Distance and Affinity Dependence of Triplex-Induced Recombination†

Melissa P. Knauert,‡ Janice A. Lloyd,‡ Faye A. Rogers,‡ Hirock J. Datta,‡ Michael L. Bennett,§ Daniel L. Weeks,|| and Peter M. Glazer*,‡

Departments of Therapeutic Radiology and Genetics, Yale University School of Medicine, P.O. Box 208040, New Haven, Connecticut 06520, Department of Microbiology, School of Medicine, University of Colorado Health Sciences Center, Denver, Colorado 80262, and Department of Biochemistry, University of Iowa, Iowa City, Iowa 52242

Received September 2, 2004; Revised Manuscript Received December 15, 2004

ABSTRACT: Triplex-forming oligonucleotides (TFOs) have the potential to serve as gene therapeutic agents on the basis of their ability to mediate site-specific genome modification via induced recombination. However, high-affinity triplex formation is limited to polypurine/polypyrimidine sites in duplex DNA. Because of this sequence restriction, careful analysis is needed to identify suitable TFO target sites within or near genes of interest. We report here an examination of two key parameters which influence the efficiency of TFO-induced recombination: (1) binding affinity of the TFO for the target site and (2) the distance between the target site and the mutation to be corrected. To test the influence of binding affinity, we compared induced recombination in human cell-free extracts by a series of G-rich oligonucleotides with an identical base composition and an increasing number of mismatches in the third strand binding code. As the number of mismatches increased and, therefore, binding affinity decreased, induced recombination frequency also dropped. There was an apparent threshold at an equilibrium dissociation constant (K_d) of 1×10^{-7} M. In addition, TFO chemical modification with *N*,*N*-diethylethylenediamine (DEED) internucleoside linkages to confer improved binding was found to yield increased levels of induced recombination. To test the ability of triplex formation to induce recombination at a distance, episomal targets with informative reporter genes were constructed to contain polypurine TFO target sites at varying distances from the mutations to be corrected. TFO-induced recombination in mammalian cells between a plasmid vector and a donor oligonucleotide was detected at distances ranging from 24 to 750 bp. Together, these results indicate that TFO-induced recombination requires high-affinity binding but can affect sites hundreds of base pairs away from the position of triplex formation.

Triplex-forming oligonucleotides (TFOs)[1] site-specifically bind duplex DNA and can induce both mutation and recombination (*1*). Because TFOs target DNA rather than RNA substrates, they are capable of producing permanent, heritable change. Accordingly, TFOs are thought to have therapeutic potential, and much effort has been spent to understand and improve upon TFO-mediated events. DNA TFOs are both constrained and powered by specific binding rules. High-affinity TFO binding requires target sites having a polypurine run on one strand of the duplex DNA and a polypyrimidine strand on the opposite; triplex formation has a specific binding code which allows the design and use of sequence-specific third strands (*2*−*4*).

TFO-induced recombination has been shown to occur in several experimental situations. Intramolecular recombination between direct repeats of selected reporter genes was found to be induced by triplex formation in both episomal (*5*) and chromosomal targets (*6*). In the episomal substrate, the TFOs were targeted to bind 4 bp downstream and 29 bp upstream of the two reporter genes, whereas in the chromosomal substrate, the target site for triplex formation was 0.8 and 0.7 kb away from upstream and downstream genes, respectively. Though these experiments suggested that triplex formation could influence DNA metabolism at a distance, they were limited to intramolecular events. In a protocol of more direct relevance to gene repair, a bifunctional oligonucleotide containing a TFO domain (30 nt) linked to a donor domain (40 nt) was found to mediate gene correction within an episomal vector in COS cells (*7*). The design of this bifunctional molecule was based on the presence of a favorable TFO binding site adjacent to the mutation to be corrected (24 bp away). Also, it appeared that linkage of the TFO and the donor domain was advantageous. However, the constraint of linkage was recently shown to be unnecessary on the basis of an assay for induced recombination in human cell-free extracts in which a 30-mer G-rich oligonucleotide, AG30, was found to induce recombination between a single-stranded donor oligonucleotide and a plasmid substrate whether or not the donor molecule was

† This work was supported by grants to P.M.G. from the NIH (R01GM54731 and R01CA64186), to M.P.K. from the NIH Medical Scientist Training Program (GM07205), to J.A.L. from the NIH National Research Service Award (CA103543), and to D.L.W. from the Cystic Fibrosis and Gene Therapy Center at the University of Iowa and NIH (P50HL062178).

* To whom correspondence should be addressed. E-mail: peter.glazer@yale.edu. Phone: (203) 737-2788. Fax: (203) 737-2630.

‡ Yale University School of Medicine.

§ University of Colorado Health Sciences Center.

|| University of Iowa.

[1] Abbreviations: DEED, *N*,*N*-diethylethylenediamine; DSB, double-strand break; HR, homologous recombination; K_d, equilibrium dissociation constant; NER, nucleotide excision repair; PNA, peptide nucleic acid; TFO, triplex-forming oligonucleotide.

10.1021/bi0481040 CCC: $30.25
Published on Web 02/12/2005

linked to the TFO (*8*). This ability of a TFO to stimulate recombination between a target substrate and an unlinked donor prompted us to test the extent to which a TFO can stimulate recombination by an unlinked donor at varying distances from the site of third strand binding. We report here that TFO-induced recombination can be detected in cells at distances from 24 to 753 bp away, thereby expanding the utility of the triplex approach for gene repair strategies.

Another important consideration in the selection of TFO binding sites and in the design of TFOs is binding affinity. The ability of a TFO to induce recombination by an unlinked donor molecule allowed us to systematically test induced recombination by a series of TFOs in conjunction with the same donor molecule and the same target plasmid in a cell-free extract system. We compared TFOs with varying numbers of mismatches (while maintaining the same base composition) and, therefore, varying affinities for target site binding. We report here that even a few mismatches in the third strand binding code substantially weakened the ability of the TFO to induce recombination. Although these results place a constraint on the identification of useful TFO target sites, they also indicate the stringent specificity of TFO-induced DNA metabolism.

In addition, we tested the effect of TFO chemical modification with *N*,*N*-diethylethylenediamine (DEED) internucleoside linkages on the ability to induce recombination. We report here that DEED modification enhances the ability of a TFO to induce recombination compared to a standard DNA TFO with an identical sequence. This increase in the level of recombination induction is attributed to the higher relative affinity of DEED TFOs for their target, especially under physiologic conditions (*9*).

