5-2006

Gene modification with triplex-forming oligonucleotides

Jennifer Melissa Kalish
Yale University.

Gene Modification with Triplex-Forming Oligonucleotides

A Dissertation
Presented to the Faculty of the Graduate School
of
Yale University
in Candidacy for the Degree of
Doctor of Philosophy

by
Jennifer Melissa Kalish

Dissertation Director: Peter M. Glazer, MD, PhD

May, 2006

Abstract
Gene Modification with Triplex-Forming Oligonucleotides
Jennifer M. Kalish
2006

Triplex-forming oligonucleotides (TFOs) bind DNA in a sequence-specific manner and mediate targeted genome modification. Triplexes bind parallel to the purine strand of the duplex in the pyrimidine motif and anti-parallel in the purine motif. TFOs in both motifs have been shown to direct psoralen adduct formation in cells, leading to mutagenesis or recombination. Only purine motif TFOs have been shown to mediate genome modification without using a DNA-adduct. In this work, we report the ability of a series of chemically modified pyrimidine motif TFOs to induce repair and recombination in the absence of any DNA-reactive conjugate. We find that a number of modified TFOs show increased formation of non-covalent triplexes under physiological conditions. However, only the N3'->P5' phosphoramidate and 5-(1-propynyl)-2'-deoxyuridine-modified TFOs mediate induced recombination in cells and stimulate repair in cell-free extracts. Additionally, we establish the ability of TFOs to induce gene correction at a single-copy chromosomal locus in mammalian cells at frequencies up to 0.1%. We demonstrate sequence, target site, and dose-dependent effects. Finally, we develop an assay to detect incisions made in repair of a triplex. This work expands the number of targets available for triplex-mediated gene targeting, shows effective targeting at a chromosomal locus, and takes first steps towards understanding the mechanism of triplex-induced gene correction.

Gene Modification with Triplex-Forming Oligonucleotides

A Dissertation
Presented to the Faculty of the Graduate School
of
Yale University
in Candidacy for the Degree of
Doctor of Philosophy

by
Jennifer Melissa Kalish

Dissertation Director: Peter M. Glazer, MD, PhD

May, 2006

Dedication

This work is dedicated to my family, each member of which had a specific role in helping me to complete my PhD. To my mother, Michele Kalish, for her tireless proof-reading and never-ending support. To my father, Murray Kalish, for always believing in me. To my sister, Danielle Kalish for a friendship with a truly kind person. To my grandmother, Teresa Marcus, and my great-grandmother, Beatrice Fine Levin for their inspiration and the beginnings of my quest for understanding a small part of the natural world. To my husband, Yechiel Schur, for being my partner in love, in life, and in the acquisition of knowledge. And to my daughter, Gavriela Beatrice Kalish-Schur, for reintroducing me to the world around us as she discovers a little more of it each day.

Acknowledgements

I would like to thank the members of the Glazer lab for their amazing support throughout my years in the lab. This work was supported by a grant to J.M.K. from NIH Medical Scientist Training Program Grant (GM07205).

Table of Contents and Figure Listing

List of Abbreviations

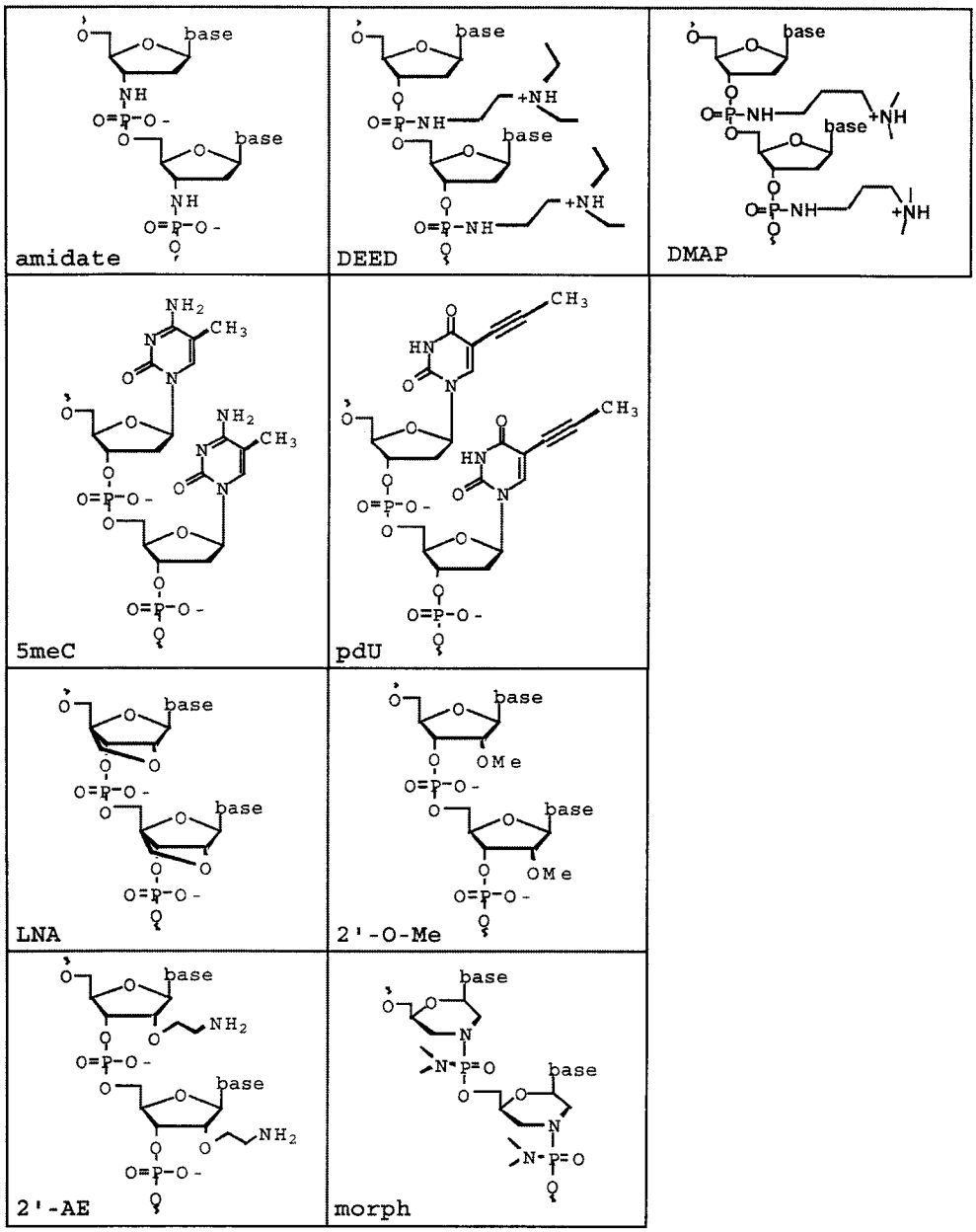
AAV	Adeno-Associated Virus
2'-AE	2'-O-(2-aminoethyl)-ribose
ATCC	American Type Culture Collection
amidate	N3'->P5' phosphoramidate
bp	base pairs
C	Cytidine
CPT	Camptothecin
CS	Cockayne syndrome
ddA	Dideoxyadenine
DEED	N, N-Diethylethylenediamine
DMAP	N, N-Dimethylaminopropylphosphoramidate
FBS	Fetal Bovine Serum
Fluc	*Firefly luciferase*
FRT	Flp Recombinase Target
GGR	Global Genome Repair
Hprt	*Hypoxanthine Guanine Phosphoribosyl Transferase*
HR	Homologous Recombination
IPTG	Isopropyl-β-D-Thiogalactopyranoside
K_d	Equilibrium Dissociation Constant
LAR II	Luciferase Assay Reagent II
LNA	2'-O, 4'-C-methylene bridged or locked nucleic acid
5meC	5-methyl-2'-deoxycytidine
2'-O-Me	2'-O-methyl-ribose
morph	morpholino
NER	Nucleotide Excision Repair
NMR	Nuclear Magnetic Resonance
nt	nucleotides
PBS	Phosphate Buffered Saline
PCV	Packed Cell Volumes
pdU	5-(1-propynyl)-2'-deoxyuridine
PNA	Peptide nucleic acid
Psoralen	4'-hydroxymethyl-4, 5', 8-trimethylpsoralen
RPA	Replication Protein A
RLUs	Relative Light Units
ssDNAs	single-stranded donor DNA oligonucleotides
SV40	Simian Virus 40
TBS	TFO Binding Site
TC-NER	Transcription-Coupled Repair
TFO	Triplex-Forming Oligonucleotide
TINA	(R)-1-O-[4-(1-pyrenylethynl)phenylmethyl]glycerol, Twisted Intercalating Nucleic Acid
TK	(Herpes Simplex Virus) Thymidine Kinase
U	Uridine

List of Abbreviations (continued)

UV	Ultraviolet Radiation
X-gal	5'-Bromo-4-Chloro-3-Indolyl-β-D-Galactopyranoside
XPA	Xeroderma Pigmentosum A
ZFN	Zinc Finger Nuclease

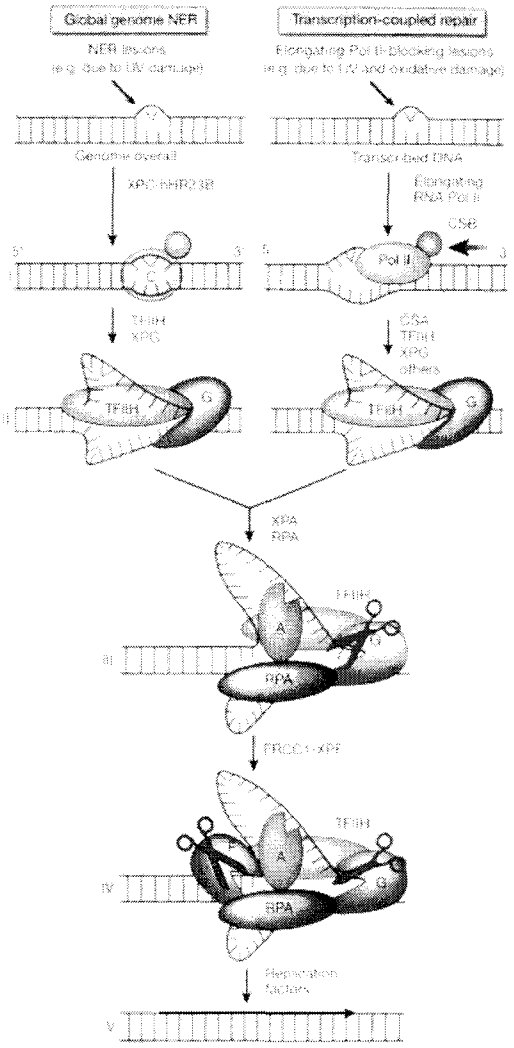
Appendix 1

Chemical Structures



Structures of the chemical modifications. N3'->P5' phosphoramidate (amidate); N, N-diethylethylenediamine (DEED); N, N-dimethylaminopropylphosphoramidate (DMAP); 5-methyl-2'-deoxycytidine (5meC); 5-(1-propynyl)-2'-deoxyuridine (pdU); 2'-O, 4'-C-methylene bridged or locked nucleic acid (LNA); 2'-O-methyl-ribose (2'-O-Me); 2'-O-(2-aminoethyl)-ribose (2'-AE); morpholino (morph).

Appendix 2



Nucleotide excision repair pathway. The parallel pathways of global genome repair (GGR) and transcription-coupled repair (TC-NER) recognize helically distorting lesions in DNA and polymerase blocking lesions in transcribed regions of DNA, respectively. In GGR, XPC-hHR23B recognizes helical distortions and in TC-NER lesions are recognized by blockage of RNA polymerase II. In GGR, TFIIH and possibly XPG are attracted to XPC-hHR23B bound to the lesion. TFIIH creates an open complex via XPB and XPD, leading to the recruitment of XPA and RPA, and to the common incision pathway. In TC-NER, CSA, CSB, TFIIH, XPG, and possibly other factors, replace RNA polymerase II, leading to the recruitment of XPA and RPA, and to the common incision pathway. The incision pathway consists of XPG making the 3' incision and ERCC1-XPF making the 5' incision. After the dual incision is completed, gap-filling DNA synthesis occurs followed by ligation. This is an adaptation of a figure from de Laat, W.L., et al., (1999). Genes and Development 13: p. 768-785.

Chapter One:
Introduction to Triplex-Forming Oligonucleotides

Triple helix (triplex) formation was first described in 1957 by Felsenfeld and colleagues, who while studying polypurine/polypyrimidine duplex interactions, noted the binding of a second polypyrimidine strand to the duplex [1]. Triplex-forming oligonucleotide (TFO) binding occurs under physiologic conditions and in a sequence specific manner [2-5]. The target specific binding of TFOs has been used to deliver DNA interacting molecules to a specific genetic location. While binding to the double helix, TFOs alter the helix and induce cellular repair mechanisms. These repair mechanisms lead to TFO alteration of gene metabolism via transcription repression, inhibition of DNA replication, and induction of site-specific mutagenesis and recombination [6-11]. Thereby, TFOs have the potential to introduce permanent changes in genomic DNA in a variety of contexts.

Triplex Chemistry and Binding Code

TFOs bind DNA in the major groove of DNA in a sequence-specific manner to homopurine sequences in two distinct motifs [2-4, 12]. Triplex formation is favored at polypurine/polypyrimidine regions of DNA and can occur with the third strand oriented either parallel or anti-parallel to the purine strand of the duplex. TFOs bind to duplex DNA using an alternative set of hydrogen bonds called Hoogsteen bonds [1]. Purine motif TFOs bind anti-parallel to the duplex via reverse-Hoogsteen hydrogen bonds and pyrimidine motif TFOs bind parallel via Hoogsteen bonds [2-4, 12]. In the anti-parallel,

purine motif, the triplets are G.G:C and A.A:T; whereas in the parallel, pyrimidine motif, the canonical triplets are C^{+}.G:C and T.A:T (Fig.1). In the pyrimidine motif, protonation of the cytidine is required for triplex formation [3, 13, 14].

Chemical Modifications

To enhance the biological activity of TFOs there has been a focus on chemical modifications (Appendix 1). Under physiologic conditions, potassium levels are high, magnesium levels are low, and pH is neutral. High potassium promotes G-quartet formation, which inhibits the activity of G-rich purine motif TFOs [15-17]. Magnesium supports third-strand binding by charge neutralization. Neutral pH disfavors cytosine protonation, which is needed for triplex formation in the pyrimidine motif, unless the cytidines are modified [3, 13, 14]. As a result, a variety of chemical modifications have been explored to offset the effects of these factors on TFO binding and thereby increase binding affinity and enhance potential biological activity of the TFOs. These modifications include sugar, backbone, and base modifications.

Purine Motif Modifications

Chemical modifications such as N, N-diethylethylenediamine (DEED) have been used to enhance binding and thereby the biological activity of TFOs in the purine motif [18-22]. The G-rich nature of purine motif TFOs allows them to form G-quartets under physiological conditions which include high levels of potassium [15-17].

7-deazaxanthine and 6-thioguanine substitutions have been used to offset the diminished TFO binding in the presence of high levels of potassium (due to the formation of G-quartets), and such molecules have shown increased effectiveness in cells [23-25].

Pyrimidine Motif Modifications

For triplex formation in the pyrimidine motif, an acidic pH is needed for N3 protonation of cytidines (C) [14]. Therefore, TFOs with several adjacent C's require a lower pH for binding to duplex DNA than TFOs with dispersed C's [14, 26]. TFOs with 5-methyl-2'-deoxycytidine (5meC) [27-29] and 5-(1-propynyl)-2'-deoxyuridine (pdU) [29] substitutions were shown to alleviate the need for an acidic pH and to improve the effectiveness of mutagenesis targeted by TFO-conjugated delivery of 4'-hydroxymethyl-4, 5', 8-trimethylpsoralen (psoralen) to specific sites. Other demonstrated base substitutions for cytidine include: pyrrolidino pseudoisocytidine [30], 8-oxoadenine [31, 32], N7-2'-deoxyguanosine [33], 7,8-dihydro-8-oxoadenine [34], 8-oxo-2'-deoxyadenosine [35, 36], 2-aminopyridine [37], and 8-aminoguanine [38]. TFOs containing these modified bases have enhanced binding affinity at neutral pH compared to unmodified TFOs [30-38].

Sugar substitutions such as 2'-O-aminoethyl (2'-AE) [39, 40] and 2'-O-methoxy (2'-OMe) [41, 42] enhance TFO stability and affinity for binding duplex DNA. Work by Michael Seidman's group showed that 2'-O-aminoethyl and 2'-O-methoxy substitutions on psoralen-conjugated TFOs in the pyrimidine motif enhance gene knock-out in mammalian cells [40]. More recent work from this group demonstrated that only a

minimal number of 2'-O-aminoethyl substitutions clustered together was sufficient for enhanced gene knock-out activity [43, 44].

Backbone modifications have also been extensively used to augment TFO binding affinity and intracellular activity through reduction of charge interactions [45]. Replacement of the bridging N3 oxygen with a nitrogen improved TFO binding *in vitro* [46] and when conjugated to psoralen these TFOs were used to target integrated HIV-1 provirus within cell nuclei [47]. The TFO binding and psoralen delivery to chronically infected cells containing the HIV-1 provirus was demonstrated using a restriction protection assay, where DNA digestion was prevented when the psoralen adduct was present at the correct target site[47]. Replacement of a non-bridging oxygen with amino groups such as N, N-diethylethylenediamine (DEED), or N, N-dimethylamino-propylphosphoramidate (DMAP) as chemical modifications on the third strand have been used to enhance triplex formation under physiologic conditions [48]. Morpholino-modified TFOs have both a sugar and a backbone modification and can also form stable triplexes [49].

