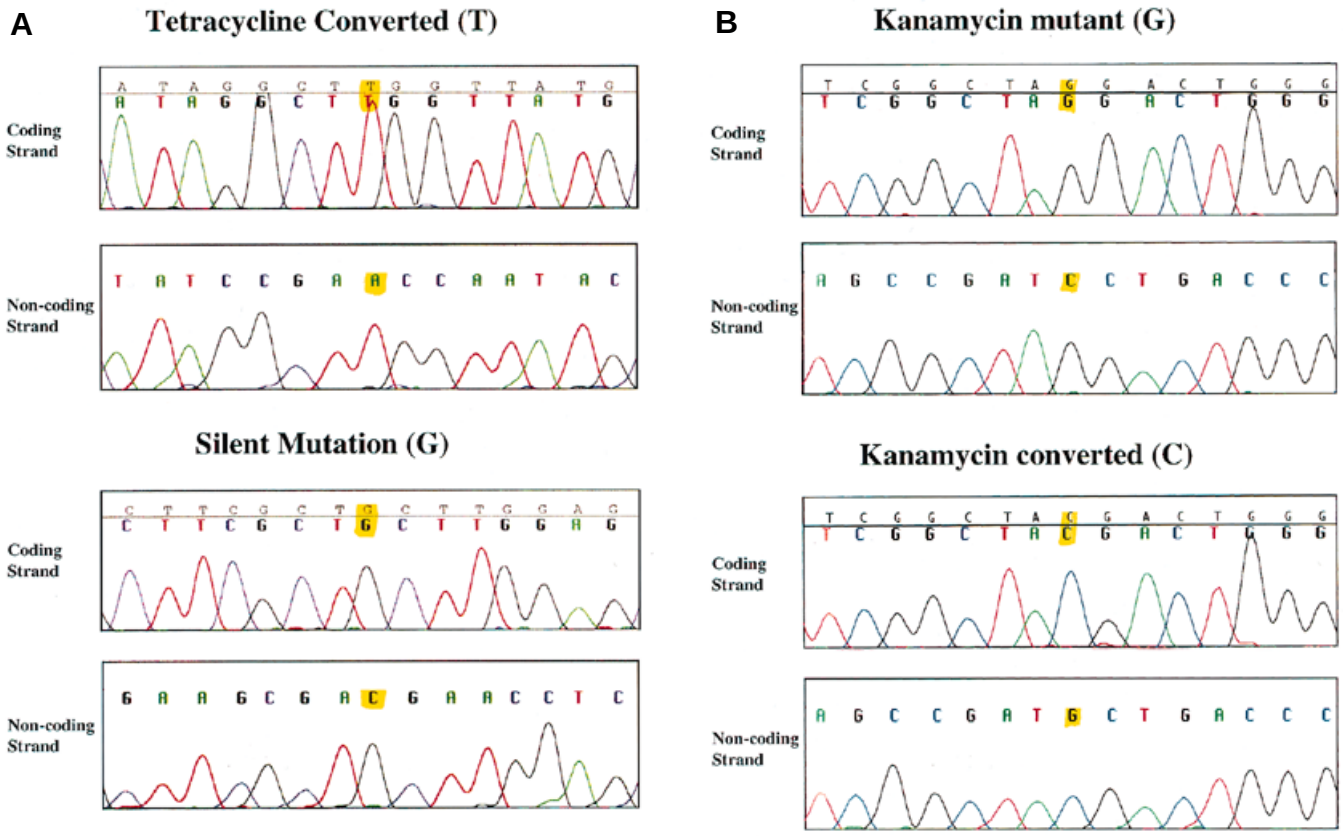

Figure 3. Confirmation of gene correction by DNA sequence analysis. The coding and non-coding strand sequences of corrected antibiotic resistance genes in plasmids $pT^{R}m153$ (**A**) and $pK^{R}m4021$ (**B**) were determined by automatic sequence analysis of PCR amplified fragments. In each panel the corrected base is highlighted as is the silent mutation in the tetracycline system. Sequences of non-coding strands confirmed the base assignments and are presented below the coding strand sequence.