EXPERIMENTAL PROCEDURES

Oligonucleotides. DNA oligonucleotides were synthesized by the Midland Certified Reagent Company Inc. (Midland, TX) and purified by RP-HPLC. DNA oligonucleotide sequences are as indicated in Figure 1A (30-mer G-rich oligonucleotides) and Figure 3A (13-mer G-rich oligonucleotide), and as follows for the *supF* and *FLuc* donor molecules: A(WT)40 (*supF*), 5′ AGGGAGCAGA CTCTAAATCT GCCGTCATCG ACTTCGAAGG 3′; FLuc13, 5′ TGGTAAAGCC ACCATGGAAG ACGCCAAAAA CATAAAGAAA GGCCCGGCGCC 3′; FLuc202, 5′ GGTTGGCAGA AGCTATGAAA CGATACGGGC TGAATACAAA TCACAGAATC 3′; Fluc713, 5′ CAAATCATTC CGGATACTGC GATTTTAAGT GTTGTTCCAT TCCATCACGGT 3′; and Fluc1215, 5′ TATGTCCGGT TATGTAAACA ATCCGGAAGC GACCAACGCC TTGATTGACAA 3′. The control oligonucleotide, Mix30, a mixed base 30-mer, has the following sequence: 5′ AGTCAGTCAG TCAGTCAGTC AGTCAGTCAG 3′. AG30, related G-rich oligonucleotides, and Mix30 were 3′ end-protected with a C7 amine group. A(WT)40 was 3′ end-protected with three phosphorothioate internucleoside linkages; *FLuc* donor molecules were 5′ and 3′ end-protected with three phosphorothioate internucleoside linkages. Synthesis and purification of the DEED oligonucleotide (sequence indicated in Figure 3A) were performed as previously described (*10*).

Plasmid Vectors. For the SV40-based shuttle vector assay in monkey COS cells, a series of plasmids was constructed from the parent shuttle vector, pSP189, described by Parris and Seidman (*11*). The pSP189 vector contains a wild-type *supF* gene. An inactivating mutation was introduced via site-directed mutagenesis of 4 out of 6 bp at positions 138−143 (Quik Change Site Directed Mutagenesis Kit, Stratagene, La Jolla, CA). The mutagenesis introduced a *Sal*I site, and this plasmid was termed pSP189(SalI). Derivatives containing AG30 duplex target sites at various distances were created. pMKsupF59 (distance from the *Sal*I site to the AG30 target of 59 bp) was created by inserting the AG30 duplex target into the *Eag*I site of pSP189(SalI). pMKsupF102 (distance from the *Sal*I site to the AG30 target of 102 bp) was created by inserting the AG30 target site into the *Mlu*I site of pSP189(SalI). pMKsupF400 and pMKsupF1000 (distances from the *Sal*I site of 400 and 1000 bp, respectively) were created by inserting unrelated stuffer sequences of 300 and 900 bp into pMKsupF102 ahead of the duplex target. The shuttle vector plasmid pSupFG1/G144C, containing a *supFG1* gene with an inactivating G·C to C·G point mutation at position 144, was described previously (*7*). The distance between the point mutation at position 144 and the start of the AG30 target is 24 bp.

Plasmids containing the firefly luciferase gene were created by subcloning of *FLuc*+ from Promega's (Madison, WI) pGL3-Basic Vector into pcDNA5/FRT (Invitrogen, La Jolla, CA). An insert containing the duplex AG30 target was introduced 40 bp upstream of the *FLuc*+ ATG start. Stop codons were created by site-directed mutagenesis at base pairs 13, 202, 713, and 1215 relative to the ATG start codon. A lack of luciferase activity was confirmed using the Promega Luciferase Assay (see below).

Determination of TFO Binding Affinity. Two complementary 57-mer oligonucleotides containing the sequence corresponding to bp 157−213 of *supFG1* were synthesized, mixed 1:1 in 25 mM NaCl, heated to 90 °C for 10 min, and allowed to cool slowly to room temperature to form duplex DNA. Following end labeling of the duplex using T4 polynucleotide kinase and [γ-^{32}P]dATP, the fragments were purified by gel electrophoresis and Centricon-3 filtration in distilled water (Amicon, Beverly, MA). Binding reaction mixtures (20 μL) containing 1×10^{-10} M duplex DNA and increasing concentrations of TFO in 10 mM Tris-HCl (pH 7.6), 10 mM $MgCl_2$, 1 mM spermine, and 10% glycerol were incubated overnight at 37 °C; gel shifts were carried out in the absence of potassium unless otherwise indicated. The products were resolved by gel electrophoresis on a 12% native acrylamide gel at 70 V in 1× TB buffer with 10 mM $MgCl_2$. A phosphorimager (Amersham Biosciences, Piscataway, NJ) was used to quantify the reaction products and determine the equilibrium dissociation constants (K_d).

In Vitro Assay for Induced Recombination. HeLa whole cell extracts were prepared as described previously (*12*). Reaction mixtures consisted of 3 μg of pSupFG1/G144C plasmid DNA, 3 μg each of selected oligonucleotides (putative TFOs and donor fragment), 60 mM NaCl, 2 mM β-mercaptoethanol, 3 mM KCl, 12 mM Tris (pH 7.4), 2 mM ATP, 0.1 mM dNTPs, 2.5 mM creatine phosphate, 1 μg of creatine phosphokinase, 12 mM $MgCl_2$, 0.1 mM spermidine, 2% glycerol, 0.2 mM DTT, and 15 μL of cell-free extract in a total reaction volume of 50 μL. After incubation for 2 h at 30 °C, the reactions were terminated by the addition of 25 μM EDTA, 0.5% SDS, and 20 μg of proteinase K. After 1

h at 37 °C, the plasmid DNA was isolated by phenol extraction and ethanol precipitation and dissolved in 10 μL of H_2O. One microliter of the resulting sample was used to transform *Escherichia coli* MBM7070 by electroporation, as described previously (*12*), followed by growth of the bacteria on indicator plates for genetic analysis of *supFG1* gene function as described previously (*13*).

In Vitro Assay for Induced Repair Synthesis. The pSupFG1 plasmid (600 ng), described previously (*14*), containing the AG30 triplex binding site, was incubated overnight at 37 °C with the selected G-rich TFO at a final concentration of 1 μM in 10 mM Tris (pH 7.6), 10 mM $MgCl_2$, and 10% glycerol. The DNA was then added to HeLa cell-free extracts supplemented with [α-^{32}P]dCTP and other components as previously described (*15*). After incubation at 30 °C for 3 h, the reactions were terminated by the addition of EDTA (20 mM), RNase A (80 μg/mL), SDS (0.5%), and proteinase K (190 μg/mL). The plasmid DNA was isolated by phenol extraction and purified using the Qiagen PCR purification kit (Valencia, CA). The DNA was linearized by *Xho*I digestion and analyzed by gel electrophoresis. Quantification of DNA repair synthesis as measured by the level of [α-^{32}P]-dCTP incorporation was performed using a phosphorimager. pSupFG1 plasmid DNA exposed to 5000 J/m^2 of short-wavelength UV radiation (254 nm) and then treated with extract was used as a positive control such that the value for this sample was set at 100% and used to normalize the raw measurements for the other samples.