Peptide Nucleic Acids

The peptide nucleic acid (PNA) resembles DNA, with purine and pyrimidine bases being attached to a polyamide backbone. PNAs can be used to mediate gene modification and bind to DNA in several different ways (Fig. 2). PNAs can bind through strand invasion to one or both of the strands of the duplex [50, 51]. This binding is via Watson-Crick base pairing between the PNA and the DNA duplex and disrupts the DNA

duplex binding to itself [52]. PNAs can also form triplexes with a strand invading PNA conjugated to a flexible linker molecule with an additional sequence that can form Hoogsteen bonds with the PNA-DNA duplex [52]. This PNA clamp binds to a single strand of duplex DNA via both Watson-Crick base pairing and Hoogsteen base pairing [52]. PNAs have a greater binding affinity as third strands than unmodified TFOs of the same sequence [53]. Mixed sequence (purine and pyrimidine) PNAs have been developed, which are an advantage over TFOs that only bind to AG-rich or TC-rich stretches of DNA [54]. The PNA clamp was demonstrated to induce XPA-dependent DNA repair and recombination [52]. In order to expand the number of target sequences for PNA binding, tail-clamp PNAs which have an additional mixed (purine and pyrimidine) sequence as part of the strand invasion sequence have been developed and were shown to inhibit transcription [55, 56].

TFO-Targeted Gene Modification

TFOs as a tool for delivery of DNA reactive agents

TFOs have been used to deliver DNA reactive molecules to specific locations on the DNA (Fig. 3A). In this regard, TFOs represent a means of effective gene modification. The DNA-reactive agents include psoralen [57, 58], camptothecin (CPT) [59, 60], padlock TFOs [61-63], [5-(125)I]dCMP [64, 65], and acridine [66, 67]. Based on the potential for molecular recognition of unique genomic sites, Havre et al. used a TFO to deliver psoralen in a sequence-specific manner [57, 58]. TFO delivery of

psoralen followed by photo-activation of the psoralen, induced site-specific mutagenesis in viral DNA within mammalian cells [57, 58]. Activation of gene expression has been demonstrated by pso-TFOs in a manner dependent on photo-crosslinking [68].

Topoisomerase I is a DNA-relaxation enzyme that is an important therapeutic target for CPTs in cancer chemotherapy. Camptothecins as agents alone do not bind to specific DNA regions, leading to drug side effects induced by the non-specific binding. CPT conjugation to a DNA targeting agent such as a TFO shows promise of increasing the specificity of drug delivery. Binding of these CPT-TFO conjugates to duplex DNA has been demonstrated *in vitro* and their cellular activity has been demonstrated [59, 60].

Padlock TFOs are TFOs with extra sequences (complementary to each other) on either side of the polypurine/polypyrimidine TFO binding sequence, the ends of which are ligated together with T4 ligase after TFO binding [62]. This end-ligation forms a loop around the plasmid DNA containing the TFO binding site (TBS) [62] and has been used to conjugate targeting peptides to plasmids [61] and to inhibit transcription elongation [63]. TFOs are also used to deliver [5-(125)I]dCMP which, after decay accumulation, causes double strand breaks in the targeted DNA [64, 65].

Binding Affinity Correlates with Intracellular Activity

Wang et al. showed that psoralen-conjugated TFOs transfected into monkey COS-7 cells can induce base pair-specific mutations within the *supF* reporter gene in a simian virus 40 (SV40) genome in cells, at a frequency in the range of 1%-2% [11]. The key

finding in this seminal work was that the binding affinity of the TFO to its target site, as measured *in vitro*, was highly correlated with its intracellular activity. TFOs with K_d's (equilibrium dissociation constants) in the range of $1x10^{-9}$ M were active; those with K_d's of $1x10^{-6}$ M were not. A more specific threshold has been determined for induced DNA recombination and repair, where a binding affinity greater than $1x10^{-7}$ M is necessary for induced recombination in cells and induced repair in cell-free extracts [18]. Binding affinity *in vitro* is influenced by Mg^{2+} and K^{+} concentrations and pH while *in vivo* chromatin structure and transcriptional activity are also important factors. Transcription through the target region substantially enhances TFO chromosomal targeting, suggesting that gene modification with TFOs may be more effective at chromosomal loci that are highly expressed [69].

The binding affinity of a TFO in the anti-parallel, purine motif is not only influenced by the number of mismatches at the binding site but is also influenced by TFO length. TFOs of length 30, 20, and 10 designed to bind to all or a portion of a 30 base pair polypurine site have K_d's of $2x10^{-8}$ M, $3x10^{-7}$ M and $3x10^{-5}$ M, respectively [70]. Correspondingly, the amount of induced repair in HeLa cell-free extracts decreased with TFO length (the 30-mer in the purine motif demonstrated the greatest induced repair, while the 20-mer and then the 10-mer showed lesser effects) [70]. The pattern of mutations produced by TFO-targeted psoralen adducts on an episomal target in COS-7 monkey cells is also influenced by the length of the TFO [71]. Additionally, excision of TFO-damaged DNA is inhibited by a linked 30-mer TFO (whereby the triplex extension

exceeds the nucleotide excision repair (NER) canonical excision patch) but not by a 10-mer TFO [71].

TFO-Inhibition of Transcription

Transcription has been inhibited by TFOs binding to a variety of promoters and other targets in and near genes of interest. TFOs were shown to reduce *c-myc* expression [6, 72, 73] and to inhibit Tat-dependent *in vitro* transcription [74]. Inhibition of transcription is dependent on the accessibility of the TFO binding site (TBS) when the DNA is bound in the nucleosome [21] and TFOs have been shown to access binding sites within the chromatin structure [47]. TFOs covalently bound to duplex DNA and PNAs through strand invasion inhibit replication of plasmid DNA *in vivo* [75]. Chemically modified TFOs were also used to inhibit transcription. 2'-O-methyl-modified TFOs binding to DNA block recognition of duplex DNA by a eukaryotic transcription factor [76] and LNA-modified TFOs inhibit transcription *in vitro* [77]. Phosphoramidate modified-TFOs have been shown to inhibit transcription of several HIV genes [78-80]. Dual conjugation to phenylacetate mustard enhanced inhibition of *Her2* gene expression [81]. Psoralen-conjugated TFOs were used to inhibit a gene involved in cutaneous imflammation [82], the human rhodopsin gene [83], and *Ets2* transcription in prostate cancer cells [84]. These successful inhibitions suggest that TFOs may be useful as agents to treat skin diseases, ocular defects, and prostate cancer.

PNAs have been shown to arrest transcription elongation by invading double-stranded DNA and forming a stable (PNA)2/DNA complex [85]. Conjugation of a PNA

to a peptide containing eight lysines increased the inhibition of transcription [86] and lactose-modified PNAs had increased cellular update and inhibition of transcription [87]. PNA triplexes presented a stable barrier to DNA polymerase extension [88]. Tail-clamp PNAs have been shown to bind duplex DNA and inhibit transcription [55]. PNAs have also been used to deliver peptides to act as transcription activators [89].

TFO-Directed Mutagenesis

Knock-out of chromosomal genes mediated by psoralen-TFOs has been demonstrated. In one study, the *supF* reporter gene integrated into the chromosome of mouse fibroblasts was used as a target [90]. Again, only high-affinity TFOs were active, achieving targeted mutagenesis frequencies of 0.1%. In the *supF* experiments, essentially unmodified G-rich oligonucleotides (except for 3'-end-capping) designed to bind in the anti-parallel motif were used. A set of experiments to target the *hprt* gene in hamster CHO cells, in contrast, used a series of T-rich psoralen TFOs (because the A-rich target favored the parallel motif) [42]. In these experiments, the unmodified TFOs were ineffective. A second conjugation of the TFO to an intercalator, either acridine or pyrene, was needed to provide additional binding affinity [42]. Such doubly modified TFOs yielded *hprt* mutagenesis at frequencies in the range of 10^{-3}. More recent studies have demonstrated that pyrimidine TFOs substituted with 2'-aminoethyl sugar residues are highly active in targeting psoralen adducts and thereby mutations to the *hprt* gene, with frequencies in the range of 0.14% [43]. Psoralen conjugated PNAs were effective in

inducing mutagenesis as well [91]. TFO-induced mutagenesis was also seen secondary to double strand breaks with [5-(125)I]dCMP labeled TFOs [64, 65].

However, the mutagenesis induced in the *supF* and *hprt* chromosomal targets, while site-specific in the sense that all the mutations clustered around the third-strand binding site, was somewhat variable. Hence, these experiments suggested that psoralen-coupled TFOs may be useful for gene-specific knock-out but not necessarily for predictable base pair-specific mutagenesis. As a research tool, however, psoralen TFOs have served as useful reagents to prove the ability of oligonucleotides to bind as third strands to chromosomal sites in living cells. The importance of these results is to establish the concept that DNA binding molecules can be used to direct site-specific genome modification and to show that the cell and nuclear membranes and the packaging of the DNA into chromatin are not insurmountable barriers to gene targeting with antigene oligonucleotides.

TFO-Directed Mutagenesis without Psoralen Conjugation

In the course of the work with psoralen TFOs, Wang et al. observed that unconjugated TFOs were also capable of inducing mutations in the target gene, at least when the binding affinity was sufficiently high (Fig. 3A) [70]. TFOs can induce mutations in a sequence-specific manner in somatic cells in adult mice [10]. Transgenic mice carrying multiple copies of lambda vector DNA in their genome were injected intraperitoneally with a G-rich TFO specific to the target site (Fig. 4) [10]. TFO-induced mutagenesis was detected in a wide variety of tissues including skin, liver, kidney, and

lung [10]. Again, the mutations were generally within the TFO target site. PNAs have also been used to target the chromosomal *supF* reporter gene, in mouse cells, yielding frequencies of mutagenesis of 0.1% (10-fold above background) [92]. PNA-induced mutagenesis was target specific with the majority of mutations falling within the PNA binding site.

TFO-Induced Recombination

Based on the concept that third-strand binding, with or without psoralen coupling, can trigger DNA repair, Faruqi et al. hypothesized that such binding might also be recombinogenic due to the production of repair-dependent DNA strand breaks [8, 9]. To test this, two distinct recombination assays were used, intramolecular and intermolecular (Fig. 3B). Using an SV40 vector containing two mutant copies of the *supF* gene (Fig. 5A), they found that both psoralen conjugated purine motif TFOs [8] and TFOs without psoralen can stimulate intramolecular recombination [9]. This induced recombination in the case of the TFO without psoralen was found to be dependent on the presence of functional XPA protein [9].

These results were extended to a chromosomal target, in which two mutant *thymidine kinase* (TK) genes were integrated as direct repeats into a single chromosomal site in mouse fibroblasts (Fig. 5B) [93]. High-affinity TFOs targeting a region in between the two TK genes yielded recombination at a frequency of approximately 10^{-4}, about 100-fold above background. When the TFOs were micro-injected into the nuclei of the cells (≈2000 copies/cell), the yield of recombinants increased to 1%-2%, 10,000-fold

over background [94]. Analysis of the recombinant clones revealed all the recombination events involved gene conversion rather than crossover recombination [93].

Repair Pathways and TFO Effects

The TFO mutagenesis effect was shown to be a consequence of the stimulation of DNA repair by the formation of a triple helix, which can be recognized by the nucleotide excision repair (NER) complex as a "lesion" (Appendix 2). TFO-induced mutagenesis was not detected in cells deficient in XPA (a key NER recognition protein) but was restored in cells that were complemented with XPA cDNA[70]. Similarly, triplex-induced mutagenesis was not detected in Cockayne's syndrome (CS) group B cells (cells deficient in transcription-coupled repair) but was detected in the cDNA corrected cell line [70]. Thus, both the overall NER pathway and the transcription-coupled repair pathway play a role in triplex-induced mutagenesis. Gene activation by TFOs was unchanged in XPA, XPD and XPG deficient cell lines, thus gene activation is dependent on other repair pathways as well [68].

Intermolecular recombination between plasmid targets and short DNA fragments can be detected in human cell-free extracts, in a process that is dependent on XPA and Rad51 [95]. TFO-induced recombination in an episomal target in cells is also dependent on the NER pathway [9]. The repair and replication complex, replication protein A (RPA) recognizes triplex structures but also forms nonspecific DNA aggregates at higher concentrations [96]. In the presence of RPA, XPA binds to the triplex structure and

decreases the non-specific binding of RPA [96]. Therefore, the synergy of RPA and XPA confers specificity and stringency for triplex lesion binding, and likely represents a key recognition step in the repair of triplex structures.

Donor Effects

The observation of the ability of third-strand binding to provoke DNA repair and stimulate recombination led to the development of a strategy to mediate targeted gene conversion using a TFO linked to a short DNA fragment homologous to the target site (except for the base pair to be corrected) (Fig. 3B) [97]. In this bi-functional molecule, the TFO domain mediates site-specific binding to target the donor molecule to the desired gene [97]. This binding also triggers repair to sensitize the target site to recombination [97]. The tethered homologous donor fragment can participate in recombination and/or gene conversion with the target gene to correct or alter the nucleotide sequence.

Using a bi-functional oligomer with a 40-mer donor domain and a 30-mer TFO domain, correction of a single base pair mutation in the *supF* reporter gene within a SV40 vector in COS-7 cells was achieved [97]. Correction frequencies were in the range of 0.1% with the full bi-functional molecule [97]. Oligonucleotides consisting of either domain above or of either domain substituted with heterologous sequences reduced activity by 10-fold or more[97]. The donor domain alone consistently did mediate some gene correction, as would be expected, based on the known ability of DNA fragments to

mediate some level of recombination. However, there was a clear synergism due to combination with the TFO domain [97].

When measuring induced recombination in HeLa cell-free extracts, tethered and untethered TFOs and donors showed similar recombination frequencies [95]. This unlinking allowed for further exploration of the distance over which TFO binding could exert an effect on recombination by a short donor DNA molecule. A similar PNA was coupled with a 40-mer DNA donor molecule and induced recombination was measured in cell-free extracts [52]. Correction of a mutated reporter gene was found to be over 60-fold above background [52] when the PNA and donor were linked but was also demonstrated without covalent linkage.

Distance and TFO Effects

Given that *in vivo* there may not always be a TFO target site adjacent to the desired site of gene modification, the distance constraints between the TFO binding site and the mutation site to be corrected by recombination were studied. The episomal target with the dual *supF* genes described above had the upstream gene 4 base pairs from the TFO binding site and the downstream gene 29 base pairs from the site (Fig. 5A) [8]. In the chromosomal TK assay, the TFO binding site was 0.8 kilo bases and 0.7 kilo bases away from the genes, respectively (Fig. 5B) [94]. Recent work demonstrated in two episomal reporter systems that TFO-induced recombination was detectable at distances (between the TFO target site and the mutation to be corrected) from 24 base pairs to 750

base pairs (Fig. 5C) [18]. In both systems, recombination between the mutated gene on the episomal target and a wild-type donor fragment in the presence of TFO was measured [18].

TFO Delivery

While the TFOs are active gene targeting molecules, delivery is a key factor in the biological activity of TFOs. Chromatinization of the DNA target, intracellular ion concentrations, and cytoplasmic versus nuclear separation of the TFO and DNA target are hurdles that have to be overcome in delivery of a TFO into a cell. TFOs have been delivered into cells via lipid mediated TFO delivery [9, 98, 99] and electroporation [11, 43]. Cells have been directly injected [93] and mice have been intraperitoneally injected with TFOs [10]. Using cationic lipids as transfection reagents, compared to microinjection, showed much lower frequencies of recombination [93]. Cell membrane permeabilization by digitonin [47, 69] or streptolysin-O [92] has also been used. Conjugation of TFOs and PNAs to a transport peptide derived from antennapedia has also been explored [100-102]. Such conjugation increased mutation targeting frequencies in a chromosomal reporter gene by 20-fold compared to TFO alone, in a dose responsive manner [101]. This effect was target specific and not due to the antennapedia peptide alone [101]. Confocal microscopy demonstrated increased cellular uptake of the conjugated TFO [101]. Taken together, these data suggest that TFO conjugation to antennapedia increases intranuclear delivery of TFOs.

Specific Aims

Both purine and pyrimidine TFOs have been used to direct psoralen to a specific target, but only purine TFOs have been shown to induce genomic changes through repair and recombination when noncovalently bound to the duplex [9]. Previous work in our lab has largely focused on the purine motif, showing that purine TFOs are active in repair and recombination in cell-free extracts, in cell culture, and in mice [8, 10, 71, 92]. Canonical nucleotides do not bind with high affinity in the pyrimidine motif under physiological conditions due to the requirement for cytosine protonation in the formation of Hoogsteen bonds [103]. In order to make pyrimidine TFOs a more effective tool, chemical modifications need to be used to enhance triplex stability *in vivo*. Modifications such as methylated cytosine, base substitutions, or sugar substitutions improve affinity in the pyrimidine motif and have been used for TFO-mediated-delivery of DNA intercalators [29, 43].

In Chapter 2, a series of chemical modifications was examined for their ability to induce recombination and repair in the pyrimidine motif. The most effective of these modifications were the N3'->P5' bridging phosphoramidate (amidate) modification and the 5-(1-propynyl)-2'-deoxyuridine base substitution. Both showed enhanced *in vitro* TFO binding, increased repair in cell-free extracts, and increased recombination in episomal targets in an XPA-dependent manner. Several other chemical modifications showed enhanced binding but did not show increased repair or recombination.

These chemically modified TFOs were compared to purine motif TFOs for their ability to induce recombination and repair in a single-copy genomic target. Previously, the ability of TFOs to induce recombination between a target locus and a donor DNA was only demonstrated with multi-copy episomal targets in mammalian cells. In Chapter 3, we report that the amidate-modified TFO showed similar induced recombination frequencies compared to the purine TFO in a single-copy chromosomal target. The induced recombination was sequence-, target-, and dose-dependent with correction frequencies up to 0.1%.

