Targeted Correction of an Episomal Gene in Mammalian Cells by a Short DNA Fragment Tethered to a Triplex-forming Oligonucleotide*

(Received for publication, November 25, 1998, and in revised form, January 27, 1999)

Phillip P. Chan‡§, Michael Lin¶, A. Fawad Faruqi‡, James Powell‡, Michael M. Seidman¶, and Peter M. Glazer¶

From the §Departments of Therapeutic Radiology and Genetics, Yale University School of Medicine, New Haven, Connecticut 06520-8040 and the ¶National Institute on Aging, National Institutes of Health, Baltimore, Maryland 21224-6825

Triplex-forming oligonucleotides (TFOs) can bind to polypurine/polypyrimidine regions in DNA in a sequence-specific manner and provoke DNA repair. We have coupled a TFO to a short donor fragment of DNA that shares homology to a selected gene as a strategy to mediate gene targeting and correction. In this bifunctional oligonucleotide, the TFO domain is designed to bind the target gene and stimulate repair and recombination, with the donor domain positioned for recombination and information transfer. A series of these tethered donor-TFO (TD-TFO) molecules with donor domains of 40–44 nucleotides and TFO domains in both the purine and pyrimidine triplex motifs were tested for their ability to mediate either gene correction or mutation of a supF reporter gene contained in a SV40 shuttle vector in mammalian cells. In vitro binding assays revealed that the attachment of the donor domain via a flexible linker did not significantly alter the binding affinity of the TFO domain for the polypurine site in the supF target DNA, with equilibrium dissociation constants in the 10−8 m range. Experiments in which the target vector and the linked TD-TFOs were pre-incubated in vitro and co-transfected into cells led to conversion frequencies approaching 1%, 4-fold greater than with the two domains unlinked. When cells that had been previously transfected with the SV40 vector were electroporated with the TD-TFOs, frequencies of base pair-specific gene correction were seen in the range of 0.04%, up to 50-fold over background and at least 3-fold over either domain alone or in unlinked combinations. Sequence conversion by the TD-TFOs was achieved using either single- or double-stranded donor domains and either triplex motif. Substitution of either domain in the TD-TFO with control sequences yielded reagents with diminished activity, as did mixtures of unlinked TFO and donor DNA segments. The boost in activity provided by the attached TFO domain was reduced in cells deficient in the nucleotide excision repair factor XPA but was restored in a subclone of these cells expressing XPA cDNA, suggesting a role for nucleotide excision repair in the pathway of triple helix-stimulated gene conversion.

The ability to correct or mutate a specific target site in mammalian cells using the TD-TFO strategy may provide a useful tool for research and possibly for therapeutic applications.

Efficient methods for site-directed genome modification are desirable for research and possibly for gene therapy applications. One approach utilizes triplex-forming oligonucleotides (TFOs), which bind as third strands to duplex DNA in a sequence-specific manner, to mediate directed mutagenesis. Such TFOs can act either by delivering a tethered mutagen, such as psoralen or chlorambucil (1–5), or by binding with sufficient affinity to provoke error-prone repair (6).

Another strategy for genome modification involves the induction of homologous recombination between an exogenous DNA fragment and the targeted gene. This approach has been used successfully to target and disrupt selected genes in mammalian cells and has enabled the production of transgenic mice carrying specific gene knockouts (7). This approach, however, relies on the transfer of selectable markers to allow isolation of the desired recombinants. Without selection, the ratio of homologous to nonhomologous integration of transfected DNA in typical gene transfer experiments is low, usually in the range of 1:1000 or less (8). This low efficiency of homologous integration limits the utility of gene transfer for gene therapy.

The frequency of homologous recombination can be enhanced by damage to the target site from UV irradiation and selected carcinogens (9) as well as by site-specific endonucleases (8, 11, 12). We and others have also demonstrated that DNA damage induced by triplex-directed psoralen photoadducts can stimulate recombination within and between extrachromosomal vectors (12, 13).

Other work has helped to define parameters that influence recombination in mammalian cells. In general, linear donor fragments are more recombinogenic than their circular counterparts (14). Recombination is also influenced by the length of uninterrupted homology between both the donor and target sites, with short fragments appearing to be inefficient substrates for recombination (15). Nonetheless, several recent efforts have focused on the use of short fragments of DNA or DNA/RNA hybrids for gene correction (10, 16).

A direct method of targeting donor DNA to the gene of interest is lacking in many of the above recombination strategies. The process by which a DNA fragment finds its site of...
homology in mammalian cells is not fully understood, but it is thought that recA-like factors (such as Rad51) catalyze a homology search and mediate DNA pairing in an intricate, energy-dependent manner. In contrast, a TFO can find its cognate site within complex DNA without the need for any associated enzyme activity (5, 17). Indeed, in previous work, we found that a TFO could find and modify its target site within mouse genomic DNA in vitro within minutes (18).