Cells. Monkey COS-7 cells were obtained from ATCC (1651-CRL) and were maintained in DMEM supplemented with 10% fetal bovine serum (Gibco BRL, Division of Invitrogen, Gaithersburg, MD). CHO cells were obtained from Invitrogen and maintained in Ham's F12 medium supplemented with 10% fetal bovine serum, 2 mM L-glutamine, and 100 μg/mL Zeocin.

Shuttle Vector Assay. Cells were transfected with the indicated *supF* plasmid DNAs using the cationic lipid formulation, Geneporter 2 (Gene Therapy Systems, San Diego, CA), according to the manufacturer's instructions; 5 $\times$ 10^5 cells were seeded into 10 cm dishes 24 h prior to transfection. At the time of transfection, a master solution of each type of DNA to be transfected was aliquoted; 1 μg each of *supF* plasmid, donor oligonucleotide, and G-rich TFO or Mix30 was used per plate of cells. Oligonucleotides to be used in the transfection were heated to 55 °C for 10 min and quenched on ice. At 48 h post-transfection, cells were harvested, and plasmid DNAs were recovered via alkaline lysis, as described previously (*7*). Recovered plasmid DNAs were assayed for *supF* gene recombination as described above.

CHO Cell Transfection and Luciferase Assay. CHO cells were transfected with plasmid, donor, and either AG30, 12AG30, or Mix30 using Geneporter 2 as described above. Briefly, 24 h prior to transfection, 12-well plates (4 cm^2 per well) were seeded with 5 $\times$ 10^4 cells per well. DNAs to be transfected were mixed with 10 mM Tris (pH 7.4) and incubated overnight at 37 °C. Transfections delivered 0.1 μg of plasmid, donor, and AG30, 12AG30, or Mix30 to each well; 48 h following transfection, cells were rinsed two times with Dulbecco's phosphate-buffered saline (1$\times$) (Gibco BRL) and then lysed with 250 μL of 1$\times$ Passive Lysis Buffer (Promega).

Luciferase activity was measured following the manufacturer's instructions using the Promega Dual Luciferase Kit (Promega) and only the LAR II substrate. The total protein in the lysate was determined via the Bradford protein assay (*16*).

RESULTS

Design of G-Rich Triplex-Forming Oligonucleotides. AG30 is a G-rich TFO used throughout these studies. AG30 binds as a third strand in the antiparallel motif to form a triple helix at the polypurine/polypyrimidine duplex target shown (Figure 1A). This duplex sequence is present at the 3′ end of the *supFG1* gene in the pSupFG1/G144C and pSupFG1 vectors as described previously (*7*). The same polypurine/polypyrimidine duplex sequence has been inserted into pSP189(SalI) at various positions. The binding characteristics of AG30 for this duplex target have been previously determined (*13*). A series of additional G-rich oligonucleotides was designed to have the same base composition as AG30 but an increasing number of mismatches in the third strand binding code (Figure 1A). There are two pyrimidine inversions in the AG30 duplex target (at positions 14 and 17 out of 30) that create unavoidable mismatches in the third strand binding code for all of the 30-mer G-rich oligonucleotides studied here.

Binding Measurements. To determine the relative binding affinities of the G-rich oligonucleotides for the target duplex, a gel mobility shift assay was used (Figure 1B,C). The K_d for each oligonucleotide was calculated as the concentration of oligonucleotide at which binding was one-half of maximal. AG30, which is optimized in the third strand binding code for the 30 bp duplex target site, binds to the target site, under the conditions used in these experiments, with high affinity ($K_d \sim 2.5 \times 10^{-8}$ M), which is consistent with previously published binding studies for this molecule (*13*). Analysis of the oligonucleotide series indicates that as the number of mismatches increases, the binding affinity decreases (Figure 1B,C). For example, the G-rich oligonucleotide containing two mismatched bases relative to AG30 (designated 2AG30) has a slightly reduced binding affinity ($K_d \sim 7.5 \times 10^{-8}$ M) compared to AG30. 8AG30 (containing eight mismatched bases) has a substantially reduced binding affinity ($K_d \sim 1 \times 10^{-6}$ M). 12AG30, which contains the most mismatches possible (while still maintaining a constant base composition), shows negligible binding to the target site. These data demonstrate that complete homology to the target site is not required for triplex formation. However, as the level of homology between the TFO and target site decreases, there is a substantial loss in affinity (Figure 1C).

TFO-Induced Recombination and Repair Depend on Binding Affinity. Using an in vitro recombination assay described previously (*8*, *15*), we tested each G-rich oligonucleotide's ability to induce recombination between the plasmid vector pSupFG1/G144C, containing an inactivating G·C to C·G mutation at position 144, and the single-stranded 40-mer oligonucleotide donor, A(WT)40, which corresponds to nucleotides 121−160 of *supFG1* with the wild-type sequence at base pair 144. Consistent with previous work, AG30 induced recombination between the donor oligonucleotide, A(WT)40, and the target plasmid at a frequency of 47 $\times$ 10^{-5} (Figure 2A). 2AG30 induced recombination