The final goal of this work was to create altered helical structures and to examine where the incisions are made in repair of these structures. Previous work in the Glazer lab and work presented in Chapter 2 determined that triplex-induced repair and recombination requires XPA, a NER recognition factor [95, 104, 105]. A 10 nucleotide TFO, conjugated to psoralen, and monoadducted to duplex DNA produces excision products, while a 30 nucleotide TFO does not produce excision products [106]. The nature of these excision products and the location of the incisions have not been previously evaluated. This idea is also consistent with the fact that the length of attachment of some TFOs (i.e. TFOs 30 nucleotides long) exceeds the canonical excision span of NER. In order to understand the mechanism of triplex action and to elucidate the pathway by which triplexes promote repair and recombination, the incision sites in repair of the bulky adduct created by triplex binding need to be mapped. TFOs of different lengths will be tested to determine the length limitation for NER to excise the triplex. In

Chapter 4, altered helical structures using a variety of substrates (linear and plasmid) were evaluated in conjunction with TFOs to help develop an assay to understand the nature of incisions made in repair of a triplex.

References

1. Felsenfeld, G., D.R. Davies, and A. Rich, *Formation of a three-stranded polynucleotide molecule.* J. Am. Chem. Soc., 1957. **79**: p. 2023-2024.

2. Moser, H.E. and P.B. Dervan, *Sequence specific cleavage of double helical DNA by triple helix formation.* Science, 1987. **238**: p. 645-650.

3. Letai, A.G., et al., *Specificity in formation of triple-stranded nucleic acid helical complexes: studies with agarose-linked polyribonucleotide affinity columns.* Biochemistry, 1988. **27**(26): p. 9108-12.

4. Beal, P.A. and P.B. Dervan, *Second structural motif for recognition of DNA by oligonucleotide-directed triple-helix formation.* Science, 1991. **251**(4999): p. 1360-3.

5. Le Doan, T., et al., *Sequence-specific recognition, photocrosslinking and cleavage of the DNA double helix by an oligo-[alpha]-thymidylate covalently linked to an azidoproflavine derivative.* Nucleic Acids Res, 1987. **15**(19): p. 7749-60.

6. Cooney, M., et al., *Site-specific oligonucleotide binding represses transcription of the human c-myc gene in vitro.* Science, 1988. **241**(4864): p. 456-9.

7. Birg, F., et al., *Inhibition of simian virus 40 DNA replication in CV-1 cells by an oligodeoxynucleotide covalently linked to an intercalating agent.* Nucleic Acids Research, 1990. **18**(10): p. 2901-8.

8. Faruqi, A.F., et al., *Recombination induced by triple-helix-targeted DNA damage in mammalian cells.* Mol Cell Biol, 1996. **16**(12): p. 6820-8.

9. Faruqi, A.F., et al., *Triple-helix formation induces recombination in mammalian cells via a nucleotide excision repair-dependent pathway.* Mol Cell Biol, 2000. **20**(3): p. 990-1000.

10. Vasquez, K.M., L. Narayanan, and P.M. Glazer, *Specific mutations induced by triplex-forming oligonucleotides in mice.* Science, 2000. **290**(5491): p. 530-3.

11. Wang, G., et al., *Targeted mutagenesis in mammalian cells mediated by intracellular triple helix formation.* Mol Cell Biol, 1995. **15**(3): p. 1759-68.

12. Francois, J.C., T. Saison-Behmoaras, and C. Helene, *Sequence-specific recognition of the major groove of DNA by oligodeoxynucleotides via triple helix formation. Footprinting studies.* Nucleic Acids Research, 1988. **16**(24): p. 11431-40.

13. Asensio, J.L., et al., *The contribution of cytosine protonation to the stability of parallel DNA triple helices.* J Mol Biol, 1998. **275**(5): p. 811-22.

14. Sugimoto, N., et al., *pH and cation effects on the properties of parallel pyrimidine motif DNA triplexes.* Biochemistry, 2001. **40**(31): p. 9396-405.

15. Cheng, A.J. and M.W. Van Dyke, *Oligodeoxyribonucleotide length and sequence effects on intramolecular and intermolecular G-quartet formation.* Gene, 1997. **197**(1-2): p. 253-60.

16. Olivas, W.M. and L.J. Maher, 3rd, *Competitive triplex/quadruplex equilibria involving guanine-rich oligonucleotides.* Biochemistry, 1995. **34**(1): p. 278-84.

17. Olivas, W.M. and L.J. Maher, 3rd, *Overcoming potassium-mediated triplex inhibition.* Nucleic Acids Res, 1995. **23**(11): p. 1936-41.

18. Knauert, M.P., et al., *Distance and Affinity Dependence of Triplex-Induced Recombination.* Biochemistry, 2005. **44**(10): p. 3856-3864.

19. Vasquez, K.M., et al., *Chromosome targeting at short polypurine sites by cationic triplex-forming oligonucleotides.* J Biol Chem, 2001. **276**(42): p. 38536-41.

20. Dagle, J.M. and D.L. Weeks, *Positively charged oligonucleotides overcome potassium-mediated inhibition of triplex DNA formation.* Nucleic Acids Res, 1996. **24**(11): p. 2143-9.

21. Bailey, C. and D.L. Weeks, *Understanding oligonucleotide-mediated inhibition of gene expression in Xenopus laevis oocytes.* Nucleic Acids Res, 2000. **28**(5): p. 1154-61.

22. Bailey, C.P., J.M. Dagle, and D.L. Weeks, *Cationic oligonucleotides can mediate specific inhibition of gene expression in Xenopus oocytes.* Nucleic Acids Res, 1998. **26**(21): p. 4860-7.

23. Rao, T.S., et al., *Incorporation of 2'-deoxy-6-thioguanosine into G-rich oligodeoxyribonucleotides inhibits G-tetrad formation and facilitates triplex formation.* Biochemistry, 1995. **34**(3): p. 765-72.

24. Faruqi, A.F., et al., *Potassium-resistant triple helix formation and improved intracellular gene targeting by oligodeoxyribonucleotides containing 7-deazaxanthine.* Nucleic Acids Res, 1997. **25**(3): p. 633-40.

25. Vasquez, K.M., et al., *High-affinity triple helix formation by synthetic oligonucleotides at a site within a selectable mammalian gene.* Biochemistry, 1995. **34**(21): p. 7243-51.

26. Volker, J. and H.H. Klump, *Electrostatic effects in DNA triple helices.* Biochemistry, 1994. **33**(45): p. 13502-8.

27. Lee, J.S., et al., *Poly(pyrimidine) . poly(purine) synthetic DNAs containing 5-methylcytosine form stable triplexes at neutral pH.* Nucleic Acids Res, 1984. **12**(16): p. 6603-14.

28. Xodo, L.E., et al., *Effect of 5-methylcytosine on the stability of triple-strand DNA-A thermodynamic study.* Nucleic Acids Res., 1991. **19**: p. 5625-5631.

29. Lacroix, L., et al., *Triplex formation by oligonucleotides containing 5-(1-propynyl)-2'-deoxyuridine: decreased magnesium dependence and improved intracellular gene targeting.* Biochemistry, 1999. **38**(6): p. 1893-901.

30. Mayer, A., A. Haberli, and C.J. Leumann, *Synthesis and triplex forming properties of pyrrolidino pseudoisocytidine containing oligodeoxynucleotides.* Org Biomol Chem, 2005. **3**(9): p. 1653-8.

31. Miller, P.S., et al., *Recognition of a guanine-cytosine base pair by 8-oxoadenine.* Biochemistry, 1992. **31**(29): p. 6788-93.

32. Miller, P.S., et al., *Triplex formation by a psoralen-conjugated oligodeoxyribonucleotide containing the base analog 8-oxo-adenine.* Nucleic Acids Res, 1996. **24**(4): p. 730-6.

33. Brunar, H. and P.B. Dervan, *Sequence composition effects on the stabilities of triple helix formation by oligonucleotides containing N7-deoxyguanosine.* Nucleic Acids Res, 1996. **24**(11): p. 1987-91.

34. Jetter, M.C. and F.W. Hobbs, *7,8-Dihydro-8-oxoadenine as a replacement for cytosine in the third strand of triple helices. Triplex formation without hypochromicity.* Biochemistry, 1993. **32**(13): p. 3249-54.

35. Ishibashi, T., et al., *Triple helix formation with oligodeoxyribonucleotides containing 8-oxo-2'-deoxyadenosine and 2'-modified nucleoside derivatives.* Nucleic Acids Symp Ser, 1995(34): p. 127-8.

36. Ishibashi, T., et al., *Properties of triple helix formation with oligodeoxyribonucleotides containing 8-oxo-2'-deoxyadenosine and 2'-modified nucleoside derivatives.* Bioorg Med Chem, 1996. **4**(12): p. 2029-34.

37. Cassidy, S.A., et al., *Recognition of GC base pairs by triplex forming oligonucleotides containing nucleosides derived from 2-aminopyridine.* Nucleic Acids Res, 1997. **25**(24): p. 4891-8.

38. Soliva, R., et al., *DNA-triplex stabilizing properties of 8-aminoguanine.* Nucleic Acids Res, 2000. **28**(22): p. 4531-9.

39. Blommers, M.J., et al., *Dual recognition of double-stranded DNA by 2'-aminoethoxy-modified oligonucleotides: the solution structure of an intramolecular triplex obtained by NMR spectroscopy.* Biochemistry, 1998. **37**(51): p. 17714-25.

40. Puri, N., et al., *Targeted gene knockout by 2'-O-aminoethyl modified triplex forming oligonucleotides.* J Biol Chem, 2001. **276**(31): p. 28991-8.

41. Escude, C., et al., *Stable triple helices are formed upon binding of RNA oligonucleotides and their 2'-O-methyl derivatives to double-helical DNA.* C R Acad Sci III, 1992. **315**(13): p. 521-5.

42. Majumdar, A., et al., *Targeted gene knockout mediated by triple helix forming oligonucleotides.* Nat Genet, 1998. **20**(2): p. 212-4.

43. Puri, N., et al., *Minimum number of 2'-O-(2-aminoethyl) residues required for gene knockout activity by triple helix forming oligonucleotides.* Biochemistry, 2002. **41**(24): p. 7716-24.

44. Puri, N., et al., *Importance of clustered 2'-O-(2-aminoethyl) residues for the gene targeting activity of triple helix-forming oligonucleotides.* Biochemistry, 2004. **43**(5): p. 1343-51.

45. Praseuth, D., A.L. Guieysse, and C. Helene, *Triple helix formation and the antigene strategy for sequence-specific control of gene expression.* Biochim Biophys Acta, 1999. **1489**(1): p. 181-206.

46. Escude, C., et al., *Stable triple helices formed by oligonucleotide N3'-->P5' phosphoramidates inhibit transcription elongation.* Proc Natl Acad Sci U S A, 1996. **93**(9): p. 4365-9.

47. Giovannangeli, C., et al., *Accessibility of nuclear DNA to triplex-forming oligonucleotides: the integrated HIV-1 provirus as a target.* Proc Natl Acad Sci U S A, 1997. **94**(1): p. 79-84.

48. Michel, T., et al., *Highly stable DNA triplexes formed with cationic phosphoramidate pyrimidine alpha-oligonucleotides.* Chembiochem, 2005. **6**(7): p. 1254-62.

49. Lacroix, L., et al., *Pyrimidine morpholino oligonucleotides form a stable triple helix in the absence of magnesium ions.* Biochem Biophys Res Commun, 2000. **270**(2): p. 363-9.

50. Nielsen, P.E., M. Egholm, and O. Buchardt, *Evidence for (PNA)2/DNA triplex structure upon binding of PNA to dsDNA by strand displacement.* J Mol Recognit, 1994. **7**(3): p. 165-70.

51. Cherny, D.Y., et al., *DNA unwinding upon strand-displacement binding of a thymine-substituted polyamide to double-stranded DNA.* Proc Natl Acad Sci U S A, 1993. **90**(5): p. 1667-70.

52. Rogers, F.A., et al., *Site-directed recombination via bifunctional PNA-DNA conjugates.* Proc Natl Acad Sci U S A, 2002. **99**(26): p. 16695-700.

53. Egholm, M., et al., *PNA hybridizes to complementary oligonucleotides obeying the Watson-Crick hydrogen-bonding rules.* Nature, 1993. **365**(6446): p. 566-8.

54. Zhang, X., T. Ishihara, and D.R. Corey, *Strand invasion by mixed base PNAs and a PNA-peptide chimera.* Nucleic Acids Res, 2000. **28**(17): p. 3332-8.

55. Kaihatsu, K., et al., *Extending recognition by peptide nucleic acids (PNAs): binding to duplex DNA and inhibition of transcription by tail-clamp PNA-peptide conjugates.* Biochemistry, 2003. **42**(47): p. 13996-4003.

56. Bentin, T., H.J. Larsen, and P.E. Nielsen, *Combined triplex/duplex invasion of double-stranded DNA by "tail-clamp" peptide nucleic acid.* Biochemistry, 2003. **42**(47): p. 13987-95.

57. Havre, P.A. and P.M. Glazer, *Targeted Mutagenesis of Simian Virus-40 DNA Mediated by a Triple Helix-Forming Oligonucleotide.* J. Virol., 1993. **67**(12): p. 7324-7331.

58. Havre, P.A., et al., *Targeted mutagenesis of DNA using triple helix-forming oligonucleotides linked to psoralen.* Proc Natl Acad Sci U S A, 1993. **90**(16): p. 7879-83.

59. Arimondo, P.B., et al., *Targeting topoisomerase I cleavage to specific sequences of DNA by triple helix-forming oligonucleotide conjugates. A comparison between a rebeccamycin derivative and camptothecin.* C R Acad Sci III, 1999. **322**(9): p. 785-90.

60. Arimondo, P.B., et al., *Exploring the cellular activity of camptothecin-triple-helix-forming oligonucleotide conjugates.* Mol Cell Biol, 2006. **26**(1): p. 324-33.

61. Roulon, T., C. Helene, and C. Escude, *Coupling of a targeting peptide to plasmid DNA using a new type of padlock oligonucleotide.* Bioconjug Chem, 2002. **13**(5): p. 1134-9.

62. Escude, C., T. Garestier, and C. Helene, *Padlock oligonucleotides for duplex DNA based on sequence-specific triple helix formation.* Proc Natl Acad Sci U S A, 1999. **96**(19): p. 10603-7.

63. Bello-Roufai, M., T. Roulon, and C. Escude, *Ligand-mediated transcription elongation control using triplex-based padlock oligonucleotides.* Chem Biol, 2004. **11**(4): p. 509-16.

64. Mezhevaya, K., T.A. Winters, and R.D. Neumann, *Gene targeted DNA double-strand break induction by (125)I-labeled triplex-forming oligonucleotides is highly mutagenic following repair in human cells.* Nucleic Acids Res, 1999. **27**(21): p. 4282-90.

65. Panyutin, I.G. and R.D. Neumann, *Sequence-specific DNA breaks produced by triplex-directed decay of iodine-125.* Acta Oncol, 1996. **35**(7): p. 817-23.

66. Kukreti, S., et al., *Extension of the range of DNA sequences available for triple helix formation: stabilization of mismatched triplexes by acridine-containing oligonucleotides.* Nucleic Acids Res, 1997. **25**(21): p. 4264-70.

67. Orson, F.M., B.M. Kinsey, and W.M. McShan, *Linkage structures strongly influence the binding cooperativity of DNA intercalators conjugated to triplex forming oligonucleotides.* Nucleic Acids Res, 1994. **22**(3): p. 479-84.

68. Song, J., et al., *Activation of gene expression by triplex-directed psoralen crosslinks.* Gene, 2004. **324**: p. 183-90.

69. Macris, M.A. and P.M. Glazer, *Transcription dependence of chromosomal gene targeting by triplex-forming oligonucleotides.* J Biol Chem, 2003. **278**(5): p. 3357-62.

70. Wang, G., M.M. Seidman, and P.M. Glazer, *Mutagenesis in mammalian cells induced by triple helix formation and transcription-coupled repair.* Science, 1996. **271**: p. 802-805.

71. Wang, G. and P.M. Glazer, *Altered repair of targeted psoralen photoadducts in the context of an oligonucleotide-mediated triple helix.* J Biol Chem, 1995. **270**(38): p. 22595-601.

72. Catapano, C.V., et al., *Inhibition of gene expression and cell proliferation by triple helix-forming oligonucleotides directed to the c-myc gene.* Biochemistry, 2000. **39**(17): p. 5126-38.

73. Kenny, M.K., et al., *The role of human single-stranded DNA binding protein and its individual subunits in simian virus 40 DNA replication.* J Biol Chem, 1990. **265**(13): p. 7693-700.

74. Arzumanov, A., et al., *Inhibition of HIV-1 Tat-dependent trans activation by steric block chimeric 2'-O-methyl/LNA oligoribonucleotides.* Biochemistry, 2001. **40**(48): p. 14645-54.

75. Diviacco, S., et al., *Site-directed inhibition of DNA replication by triple helix formation.* Faseb J, 2001. **15**(14): p. 2660-8.

76. Maher, L.J.d., B. Wold, and P.B. Dervan, *Inhibition of DNA binding proteins by oligonucleotide-directed triple helix formation.* Science, 1989. **245**(4919): p. 725-30.

77. Obika, S., et al., *2'-O,4'-C-Methylene bridged nucleic acid (2',4'-BNA): synthesis and triplex-forming properties.* Bioorg Med Chem, 2001. **9**(4): p. 1001-11.