The sequence-specific binding properties of TFOs have been used to deliver a series of different molecules to target sites in DNA. For example, a diagnostic method for examining triplex interactions utilized TFOs coupled to Fe-EDTA, a DNA cleaving agent (19). Others have linked biologically active enzymes like micrococcal nuclease and streptococcal nuclease to TFOs and demonstrated site-specific cleavage of DNA (20, 21). We and others have previously shown that site-directed DNA damage and mutagenesis can be achieved using TFOs conjugated to either psoralen (1, 4) or alkylating agents (5, 22).

Because TFOs can efficiently deliver reactive molecules to specific sites within target genes, we asked whether a TFO could also be used to guide a homologous donor DNA fragment to its intended target site and position it for efficient information transfer via recombination and/or gene conversion. Thus, we designed an oligonucleotide that covalently tethers a TFO to a DNA donor fragment. In addition, because triplex formation itself can provoke repair (6) activity, we hypothesized that the TFO, in addition to positioning the donor fragment, may also enhance strand transfer and recombination by inducing repair activity at the target site.

Here we describe the design of several TD-TFO reagents. These bind specifically and with high affinity to DNA via a triplex-forming domain while simultaneously providing DNA sequence information through an attached donor domain to revert or induce a mutation in a target gene (Fig. 1). Using a SV40-based shuttle vector assay, we demonstrate that these TD-TFOs can mediate specific and directed sequence changes within an extrachromosomal supF reporter gene in mammalian cells. Furthermore, we show that the intact TD-TFO molecule is more active than either of its parts alone or unlinked combinations thereof and is also more active than control oligonucleotides that substitute either domain with a control sequence. Using human mutant cell lines, we also demonstrate that the enhanced activity of the combined TD-TFO is diminished in cells deficient in nucleotide excision repair (NER), suggesting that the TFO domain may stimulate gene conversion in part through the ability of triplex helices to provoke repair.

**EXPERIMENTAL PROCEDURES**

**Oligonucleotides—**Oligonucleotides were synthesized by standard phosphoramidite chemistry using materials from Glen Research (Sterling, VA) and purified by either gel electrophoresis or high pressure liquid chromatography, followed by Centricron-3 filtration in distilled water (Amicon, Beverly, MA). The oligonucleotides consisted primarily of phosphodiester linkages but were modified at the 3′ end to resist 3′ exonuclease activity by the inclusion of phosphorothioate linkages at the terminal three residues. Exceptions were the oligonucleotides A, B, B(144), B(Sal), and Y(Rndm), which included a 3′ propylamine (Glen Research) as their only modification. The linker segment between the donor fragment and the TFO domain consisted of the sequence 9TT9TT9, in which 9 indicates a 9-carbon polyethylene glycol linker (Spacer 9; Glen Research). The pyrimidine oligonucleotide UC30 was composed of the RNA nucleosides 2′-O-methyluridine and 5-methyl-2′-O-methylcytidine.

**Vectors—**The mammalian shuttle vectors pSupFG1 and pSupFG1/G144C (Fig. 2A), which contain a wild-type supFG1 gene or a supFG1 gene with an inactivating G-to-C point mutation at position 144, respectively, were prepared by large-scale DNA preparation (Qiagen, Santa Clarita, CA). The pSupFG1 plasmid has been described previously (3). The mutant pSupFG1/G144C plasmid was isolated in the course of mutagenesis experiments with pSupFG1 (3).

The shuttle vector pSupFG5/G115T was used in experiments involving TD-TFOs in the pyrimidine triplex motif. This vector is distinct from pSupFG1; it contains a modified supFG gene in which a 30-bp A-rich site amenable to triplex formation in the pyrimidine motif has been placed approximately 5 bp to the supFG gene (Fig. 2D). The gene also contains an inactivating G-to-T point mutation at position 115. 

**RESULTS**

**Design and Rationale of the TD-TFO Strategy—**In an endeavor to promote targeted recombination, we have designed a TD-TFO molecule that tethers a TFO to a donor DNA fragment homologous to a region of the target gene via a linker segment.
Triplex-directed Gene Correction

![Figure 1](image1.png)

**Fig. 1.** Schematic diagram of a TD-TFO binding to DNA via triplex formation, delivering donor sequence information to a desired region and leading to the change of the original DNA sequence. X, the base pair to be changed.

(Fig. 1). This arrangement facilitates target site recognition via triplex formation while positioning the donor fragment for possible recombination and information transfer. This strategy also attempts to exploit the ability of the triple helix itself to provoke DNA repair at the binding site, potentially increasing the probability of recombination with the tethered donor DNA.

The donor domain is identical in sequence to the target gene, except for a desired sequence change that may either restore or abolish gene activity. In designing the tethered donor, we hypothesized that either a single- or double-stranded donor could potentially transfer sequence information to the target gene through either direct assimilation into DNA via recombination or as a template for replicative or reparative DNA synthesis in a gene conversion-type event.