Table 4. hMSH2 may be required for chimera-directed gene correction

	Tet Substitutory System Plasmid	Chimera	Extract	Antibody	% Conversion
1	$pT^{S}m153$	Tet153T	HuH-7	-	100
2	↓	↓	↓	IgG	93±19
3	↓	↓	↓	α-hMSH2	39±4
4	↓	↓	LoVo	-	13
5	↓	↓	HuH-7/LoVo	-	119

Each reaction contained 1 µg of plasmid, 1.4 µg of chimera and 30 µg of extract except for the reaction described in lane 5. Here, 15 µg of HuH-7 and 10 µg of LoVo cell-free extracts were used. All other components are as listed. Genetic readout in DH10B cells was used to calculate conversion frequency. The 100% result reflects 200 tet^{R} colonies/10^{5} amp colonies.

however, rely primarily on homologous recombination pathways which are known to be inefficient in mammalian cells. The discrepancy in frequency levels has led to speculation about the nature of the correction events observed in our system.

The results reported here demonstrate that the chimeric oligonucleotide uses the process of DNA repair to fix a point mutation or a single base deletion. The repair of plasmid DNA occurs in a dose-dependent fashion with respect to increasing levels of extract. Unlike other *in vitro* mismatch repair assays (26,27), the template does not need to be pre-formed; as in this system, the plasmid and chimera are added simultaneously to a reaction mixture containing the cell-free extract. With confirmation of base changes at the DNA level by RFLP analyses and DNA sequencing, the development of such a system has helped address questions of phenotypic response and genetic inheritance.

To avoid the possibility that wild-type plasmid DNA contaminated the reaction mixtures, safeguards were engineered into both the kan and tet systems. In the kan system, the mutant G base is converted to a C, making a conservative change from the wild-type T. Sequencing data confirm that only C residues are observed in kan-resistant bacterial colonies. In the tet system, a silent mutation at position 325 was engineered so that plasmid DNA isolated from tet^{R} colonies is cleaved by the restriction enzyme *Tse*I. In all cases, where the targeted base was changed, the *Tse*I site remained intact. The insertion of a single base in $pT^{S}\Delta 208$ was also observed in the same cell-free extract, albeit to a lower extent. The reduction in the rescue frequency was ~30-fold, but in each resistant colony, the insertion was at the exact position expected and no other alterations within the targeted plasmid were observed.

Mammalian cells have been shown previously to catalyze targeted mutagenesis of a plasmid substrate. In one case, triplex-forming oligonucleotides (TFOs) linked to psoralen (28) were used to mutagenize a sequence in the *SupF* gene of a shuttle