78. Faria, M., et al., *Targeted inhibition of transcription elongation in cells mediated by triplex-forming oligonucleotides.* Proc Natl Acad Sci U S A, 2000. **97**(8): p. 3862-7.

79. Giovannangeli, C., et al., *Efficient inhibition of transcription elongation in vitro by oligonucleotide phosphoramidates targeted to proviral HIV DNA.* J Mol Biol, 1996. **261**(3): p. 386-98.

80. Boulme, F., et al., *Modified (PNA, 2'-O-methyl and phosphoramidate) anti-TAR antisense oligonucleotides as strong and specific inhibitors of in vitro HIV-1 reverse transcription.* Nucleic Acids Res, 1998. **26**(23): p. 5492-500.

81. Ziemba, A.J., et al., *A bis-alkylating triplex forming oligonucleotide inhibits intracellular reporter gene expression and prevents triplex unwinding due to helicase activity.* Biochemistry, 2003. **42**(17): p. 5013-24.

82. Besch, R., et al., *Triple helix-mediated inhibition of gene expression is increased by PUVA.* J Invest Dermatol, 2004. **122**(5): p. 1114-20.

83. Intody, Z., et al., *Blocking transcription of the human rhodopsin gene by triplex-mediated DNA photocrosslinking.* Nucleic Acids Res, 2000. **28**(21): p. 4283-90.

84. Carbone, G.M., et al., *Triplex DNA-mediated downregulation of Ets2 expression results in growth inhibition and apoptosis in human prostate cancer cells.* Nucleic Acids Res, 2004. **32**(14): p. 4358-67.

85. Larsen, H.J. and P.E. Nielsen, *Transcription-mediated binding of peptide nucleic acid (PNA) to double-stranded DNA: sequence-specific suicide transcription.* Nucleic Acids Res, 1996. **24**(3): p. 458-63.

86. Zhao, X., K. Kaihatsu, and D.R. Corey, *Inhibition of transcription by bisPNA-peptide conjugates.* Nucleosides Nucleotides Nucleic Acids, 2003. **22**(5-8): p. 535-46.

87. Zhang, X., C.G. Simmons, and D.R. Corey, *Liver cell specific targeting of peptide nucleic acid oligomers.* Bioorg Med Chem Lett, 2001. **11**(10): p. 1269-72.

88. Ziemba, A.J., et al., *Targeting and regulation of the HER-2/neu oncogene promoter with bis-peptide nucleic acids.* Oligonucleotides, 2005. **15**(1): p. 36-50.

89. Liu, B., et al., *Toward synthetic transcription activators: recruitment of transcription factors to DNA by a PNA-peptide chimera.* J Am Chem Soc, 2002. **124**(9): p. 1838-9.

90. Vasquez, K.M., et al., *Chromosomal mutations induced by triplex-forming oligonucleotides in mammalian cells.* Nucleic Acids Res, 1999. **27**(4): p. 1176-81.

91. Kim, K.H., P.E. Nielsen, and P.M. Glazer, *Site-Specific Gene Modification by PNAs Conjugated to Psoralen.* Biochemistry, 2006. **45**(1): p. 314-323.

92. Faruqi, A.F., M. Egholm, and P.M. Glazer, *Peptide nucleic acid-targeted mutagenesis of a chromosomal gene in mouse cells.* Proc Natl Acad Sci U S A, 1998. **95**(4): p. 1398-403.

93. Luo, Z., et al., *High-frequency intrachromosomal gene conversion induced by triplex-forming oligonucleotides microinjected into mouse cells.* Proc Natl Acad Sci U S A, 2000. **97**(16): p. 9003-8.

94. Luo, Z., et al., *High-frequency intrachromosomal gene conversion induced by triplex-forming oligonucleotides microinjected into mouse cells.* Proc Natl Acad Sci U S A, 2000. **97**(16): p. 9003-8.

95. Datta, H.J., et al., *Triplex-induced Recombination in Human Cell-free Extracts. Dependence on XPA and HsRad51.* J Biol Chem, 2001. **276**(21): p. 18018-23.

96. Vasquez, K.M., et al., *Human XPA and RPA DNA repair proteins participate in specific recognition of triplex-induced helical distortions.* Proc Natl Acad Sci U S A, 2002. **99**(9): p. 5848-53.

97. Chan, P.P., et al., *Targeted correction of an episomal gene in mammalian cells by a short DNA fragment tethered to a triplex-forming oligonucleotide.* J Biol Chem, 1999. **274**(17): p. 11541-8.

98. Porumb, H., et al., *Temporary ex vivo inhibition of the expression of the human oncogene HER2 (NEU) by a triple helix-forming oligonucleotide.* Cancer Res, 1996. **56**(3): p. 515-22.

99. Hobbs, C.A. and K. Yoon, *Differential regulation of gene expression in vivo by triple helix-forming oligonucleotides as detected by a reporter enzyme.* Antisense Res Dev, 1994. **4**(1): p. 1-8.

100. Derossi, D., et al., *Cell internalization of the third helix of the Antennapedia homeodomain is receptor-independent.* J Biol Chem, 1996. **271**(30): p. 18188-93.

101. Rogers, F.A., et al., *Peptide conjugates for chromosomal gene targeting by triplex-forming oligonucleotides.* Nucleic Acids Res, 2004. **32**(22): p. 6595-604.

102. Derossi, D., et al., *The third helix of the Antennapedia homeodomain translocates through biological membranes.* J Biol Chem, 1994. **269**(14): p. 10444-50.

103. Chan, P.P. and P.M. Glazer, *Triplex DNA: fundamentals, advances, and potential applications for gene therapy.* J Mol Med, 1997. **75**(4): p. 267-82.

104. Faruqi, A.F., et al., *Triple-helix formation induces recombination in mammalian cells via a nucleotide excision repair-dependent pathway.* Mol Cell Biol, 2000. **20**(3): p. 990-1000.

105. Kalish, J.M. and P.M. Glazer, *Targeted Genome Modification via Triple Helix Formation.* Ann NY Acad Sci, 2005. **1058**: p. 151-61.

106. Wang, G. and P.M. Glazer, *Altered repair of targeted psoralen photoadducts in the context of an oligonucleotide-mediated triple helix.* The Journal of Biological Chemistry, 1995. **270**(39): p. 22595-22601.

Figures

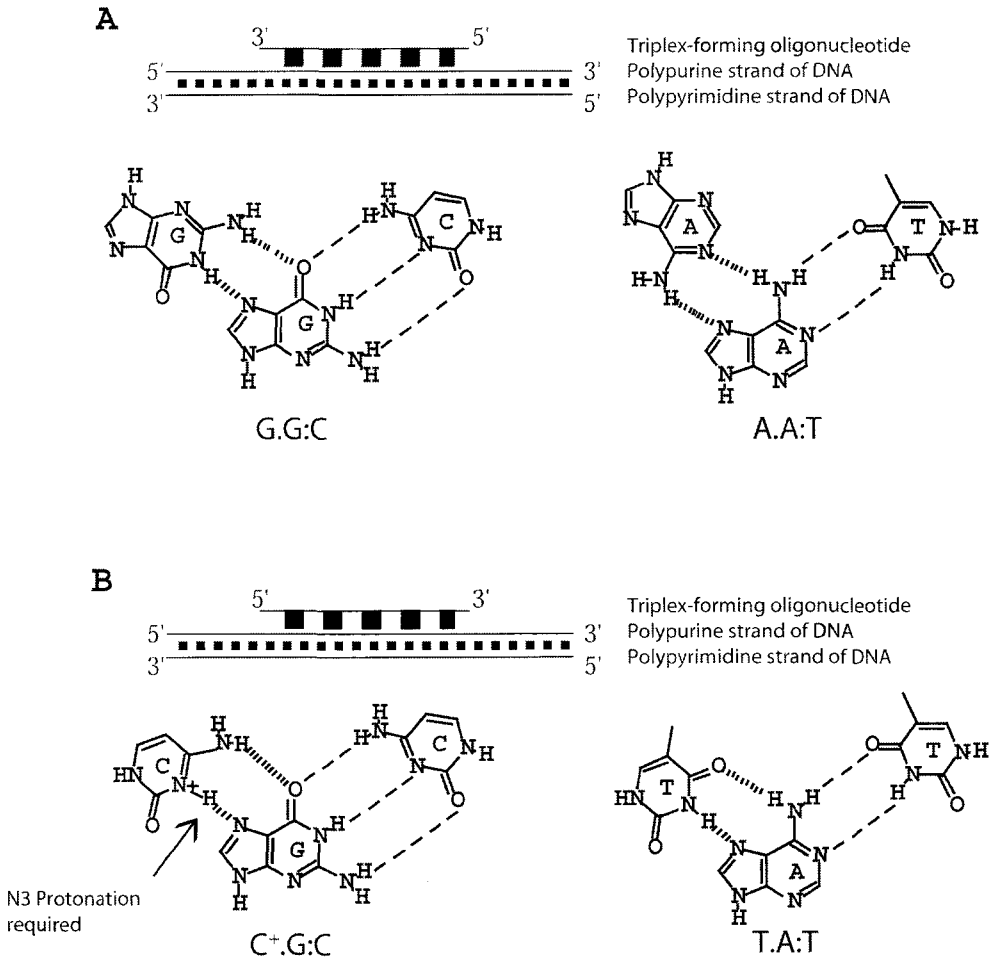


Figure 1. Triplex-Binding Code and Orientation. (A) In the purine motif, triplex-forming oligonucleotides bind in an anti-parallel orientation via reverse Hoogsteen bonds. (B) In the pyrimidine motif, the third strand binds in a parallel orientation via Hoogsteen bonds. The canonical triplets are shown for each motif.

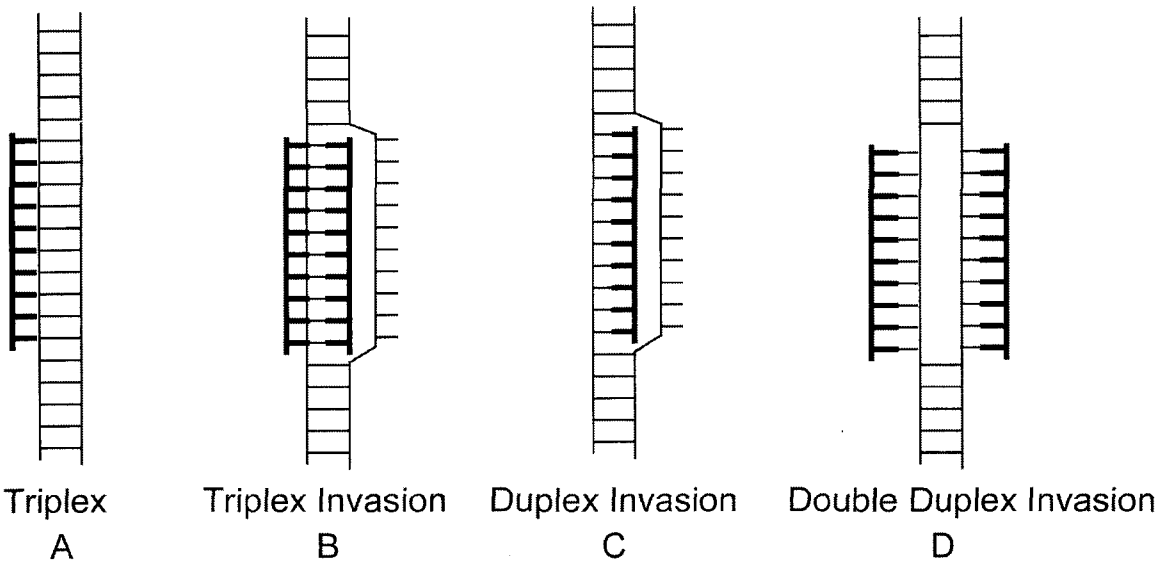


Figure 2. Peptide nucleic acid (PNA) binding to duplex DNA. (A) PNAs consisting of cytidine-rich polypyrimidine sequences bind as third strands through Hoogsteen bonds. (B) PNAs can also bind as third strands in conjunction with a PNA that strand invades (binding via Watson-Crick base pairing to one strand of the duplex DNA). When the strand invading and triplex-forming strands are tethered together, they form a PNA clamp. Homopurine PNAs bind through single strand duplex invasion (C) or double strand (D) duplex invasion of the DNA. This figure was adapted from Nielson, P.E., (2001). Current Opinion in Biotechnology, 12: p. 16-20.

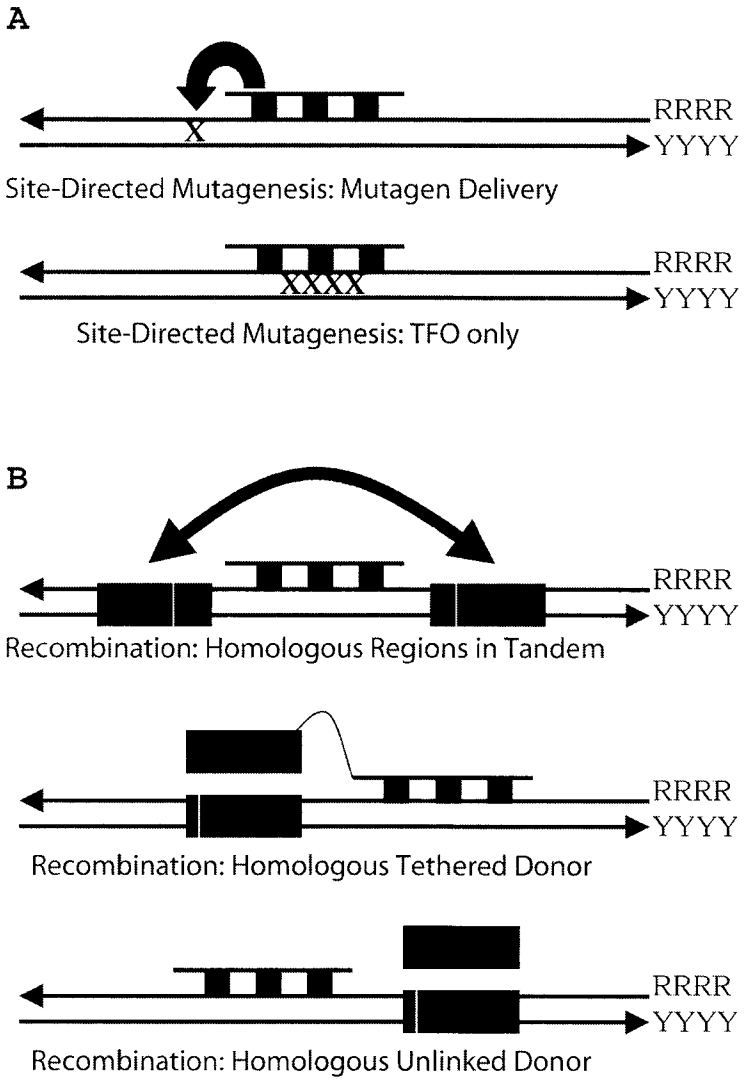


Figure 3. TFO-induced gene modification. (A) TFO-directed mutagenesis. TFOs can be used for site-specific delivery of DNA reactive conjugates such as psoralen. TFOs also induce mutations around the TFO binding site independent of conjugation to reactive agents. (B) TFO-induced recombination. TFOs have been used in a variety of intramolecular and intermolecular targeting. Intramolecular targeting involves gene correction through recombination between tandem differentially mutated genes located on opposite sides of a TFO binding site. Intermolecular targeting relies on either a tethered or untethered donor molecule for the wild-type correction of the mutation.

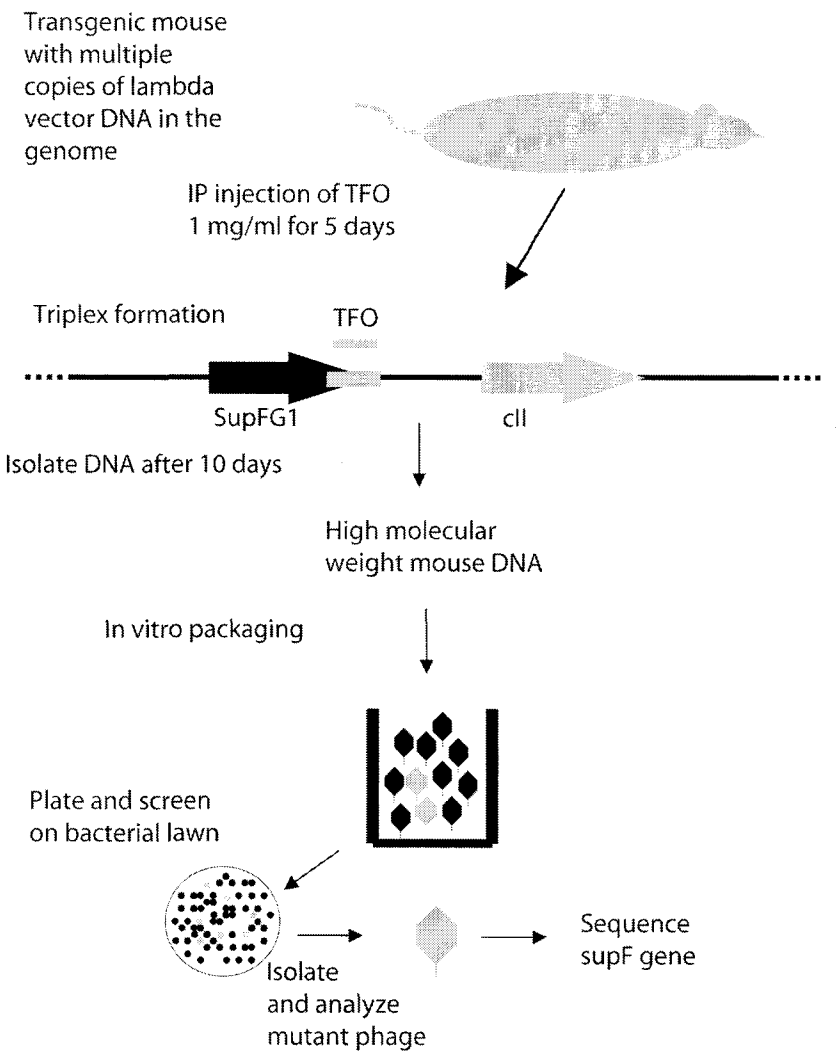


Figure 4. Protocol for detecting chromosomal mutations in TFO treated mice. Transgenic reporter mice containing multiple copies of chromosomally integrated lambda supFG1 (containing a triplex target site) and a control gene cII (lacking a triplex target site) were intraperitoneally (IP) injected with TFO for 5 consecutive days. Ten days later genomic DNA was isolated from a variety of tissues. Mutagenesis was detected by *in vitro* packaging of the phage followed by plating on a bacterial lawn (on plates containing IPTG and X-Gal). Mutant plaques were white while wild-type plaques were blue. Mutant plaque DNA was isolated and sequenced.