A series of TD-TFOs were constructed to target two different modified supF tRNA reporter genes. The first target gene, supFG1, contains a 30-bp G-rich site at the 3′ end of the gene (Fig. 2A) to which the purine-rich 30-mer TFO, AG30, binds and forms a triple helix in the anti-parallel motif (3). The other target gene, supF5/G115T, contains a 30-bp A-rich site at the 5′ end of the gene suitable for triplex formation in the parallel pyrimidine motif by the TFO designated UC30 (Fig. 3B).

For each target, a series of experimental and control oligonucleotides or oligonucleotide mixtures were designed (Fig. 3). These include TD-TFOs with either single- or double-stranded donor domains, TD-TFOs with either domain substituted with control sequences to prevent function, and lastly, separate TFO and donor domains, either alone or in unlinked combinations.

For each TD-TFO, a 40- or 44-nt donor domain was connected to a 30-nt triplex-forming domain by a mixed linker and donor domains, either alone or in unlinked combinations.

For each TD-TFO, a 40- or 44-nt donor domain was connected to a 30-nt triplex-forming domain by a mixed linker equivalent to approximately 13 nt. Our nomenclature describes each TD-TFO by its donor and TFO domains written in the 5′ to 3′ direction. For TD-TFOs targeting the supFG1 gene, the donor domains are referred to as A (corresponding, from 5′ to 3′, to nt 121–160 of the supFG1 gene) and its complement, B (nt 160–121 of supFG1). The notation also indicates whether the donor region contains any sequence changes. For example, A/B(144)-AG30 describes a double-stranded donor fragment containing a sequence change at position 144, along with AG30 as the TFO. In targeting the supFG1 gene, C (nt 100–143 of supF5) and its complement, D (nt 143–100 of supF5), represent the individual strands of the donor domain.

**Triplex Formation Is Not Altered by the Presence of an Attached Donor Domain**—Cheng and Van Dyke (24) reported the destabilizing effect of a 3′ or 5′ nucleotide tail on triplex formation in certain sequence contexts, with a 3′ tail being most detrimental. We were therefore concerned that an attached donor fragment might substantially destabilize TFO binding. However, unlike the oligonucleotides tested in the Cheng and Van Dyke study, the TD-TFO molecules were designed to incorporate a highly flexible linker region between the TFO and the donor tail, with the linker and donor segments extending from the 5′ end of the TFO (AG30). To determine the influence of the attached linker and donor domains on the affinity of third-strand binding by the TFO region, the binding of selected oligonucleotides to the target sequence in the supFG1 gene was measured by gel mobility shift assay (Fig. 4). As shown, the single-stranded purine TD-TFOs, A-AG30 and A(144)-AG30, bind with high affinity to a 57-bp duplex containing the poly-purine target site in the supFG1 gene. Calculation of the equilibrium dissociation constants (Kd values) for third-strand binding from the quantified band intensities revealed only a very small difference in binding affinity between A-AG30 (Kd = 5 × 10−8 M) and AG30 alone (Kd = 3 × 10−8 M). These measurements demonstrate that a single-stranded donor fragment attached via a flexible linker does not significantly alter the binding affinity of the 30-nt AG30 TFO domain to the target site. Similar results were observed using double-stranded TD-TFOs (data not shown). Further gel shift analyses also show, as expected, that oligonucleotides lacking the specific TFO domain, AG30, do not bind to the duplex target site (A-MX30 (Fig. 4) or A and A/B (data not shown)). In the pyrimidine motif, some decrease in the binding affinity of C-UC30 was seen as compared with UC30, although the Kd values were still within the 10−8 M range (see the Kd values given in Fig. 3).

**Reversion of a Point Mutation at Position 144 in the supFG1 Gene**—The single-stranded TD-TFO, A-AG30, and its double-
stranded counterpart, A/B-AG30, were designed to correct a single point mutation (G-to-C point mutation at position 144) in the supFG1 reporter gene of the shuttle vector, pSupFG1/G144C. Within the TD-TFO, position 144 corresponds to the center of the 40-nt donor domain, which is otherwise homologous to bp 121–160 of the gene (Fig. 2A). Shuttle vector DNA was incubated in vitro with selected oligonucleotides, transfected into COS-7 cells, and recovered after 48 h for analysis of the supFG1 gene in indicator bacteria. Reversion frequencies (Fig. 5A) were calculated as the number of blue colonies (revertants) among white ones. The presence of a corrected wild-type supFG1 sequence was confirmed by DNA sequencing in randomly selected blue colonies (n = 15 for A-AG30; n = 15 for A/B-AG30).