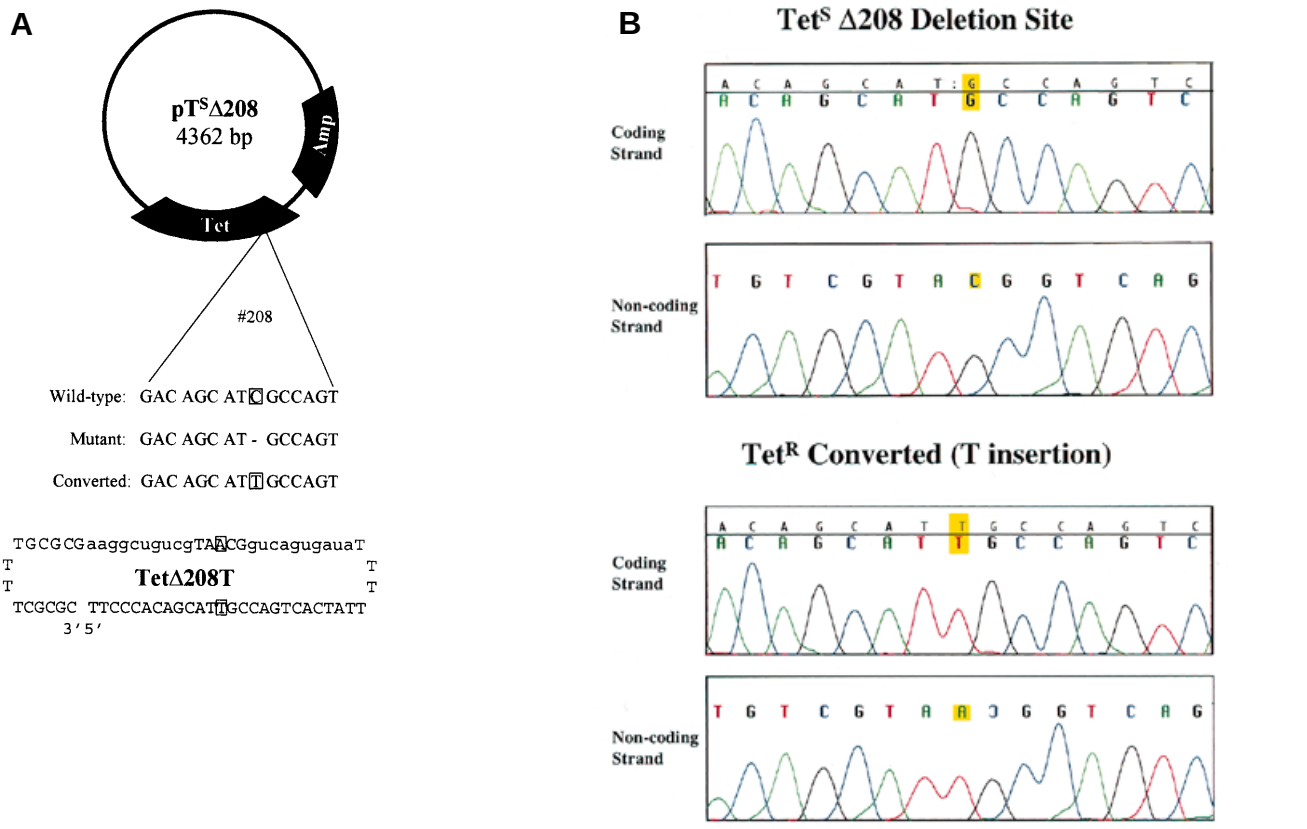

Figure 4. (**A**) Tet^{S}Δ208 deletion plasmid (pT^{S}Δ208) and targeting chimera. Model system for detecting the single base insertion mediated by chimeric oligonucleotide TetΔ208T. Note that the chimera is designed to insert a T residue at the deletion site rather than the wild-type C nucleotide. (**B**) The coding and non-coding strand sequence of the mutated and repaired Δ208 site. DNA sequences were determined by automated sequence analyses of PCR amplified fragments from the appropriate sites within plasmids pT^{S}Δ208 and pT^{R}Δ208.

vector. Analysis of the vector in *E.coli* showed that specific targeting events occurred up to a frequency of 2%. Wang *et al.* (29) extended these observations by showing that unmodified TFOs could also induce targeted mutagenesis with frequencies ranging from 0.2 to 0.3%. Using a variety of extracts from cells, such as xeroderma pigmentosum or Cockayne's syndrome group B, these workers demonstrated that excision repair was a necessary component of TFO-directed mutagenesis.

Extracts from mammalian cells and other sources have also been used extensively to study the biochemical mechanism of DNA mismatch repair (see 30 for review). Some differences in correction efficiency as a function of the type of mismatched base pair have been observed but, for the most part, nick-directed mismatch repair appears mutation-independent (31). The correction is dependent, however, on homologs of MutS and MutL (20,31). Furthermore, a new fragment of between 90 and 170 nt in length surrounding the region of damaged DNA is generated by DNA synthesis. Hence, mismatch repair requires multiple steps that involve recognition, incision, removal and re-synthesis of the affected base pair. Should any of the steps be blocked, the overall process is largely disabled. As an example, certain inhibitors such as aphidicolin are known to block the action of DNA polymerase δ, an enzyme required for mismatch repair synthesis in human cells (32). Although aphidicolin did not inhibit gene correction when added to the cell-free extract, it is possible that repair intermediates which accumulated in its presence were processed after introduction into *E.coli*. The involvement of hMSH2 in the correction process seems likely based on the low conversion rate catalyzed by the LoVo cell-free extract and inhibition of repair by anti-hMSH2 antibodies in the HuH-7 extract.

Although elucidation of the mechanism is far from complete, we envision the following pathway. The first step is likely to be the pairing of the chimera with its plasmid target, a process most likely governed by a search for homology through the action of DNA pairing enzymes or complexes. The pairing of DNA molecules is a recombination function dependent, for example, on the RecA protein in *E.coli*, and R.Metz (personal communication) demonstrated recently that chimera-directed repair in *E.coli* is, in fact, dependent on functional RecA protein. In mammalian cells, however, it is still not clear which enzyme(s) is providing the pairing activity, but preliminary data suggest that both strands of the chimera hybridize to the two strands of the helix. The repair machinery would subsequently encounter an unusual structure at the targeted site, perhaps consisting of a complement-stabilized or double D-loop (33,34; H. Gamper, unpublished observations). The repair process may initiate correction based on the identification of a chimera/plasmid complex bearing unusual structural features. The DNA-rich A-form duplex of the chimera (35) is a structure not normally seen in the cell, providing more support to the notion of structure activating function. The mismatched base pair is then

exchanged for the correct one using the sequence of the chimera as template.