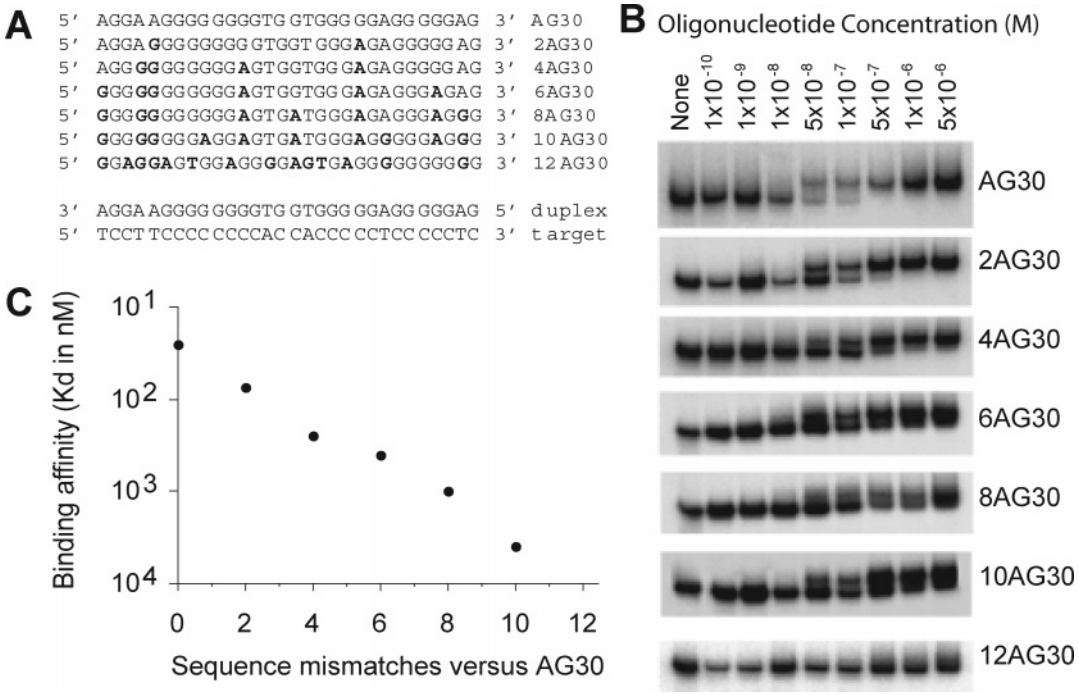

FIGURE 1: In vitro binding affinity of matched and mismatched G-rich oligonucleotides for triplex formation. (A) Design of 30-mer G-rich oligonucleotides. The base composition is constant; nucleotides in bold type are mismatched to the target site. The duplex target site is listed below oligonucleotide sequences. (B) In vitro gel shift analyses to quantify third strand binding. The labeled duplex was incubated in the presence of excess 30-mer G-rich oligonucleotides at varying concentrations and analyzed by electrophoresis on a nondenaturing polyacrylamide gel. K_d values were estimated as the concentration at which the binding was half-maximal, as quantified by phosphorimager analysis. (C) Inverse correlation of binding affinity with the number of mismatches in the third strand binding code.

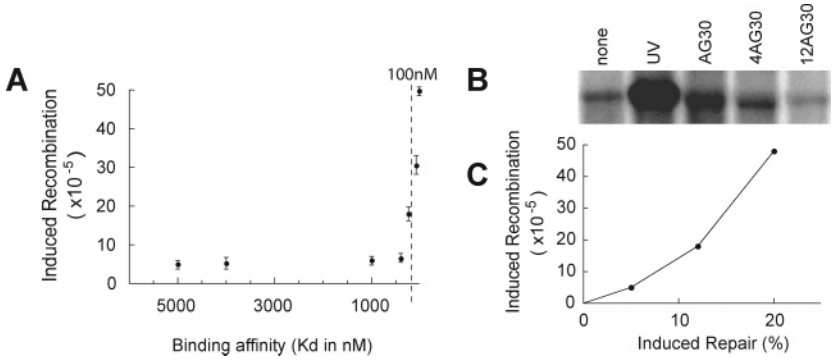

FIGURE 2: TFO-mediated induction of recombination and repair in human cell-free extracts. (A) Binding affinity of 30-mer oligonucleotides (K_d in nanomolar) vs induction of recombination. Error bars indicate one standard error above and below the average value of all experimental replicates. (B) Oligonucleotide-induced repair synthesis in human cell-free extracts. pSupFG1, containing the AG30 duplex target site, was treated as indicated [none, plasmid with extract treatment only; UV, 5000 J/m^2 of short-wavelength UV radiation (254 nm); AG30, 1 μM AG30 incubated overnight at 37 °C; 4AG30, 1 μM 4AG30 incubated overnight at 37 °C; 12AG30, 1 μM 12AG30 incubated overnight at 37 °C] and then added to HeLa cell-free extracts in the presence of [α-^{32}P]dCTP. The plasmid substrates were isolated, resolved by gel electrophoresis, and visualized and quantified for incorporation of radioactivity by the phosphorimager; the repair band intensity for the UV-damaged plasmid was set at 100%. (C) Relative induction of repair by oligonucleotides vs induction of recombination.

at a frequency of 43×10^{-5}, similar to that of AG30. However, as the number of mismatches between the 30-mer G-rich oligonucleotides and the target duplex increased beyond two, the induced recombination frequency decreased. For example, 4AG30 yielded a frequency of 28×10^{-5}, and 12AG30 yielded a frequency of just 3×10^{-5}, a level similar to that of the donor alone in the assay. These results suggest that the ability of the TFOs to induce recombination is closely correlated with the affinity of the TFO for the target site. The mismatch studies further suggest that while 100% homology between the target sequence and the TFO is not absolutely necessary to induce recombination, there is a sharp threshold at a K_d of approximately 1×10^{-7} M (correlating to four to six mismatches in this system) beyond which the extent of triplex-induced recombination is substantially reduced.

In previous work, we have shown that triplex formation is able to stimulate repair synthesis on an otherwise undamaged plasmid in HeLa cell-free extracts (*13*) and that this induction of repair is dependent upon the nucleotide excision repair (NER) pathway (*15*). To assess the dependence of induced repair on binding affinity, we tested several of the 30-mer oligonucleotides in a repair assay to correlate the ability to stimulate repair synthesis with the binding affinity of the TFOs (Figure 2B). We found a direct correlation between TFO binding affinity and the ability to induce repair synthesis, with the level of induced repair by AG30 greater than that by 4AG30 and greater than that by 12AG30. For

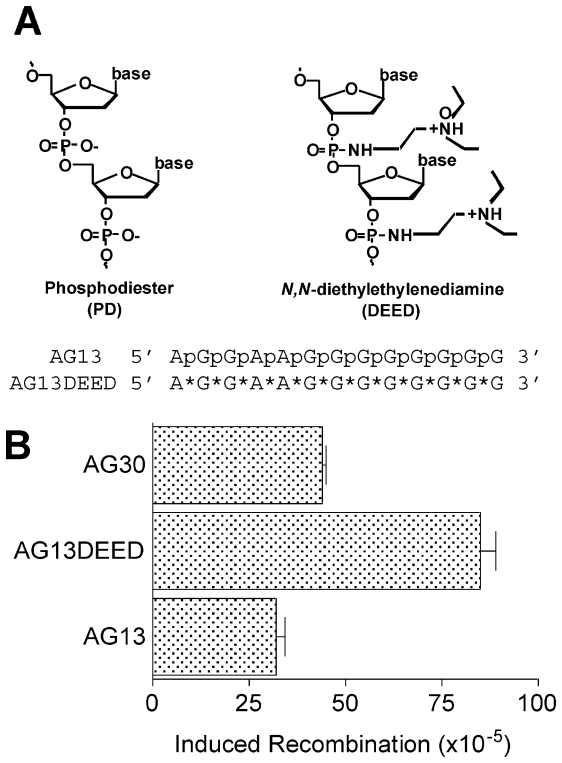

FIGURE 3: Induced recombination by a natural DNA phosphodiester oligonucleotide vs a DEED-modified oligonucleotide. (A) Backbone structure of a phosphodiester vs a DEED-modified oligonucleotide. Note the replacement of a nonbridging oxygen with a *N*,*N*-diethylethyleneamine group. Sequences of 13-mer G-rich oligonucleotides, AG13 [phosphodiester (p) internucleoside linkages] and AG13DEED [DEED (∗) internucleoside linkages] are as indicated. These oligonucleotides are designed to bind to the first 13 bp of the duplex target shown in Figure 1. (B) Induction of recombination in HeLa cell-free extracts by phosphodiester vs DEED-modified G-rich oligonucleotides. All samples include a plasmid, a donor, and the oligonucleotide indicated at the left of each bar.

quantification, the level of induced repair was normalized to that induced by UV damage (set at 100%) and plotted against the frequency of induced recombination for the oligonucleotides that were tested. As shown, there is a direct correlation between induction of repair and induction of recombination (Figure 2C).