The data show that the frequencies of reversion due to the single-stranded A-AG30 (0.17%) and the double-stranded A/B-AG30 (0.68%) are significantly above the frequency of spontaneous reversion (0.003%), representing 57-fold and 227-fold increases, respectively (Fig. 5A). In comparison, lower levels of reversion were produced by the single- and double-stranded donor segments alone in the absence of an associated TFO (A, 0.051%; B, 0.062%; A/B, 0.17%), and there is essentially no reversion above background mediated by the TFO domain alone (AG30, 0.014%). Furthermore, A/B-MX30, which has the same donor domain as A/B-AG30 but cannot form a triplex with supFG1/G144C due to substitution of AG30 with a non-triplex-forming mixed sequence (MX30), also showed a lower reversion frequency (0.15%) than A/B-AG30 and almost the same reversion frequency as the A/B donor alone. As another control, X/Y(Rndm)-AG30, whose donor domain is replaced with a randomized sequence unrelated to supF, also demonstrates little activity (<0.006%). Additionally, to rule out the potential for bacterially derived recombinants, 210 pmol of A-AG30 or A/B-AG30 were electroporated directly into bacteria along with supFG1/G144C DNA recovered from COS-7 cells. These experiments showed reversion frequencies that were not significantly above background (background, 0.003±0.003%; A-AG30, 0.009±0.006%; A/B-AG30, 0.008±0.005%). Also, when pSupFG1/G144C plasmid DNA prepared in Escherichia coli was incubated with a 233-fold excess of either A/B-AG30 or A-AG30 and electroporated directly into bacteria, reversion frequencies of <0.003% (0 blue colonies/39,156 colonies) and <0.002% (0 blue colonies/44,859 colonies) were seen, respectively.

These results demonstrate that both the TD-TFO composite molecules and the corresponding unlinked donor fragments are active in the assay. However, whereas the single- and double-stranded donor fragments each yield revertants on their own,
Triplex-directed Gene Correction

Introduction of Specific Mutations into Wild-type supFG1—
We further tested the activity of TD-TFO molecules by assaying their ability to introduce a variety of sequence changes into a wild-type supFG1 gene in a forward mutation assay. One set of TD-TFOs was designed to introduce a specific inactivating G-to-C point mutation at position 144 in the wild-type supFG1 gene. Using the protocol described above, we found that A/B(144)-AG30 induced a mutation frequency of 0.64% (Fig. 5B), approximately 9-fold higher than the spontaneous forward mutation frequency of 0.07% and also higher than that of the single-stranded A(144)-AG30 (0.30%). (Note that the background frequency in this forward mutation assay is much higher than that seen in the reversion assay because many nonspecific sequence changes can inactivate the gene.) Furthermore, A/B(144)-AG30 generated a higher mutation frequency than any of the components alone (AG30, 0.07%; A/B(144), 0.15%; A/B(144) + AG30, 0.31%). Whereas the majority of spontaneous mutations in the absence of oligonucleotide exposure were deletions upon sequencing, the mutants induced by A/B(144)-AG30 carried the specific 144 G-to-C base change in 79% of the mutants tested (n = 15). Random deletions and scattered point mutations were found in 20%, similar to the type and frequency of mutations produced by AG30 alone (data not shown) and consistent with our previous observation of triplex-induced mutagenesis (6). The remaining mutation was a short tandem duplication (11 bp; nt 121–131) corresponding to the 5’ end of the donor domain, suggesting a possible abortive strand-exchange event.

To determine whether a more complex mutation could be reproducibly introduced into the target gene by this strategy, we also designed a TD-TFO to insert a SalI restriction site into the supFG1 gene. This requires changing 5 of 7 bases from position 138 to 144. With this reagent, A/B(Sal)-AG30, we observed a mutation frequency of 0.40% (versus 0.07% background and 0.11% for A/B(Sal); Fig. 5B). Of six randomly picked colonies, five were found by restriction analysis and DNA sequencing to have incorporated the SalI site; the sixth was a deletion mutant, again consistent with the presence of some background-related or AG30-induced mutants in the sample. The introduction of a complex, novel mutation such as this expands the potential utility of the TD-TFO approach. It also provides evidence that the results cannot be attributed to
simple contamination, because the SalI site-containing plasmid did not previously exist in the laboratory.

Dependence of the Activity of the TD-TFOs on Nucleotide Excision Repair—In previous work, we had found that triplex formation can induce mutagenesis in a pathway dependent in part on NER (6). In addition, in vitro studies in human cell extracts revealed that high affinity intermolecular triple helices could provoke DNA repair of an otherwise undamaged substrate as measured by the stimulation of DNA repair synthesis (6). Based on these observations, we hypothesized that triplex-provoked NER might play a role in TD-TFO activity. To test this, we assayed for reversion of the pSupFG1/G144C vector after co-transfection of selected oligonucleotide/plasmid complexes into a human mutant cell line (XP2OS) derived from a patient with Xeroderma pigmentosum, group A, and deficient in the NER damage recognition factor, XPA (23). In comparison, we also transfected the complexes into a subclone of this cell line stably expressing the XPA cDNA after gene transfer (23). The results of these reversion assays (Fig. 6) show that the activity of A/B-AG30 is reduced to 0.29% in the XPA-deficient cells, which is only marginally greater than the activity of the non-triplex-forming control A/B-MX30 (0.19%) in these cells. In contrast, in the corrected cell line, the differential between A/B-AG30 (0.48%) and A/B-MX30 (0.15%) is increased and approaches that seen in the COS-7 cells (Fig. 5). Hence, much of the extra stimulation provided by the TFO domain depends on XPA function and presumably on NER activity. Because A/B-MX30 is somewhat active in both the mutant and corrected cell lines, there otherwise appears to be near normal recombination activity in the XPA mutant, a result consistent with other studies of recombination in such cells (9). Thus, whereas the XPA cells can carry out recombination, the extra stimulation provided by the TFO domain is reduced in these cells.