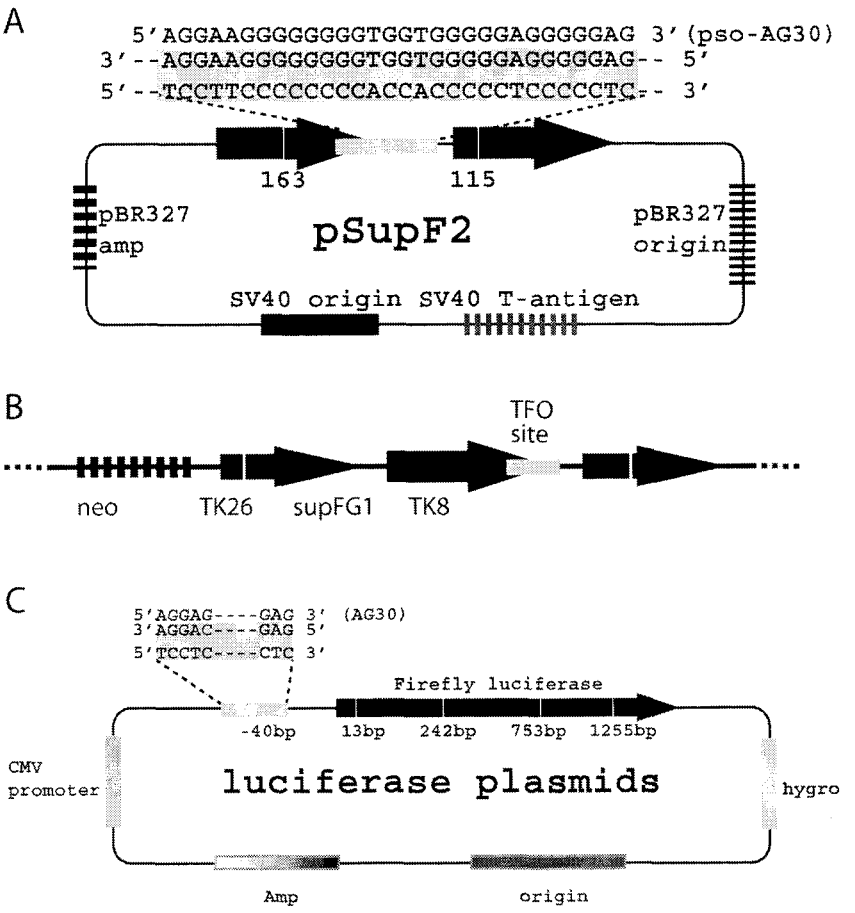


Figure 5. Reporter gene constructs for recombination. (A) Tandem *supF* genes for psoralen-AG30 TFO targeting. Episomal target for transient transfection in COS-7 cells. The *SupF* gene upstream of TFO target site was mutated at base pair (bp) 163, the downstream *SupF* gene was mutated at bp 115. (B) Dual thymidine kinase genes. A chromosomal target in mouse cells for induced recombination with two tandem mutated copies of the *herpes simplex thymidine kinase* (TK) gene. (C) The luciferase reporter construct used in the distance experiments. The TFO target site is 40 bp upstream of the gene start site and mutations in the firefly luciferase gene at the different distances are indicated (13 bp, 242 bp, 753 bp, 1255 bp).

Chapter Two:
Triplex-Induced Recombination and Repair in the Pyrimidine Motif

Introduction

TFOs provide potential tools for altering gene function by either repressing transcription, inhibiting DNA replication, or inducing site-specific mutagenesis and recombination [1-6]. TFOs bind in the major groove of DNA in a sequence-specific manner to polypurine/polypyrimidine sequences, in two distinct motifs [7-10]. TFOs in the purine motif bind anti-parallel to the purine strand of the duplex via reverse-Hoogsteen hydrogen bonds and TFOs in the pyrimidine motif bind in a parallel orientation via Hoogsteen bonds [7-10]. The third-strand binding code specifies that in the purine motif G and A in the third strand bind to the duplex G and A, respectively; while in the pyrimidine motif C and T bind to the duplex G and A, respectively [11].

Binding affinity in the two motifs is differentially affected by pH and ion concentrations. In the pyrimidine motif, the need for protonation of cytosines at the N3 position (favored at low pH) limits the binding affinity of TFOs composed of unmodified DNA under physiological conditions [12]. In order to make TFOs more effective tools, chemical modifications have been developed to enhance third-strand binding under *in vivo* conditions. Base modifications such as 5-methyl-2'-deoxycytidine (5meC) and 5-(1-propynyl)-2'-deoxyuridine (pdU) and sugar modifications such as 2'-O-(2-aminoethyl)-ribose (2'-AE) substitution improve binding affinity and psoralen delivery in the pyrimidine motif [13, 14]. Pyrimidine TFOs with a N3'->P5' phosphoramidate (amidate) backbone (in which nitrogen replaces the bridging 3' oxygen of the diester) have also been shown to form stable triplexes at neutral pH [15].

In addition, conjugation of intercalators such as pyrene and acridine to TFOs has been shown to further increase their biological activity [16]. Although purine TFOs consisting of standard DNA (except for 3'-end-capping [4, 17]) have been shown to be active in cells and animals [16-18], fewer strategies have been established for effective modification of purine TFOs. The most effective option that has been demonstrated is backbone substitution with N, N-diethylethylenediamine (DEED) internucleoside linkages which reduce the Mg^{2+}-dependence and enhance intracellular activity of polypurine TFOs [18].

Both purine and appropriately modified pyrimidine TFOs have been used to direct psoralen intercalation and photoadduct formation to specific sites in mammalian cells and yeast [16, 17, 19]. In the absence of psoralen, however, only purine TFOs have been shown to induce genomic changes (either mutagenesis or recombination) via the effects of third-strand binding alone [4]. Previous work has demonstrated that non-covalent triplexes formed by purine TFOs are active in stimulating repair and recombination in cell-free extracts, in cell culture, and even in mice [3, 5, 20, 21]. Studies in human cell-free extracts and in DNA repair-deficient human cell lines have revealed the ability of non-covalent purine motif triplexes to provoke recombination in a manner dependent on the nucleotide excision repair (NER) pathway [4]. However, the ability of non-covalent triplexes formed by pyrimidine TFOs to induce recombination has not been established.

In the work reported here, we sought to determine the extent to which triplexes formed by pyrimidine TFOs can stimulate recombination and repair. Because standard

DNA TFOs bind poorly in the pyrimidine motif, we evaluated TFOs containing a series of modifications to determine which, if any, are effective in provoking DNA metabolism via third-strand binding alone, in the absence of a DNA-reactive conjugate. Interestingly, although several modifications did support high affinity binding under physiological conditions *in vitro*, only amidate and pdU-modified TFOs induced recombination in episomal targets in cells and repair in cell-free extracts. These results demonstrate that certain pyrimidine TFOs can provoke recombination without the need for covalent damage to the DNA. Because pyrimidine triplexes are favored at A:T bp-rich sites (in contrast to purine triplexes, which are favored at G:C rich sites), these findings increase the number of potential genomic target sites suitable for modification by non-covalent triplexes. In addition, we propose a model in which TFO sugar substitution at the 2' position may influence the extent to which the resulting triplex can provoke DNA metabolism, with reference to NMR studies showing that 2' substitution of the TFO influences the degree of structural distortion of the duplex to which the TFO is bound [22].

Materials and Methods

Oligonucleotides and Vectors

Oligonucleotides were obtained from Transgenomic (amidate18), Gilead Sciences and Eurogentec (G3pUmC18), Oligos Etc. (G3TmC18, 2'-O-Me18), Proligo (LNA18, LNAcontrol) and Eurogentec (LNA18), Eurogentec (UmCcontrol), Gene Tools

(morph18), Midland Certified Reagents Company (AG30, SCR30, MIX30, as51), Rowshon Alam and Michael Seidman (2'-AE), Chip Greaves and Dan Weeks (DEED), and Keck Oligonucleotide Synthesis Facility (Yale University) (AG19, DNA18). Duplex target oligonucleotides (G3Y, G3R) were synthesized by Oligos Etc. When the chemical modification did not provide protection against exonuclease degradation, which is necessary for cell targeting experiments, 3'-ends were synthesized to contain a 3' propylamine group, using the C3 amino-CPG from Glen Research. The following oligonucleotides contain a 3' propylamine group: G3TmC18, G3pUmC18, AG30, SCR30, and MIX30. Three terminal phosphorothioate linkages were synthesized on the 3'-end of 2'-O-Me18 and as51 has terminal phosphorothioate linkages on both ends. Pyrimidine TFO sequences and modifications are listed in Fig. 1. The sequences for the control oligonucleotides are as follows: AG30, 5'AGGAAGGGGG GGGTGGTGGG GGAGGGGGAG 3'; SCR30, 5' GGAGGAGTGG AGGGGAGTGA GGGGGGGGGG 3'; MIX30, 5' AGTCAGTCAG TCAGTCAGTC AGTCAGTCAG 3'; UmCcontrol, 5' cucucucucuuuuuuu 3' (c = 5meC, u = pdU); LNAcontrol , 5' TtTTCtTTTtTCTtTTCt 3' (t = LNA); AG19, 5'AGGAAGGGGG GGGTGGTGGG 3'; DNA18, 5' TTTTCTTTTT TCTTTTCT 3'. And the single-stranded donor, as51, had the sequence: 5' TGGTAAAGCC ACCATGGAAG ACGCCAAAAA CATAAAGAAA GGCCCGGCGCC 3'.

The pSupFTC18 shuttle vector, derived from pSP189 [23], carries two copies of the *supF* gene, each with a distinct mutation, flanking an 18-bp A-rich

homopurine/homopyrimidine target site (the TC18 site from pLSG3 [13]). These mutations were created using the QuikChange Site-Directed Mutagenesis Kit (Stratagene). These mutations result in inactive tRNAs which lack the ability to suppress the amber mutation in the *lacZ* gene in MBM7070 E. coli which are used as indicator hosts for this recombination assay [24].

The plucTC18 vector was derived from a plasmid constructed by subcloning the *firefly luciferase* gene *Fluc+* (pGL3-Basic Vector, Promega) into pcDNA5/FRT (Invitrogen) [25]. The TC18 target site was inserted 40 bp upstream of the *Fluc+* start site. Site-specific mutagenesis was used to create a stop codon at bp 13 downstream from the start codon.

The pRLTC18 vector was constructed by adding the TFO target site (the TC18 site from pLSG3 [13]) into the multiple cloning region of the pRL-CMV vector (Promega).

Third-strand binding assays

Third-strand binding was measured using gel mobility shift assays under native conditions. Two complementary 30-mers (G3Y, G3R) containing the TC18 target from pLSG3 (representing bp 76 to 105) were synthesized. Duplex DNA was prepared by mixing 1000 pmol of each 30-mer together with 50 mM NaCl and incubating at 85°C for 20 minutes and cooling to room temperature before end-labeling with T4 polynucleotide kinase (New England BioLabs) and [γ-^{32}P]ATP (Amersham Biosciences). Duplex was gel purified, electroeluted, and filtered by Centricon (Millipore). A fixed concentration

of duplex at 5 x 10^{-8} M was added to binding reactions with increasing concentrations of oligonucleotides in 20 µl of 10 mM Tris (pH 7.2) and selected concentrations of $MgCl_2$ (0.1 mM or 10 mM) and KCl (0 mM or 140 mM). Binding for full DEED18 was carried out at pH 7.6 with the addition of 1 mM spermine. Binding assays were also carried out at pH 5.4, in which case 40 mM Tris-acetic acid was substituted for Tris base at pH 5.4. Samples were incubated at 37°C for 16 hours unless otherwise specified.

Samples bound at pH 7.2 were loaded onto 15% polyacrylamide gels [acrylamide/bisacrylamide (19:1)] containing 17.8 mM Tris and 17.8 mM boric acid (pH 7.2) and 10 mM $MgCl_2$ and the gels were run in the same concentration of Tris-boric acid and $MgCl_2$ at 60 V for 6 hours at 23°C. For samples bound at pH 5.4, 15% polyacrylamide gels containing 40 mM Tris-acetate and 10 mM $MgCl_2$ were used and gels were run in buffers of the same concentration of Tris-acetate and $MgCl_2$ at 60 V for 6 hours at 23°C.

Shuttle vector recombination assay

Monkey COS-7 cells were obtained from ATCC (1651-CRL) and grown in DMEM/10% fetal bovine serum (FBS) (Gibco). Cells were grown to 50-70% confluency and cell transfection was performed with GenePORTER2 (Gene Therapy Systems, Inc.). Cells were transfected with 1 µg of pSupFTC18 followed 24 hours later by a second, separate transfection with 1 µg of oligonucleotide. Cells were harvested at 48 hours after oligonucleotide transfection and shuttle vector DNA was isolated by a modified alkaline

lysis procedure, as previously described [6]. Vector samples were used to transform indicator bacteria (MBM7070 [lacZ(Am)] [23]) by electroporation (Bio-Rad gene pulser; settings: 25 μF, 250 W, and 1800 V; 0.1 cm electrode gap cuvette), and colonies were screened for *supF* function by growth on plates supplemented with 75 μg/ml ampicillin, 210 μg/ml 5'-bromo-4-chloro-3-indolyl-β-D-galactopyranoside (X-Gal), and 200 μg/ml IPTG (isopropyl-β-D-thiogalactopyranoside). Recombination events were indicated by wild-type blue colonies in a background of white colonies.

CHO cell transfection and luciferase assay

CHO cells were obtained from Invitrogen and grown in F12/10% FBS (Gibco) supplemented with 2 mM L-glutamine (Gibco) and 100 μg/mL Zeocin (Invitrogen). CHO cells were transfected and the *luciferase* activity was assayed as previously described [25]. Briefly, 24 hours prior to transfection, 12-well plates were seeded with 5×10^4 cells/well. plucTC18 was preincubated overnight with as51 (an oligonucleotide containing the wild-type *luciferase* sequence for the mutated region) and the TFO to be tested with 10 mM $MgCl_2$ and 10 mM Tris pH 7.2 at 37°C. Cells were transfected with GenePORTER2 as described above. Each transfection delivered 0.3 μg of plasmid, donor, and TFO per well. Cells were collected 48 hours post-transfection by rinsing twice with PBS followed by lysis with 1x Passive Lysis Buffer (Promega). *Luciferase* activity was measured using the Promega Dual *Luciferase* Kit and only the LAR II

substrate. The total protein in the cell lysate was determined via Bradford protein assay [26].

In vitro repair synthesis assay

Repair reactions were carried out as described [27], with several modifications. HeLaScribe Nuclear Extract in vitro Transcription Grade (Promega) was used. The pRLTC18 plasmid (500 ng) containing the TC18 target sequence was preincubated with 1 μM TFO (37°C, 18 hours, 10 mM Tris, pH 7.2, 10 mM $MgCl_2$). As a positive control, one sample was incubated with sterile water in place of oligonucleotide, this sample was then damaged by UVC radiation treatment (2 inches, 30 seconds). The DNA was added to HeLa cell-free extracts with [α-^{32}P]dCTP (Amersham Biosciences) and supplemented as previously described [27]. Samples were incubated for 3 hours at 30°C and the reactions were stopped as previously described [27]. Substrate plasmid DNAs were isolated from the reaction by phenol/chloroform/iso-amyl alcohol extraction and ethanol precipitation. The isolated vector DNAs were linearized by Xho I digestion and the samples were analyzed by gel electrophoresis. DNA repair synthesis was quantified by measure of [α-^{32}P]dCTP incorporation using a phosphorimager (Storm 860, Molecular Dynamics, Amersham Pharmacia Biotech). The quantified value of the UVC treated sample was set to 100% and used to normalize the measurements of the other samples.

Shuttle vector assay in repair-deficient and corrected cells

XP12BE cells (deficient in the nucleotide excision repair factor, XPA) were obtained from Coriell Cell Repositories (GM04429) and grown in MEM/10% FBS

(Gibco). The shuttle vector recombination assay was performed as described above except that in some cases the cells were also transfected with 1 μg of a plasmid containing XPA cDNA (XPA cDNA with a His tag on the C-terminus cloned into the multiple cloning site of pcDNA3.1 (Invitrogen)) at the same time the cells were transfected with the psupTC18. XPA expression was verified by Western blot. Briefly, cells were lysed (at 48 and 72 hours post-transfection) with RIPA buffer (150 mM NaCl, 0.1% SDS) and 100 μg of total protein per sample was resolved on 12% SDS/PAGE gels. Proteins were detected by standard immunoblot procedures using XPA [28] and tubulin (clone B-5-1-2; Sigma) primary antibodies.