As a further examination of the correlation of TFO binding with induced recombination, we tested TFO chemical modification with DEED internucleoside linkages (Figure 3A). In previous work, we had determined that DEED TFOs have higher binding affinities than phosphodiester DNA TFOs of the same sequence, especially under physiologic conditions with a low level of magnesium and a high level of potassium (*9*). We compared a pair of oligonucleotides, AG13 and AG13DEED, both designed to bind to the first 13 bp of the AG30 binding site in *supFG1* (Figure 3A). AG13DEED was DEED substituted at all 12 possible internucleoside linkages. AG13DEED shows a binding affinity for the *supFG1* site 3-fold greater than that of AG30 (data not shown) and was found to induce recombination at a frequency of 85×10^{-5} (Figure 3B). In contrast, the phosphodiester DNA TFO, AG13, which has the same base sequence as AG13DEED, has a K_d of approximately 5×10^{-7} M (similar to that of 4AG30) and a recombination induction frequency of 32×10^{-5} (also similar to that of 4AG30). These results further support the correlation of TFO binding affinity and induced recombination, and they suggest that improvements in affinity mediated by chemical modification may lead to additional increases in the level of TFO-induced recombination.

Distance Dependence of Triplex-Induced Recombination in Cells. In previous studies of triplex-induced recombination in mammalian cells, we used the pSupFG1/G144C shuttle vector as a target substrate. We found that a bifunctional oligonucleotide consisting of the donor molecule, A(WT)-40, covalently linked to the TFO, AG30, mediated gene correction in the plasmid target (following cotransfection of the oligonucleotide and plasmid) at a frequency in the range of 0.17% (*7*). Subsequent in vitro work suggested that the two domains did not need to be covalently linked for the TFO to stimulate recombination (*8*). To confirm this in cells, we cotransfected AG30 and A(WT)40 (unlinked) into COS cells along with pSupFG1/G144C. In this plasmid vector, the AG30 binding site is situated 24 bp from the inactivating mutation at base pair 144 (Figure 4A). Upon shuttle vector rescue and analysis in indicator bacteria, we found a frequency of induced gene correction of 0.14%, similar to the value of 0.17% produced by the bifunctional, linked oligonucleotide. Neither the mismatched G-rich oligonucleotide, 12AG30, nor a mixed sequence 30-mer, Mix30, stimulated recombination by the unlinked A(WT)40 donor molecule above what is achieved by A(WT)40 alone (Figure 4B).

To test the ability of AG30 to induce recombination at a distance greater than 24 bp, a series of plasmids was constructed as described above; these plasmids are derived from the parent vector, pSP189(SalI), and vary in the placement of the AG30 duplex target 59, 102, 400, and 1000 bp from the site of the inactivating *supF* mutation (a mutation of 4 bp of 6 bp at positions 138−143 to create a *Sal*I restriction site). The parent pSP189(SalI) plasmid lacks the extended polypurine target site. There is a 10 bp sequence at base pairs 167−176 which matches the first 10 bp of the AG30 site, but beyond this, there are 16 mismatches against binding by AG30. This substantially mismatched site starts 24 bp from the *Sal*I mutation. The other plasmid in the set is pSupFG1/G144C, which, as described above, has the complete AG30 binding site beginning 24 bp from an inactivating G to C mutation.

To measure the level of induced recombination, the DNA components [plasmid substrate, A(WT)40 donor, and AG30] were cotransfected via cationic lipids into COS cells. For comparison in each experiment, the plasmid and donor DNAs were also cotransfected with 12AG30 (or, in some cases, with the mixed sequence control oligonucleotide, Mix30). On the basis of Figure 4B and previous work, we know that 12AG30 (also designated Scr30 in some prior publications from our group) does not detectably induce repair or recombination; in addition, gel shift assays have established that Mix30 does not bind the AG30 duplex target (*7*). The samples including one of the nonbinding 30-mer oligonucleotides or the other serve to control for nonspecific effects that might result from the addition of extra nucleic acid to the donor DNA in the transfection mix. Following transfection, the episomal DNA was recovered 48 h later and used to transform indicator bacteria to determine the frequency of recombination within the *supF* gene. For each experiment, a minimum of 45000 colonies were analyzed and between