Currently, a variety of structural changes in the standard chimera are being made in an effort to improve the frequency of correction. One such change, removing the intervening DNA stretch in the molecule, appears to achieve this goal. In the original concept of targeted repair by chimeric oligonucleotides, we envisioned the RNA regions as stabilizers of the reaction intermediate and it is possible that an all RNA strand facilitates this binding even more. Furthermore, this data and preliminary results using other structural modifications (H.Gamper *et al*., in preparation) lead us to believe that the initial correction event is directed by the strand *lacking* RNA residues.

The choice of the transformed HUH-7 liver cell line as a source of the extracts was predicated by the highly successful chimera-based targeting reported for liver cells (14,15). The frequency of conversion is quite high ranging from 20 to 40%. Unfortunately, we have not been able to achieve levels >0.2% reproducibly in the cell-free extract system. Hence we are unable to unequivocally state that the mechanism that we have uncovered is operational in liver cells, but are satisfied that this mechanism can explain gene repair in other cell types. Among the many possible reasons for such a large discrepancy is that a complex of proteins, crucial to the gene repair process, is rendered dysfunctional during the preparation of the cell-free extract. Alternatively, the pairing reaction proposed above as the first step in gene repair may be less efficient *in vitro*. Furthermore, a significant difference in the frequency of repair of point mutations versus deletions exists supporting the observations reported by Ye *et al*. (36). It is clear that, in our hands, the repair of a deletion mediated by the chimeric oligonucleotide is at least an order of magnitude less than point mutation correction even in cultured cells (S.Ye *et al*., in preparation).

The present work establishes a biochemical system for studying chimera-directed gene repair. The reaction involves the hMSH2 protein and can facilitate point mutation and single base insertion corrections. It is important to note that the *in vitro* system itself may be useful as a research tool for site-directed mutagenesis studies. These results also provide a biochemical basis for the gene repair observations in mammalian cells and enable detailed mechanistic studies of this process in a controlled environment as well as improving the frequency for applications such as functional genomics and pharmacogenetics.

ACKNOWLEDGEMENTS

We thank members of the Kmiec laboratory for their help during the course of this work. We are also grateful to Drs Ramesh Kumar and Rick Metz (Kimeragen Inc., Newtown, PA) for developing the bacterial readout system. We thank our colleagues in the repair community for critical reading of the manuscript. This work was supported by a NIH grant (RO1 HL58563–01A1) and by funds from Kimeragen Inc.