However, these data should not be taken to suggest that the full benefit of the tethered TFO domain is mediated through the ability of triplexes to stimulate repair. These experiments were carried out by transfection of pre-incubated oligonucleotide/plasmid complexes, thus they may underestimate the potential role of the TFO in mediating a homology search in cells.

Intracellular Targeting of the pSupFG1/G144C Vector and Correction of a Point Mutation at Position 144—To demonstrate in vivo targeting of the supF gene by the TD-TFO molecules, cells were pre-transfected with the shuttle vector DNA, followed 34 h later by electroporation of the oligonucleotides into the cells. After another 48 h, the shuttle vector DNA was harvested for genetic analysis of the supF gene (Table I). In this protocol, A/B-AG30 was found to mediate reversion at a frequency of 0.014% or 20-fold over background (0.0007%; Table I). The TD-TFO with the single-stranded donor domain A-AG30 was also active, yielding a reversion frequency of 0.036%, 51-fold above background. The presence of a wild-type supFG1 sequence in randomly selected blue colonies (n = 10 for each) was confirmed by DNA sequencing. The controls yielded few, if any, revertants, indicating that the triplex domain is important for intracellular targeting (Table I; compare A, A-MX30, and A/B-AG30). Also, the covalent attachment of the donor and triplex-forming domains is required, because A + AG30 and A/B + AG30 showed little activity in this assay. These results suggest that target gene recognition and information transfer can be mediated by the TD-TFOs within cells.

Considering the 34-h interval between the shuttle vector transfection into COS-7 cells and the subsequent TD-TFO transfection, it is highly likely that the interaction of the TD-TFO with the shuttle vector is occurring intracellularly. Also, because SV40 replicons become rapidly covered with histones to form mini-chromosomes (25), these data suggest that the TD-TFOs can target chromatinized DNA. However, the chromatin structure of the SV40 vectors is quite dynamic, because they replicate actively in COS-7 cells and are highly transcribed. This feature of the vectors may be important in facilitating targeting by the TD-TFOs.

Whereas both A/B-AG30 and A-AG30 were more active in the intracellular targeting assay than any of the controls (including the donor fragments alone), A-AG30 was found to be 2.5-fold more active than its double-stranded counterpart, A/B-AG30. However, in the co-transfection assay (Fig. 5), A/B-AG30 was 4-fold more active than A-AG30. Also, in intracellular targeting experiments with the pSupF5 target (see below), the TD-TFO with the double-stranded donor domain was found to be more active than that with the single-stranded one. Whether these differences between the single- and double-stranded TD-TFOs have functional significance is not yet clear, and further studies are needed to optimize the design of the donor domain.

Table I  Specific reversion of supF gene mutations mediated by transfection of TD-TFO conjugates into cells that already contained the shuttle vector target

| Targeted vector | Oligonucleotide | No. of blue colonies/total | Reversion frequency (%) |
|-----------------|-----------------|---------------------------|-------------------------|
| pSupFG1/G144C   | No oligonucleotide | 1/144,746 | 0.0007 ± 0.0007* |
| A-AG30          | 40/111,624      | 0.006 ± 0.006* |
| X(Y/Rndm)-AG30  | 0/25,136        | <0.002        |
| A-MX30          | 5/1,939         | <0.002        |
| A + AG30        | 1/29,937        | 0.003 ± 0.003 |
| A/B-AG30        | 18/131,226      | 0.014 ± 0.003 |
| X(Y/Rndm)-AG30  | 0/54,791        | <0.002        |
| A/B-MX30        | 0/60,707        | <0.002        |
| A/B             | 1/52,843        | 0.002 ± 0.002 |
| A/B + AG30      | 0/25,136        | <0.004        |
| No oligonucleotide | 0/545,000      | <0.003        |
| C-UC30          | 5/1,188,000     | 0.0004 ± 0.0002 |
| D               | 0/1,292,000     | <0.0001       |
| UC30            | 0/1,380,000     | <0.0001       |
| C/D-UC30        | 67/310,000      | 0.0051 ± 0.0006 |
| C/D             | 6/1,400,000     | 0.0004 ± 0.0002 |

* Standard error of proportions.
It is clear, however, that the composite TD-TFO molecules are consistently more active than the corresponding controls in all of the experiments conducted thus far.