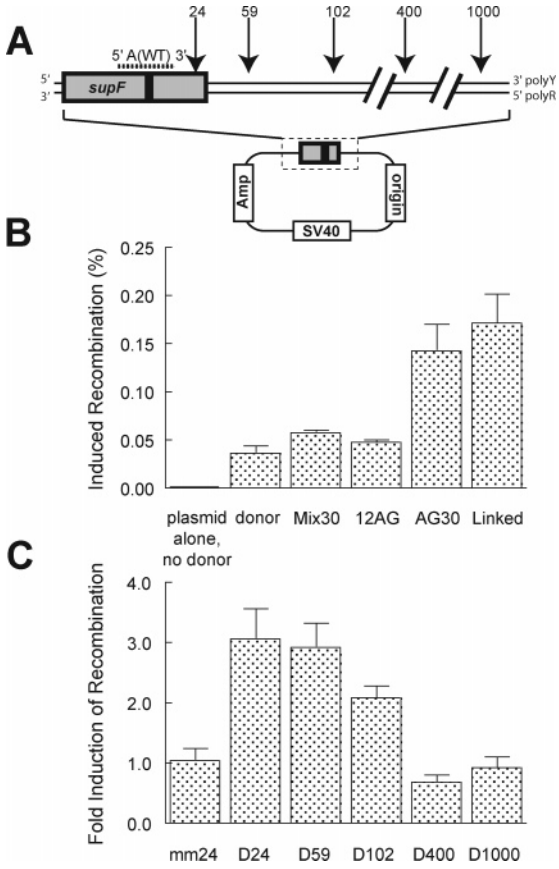

FIGURE 4: TFO-induced recombination in the *supF* gene at various distances. (A) Schematic of plasmid vector targets. The *SupF* gene is represented by the gray box. The *Sal*I site at positions 138−143 is indicated by the black bar. TFO binding sites 24, 59, 102, 400, and 1000 bp from the inactivating *supF* mutation are indicated by downward-pointing arrows. Target site 24mm (24 mismatch) is not depicted, but is colocated with target 24. The A(WT) donor is drawn above the strand to which it is homologous. The orientation of the polypurine/polypyrimidine duplex target for AG30 (at all distances) is indicated by the polyY and polyR labels. The polypyrimidine strand of the duplex AG30 target is on the top strand as drawn here (polyY). The polypurine strand is on the bottom strand as drawn here (polyR). (B) Induction of recombination in the psupFG1/G144C plasmid target by selected oligonucleotides. Plasmid alone, no donor indicates psupFG1/G144C only samples. Donor indicates plasmid and donor oligonucleotide alone without G-rich oligonucleotide. Mix30 indicates samples containing plasmid, donor, and Mix30. 12AG30 indicates samples containing plasmid, donor oligonucleotide, and 12AG30. AG30 indicates samples containing plasmid, donor, and AG30. Linked indicates samples containing plasmid and the bifunctional oligonucleotide consisting of AG30 linked covalently to a donor. Error bars indicate one standard error. (C) Induction of recombination by AG30 at various distances between the site of *supF* gene mutation and the third strand binding site. For each plasmid substrate (mm24, D24, D59, D102, D400, and D1000), the average ratio of AG30 samples to 12AG30 samples was calculated and is presented as the fold induction of recombination. The numbers (used to designate each plasmid and preceded by the letter D) indicate the distance (D) between the mutation site and the TFO-binding site. 24mm indicates the plasmid with a mismatched target site at a distance of 24 bp. Error bars indicate one standard error.

3 and 10 independent experiments were conducted for each plasmid in the series. For purposes of standardization, data are presented as the ratio of the AG30-induced recombination frequency over the baseline recombination frequency in the 12AG30 samples. We found significant induction over background 24, 59, and 102 bp from the site of gene correction (Figure 4C). The highest levels of induction, up to approximately 3-fold, were seen at the closer distances of 24 and 59 bp; a 2-fold induction was seen at the distance of 102 bp. Neither the plasmid with a mismatched target nor the plasmids with the TFO binding targets 400 and 1000 bp away showed induction by the sequence-specific TFO (Figure 4C). As a further control, we found that there was no induction of recombination above baseline in any of the plasmids by Mix30 (data not shown). Hence, TFO binding in this assay can influence recombination by a 40-mer donor oligonucleotide over distances of up to 100 bp.

To further investigate TFO-induced recombination at a distance, we set up an alternative assay system using the firefly luciferase gene as a reporter; firefly luciferase has previously been shown to be a sensitive marker of gene repair (*17*). In this system, the TFO binding site was fixed at a single location, but the position of the inactivating mutation in the gene was varied. (Because the *supF* coding region is only 93 bp long, this strategy was not possible using the *supF* gene.) An AG30 binding site was inserted within the 5′ untranslated region of the gene 40 bp upstream of the coding region start site. In this construct, luciferase gene expression is driven by the CMV immediate early promoter. Variants of this plasmid were produced in which stop codons were created by mutations at base pairs 13, 202, 713, and 1215 in the coding region (out of a total of 1652 bp), yielding TFO binding site-to-mutation distances of 53, 242, 753, and 1255 bp, respectively (Figure 5A). It was confirmed that each of these nonsense mutations abrogates luciferase activity to undetectable levels in a transient transfection assay (data not shown).

As described above, vectors were cotransfected with the donor oligonucleotides (in this case 50-mer oligonucleotides were used) and either AG30, 12AG30, or Mix30. Again, the 12AG30 and Mix30 samples served as mismatched oligonucleotide controls for each vector that was tested. In all cases, luciferase activity was standardized via normalization of reaction values by setting the average activity of the plasmid/donor/12AG30 samples equal to 1000 relative light units and then adjusting all other samples accordingly.

In vectors with mutations at distances of 53, 242, and 753 bp, AG30 was seen to induce recombination at frequencies 2.4-, 2.2-, and 1.7-fold, respectively, above the baseline of the donor with mismatched 12AG30 (Figure 5B). However, there was no detectable effect of AG30 at the furthest distance of 1255 bp. In addition, no induced recombination was seen in any of the samples with the Mix30 control oligonucleotides. Overall, results with this vector set indicate that a sequence-specific effect of AG30 on induced recombination can be detected up to 753 bp from the site of third strand binding.

DISCUSSION

The work described here explores the fundamental parameters of binding affinity and distance as they relate to the ability of TFOs to exert biological effects. As such, this work is intended to guide the choice of TFO target sites and to aid in the design of reagents for TFO-induced recombination and genome modification. Using a series of G-rich 30-mer oligonucleotides, we measured the effects of incremental increases in target site mismatches and therefore decreasing

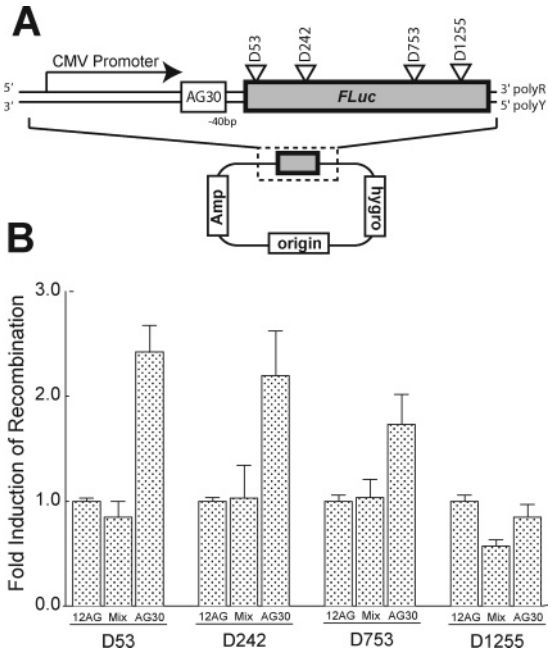

FIGURE 5: TFO-induced recombination in the firefly luciferase gene at various distances. (A) Schematic of plasmid targets. Firefly luciferase is represented by the gray box. The AG30 target site is represented by the boxed AG30 upstream of the firefly luciferase gene. Nonsense mutations were created at base pairs 13, 202, 713, and 1215, at distances of 53, 242, 753, and 1255 bp from the AG30 binding site, as indicated in the coding region of the firefly luciferase gene. The orientation of the polypurine/polypyrimidine duplex target site is indicated by the polyY and polyR labels. The polypyrimidine strand of the duplex AG30 target is on the bottom strand as drawn here (polyY); the polypurine strand is on the top strand as drawn here (polyR). (B) Induction of recombination by AG30 at various distances between the site of *FLuc* gene mutation and the third strand binding site. For each plasmid substrate (D53, D242, D753, and D1255), the samples included donor oligonucleotides mixed with either the specific TFO, AG30, or one of the mismatched oligonucleotides, 12AG30 (labeled 12AG) or Mix30 (labeled Mix). The values are normalized to the results for the 12AG30 samples. The numbers (used to designate each plasmid and preceded by the letter D) indicate the distance (D) between the mutation site and the TFO-binding site. Error bars indicate one standard error.