Serum stability assay

Each oligonucleotide was 5'-end-labeled at a concentration of 1 μM with [γ-^{32}P]ATP using T4 polynucleotide kinase (New England BioLabs) for 1 hour, followed by heat inactivation for 20 minutes at 70°C. 5 μl of end-labeled oligonucleotide were added to 100 μl of FBS (Gibco) and incubated at 37°C. Aliquots were removed at 0 to 24 hours and added to an equal volume of formamide dye, before being run on a 20% denaturing polyacrylamide gel. Visualization and quantification of intact oligonucleotide was done using a phosphorimager.

Results

TFO target site and oligonucleotide modifications

To examine the ability of triplexes formed in the pyrimidine, parallel motif to induce recombination and repair, we evaluated a series of chemically modified pyrimidine TFOs each designed to bind to an 18 bp A-rich duplex target site (Fig. 1). The modifications included amidate and DEED internucleoside linkages and 5meC and pdU base substitutions. Sugar modifications included 2'-O, 4'-C-methylene bridged or locked nucleic acid (LNA), 2'-O-methyl-ribose (2'-O-Me), and 2'-AE substitutions. Morpholino (morph) TFOs consist of both sugar and backbone changes (Fig. 1).

We chose these modifications because either they have been shown to enhance intracellular delivery of psoralen conjugates to target genes [amidate [29-33], DEED [18], 5meC [13, 34], pdU [13, 34, 35], 2'-O-Me [14], and 2'-AE [14]] or they have been found to provide enhanced third-strand binding *in vitro* [LNA [36-38] and morpholino [39, 40]]. In the case of the DEED modification, enhanced chromosome targeting has been demonstrated only in the purine motif. However, the ability of the DEED internucleoside substitution to confer a potentially advantageous positive charge on the TFO prompted us to consider pyrimidine TFOs with either full or partial DEED modifications.

Unless otherwise specified, all TFOs contained phosphodiester backbones. In some cases all nucleotides were modified in a given TFO while in other cases only selected ones were modified, as indicated (Fig. 1). The C's in G3TmC18, G3pUmC18,

2'-O-Me18, and 2'-AE18 were substituted with 5meC, because this modification has been shown to enhance third-strand binding at physiological pH [41, 42] and U's in 2'-AE18 contained 5-methyl-2'-deoxyuridine modifications. Nucleotides in 2'-AE18 not containing the 2'-AE modification were synthesized to contain 2'-O-Me modifications.

In vitro binding measurements

The binding affinities of TFOs with the modifications described above were compared under a series of conditions as determined by gel mobility shift assays in non-denaturing gels. Binding was quantified and equilibrium dissociation constants (K_d's, representing TFO concentrations yielding half-maximal binding) were determined. Because stable triplex formation is dependent on Mg^{2+} and sensitive to K^+ [43, 44], we tested binding in either 0.1 or 10 mM $MgCl_2$ and either 0 or 140 mM KCl (both conditions at pH 7.2). In recognition of the possible pH dependence of binding by certain TFOs, we also tested binding at pH 5.4 (in 10 mM $MgCl_2$ and without KCl).

Under conditions where the minimally modified TFO, G3TmC18, formed triplexes (pH 5.4, 10 mM $MgCl_2$, 0 mM KCl), all of the chemically modified TFOs formed triplexes (Fig. 2). However, there was a range of binding affinities (Fig. 2 and Table 1). The best binding was seen with the amidate, LNA, 2'-AE, pdU, and 2'-O-Me-modified oligonucleotides, with K_d's of $2.8x10^{-10}$, $3.0x10^{-10}$, $4x10^{-10}$, $4.5x10^{-10}$, and $4.6x10^{-10}$ M, respectively (Fig. 2 and Table 1). G3TmC18, morph18, and partially modified DEED oligonucleotides showed lower binding affinities, and the full DEED18 showed the lowest binding affinity (Fig. 2 and Table 1).

At neutral pH with 10 mM $MgCl_2$, the 2'-O-Me, LNA, 2'-AE, and amidate-modified TFOs showed strong binding with K_d's of $1.5x10^{-10}$, $2.7x10^{-10}$, $6.0x10^{-10}$, and $8.0x10^{-10}$ M, respectively (Fig. 2 and Table 1). The pdU and 5meC-modified TFOs showed approximately 10-fold weaker binding and the morpholino-modified TFO showed 1000-fold weaker binding. Neither of the DEED-modified oligonucleotides (partial or full) showed any binding at neutral pH.

Under physiological conditions (pH 7.2, 0.1 mM $MgCl_2$, and 140 mM KCl), the amidate, 2'-AE, LNA, and pdU–modified TFOs showed the strongest binding with K_d's of $1.0x10^{-8}$, $2.0x10^{-8}$, $2.5x10^{-8}$, and $9.0x10^{-8}$ M, respectively. G3TmC18 showed slightly weaker binding, and the morpholino and the 2'-O-Me-modified TFOs showed even weaker binding. The DEED-modified TFOs showed no detectable binding under these conditions.

TFO-induced recombination in COS-7 cells

To test triplex-induced recombination by pyrimidine TFOs, we constructed a SV40-based shuttle vector similar to pSupFAR used in previous work to test purine TFOs [20]. However, in this case, we designed the vector to contain the 18 bp A-rich target site used in the binding assays above. This site is inserted between two copies of the *supF* gene, each with an inactivating point mutation, at bp 163 in the upstream gene and at bp 115 in the downstream gene (Fig. 3A). The 5' mutation is 33 bp from the beginning of the TFO binding site while the 3' mutation is 24 bp away from the end of the site. Recombination between the two genes can generate a functional *supF* gene, which can be

scored upon shuttle vector transformation into indicator bacteria following vector DNA rescue from the COS-7 cells.

The plasmid vector was pre-transfected into the COS-7 cells to establish the episomal target in these cells. One day later, the cells were transfected with the TFOs and maintained in culture for 48 hours to allow for triplex formation and possible induced recombination. As a negative control to establish the background frequency of recombination in this assay, we included in each experiment cells transfected with vector but mock transfected the next day without TFO. The vector DNAs were harvested from the cells and analyzed for recombination events (Fig. 3B). The results were normalized between experiments by comparison to the plasmid only (no oligonucleotide) background and presented as recombination frequency. Transfection of cells with the amidate18 and G3pUmC18 TFOs gave the highest levels of induced recombination, at 0.37% and 0.31%, respectively; whereas G3TmC18, 2'-AE18, and morph18 gave lower frequencies of 0.14%, 0.14%, and 0.13%, respectively. The 2'-O-Me, LNA, and DEED TFOs had minimal effects over background. The control TFOs, UmCcontrol, LNAcontrol, AG30, SCR30, and MIX30 showed little effect over background.

To further control for sequence specificity, we tested induced recombination by the amidate18 and G3pUmC18 TFOs using a different shuttle vector, pSupFAR, as a target in the cells. pSupFAR contains two mutant *supF* genes flanking a 30 bp G-rich polypurine target site suitable for triplex formation in the purine motif [3]. In previous work, we have shown that a G-rich TFO, AG30, can bind with high affinity to the third-

strand binding site in pSupFAR [3]. AG30 shows no detectable binding to the A-rich polypurine site found in pSupFTC18, and conversely, TC-rich pyrimidine TFOs show no binding to the G-rich site in pSupFAR (data not shown). We found that while transfection of amidate18 and G3pUmC18 into COS-7 cells can induce recombination in pSupFTC18, neither has an effect on recombination in pSupFAR in the same cell assay (nor do any of the other modified pyrimidine TFOs tested) (Fig. 4). In contrast, AG30 can induce recombination in pSupFAR in COS-7 cells but has no effect on recombination in pSupFTC18 (Fig. 4). Hence, the effects of the amidate18, G3pUmC18, and AG30 TFOs are sequence-specific, based on both oligonucleotide and target site controls.

TFO-induced recombination in CHO cells

To test pyrimidine induced recombination in a different episomal target, we constructed the plucTC18 plasmid, which is a vector containing a pyrimidine TFO target adjacent to a mutated *firefly luciferase* reporter gene, similar to a vector used in previous work studying targeting in the purine motif [25]. The TFO binding site is 40 bp upstream from the start site of the *luciferase* gene, and a nonsense mutation that inactivates the *luciferase* gene is 13 bp downstream from the start site. In this assay, recombination induced by TFO binding occurs between the plasmid and a separate donor oligonucleotide (a 51-mer) containing the wild-type *luciferase* sequence in the region of the mutation. The amidate18 was found to induce recombination at a frequency of 0.11% (Fig. 5B), whereas the 2'-O-Me and 2'-AE-modified TFOs showed minimal levels of induced recombination above background.

Pyrimidine TFO-induced repair in HeLa cell-free extracts

Because previous work had shown a correlation between the ability of purine TFOs to induce recombination and the ability of the triplexes they form to induce repair, we assayed for induced repair synthesis in HeLa cell-free extracts on plasmid substrates containing pyrimidine motif triplexes. The substrate for DNA repair was a plasmid incubated in the presence or absence of TFO. The amount of DNA repair-associated synthesis was measured based on the level of incorporation of [α-^{32}P]dCTP into the plasmid. This is a well-established assay for measuring repair provoked by a variety of DNA lesions [45-47]. Amidate18 and G3pUmC18 showed the highest level of induced repair activity at 42% and 66%, respectively, as compared to UVC damage set at 100% activity (Fig. 6A and 6B). G3TmC18 and LNA18 produced lower repair activity (Fig. 6A and 6B).

Role of nucleotide excision repair in induced recombination

To determine if the induced recombination in cells by the amidate18 TFO is due to the NER pathway, a shuttle vector assay similar to that used in COS-7 cells, above, was performed in XP12BE cells. These cells are SV40 transformed fibroblasts isolated from a patient lacking the Xeroderma Pigmentosum, complementation group A protein (XPA, a recognition protein in the NER pathway). The amidate18 TFO was unable to induce recombination in XP12BE cells (Fig. 6C). However, with the expression of XPA cDNA, induced recombination by amidate18 was detected (Fig. 6C). The expression of XPA in the cDNA-transfected cells was confirmed by Western blot (Fig. 6D).

Nuclease resistance of pyrimidine TFOs

To test the possibility that differences in resistance to nucleases of the various TFOs could be responsible for the observed differences in induced recombination frequencies, we tested the stability of selected TFOs in serum. The 5'-end-labeled TFOs were added to FBS and aliquots were removed at time points from 0 to 24 hours to measure the quantity of full-length oligonucleotide by gel electrophoresis and autoradiography. After 1 hour in serum, most of the chemically modified pyrimidine TFOs were intact, except for the unmodified-TFO DNA18 and the 5meC-modified TFO G3TmC18 (Fig. 7). After 3 hours, the amidate18, LNA18, 2'-O-Me18, 2'-AE18, and the G-rich TFOs all showed greater than 50% remaining intact. The G3pUmC18 TFO showed around 36% oligonucleotide intact at 3 hours. By 6 hours only the 2'-AE18 and the amidate18 pyrimidine motif TFOs showed greater than 50% intact; the G3pUmC18 TFO showed about 25% intact. In all cases, less than 10% of the pyrimidine TFOs tested remained intact at 12 hours. Interestingly, two G-rich TFOs, AG30 and AG19 showed higher levels of nuclease resistance, with greater than 50% oligonucleotide remaining intact at 24 hours (Fig. 7) and persisting at this level even to 72 hours (data not shown). This serum stability shown by the G-rich TFOs may, in part, account for their documented effectiveness for gene targeting *in vivo* following systemic administration in mice [5].

Discussion

In the work reported here, we tested a series of TFOs with selected chemical modifications for their ability to bind as third strands in the pyrimidine motif under a variety of *in vitro* conditions, for their ability to induce recombination in two different cell-based episomal targeting assays, and for their ability to induce repair in cell-free extracts. We found that high affinity *in vitro* binding was necessary but not sufficient for increased intracellular activity. Certain modifications that provide for enhanced *in vitro* binding were only minimally effective at increasing the ability of TFOs to induce recombination in cells or to stimulate repair in cell-free extracts, while others (amidate18 and G3pUmC18) stood out in the initial screen and their activity held up under close evaluation (showing both high third-strand binding affinity and effective intracellular activity). These differences cannot simply be attributed to differential nuclease resistance.

The modified TFOs with the highest binding affinities under physiological conditions *in vitro* (low Mg^{2+}, high K^{+}, and neutral pH) were amidate18, G3pUmC18, 2'-AE18, and LNA18. Of these, only the amidate and pdU–modified TFOs mediated increases in recombination frequencies over background, 3.4 and 2.8-fold, respectively. The recombination frequencies in the COS-7 episomal shuttle vector assay produced by amidate18 (0.37%) and G3pUmC18 (0.31%) are comparable to the recombination frequency induced by AG30 in a similar shuttle vector assay (0.40%). These recombination frequencies are based on targeting an episomal not a chromosomal target.

Transfection of the episomal target 24 hours prior to oligonucleotide transfection allowed for chromatin assembly on the target plasmid [48]. In the purine motif, similar mutagenesis and recombination frequencies have been established in episomal and chromosomal targets using G-rich TFOs with both targets [4, 6, 17, 49].

Using a second episomal targeting assay based on intermolecular recombination in a *luciferase* reporter gene, similar results were seen with the amidate18 TFO, with a recombination frequency of 0.11% (which is a 48-fold induction of recombination compared to recombination by donor DNA alone in this assay). Thus, in two distinct episomal targeting systems in cells, the amidate18 TFO was seen to induce recombination in the absence of covalent adduct formation.

Additionally, the amidate18 and the G3pUmC18 TFOs induced repair in cell-free extracts. In cells, the induction of recombination in the shuttle vector assay by amidate18 was found to be dependent on XPA, consistent with the *in vitro* repair results and further suggesting that triplexes formed by amidate18 induce recombination in a repair-dependent manner.

Hence, while several of the chemically modified TFOs in the pyrimidine motif were found to bind well as third strands under physiologic conditions *in vitro*, only the amidate18 and G3pUmC18 TFOs were seen to induce recombination and repair. One explanation consistent with our results would be that the triplexes formed by the 2'-AE and LNA-modified TFOs are less effective at provoking DNA repair and thereby less effective at inducing recombination. Our prior work with purine motif TFOs indicated

that triplexes can be recognized as "lesions" in the DNA by factors in the NER pathway, leading to induced repair and recombination [4, 20, 27]. Our current results suggest that the same is true for the pyrimidine motif triplexes formed by amidate18 or G3pUmC18 TFOs.

Conceptually, this recognition is likely to result from helical alterations caused by triplex formation [50]. Interestingly, Asensio et al. [22] found in an NMR-based structural analysis of pyrimidine motif triplexes that the degree of duplex helical distortion can be influenced by substitutions at the 2'-position of the ribose on the TFO. Comparing TFOs with 2'-deoxy (DNA), 2'-OH (RNA), and 2'-O-methyl (2'-O-Me) ribose moieties, they found that the relative degrees of distortion were: DNA > RNA > 2'-O-Me. Hence, RNA-like substitution at the 2'-position in the TFO, which tends to promote a C3-endo sugar conformation [51], appears to reduce the helical distortion that occurs upon third-strand binding.

The TFOs showing the highest levels of induced recombination were amidate18 and G3pUmC18 in the pyrimidine motif and AG30 in the purine motif. All of these contain DNA-like deoxyribose. In contrast, the 2'-AE18 and LNA18 TFOs have RNA-like 2'-substituted sugars. In fact, the LNA 2' substitution (consisting of the 2'-O, 4'-C-methylene bridge) locks the ribose into the RNA-like C3-endo conformation [52]. This is proposed to be an advantage for triplex formation because it may serve to pre-organize the TFO into a uniform conformation favorable to third-strand binding [52], but it may also serve to minimize the resulting helical distortion. Therefore, based on this model, it

is possible that the amidate18, G3pUmC18, and AG30 TFOs may create triplexes that are more distorting than those produced by the 2'-AE18 and LNA18, resulting in an increased ability of the resulting triplexes to provoke DNA repair (as shown) and thereby stimulate recombination.

In conclusion, the work reported here demonstrates that non-covalent pyrimidine TFOs, when appropriately modified, can mediate high affinity binding, induction of recombination in cells, and induction of repair in cell-free extracts, at levels similar to those seen with G-rich purine TFOs at their respective target sites. These results expand the number of possible reagents for targeted genome modification.

The majority of this work has been published and the reference is: Kalish, J.M., et al., *Triplex-induced recombination and repair in the pyrimidine motif.* Nucleic Acids Research, 2005. 33(11): p. 3492-502.

References

1. Cooney, M., et al., *Site-specific oligonucleotide binding represses transcription of the human c-myc gene in vitro.* Science, 1988. **241**(4864): p. 456-9.
2. Birg, F., et al., *Inhibition of simian virus 40 DNA replication in CV-1 cells by an oligodeoxynucleotide covalently linked to an intercalating agent.* Nucleic Acids Research, 1990. **18**(10): p. 2901-8.
3. Faruqi, A.F., et al., *Recombination induced by triple-helix-targeted DNA damage in mammalian cells.* Mol Cell Biol, 1996. **16**(12): p. 6820-8.
4. Faruqi, A.F., et al., *Triple-helix formation induces recombination in mammalian cells via a nucleotide excision repair-dependent pathway.* Mol Cell Biol, 2000. **20**(3): p. 990-1000.
5. Vasquez, K.M., L. Narayanan, and P.M. Glazer, *Specific mutations induced by triplex-forming oligonucleotides in mice.* Science, 2000. **290**(5491): p. 530-3.
6. Wang, G., et al., *Targeted mutagenesis in mammalian cells mediated by intracellular triple helix formation.* Mol Cell Biol, 1995. **15**(3): p. 1759-68.
7. Moser, H.E. and P.B. Dervan, *Sequence specific cleavage of double helical DNA by triple helix formation.* Science, 1987. **238**: p. 645-650.