In addition, the differences in reversion frequencies seen when the oligonucleotides are incubated with the target vector \textit{in vitro} and co-transfected (Fig. 5) \textit{versus} when they are introduced into cells already containing the vector (Table I) reflect the challenges involved in achieving intracellular oligonucleotide delivery and triplex formation under physiologic conditions. Furthermore, the \textit{in vivo} protocol may underestimate the true frequency of intracellular targeting because electroporation delivers oligonucleotides into only a fraction of the cells containing the \textit{supF} plasmid (transfected 34 h earlier using cationic lipids). Hence, much of the rescued shuttle vector DNA used for analysis is derived from cells that did not receive the oligomers and in which no recombination could have occurred. Nonetheless, revertants were detected in the \textit{in vivo} protocol, suggesting that the TD-TFO molecules can successfully target an episomal gene inside cells. These initial results raise the possibility that improvements in oligonucleotide delivery and reagent design may enhance the effectiveness of this approach.

**Targeting by a TD-TFO Molecule in the Pyrimidine Triple Helix Motif**—The work presented above focuses on TD-TFOS that form triplexes in the purine motif, where the third strand is typically G-rich and binds anti-parallel to the purine-rich strand of the duplex. For comparison, we also tested the activity of a TD-TFO whose pyrimidine-rich TFO domain was designed to bind to a different target site in the pyrimidine motif. In this motif, the third strand binds in a parallel orientation relative to the purine strand of the duplex, and we considered the possibility that these differences in target site and polarity might influence our results.

To this end, TD-TFOS C-UC30 and C/D-UC30 were designed to correct a G-to-T point mutation at position 115 in the \textit{supF} gene in vector pSupF5/G115T (Fig. 2B). The \textit{supF5/G115T} gene contains an A-rich polypurine site at its 5’ end, suitable for third-strand binding in the pyrimidine motif. This vector has an extremely low spontaneous rate of reversion in our assay (≤0.0003%). In the reagents tested, the triplex-forming segment (UC30) was synthesized to contain the RNA residues 2′-O-methyluridine and 5-methyl-2′-O-methylcytidine. These modifications provide for increased nuclease resistance and decreased pH sensitivity of third-strand binding.

When the oligonucleotides and target vector were pre-incubated \textit{in vitro} and transfected into cells, results similar to those with the purine motif TD-TFOS were obtained, with reversion frequencies in the 0.1% range (data not shown). In experiments in which pSupF5/G115T was pre-transfected into the cells and electroporated 24 h later with the oligonucleotides, C/D-UC30 led to a reversion frequency of 0.0051% as compared with a background of ≤0.0003% (Table I), representing at least a 17-fold increase in activity over background. Also, UC30 by itself was ineffective (≤0.0001%), as were the donor fragments alone (C, ≤0.0001%; D, 0.0004%) or together as a duplex (C/D, 0.0004%). The presence of the wild-type \textit{supF} gene was confirmed by sequencing randomly selected revertants (n = 18 for C/D-UC30). In these experiments, the double-stranded C/D-UC30 was more active than the single-stranded C-UC30 (0.0004%), consistent with the trend seen when the oligonucleotides and the shuttle vector DNA were co-incubated \textit{in vitro} and subsequently co-transfected into cells (Fig. 5).

In general, the pyrimidine motif TD-TFOS were not as effective as their purine counterparts (A/B-AG30 was nearly 3-fold more active than C/D-UC30). This may be due to differences in several factors, including strand polarity relative to the target duplex, binding characteristics \textit{in vivo} (such as pH and ion dependence), intracellular stability of the TD-TFOS, the experimental protocol, the mechanisms of information transfer, and the suitability of the target site. For example, C-UC30 was designed to target a site in the \textit{supF} gene where the sequences corresponding to the TFO and donor domains are directly adjacent (Fig. 2B). In contrast, the regions of the \textit{supFG1} gene that correspond to the donor and TFO domains of A-AG30 or A/B-AG30 are separated by 6 bp (Fig. 2A). This spacing may ultimately be important for efficient information transfer by TD-TFO-type molecules. As a result, caution should be exercised in making a direct comparison between the efficiency of gene correction by A/B-AG30 and C/D-UC30. However, within each series, the TD-TFOS are consistently more active than their component parts.