binding affinity on TFO-induced recombination and repair. In these studies, we demonstrated a direct correlation between binding affinity and induction of both recombination and repair events. In the case of recombination induction, there appears to be a threshold with respect to TFO binding affinity at a K_d of approximately 1×10^{-7} M (under the conditions of our assay). Molecules with weaker binding affinities exhibited minimal or no biological effect. A similar relationship was seen between TFO binding affinity and induced repair. This stringent requirement for high binding affinity limits the number of putative target sites for TFO binding in efforts to achieve genomic modification. However, this stringency also highlights the specificity of TFO action. It should also be noted that all of the oligonucleotides tested had a high G content (23 of 30 nucleotides), yet only those closely matched to the duplex target site sequence had a measurable biological effect. Hence, the induction of recombination and repair cannot simply be attributed to a nonspecific effect of G-rich polynucleotides on DNA metabolism.

In addition, work with a chemically modified DEED TFO further suggests that improved binding affinity may translate into further increases in the efficiency of induced recombination. In the case of DEED-modified oligonucleotides, these molecules have a particular binding advantage at physiologic conditions of low magnesium and high potassium concentrations (*9*).

Using two different episomal reporter vectors, we also examined the ability of TFOs to induce recombination between a plasmid substrate and a single-stranded oligonucleotide donor molecule over varying distances. In one assay, the reporter gene mutation to be corrected was in a fixed position and the TFO target site was varied. In the other, the TFO target site was fixed and the position of the reporter gene mutation was varied. In the first case, the same donor oligonucleotide was used in all samples; in the latter, a different donor molecule was required to provide homology to the varying regions of the target gene. Nonetheless, the two assays yielded similar results, with both suggesting that TFO-induced recombination can be detected over distances ranging from approximately 25 to 750 bp. This raises the important possibility that TFO-induced recombination may eventually be applied to disease-related genes even in cases where suitable polypurine sites for triplex formation are not directly adjacent to or overlapping with the disease-causing mutation. Genomic analyses have suggested that polypurine sites are over-represented in the genome, with sites amenable to triplex formation occurring roughly every 1000 bp (*18*−*20*). Hence, if TFO-induced recombination can take place over several hundred base pairs, then many human disease-associated mutations may be amenable to triplex-mediated modification. A key remaining challenge to be addressed is the current low frequency of TFO-induced recombination. In other studies, we have found that this frequency can vary substantially with the method of transfection, and as seen in the work described above, frequency can also be influenced by chemical modification of the TFO or by the use of TFO analogues such as peptide nucleic acids (PNAs) (*6*, *9*, *15*, *21*, *22*). We expect that continued improvement in methods for intracellular delivery and in the binding properties of TFOs will lead to further increases in the efficiency of TFO-mediated genome modification.

Importantly, the concept that DNA binding molecules such as TFOs can induce recombination at a distance has been established by this work. At this point, the mechanism for this effect over several hundred base pairs is not fully established. We know from previous work that the NER pathway is required for TFO- and PNA-induced recombination (*8*, *15*, *23*, *24*). This suggests that NER-mediated strand breaks and gaps may play a key initiating role in the induction of recombination between a targeted mutation and a homologous oligonucleotide donor. NER is known to recognize and excise bulky adducts and has been shown to act on triplex structures (*8*, *15*, *23*, *24*); however, the canonical NER repair patch is only 24−32 bases long (*25*, *26*). Therefore, a mechanism for extending the effect of the TFO-triggered NER reaction across tens or hundreds of base pairs to the site of gene correction must exist. The single-stranded gap produced by NER could evolve into a double-strand break (DSB) (*27*−*29*) and thereby trigger homologous recombination (HR) (*30*, *31*). It is also possible that nicked intermediates or the NER patch itself may be sufficient for the triggering of HR reactions (*32*). In particular, as in the single-strand annealing pathway (one of several cellular

pathways used in the resolution of genomic DSBs), the single-stranded patch produced by NER could provide localized access to exonucleases that degrade one strand to expose a region homologous to the donor DNA (*29*). Alternatively, the recruited repair activity stimulated by high-affinity TFOs such as AG30 may simply increase the accessibility of the duplex DNA to both the donor oligonucleotide and factors that play a role in HR (*33*, *34*). This could be mediated by alterations in chromatin structure or by conformational changes in the duplex DNA itself. Hence, we propose that either by provoking repair-mediated recombinagenic intermediates or by enhancing local chromosomal accessibility, triplex formation can recruit or promote the activity of factors involved in HR. In this regard, it has already been established that TFO-induced recombination requires RAD51, a member of the RAD52 epistasis group, that is well known to mediate HR (*8*).

In summary, the work presented here indicates that triplex formation of a sufficiently high affinity can stimulate DNA metabolism and that this effect can extend over a range of several hundred base pairs. These findings open up additional possibilities in the use of TFOs for targeted genomic modification.

ACKNOWLEDGMENT

We thank Chip Greaves for synthesis of the DEED oligonucleotide. We thank D. Campisi-Hegan, E. Kosowsky, L. Cabral, and N. Brown for their help. We thank F. Knauert in memoriam.