8. Francois, J.C., T. Saison-Behmoaras, and C. Helene, *Sequence-specific recognition of the major groove of DNA by oligodeoxynucleotides via triple helix formation. Footprinting studies.* Nucleic Acids Research, 1988. **16**(24): p. 11431-40.

9. Letai, A.G., et al., *Specificity in formation of triple-stranded nucleic acid helical complexes: studies with agarose-linked polyribonucleotide affinity columns.* Biochemistry, 1988. **27**(26): p. 9108-12.

10. Beal, P.A. and P.B. Dervan, *Second structural motif for recognition of DNA by oligonucleotide-directed triple-helix formation.* Science, 1991. **251**(4999): p. 1360-3.

11. Knauert, M.P. and P.M. Glazer, *Triplex forming oligonucleotides: sequence-specific tools for gene targeting.* Hum Mol Genet, 2001. **10**(20): p. 2243-51.

12. Chan, P.P. and P.M. Glazer, *Triplex DNA: fundamentals, advances, and potential applications for gene therapy.* J Mol Med, 1997. **75**(4): p. 267-82.

13. Lacroix, L., et al., *Triplex formation by oligonucleotides containing 5-(1-propynyl)-2'-deoxyuridine: decreased magnesium dependence and improved intracellular gene targeting.* Biochemistry, 1999. **38**(6): p. 1893-901.

14. Puri, N., et al., *Minimum number of 2'-O-(2-aminoethyl) residues required for gene knockout activity by triple helix forming oligonucleotides.* Biochemistry, 2002. **41**(24): p. 7716-24.

15. Sachsenmaier, C., et al., *Involvement of growth factor receptors in the mammalian UVC response.* Cell, 1994. **78**(6): p. 963-72.

16. Majumdar, A., et al., *Targeted gene knockout mediated by triple helix forming oligonucleotides.* Nat Genet, 1998. **20**(2): p. 212-4.

17. Vasquez, K.M., et al., *Chromosomal mutations induced by triplex-forming oligonucleotides in mammalian cells.* Nucleic Acids Res, 1999. **27**(4): p. 1176-81.

18. Vasquez, K.M., et al., *Chromosome targeting at short polypurine sites by cationic triplex-forming oligonucleotides.* J Biol Chem, 2001. **276**(42): p. 38536-41.

19. Barre, F.X., et al., *Covalent crosslinks introduced via a triple helix-forming oligonucleotide coupled to psoralen are inefficiently repaired.* Nucleic Acids Res, 1999. **27**(3): p. 743-9.

20. Wang, G. and P.M. Glazer, *Altered repair of targeted psoralen photoadducts in the context of an oligonucleotide-mediated triple helix.* J Biol Chem, 1995. **270**(38): p. 22595-601.

21. Faruqi, A.F., M. Egholm, and P.M. Glazer, *Peptide nucleic acid-targeted mutagenesis of a chromosomal gene in mouse cells.* Proc Natl Acad Sci U S A, 1998. **95**(4): p. 1398-403.

22. Asensio, J.L., Carr R., Brown, T., Lane, A.N., *Conformational and Thermodynamic Properties of Parallel Intramolecular Triple Helices Containing a DNA, RNA, or 2'-OMeDNA Third Strand.* J. Am. Chem. Soc., 1999. **121**: p. 11063-11070.

23. Parris, C.N. and M.M. Seidman, *A signature element distinguishes sibling and independent mutations in a shuttle vector plasmid.* Gene, 1992. **117**(1): p. 1-5.

24. Kraemer, K.H. and M.M. Seidman, *Use of supF, an Escherichia coli tyrosine suppressor tRNA gene, as a mutagenic target in shuttle-vector plasmids.* Mutat Res, 1989. **220**(2-3): p. 61-72.

25. Knauert, M.P., et al., *Distance and Affinity Dependence of Triplex-Induced Recombination.* Biochemistry, 2005. **44**(10): p. 3856-3864.

26. Bradford, M.M., *A rapid and sensitive method for the quantitation of microgram quantities of protein utilizing the principle of protein-dye binding.* Anal Biochem, 1976. **72**: p. 248-54.

27. Datta, H.J., et al., *Triplex-induced Recombination in Human Cell-free Extracts. Dependence on XPA and HsRad51.* J Biol Chem, 2001. **276**(21): p. 18018-23.

28. Rogers, F.A., et al., *Site-directed recombination via bifunctional PNA-DNA conjugates.* Proc Natl Acad Sci U S A, 2002. **99**(26): p. 16695-700.

29. Gryaznov, S.M., et al., *Oligonucleotide N3'-->P5' phosphoramidates.* Proc Natl Acad Sci U S A, 1995. **92**(13): p. 5798-802.

30. Testa, S.M., S.M. Gryaznov, and D.H. Turner, *Antisense binding enhanced by tertiary interactions: binding of phosphorothioate and N3'-->P5' phosphoramidate hexanucleotides to the catalytic core of a group I ribozyme from the mammalian pathogen Pneumocystis carinii.* Biochemistry, 1998. **37**(26): p. 9379-85.

31. Skorski, T., et al., *Antileukemia effect of c-myc N3'-->P5' phosphoramidate antisense oligonucleotides in vivo.* Proc Natl Acad Sci U S A, 1997. **94**(8): p. 3966-71.

32. Gryaznov, S., et al., *Oligonucleotide N3'-->P5' phosphoramidates as antisense agents.* Nucleic Acids Res, 1996. **24**(8): p. 1508-14.

33. Giovannangeli, C., et al., *Accessibility of nuclear DNA to triplex-forming oligonucleotides: the integrated HIV-1 provirus as a target.* Proc Natl Acad Sci U S A, 1997. **94**(1): p. 79-84.

34. Matteucci, M., et al., *Sequence-Sepcific Targeting of Duplex DNA Using a Camptothecin-Triple Helix Forming Oligonucleotide Conjugate and Topoisomerase I.* J. Am. Chem. Soc., 1997. **119**: p. 6939-6940.

35. Fenster, S.D., et al., *Inhibition of human immunodeficiency virus type-1 env expression by C-5 propyne oligonucleotides specific for Rev-response element stem-loop V.* Biochemistry, 1994. **33**(28): p. 8391-8.

36. Wahlestedt, C., et al., *Potent and nontoxic antisense oligonucleotides containing locked nucleic acids.* Proc Natl Acad Sci U S A, 2000. **97**(10): p. 5633-8.

37. Torigoe, H., et al., *2'-O,4'-C-methylene bridged nucleic acid modification promotes pyrimidine motif triplex DNA formation at physiological pH: thermodynamic and kinetic studies.* J Biol Chem, 2001. **276**(4): p. 2354-60.

38. Arzumanov, A., et al., *Inhibition of HIV-1 Tat-dependent trans activation by steric block chimeric 2'-O-methyl/LNA oligoribonucleotides.* Biochemistry, 2001. **40**(48): p. 14645-54.

39. McCaffrey, A.P., et al., *A potent and specific morpholino antisense inhibitor of hepatitis C translation in mice.* Hepatology, 2003. **38**(2): p. 503-8.

40. Lacroix, L., et al., *Pyrimidine morpholino oligonucleotides form a stable triple helix in the absence of magnesium ions.* Biochem Biophys Res Commun, 2000. **270**(2): p. 363-9.

41. Lee, J.S., et al., *Poly(pyrimidine) . poly(purine) synthetic DNAs containing 5-methylcytosine form stable triplexes at neutral pH.* Nucleic Acids Res, 1984. **12**(16): p. 6603-14.

42. Xodo, L.E., et al., *Effect of 5-methylcytosine on the stability of triple-strand DNA-A thermodynamic study.* Nucleic Acids Res., 1991. **19**: p. 5625-5631.

43. Vasquez, K.M., et al., *High-affinity triple helix formation by synthetic oligonucleotides at a site within a selectable mammalian gene.* Biochemistry, 1995. **34**(21): p. 7243-51.

44. Cheng, A.J. and M.W. Van Dyke, *Monovalent cation effects on intermolecular purine-purine-pyrimidine triple-helix formation.* Nucleic Acids Res, 1993. **21**(24): p. 5630-5.

45. Wang, G., M.M. Seidman, and P.M. Glazer, *Mutagenesis in mammalian cells induced by triple helix formation and transcription-coupled repair.* Science, 1996. **271**: p. 802-805.

46. Sibghatullah, et al., *Human nucleotide excision repair in vitro: repair of pyrimidine dimers, psoralen and cisplatin adducts by HeLa cell-free extract.* Nucleic Acids Res, 1989. **17**(12): p. 4471-84.

47. Wood, R.D., P. Robins, and T. Lindahl, *Complementation of the xeroderma pigmentosum DNA repair defect in cell-free extracts.* Cell, 1988. **53**(1): p. 97-106.

48. Cereghini, S. and M. Yaniv, *Assembly of transfected DNA into chromatin: structural changes in the origin-promoter-enhancer region upon replication.* Embo J, 1984. **3**(6): p. 1243-53.

49. Luo, Z., et al., *High-frequency intrachromosomal gene conversion induced by triplex-forming oligonucleotides microinjected into mouse cells.* Proc Natl Acad Sci U S A, 2000. **97**(16): p. 9003-8.

50. Vasquez, K.M., et al., *Human XPA and RPA DNA repair proteins participate in specific recognition of triplex-induced helical distortions.* Proc Natl Acad Sci U S A, 2002. **99**(9): p. 5848-53.

51. Holland, J.A. and D.W. Hoffman, *Structural features and stability of an RNA triple helix in solution.* Nucleic Acids Res, 1996. **24**(14): p. 2841-8.

52. Sun, B.W., et al., *Sequence and pH effects of LNA-containing triple helix-forming oligonucleotides: physical chemistry, biochemistry, and modeling studies.* Biochemistry, 2004. **43**(14): p. 4160-9.

Figures and Tables

Table 1. Binding of chemically modified pyrimidine oligonucleotides to an 18 bp A-rich site.[a]

oligonucleotide	pH 5.4 10 mM $MgCl_2$ 0 mM KCl	pH 7.2 10 mM $MgCl_2$ 0 mM KCl	pH 7.2 0.1 mM $MgCl_2$ 140 mM KCl
amidate18	2.8×10^{-10}	8.0×10^{-10}	1.0×10^{-8}
pDEED18	3.0×10^{-8}	none	none
full DEED18	1.0×10^{-6}	none[b]	none
G3TmC18	3.0×10^{-9}	7.0×10^{-9}	3.0×10^{-7}
G3pUmC18	4.5×10^{-10}	7.0×10^{-9}	9.0×10^{-8}
2'-O-Me18	4.6×10^{-10}	1.5×10^{-10}	8.0×10^{-6}
2'-AE18	4.0×10^{-10}	6.0×10^{-10}	2.0×10^{-8}
LNA18	3.0×10^{-10}	2.7×10^{-10}	2.5×10^{-8}
morph18[c]	3.0×10^{-8}	3.5×10^{-6}	5.1×10^{-6}

[a] Binding affinities were calculated as equilibrium dissociation constants (K_d s) under the indicated conditions, as determined by gel mobility shift assay under non-denaturing conditions. K_d values are given in molarity. Binding was done for 16 hours at 37°C unless otherwise specified.
[b] Binding was done at pH 7.6 with 1 mM spermine.
[c] Binding was done at 23°C.

Duplex target and oligonucleotides

```
     G3Y 5'tcgaggTCTTTTCTTTTTTCTTTTaggggggg3'
     G3R 3'agctccAGAAAAGAAAAAAGAAAAtccccccc5'

amidate18      3' tcttttctttttttctttt 5'
pDEED18        3' TcTTTTcTTTTTtCTTTt 5'
full DEED18    3' tcttttctttttttctttt 5'
G3TmC18        3' TcTTTTcTTTTTTcTTTT 5'
G3pUmC18       3' ucuuuucuuuuuucuuuu 5'
LNA18          3' TcTtTtCtTtTtTcTtTt 5'
2'-O-Me18      3' tcttttctttttttctttt 5'
2'-AE18        3' UcuuuUCUUUUUUCUUUU 5'
morph18        3' tcttttctttttttctttt 5'
```

Figure 1A. Duplex target and oligonucleotide modifications. The third-strand binding site is indicated in capital letters in the duplex (G3Y/G3R) used for binding studies. For the oligonucleotides, lower case letters indicate modified nucleotides with the specific chemistry as indicated in the name of the oligonucleotide. Upper case letters indicate unmodified nucleotides. All internucleoside linkages in amidate18 are N3'->P5' phosphoramidate (amidate)-modified. N, N-diethylethylenediamine (DEED)-modified nucleotides are present in partial DEED (pDEED18) and full DEED18 as indicated. G3TmC18 contains 5-methyl-2'-deoxycytidines (5meC) in place of cytidines. G3pUmC18 contains 5meC-modifications on all cytidines and all thymines are replaced by 5-(1-propynyl)-2'-deoxyuridines (pdU). LNA18 contains unmodified nucleotides alternating with 2'-O, 4'-C-methylene bridged or locked nucleic acid (LNA)-modified nucleotides. 2'-O-Me18 contains 2'-O-methyl-ribose (2'-O-Me)-modified nucleotides and all cytidines have 5meC-modifications. 2'-AE18 contains 2'-O-(2-aminoethyl)-ribose (2'-AE)-modified nucleotides at the four lower case letter positions, while the upper case positions have 2'-O-Me-modifications and all positions also have either 5meC-modifications or 5-methyl-2'-deoxyuridine modifications. All nucleotides in morph18 are morpholino (morph)-modified.

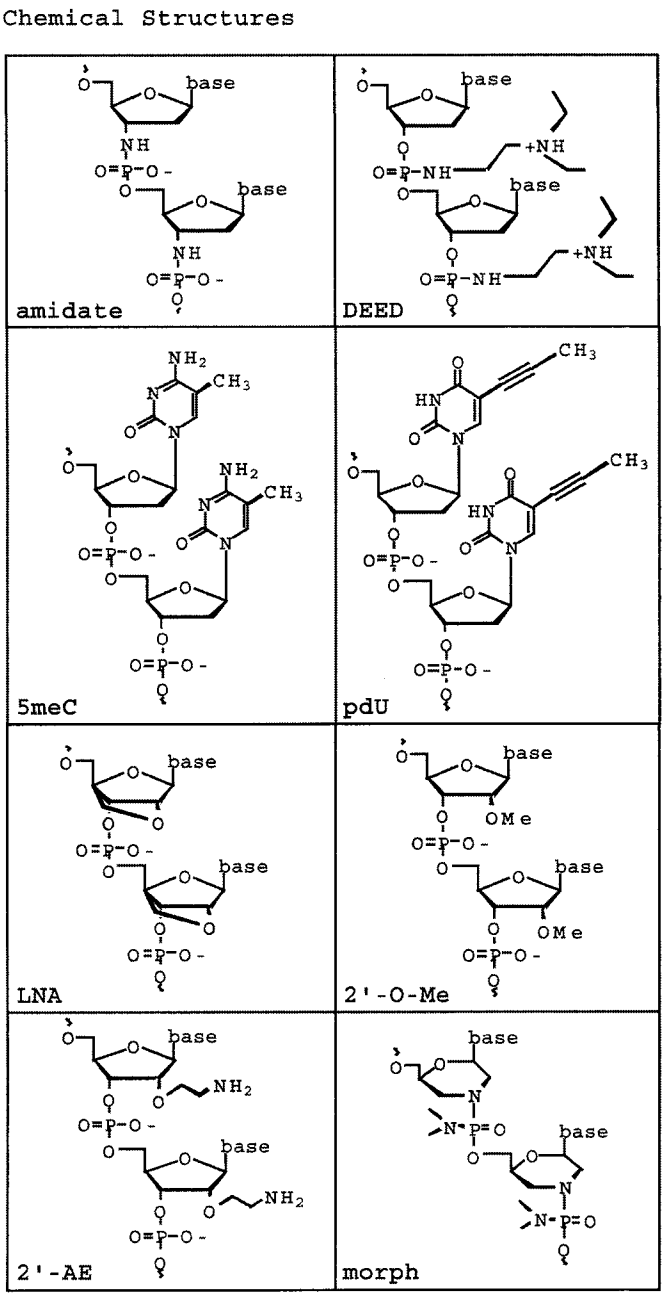


Figure 1B. Structures of the chemical modifications. N3'->P5' phosphoramidate (amidate); N, N-diethylethylenediamine (DEED); 5-methyl-2'-deoxycytidine (5meC); 5-(1-propynyl)-2'-deoxyuridine (pdU); 2'-O, 4'-C-methylene bridged or locked nucleic acid (LNA); 2'-O-methyl-ribose (2'-O-Me); 2'-O-(2-aminoethyl)-ribose (2'-AE); morpholino (morph).