**DISCUSSION**

The work presented here demonstrates the ability of selected TD-TFO molecules in both the purine and pyrimidine triple helix motifs to mediate targeted sequence alterations within a SV40 shuttle vector in mammalian cells. Successful reversion of the \textit{supF} target genes was seen after \textit{in vitro} co-incubation of the target vector with the TD-TFOS and also after an \textit{in vivo} protocol in which cells already containing the shuttle vector were transfected with the oligonucleotides. Oligonucleotides consisting of covalently linked donor and triplex-forming domains were more effective than either of the individual segments alone and more effective than the unjoined segments mixed together, suggesting that the two domains work synergistically when present in a single molecule. Furthermore, we found that a TD-TFO could introduce a complex mutation into the wild-type \textit{supFG1} gene to generate a novel \textit{SalI} restriction site, demonstrating the potential ability of TD-TFOS to correct mutations involving more than a single base pair substitution. To our knowledge, this has not been reported previously using such short fragments (40 bp or less).

The tethered donor approach differs from our previous efforts to use triplex-targeted DNA damage to induce homologous recombination (13). In that work, the TFO was used to introduce site-specific psoralen photoadducts to stimulate recombination between two separate \textit{supF} genes. In the work presented here, no mutagen other than the triplex itself is involved, and the recombination is intended to occur between two intact genes but between a target gene and a donor fragment tethered to the TFO.

The data from the \textit{in vitro} co-incubation and transfection experiments (Fig. 5) show that sequence alteration of the \textit{supF} gene in the SV40 vector can also be mediated by short donor fragments alone. The unlinked single- or double-stranded donor fragments achieved reversion frequencies that were 17- and 57-fold above background, respectively. These results are somewhat consistent with other reports that short nucleic acid constructs can mediate directed sequence changes (10, 16). However, we found that short donor molecules alone were not effective in the \textit{in vivo} protocol, suggesting that, at least for the SV40/\textit{supF} target, the attached TFO domain is functionally important in cells.

We propose two mechanisms to explain the apparent synergy between the donor and TFO domains. The first mechanism relies on the ability of an attached TFO to conduct a functional homology search for a target site in an ATP- and enzyme-independent manner and bind specifically to the DNA as a third strand, thereby positioning the donor domain near the sequence to be changed. This increases the local concentration of the donor fragment and may increase the likelihood of an information transfer event from the TD-TFO to the gene. Supporting this hypothesis is the 9-fold increase in activity of A-AG30 over A alone in the \textit{in vivo} protocol in which the
oligonucleotides must find the target plasmid that was previously transfected into the cells. In contrast, in the in vitro co-incubation and co-transfection protocol in which the oligonucleotides and the plasmid are already mixed together in solution at relatively high concentrations, there is less difference between A-AG30 and A alone (3.4-fold). Further supporting this hypothesis is the demonstrated ability of TFOs to deliver other moieties, such as a psoralen conjugate, to specific DNA sites in cells (3, 26). Theoretically, an alternative mechanism for the activity of the TFO domain could be strand invasion and the formation of a Watson-Crick double helix with one of the strands of the target DNA. We consider this unlikely, because the AG30 sequence is anti-parallel in polarity to the purine strand of the duplex and thus is not complementary to the pyrimidine-rich strand of the target site. Also, no species other than the expected triplexes were seen in the gel mobility shift experiments (Fig. 4).

The second mechanism involves the finding that triplex formation can provoke DNA repair at or around the triplex target site (6). This prompted the experiments reported here to test whether part of the benefit of the TFO domain in the TD-TFO molecules might come from the stimulation of DNA repair. Such induced repair could potentially lead to recombinogenic strand breaks or could possibly involve the recruitment of proteins important in homologous pairing, strand exchange, and/or recombination. The results of experiments using the XPA-deficient cell line and its corrected derivative are consistent with this model in that the activity of the TD-TFOs was diminished in the absence of NER but was restored when the NER defect was corrected.

Hence, taken together, the TD-TFO molecules may enhance the frequency and specificity of oligonucleotide-mediated mutagenesis by making specific sequence information available during a process of induced repair, facilitating a pathway of recombinational repair. The coupling of repair and recombination as well as the ability of the third strand to mediate target site recognition would therefore be a plausible mechanism to explain the effectiveness of the tethered donor molecules.

The TD-TFO molecules described here represent a first-generation design. Improvements may be possible by elucidating a mechanism of action and then exploiting this knowledge to guide refinements in reagent structure and composition. For example, the data on mutation induction using TD-TFOs indicate that these reagents can introduce a specific mutation at an efficiency of nearly 75%. However, in 20% of the sequenced mutations, we observed deletions and other point mutations that are consistent with mutagenesis induced by the triplex alone (6). We previously observed that triplex-induced mutagenesis may be related to either TFO length or affinity or both (6) and that triplexes of different length may differentially interfere with NER (27). We are currently exploring means of reducing nonspecific mutagenesis by changing these TFO parameters and by trying to improve the efficiency of the information transfer from the donor.

Also, the work presented here involves targeting an extra-chromosomal SV40-based vector. Whereas such a vector provides a chromatinized target, it may not fully reflect a chromosomal gene. The ability of psoralen-conjugated TFOs to target a chromosomal site has recently been demonstrated (26), but the effectiveness of TD-TFOs to mediate sequence changes at chromosomal sites remains to be determined.