REFERENCES

1. Knauert, M. P., and Glazer, P. M. (2001) Triplex forming oligonucleotides: Sequence-specific tools for gene targeting, *Hum. Mol. Genet. 10*, 2243−2251.
2. Letai, A. G., Palladino, M. A., Fromm, E., Rizzo, V., and Fresco, J. R. (1988) Specificity in formation of triple-stranded nucleic acid helical complexes: Studies with agarose-linked polyribonucleotide affinity columns, *Biochemistry 27*, 9108−9112.
3. Beal, P. A., and Dervan, P. B. (1991) Second structural motif for recognition of DNA by oligonucleotide-directed triple-helix formation, *Science 251*, 1360−1363.
4. Dervan, P. B., and Burli, R. W. (1999) Sequence-specific DNA recognition by polyamides, *Curr. Opin. Chem. Biol. 3*, 688−693.
5. Faruqi, A. F., Seidman, M. M., Segal, D. J., Carroll, D., and Glazer, P. M. (1996) Recombination induced by triple-helix-targeted DNA damage in mammalian cells, *Mol. Cell. Biol. 16*, 6820−6828.
6. Luo, Z., Macris, M. A., Faruqi, A. F., and Glazer, P. M. (2000) High-frequency intrachromosomal gene conversion induced by triplex-forming oligonucleotides microinjected into mouse cells, *Proc. Natl. Acad. Sci. U.S.A. 97*, 9003−9008.
7. Chan, P. P., Lin, M., Faruqi, A. F., Powell, J., Seidman, M. M., and Glazer, P. M. (1999) Targeted correction of an episomal gene in mammalian cells by a short DNA fragment tethered to a triplex-forming oligonucleotide, *J. Biol. Chem. 274*, 11541−11548.
8. Datta, H. J., Chan, P. P., Vasquez, K. M., Gupta, R. C., and Glazer, P. M. (2001) Triplex-induced recombination in human cell-free extracts. Dependence on xpa and hsrad51, *J. Biol. Chem. 276*, 18018−18023.
9. Vasquez, K. M., Dagle, J. M., Weeks, D. L., and Glazer, P. M. (2001) Chromosome targeting at short polypurine sites by cationic triplex-forming oligonucleotides, *J. Biol. Chem. 276*, 38536−38541.
10. Dagle, J. M., and Weeks, D. L. (1996) Positively charged oligonucleotides overcome potassium-mediated inhibition of triplex DNA formation, *Nucleic Acids Res. 24*, 2143−2149.
11. Parris, C. N., and Seidman, M. M. (1992) A signature element distinguishes sibling and independent mutations in a shuttle vector plasmid, *Gene 117*, 1−5.
12. Glazer, P. M., Sarkar, S. N., Chisholm, G. E., and Summers, W. C. (1987) DNA mismatch repair detected in human cell extracts, *Mol. Cell. Biol. 7*, 218−224.
13. Wang, G., and Glazer, P. M. (1995) Altered repair of targeted psoralen photoadducts in the context of an oligonucleotide-mediated triple helix, *J. Biol. Chem. 270*, 22595−22601.
14. Wang, G., Seidman, M. M., and Glazer, P. M. (1996) Mutagenesis in mammalian cells induced by triple helix formation and transcription-coupled repair, *Science 271*, 802−805.
15. Rogers, F. A., Vasquez, K. M., Egholm, M., and Glazer, P. M. (2002) Site-directed recombination via bifunctional PNA-DNA conjugates, *Proc. Natl. Acad. Sci. U.S.A. 99*, 16695−16700.
16. Bradford, M. M. (1976) A rapid and sensitive method for the quantitation of microgram quantities of protein utilizing the principle of protein-dye binding, *Anal. Biochem. 72*, 248−254.
17. Bennett, M., and Schaack, J. (2003) Development of a dual-luciferase fusion gene as a sensitive marker for site-directed DNA repair strategies, *J. Gene Med. 5*, 723−732.
18. Behe, M. J. (1995) An overabundance of long oligopurine tracts occurs in the genome of simple and complex eukaryotes, *Nucleic Acids Res. 23*, 689−695.
19. Manor, H., Rao, B. S., and Martin, R. G. (1988) Abundance and degree of dispersion of genomic d(GA)n·d(TC)n sequences, *J. Mol. Evol. 27*, 96−101.
20. Vasquez, K. M., Marburger, K., Intody, Z., and Wilson, J. H. (2001) Manipulating the mammalian genome by homologous recombination, *Proc. Natl. Acad. Sci. U.S.A. 98*, 8403−8410.
21. Faruqi, A. F., Krawczyk, S. H., Matteucci, M. D., and Glazer, P. M. (1997) Potassium-resistant triple helix formation and improved intracellular gene targeting by oligodeoxyribonucleotides containing 7-deazaxanthine, *Nucleic Acids Res. 25*, 633−640.
22. Lacroix, L., Lacoste, J., Reddoch, J. F., Mergny, J. L., Levy, D. D., Seidman, M. M., Matteucci, M. D., and Glazer, P. M. (1999) Triplex formation by oligonucleotides containing 5-(1-propynyl)-2′-deoxyuridine: Decreased magnesium dependence and improved intracellular gene targeting, *Biochemistry 38*, 1893−1901.
23. Faruqi, A. F., Datta, H. J., Carroll, D., Seidman, M. M., and Glazer, P. M. (2000) Triple-helix formation induces recombination in mammalian cells via a nucleotide excision repair-dependent pathway, *Mol. Cell. Biol. 20*, 990−1000.
24. Vasquez, K. M., Christensen, J., Li, L., Finch, R. A., and Glazer, P. M. (2002) Human XPA and RPA DNA repair proteins participate in specific recognition of triplex-induced helical distortions, *Proc. Natl. Acad. Sci. U.S.A. 99*, 5848−5853.
25. Huang, J. C., Svoboda, D. L., Reardon, J. T., and Sancar, A. (1992) Human nucleotide excision nuclease removes thymine dimers from DNA by incising the 22nd phosphodiester bond 5′ and the 6th phosphodiester bond 3′ to the photodimer, *Proc. Natl. Acad. Sci. U.S.A. 89*, 3664−3668.
26. Moggs, J. G., Yarema, K. J., Essigmann, J. M., and Wood, R. D. (1996) Analysis of incision sites produced by human cell extracts and purified proteins during nucleotide excision repair of a 1,3-intrastrand d(GpTpG)-cisplatin adduct, *J. Biol. Chem. 271*, 7177−7186.
27. Strathern, J. N., Weinstock, K. G., Higgins, D. R., and McGill, C. B. (1991) A novel recombinator in yeast based on gene II protein from bacteriophage f1, *Genetics 127*, 61−73.
28. Arcangioli, B. (1998) A site- and strand-specific DNA break confers asymmetric switching potential in fission yeast, *EMBO J. 17*, 4503−4510.
29. Obe, G., Pfeiffer, P., Savage, J. R., Johannes, C., Goedecke, W., Jeppesen, P., Natarajan, A. T., Martinez-Lopez, W., Folle, G. A., and Drets, M. E. (2002) Chromosomal aberrations: Formation, identification and distribution, *Mutat. Res. 504*, 17−36.
30. Szostak, J. W., Orr-Weaver, T. L., Rothstein, R. J., and Stahl, F. W. (1983) The double-strand-break repair model for recombination, *Cell 33*, 25−35.
31. Sun, H., Treco, D., Schultes, N. P., and Szostak, J. W. (1989) Double-strand breaks at an initiation site for meiotic gene conversion, *Nature 338*, 87−90.
32. Lee, G. S., Neiditch, M. B., Salus, S. S., and Roth, D. B. (2004) RAG proteins shepherd double-strand breaks to a specific pathway, suppressing error-prone repair, but RAG nicking initiates homologous recombination, *Cell 117*, 171−184.

33. Macris, M. A., and Glazer, P. M. (2003) Transcription dependence of chromosomal gene targeting by triplex-forming oligonucleotides, *J. Biol. Chem. 278*, 3357−3362.
34. Majumdar, A., Puri, N., Cuenoud, B., Natt, F., Martin, P., Khorlin, A., Dyatkina, N., George, A. J., Miller, P. S., and Seidman, M. M. (2003) Cell cycle modulation of gene targeting by a triple helix-forming oligonucleotide, *J. Biol. Chem. 278*, 11072−11077.

BI0481040