REFERENCES

1 Yanez,R.J. and Porter,A.G. (1998) *Gene Ther.*, **5**, 149–159.
2 Smithies,O., Gregg,R.G., Boggs,S.S., Koralewski,M.A. and Kucherlapati,R.S. (1985) *Nature*, **317**, 230–234.
3 Smithies,O., Koralewski,M.A., Song,K.Y. and Kucherlapati,R.S. (1984) *Cold Spring Harbor Symp. Quant. Biol.*, **49**, 161–170.
4 Bethke,B.D. and Sauer,B. (1998) In *Gene Targeting Protocols*. Humana Press, Totowa, NJ, in press.
5 Smyth-Templeton,N. (1998) In *Gene Targeting Protocols*. Humana Press, Totowa, NJ, in press.
6 Yuan,Z.M., Huang,Y., Ishiko,T., Nakada,S., Utsugisawa,T., Kharbanda,S., Wang,R., Sung,P., Shinohara,A., Weichselbaum,R. and Kufe,D. (1998) *J. Biol. Chem.*, **273**, 3799–3802.
7 Ferguson,D.O., Rice,M.C., Rendi,M.H., Kotani,H., Kmiec,E.B. and Holloman,W.K. (1997) *Genetics*, **145**, 243–252.
8 Kotani,H., Germann,M.W., Andrus,A., Vinayak,R., Mullah,B. and Kmiec,E.B. (1996) *Mol. Gen. Genet.*, **250**, 626–634.
9 Kotani,H. and Kmiec,E.B. (1994) *Mol. Cell. Biol.*, **14**, 6097–6106.
10 Kmiec,E.B., Cole,A. and Holloman,W.K. (1994) *Mol. Cell. Biol.*, **14**, 7163–7172.
11 Yoon,K., Cole-Strauss,A. and Kmiec,E.B. (1996) *Proc. Natl Acad. Sci. USA*, **93**, 2071–2076.
12 Bartlett,R.J., Denis,M., Inverardi,L., Tremblay,J., Bartlett,W., Marques,M., Sharma,K. and Ricordi,C. (1998) In *Proceedings from the 1998 Miami Nature Biotechnology Winter Symposium*. Nature Biology Short Reports, Miami, FL, pp. 163–164.
13 Cole-Strauss,A., Yoon,K., Xiang,Y., Byrne,B.C., Rice,M.C., Gryn,J. and Kmiec,E.B. (1996) *Science*, **273**, 1386–1389.
14 Kren,B.T., Bandyopadhyay,P. and Steer,C.J. (1998) *Nature Med.*, **4**, 1–6.
15 Kren,B.T., Cole-Strauss,A., Kmiec,E.B. and Steer,C.J. (1997) *Hepatology*, **25**, 1462–1468.
16 Lai,L.W., O'Connor,H.M. and Lien,Y.H. (1998) In *Conference Proceedings, 1st Annual Meeting of the American Society of Gene Therapy*. American Society of Gene Therapy, Seattle, WA, p. 183a.
17 Santana,E., Peritz,A.E., Iyer,S., Uitto,J. and Yoon,K. (1998) *J. Invest. Dermatol.*, **111**, 1172–1177.
18 Alexeev,V. and Yoon,K. (1998) *Nature Biotechnol.*, **16**, 1343–1346.
19 Li,J.J. and Kelly,T.J. (1985) *Mol. Cell. Biol.*, **5**, 1238–1246.
20 Umar,A., Boyer,J.C., Thomas,D.C., Nguyen,D.C., Risinger,J.I., Boyd,J., Ionov,Y., Perucho,M. and Kunkel,T.A. (1994) *J. Biol. Chem.*, **269**, 14367–14370.
21 Woolf,T.M. (1998) *Nature Biotechnol.*, **16**, 341–344.
22 Kole,R. (1997) *Acta Biochim. Pol.*, **44**, 231–237.
23 Jones,J.T. and Sullenger,B.A. (1997) *Nature Biotechnol.*, **15**, 902–905.
24 Fujioka,K., Aratani,Y., Kusano,K. and Koyama,H. (1993) *Nucleic Acids Res.*, **21**, 407–412.
25 Hunger-Bertling,K., Harrer,P. and Bertling,W. (1990) *Mol. Cell. Biochem.*, **92**, 107–116.
26 Holmes,J.J., Clark,S. and Modrich,P. (1990) *Proc. Natl Acad. Sci. USA*, **87**, 5837–5841.
27 Thomas,D.C., Roberts,J.D. and Kunkel,T.A. (1991) *J. Biol. Chem.*, **266**, 3744–3751.
28 Wang,G., Levy,D.D., Seidman,M.M. and Glazer,P.M. (1995) *Mol. Cell. Biol.*, **15**, 1759–1768.
29 Wang,G., Seidman,M.M. and Glazer,P.M. (1996) *Science*, **271**, 802–805.
30 Modrich,P. and Lahue,R. (1996) *Annu. Rev. Biochem.*, **65**, 101–133.
31 Kunkel,T.A. (1985) *J. Biol. Chem.*, **260**, 5787–5796.
32 Papadopoulos,N., Nicolaides,N.C., Wei,Y.F., Ruben,S.M., Carter,K.C., Rosen,C.A., Haseltine,W.A., Fleischmann,R.D., Fraser,C.M. and Adams,M.D. (1994) *Science*, **263**, 1625–1629.
33 Jayasena,V.K. and Johnston,B.H. (1993) *J. Mol. Biol.*, **230**, 1015–1024.
34 Sena,E.P. and Zarling,D.A. (1993) *Nature Genet.*, **3**, 365–372.
35 Andrus,A., Kuimelis,R., Vinayak,R. and Mullah,B. (1997) *Nucleosides Nucleotides*, **16**, 1183–1194.
36 Ye,S., Cole-Strauss,A., Frank,B. and Kmiec,E.B. (1998) *Mol. Med. Today*, **4**, 431–437.