Finally, the TD-TFO approach remains constrained by the restriction of high-affinity triplex formation to homopurine regions of DNA. Expansion of the third strand binding code through either the development of nucleotide analogs or the use of novel oligonucleotide backbones may greatly enhance the effectiveness of these reagents. Nonetheless, the TD-TFO approach as a method for DNA sequence modification has the potential to be a useful research tool and may eventually provide the basis of a gene therapy strategy.

Acknowledgments—We thank G. Wang, L. Narayanan, J. George, P. S. Chan, S. J. Baerga, F. Franklin, C. Chan, A. Chew, and L. Cabral for help and the members of the Glazer laboratory and J. Hearst for helpful discussions.

REFERENCES
1. Havre, P. A., Gunther, E. J., Gasparro, F. P., and Glazer, P. M. (1993) Proc. Natl. Acad. Sci. U. S. A. 90, 7879–7883
2. Havre, P. A., and Glazer, P. M. (1993) J. Virol. 67, 7324–7331
3. Wang, G., Levy, D. D., Seidman, M. M., and Glazer, P. M. (1995) Mol. Cell. Biol. 15, 1759–1768
4. Takasugi, M., Guendouz, A., Chassignol, M., Decout, J. L., Lhomme, J., Thuong, N. T., and Helene, C. (1991) Proc. Natl. Acad. Sci. U. S. A. 88, 5602–5606
5. Belousova, E. S., Afonina, I. A., Podyminogin, M. A., Gamper, H. B., Reed, M. W., Wydro, R. M., and Meyer, R. B. (1997) Nucleic Acids Res. 25, 3440–3444
6. Wang, G., Seidman, M. M., and Glazer, P. M. (1996) Science 271, 802–805
7. Capecechi, M. R. (1989) Science 244, 1288–1292
8. Sedivy, J. M., and Joyner, A. L. (1992) Gene Targeting, W. H. Freeman and Co., New York
9. Wang, Y. Y., Maher, V. M., Linskay, R. M., and McCormick, J. J. (1988) Mol. Cell. Biol. 8, 196–202
10. Kuzmin, M. A., Legendre, J. Y., Kneed, D. L., Escobar, L. C., Xu, Z., and Gruenert, D. C. (1996) Gene Ther. 3, 859–867
11. Rouet, P., Smith, F., and Jasin, M. (1994) Proc. Natl. Acad. Sci. U. S. A. 91, 6064–6068
12. Segal, D. J., and Carroll, D. (1995) Proc. Natl. Acad. Sci. U. S. A. 92, 806–810
13. Faruqui, A. F., Seidman, M. M., Sigal, D. J., Carroll, D., and Glazer, P. M. (1996) Mol. Cell. Biol. 16, 6820–6826
14. Folger, K. R., Wong, E. A., Wahli, G., and Capecechi, M. R. (1982) Mol. Cell. Biol. 2, 1372–1387
15. Rubnitz, J., and Subramani, S. (1984) Mol. Cell. Biol. 4, 2253–2258
16. Cole-Straus, A., Yoon, K., Xiang, Y., Byrne, B. C., Rice, M. C., Geyn, J., Holloman, W. K., and Kmiec, E. B. (1996) Science 273, 1386–1389
17. Strobel, S. A., Moser, H. E., and Dervan, P. E. (1988) J. Am. Chem. Soc. 110, 7927–7929
18. Gunther, E. J., Havre, P. A., Gasparro, F. P., and Glazer, P. M. (1996) Photochem. Photobiol. 63, 207–212
19. Moser, H. E., and Dervan, P. B. (1987) Science 238, 645–650
20. Pei, D., Corey, D. R., and Schultz, P. G. (1990) Proc. Natl. Acad. Sci. U. S. A. 87, 9858–9862
21. Landgraf, R., Chen, C. H., and Sigman, D. S. (1994) Biochemistry 33, 10673–10675
22. Poyart, T. J., and Dervan, P. B. (1996) J. Am. Chem. Soc. 118, 9428–9430
23. Levy, D. D., Saito, M., Tanaka, K., and Kraemer, K. H. (1995) Carcinogenesis 16, 1557–1563
24. Cheng, A. J., and Van Dyke, M. W. (1994) Nucleic Acids Res. 22, 4742–4747
25. Reeves, R., Gorman, C. M., and Howard, B. (1985) Nucleic Acids Res. 13, 3599–3615
26. Majmudar, A., Khorlin, A., Dyatkina, N., Lin, F.-L. M., Powell, J., Liu, J., Fei, Z., Khrupine, Y., Watanabe, K. A., George, J., Glazer, P. M., and Seidman, M. M. (1998) Nat. Genet. 20, 212–214
27. Wang, G., and Glazer, P. M. (1995) J. Biol. Chem. 270, 22595–22601