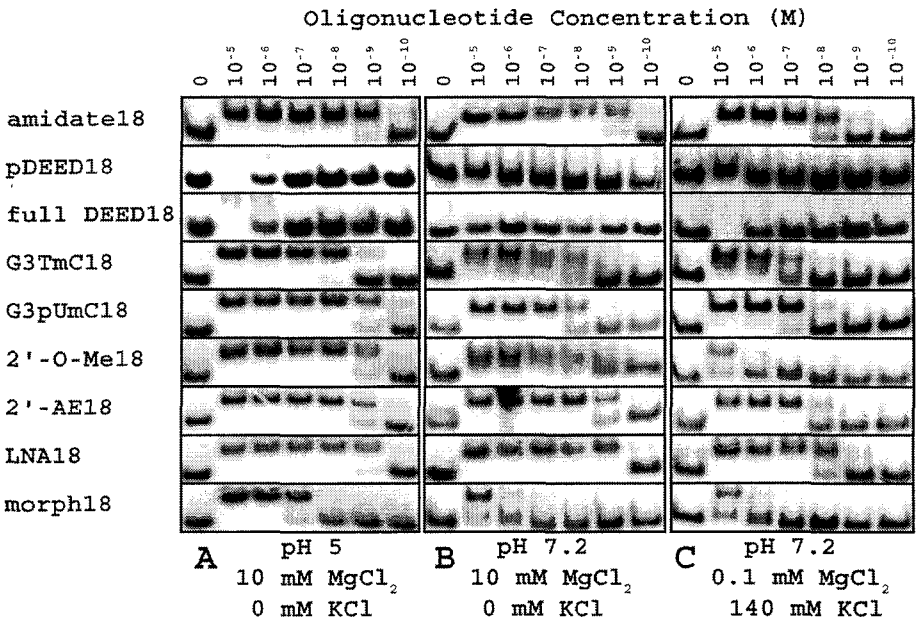


Figure 2. Third-strand binding under selected conditions as determined by gel mobility shift analysis. The first lane in each panel contains duplex alone (0 concentration of TFO), the second lane contains 10^{-5} M oligonucleotide, and each subsequent column contains a ten-fold dilution. Third-strand binding conditions: (A) pH 5.4, 10 mM $MgCl_2$, and 0 mM KCl; (B) pH 7.2, 10 mM $MgCl_2$, and 0 mM KCl; (C) pH 7.2, 0.1 mM $MgCl_2$, and 140 mM KCl. The bands of reduced mobility relative to the duplex alone represent triplex formation.

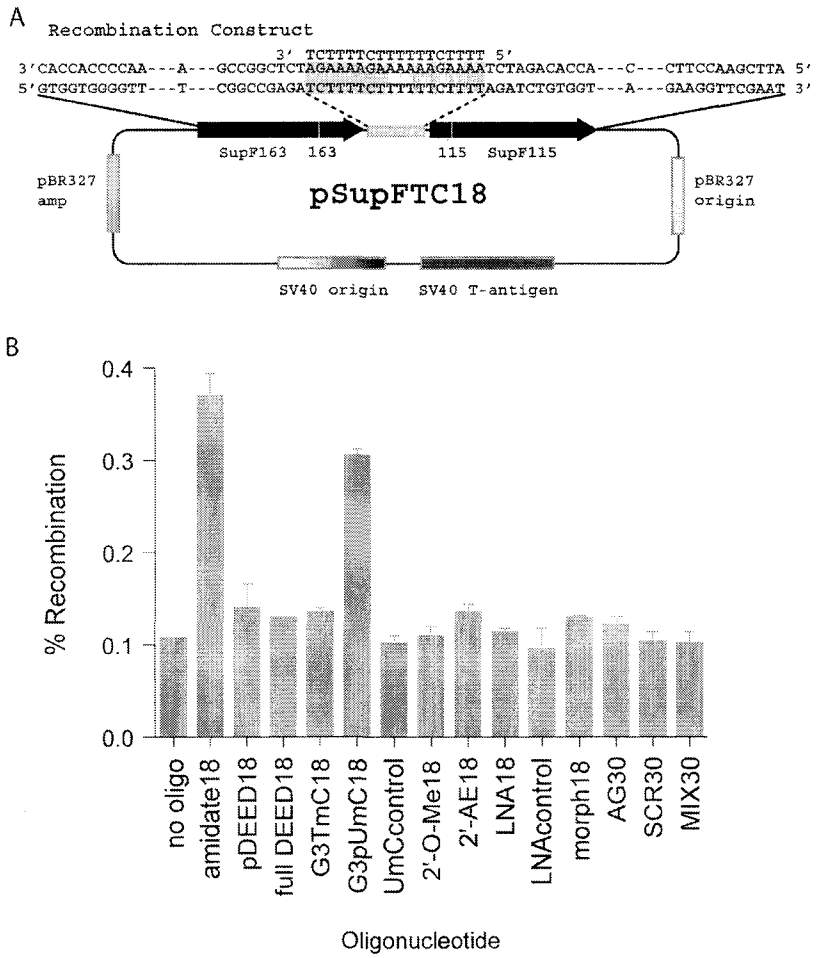


Figure 3. Induced recombination mediated by pyrimidine TFOs. Schematic representation of the pSupFTC18 vector. This SV40-based vector contains two mutant *supF* genes in tandem. The upstream mutant *supF* gene, *supF163*, contains a C-to-T point mutation at nucleotide position 163. The downstream mutant *supF* gene, *supF115*, contains a G-to-A point mutation at nucleotide position 115. In between the two mutant *supF* genes is an 18 bp A-T rich homopurine/homopyrimidine target site. The chemically modified oligonucleotides described in Fig. 1A were designed to bind to this target sequence. (B) Triplex-induced recombination in COS-7 cells. The plasmid described in (A) was pre-transfected into COS-7 cells followed one day later by transfection of the indicated oligonucleotides or mock transfection. After 48 hours, shuttle vector DNA was isolated and transformed into indicator bacteria to analyze *supF* gene function and allow quantification of recombination events. Percent recombination was calculated by the ratio of the number of blue colonies to total colonies between cells transfected with the modified TFOs and mock-transfected cells.

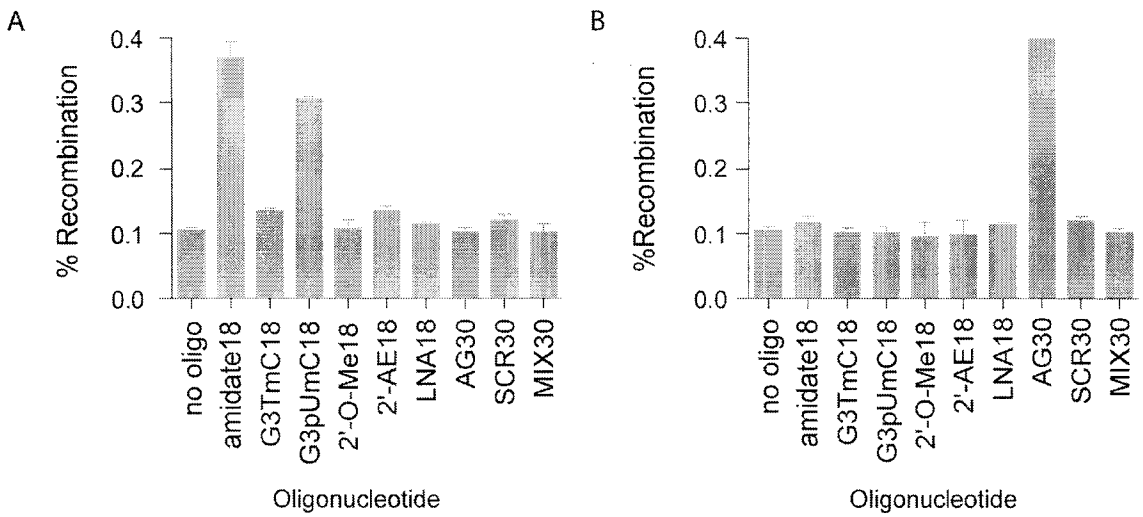


Figure 4. Sequence-specificity of the TFO-induced recombination. (A) Triplex-induced recombination with a pyrimidine motif target. Cells pre-transfected with the dual *supF* plasmid containing the 18 bp A-T rich site, pSupFTC18, were transfected the next day with lipid reagent alone or with the indicated TFOs. AG30 and SCR30 are G-rich TFOs and MIX30 has a mixed sequence. Shuttle vector rescue and analysis were performed as for Fig. 3. (B) Triplex-induced recombination with a purine motif target. Cells pre-transfected with a dual *supF* plasmid containing a 30 bp G-rich site, pSupFAR, were transfected one day later with lipid reagent alone or with the indicated TFOs. Analysis of recombination events was performed as above. Amidate18 and G3pUmC18 bind with high affinity to the A-rich target site in the intergenic region in pSupFTC18 but not to the G-rich site in pSupFAR. AG30 binds well to the site in pSupFAR but not to that in pSupFTC18.

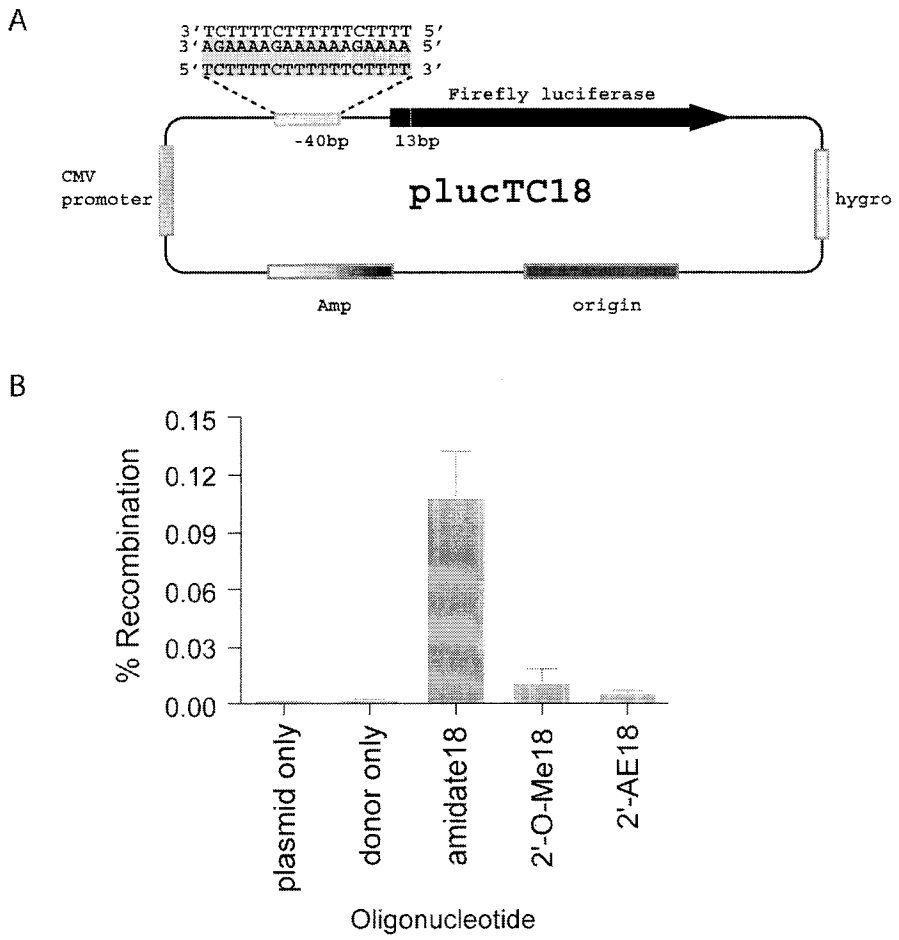


Figure 5. Intermolecular induced recombination mediated by pyrimidine TFOs. (A) Schematic of the plucTC18 construct. The plasmid contains the pyrimidine triplex binding site located 40 bp upstream of the start codon for the *firefly luciferase* gene. The *luciferase* gene has a nonsense mutation at 13 bp downstream from the start codon. (B) Percent recombination in CHO cells. Plasmid, TFO, and the donor oligonucleotide (as51) were co-transfected into CHO cells. After 48 hours, cells were lysed and recombination was measured based on relative light units and normalized to protein.

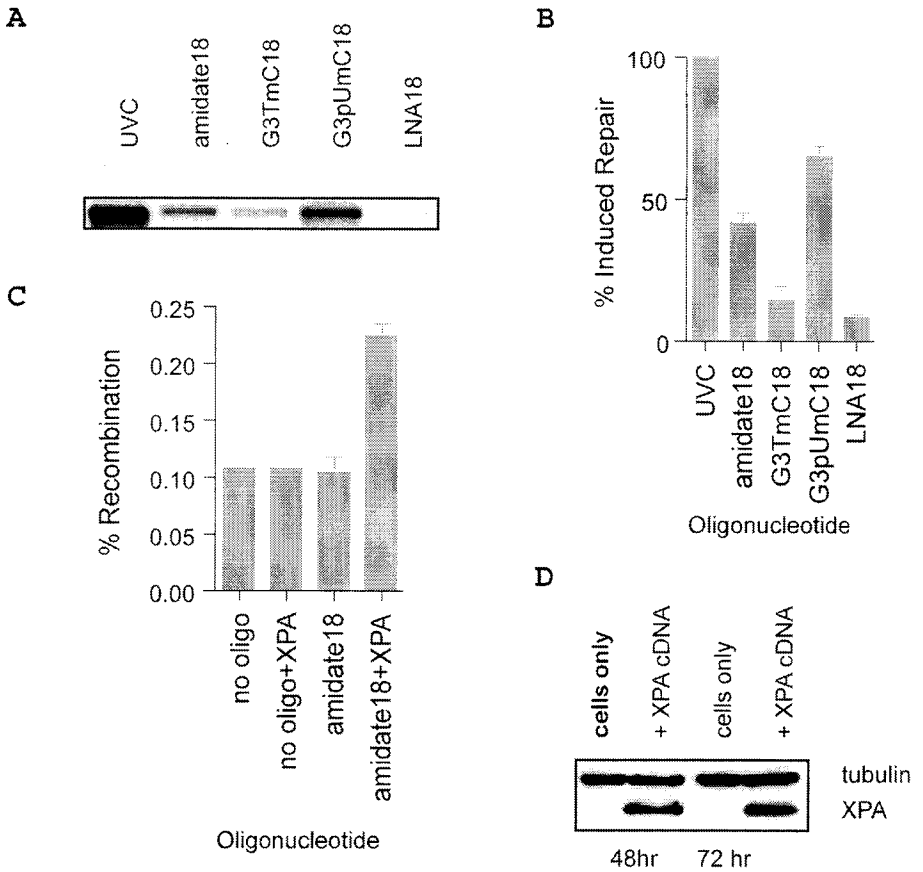


Figure 6. Pyrimidine motif TFOs and DNA repair. (A) Pyrimidine triplex-induced DNA repair in HeLa cell-free extracts. TFOs and plasmids were co-incubated in HeLa cell-free extracts supplemented with [α-^{32}P]dCTP. Induced DNA repair synthesis was measured by visualization and quantification of incorporation of [α-^{32}P]dCTP into the plasmid. As a positive control, plasmid alone was treated with UVC irradiation. (B) Quantification of induced repair as measured by [α-^{32}P]dCTP incorporation in (A); normalized to UVC-treated plasmid representing 100% repair activity. (C) Dependence of triplex-induced recombination on XPA. XP12BE (XPA-deficient) cells were transfected with the dual *supF* reporter construct containing the pyrimidine motif triplex binding site and then transfected or not with the amidate18 TFO. Parallel samples were also transfected with XPA cDNA at the same time as the dual *supF* vector. Induced recombination was assayed as in Fig. 3 and Fig. 4. (D) Levels of XPA expression in XP12BE cells versus XP12BE cells transfected with a XPA cDNA construct. XP12BE cells and XP12BE cells transfected with XPA cDNA were collected at 48 and 72 hours post-transfection and expression of XPA was detected by Western blot.

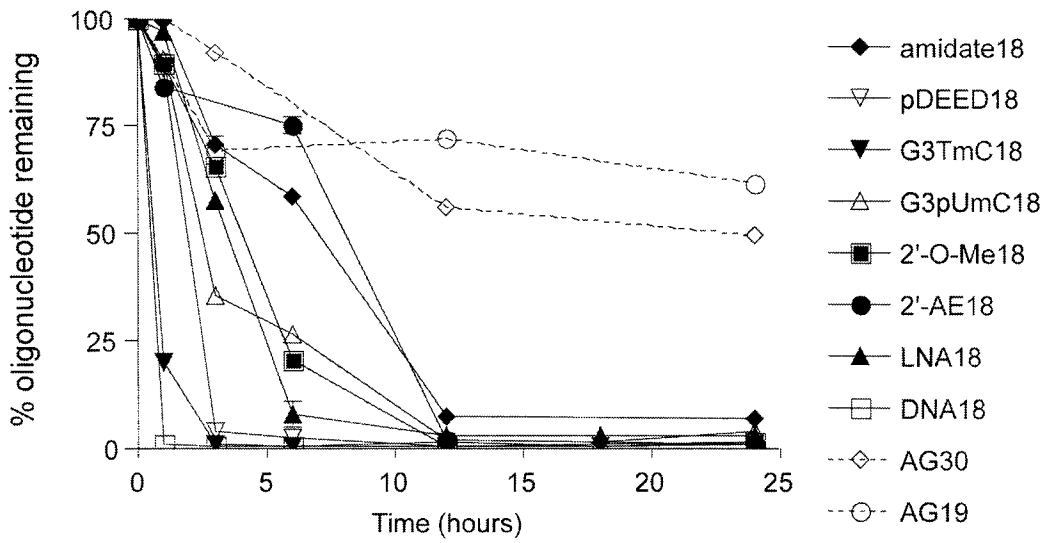


Figure 7. Serum stability of pyrimidine TFOs. [γ-^{32}P]ATP-labeled TFOs were added to FBS and aliquots were removed at time points from 0 to 24 hours. The percent oligonucleotide remaining represents the amount of full-length TFO still intact at the given time point.

Chapter Three:
Triplex-Stimulated Intermolecular Recombination at a Single-Copy Genomic Target

Introduction

Gene targeting via homologous recombination (HR) is well established as a research tool to generate knock-out or genetically modified mammalian cells and mice [1]. However, the technology for gene knock-out depends on selection techniques to enrich for cells in which HR events have occurred and to eliminate cells in which off-target vector integration has taken place [1, 2]; it does not alter the absolute frequency of targeted HR. For gene therapy applications, such techniques are less useful because the ability to apply clonal selection to human stem cells is limited. Consequently, most gene therapy efforts have focused on the introduction of cDNAs within expression constructs, typically using viral vectors, to complement inherited genetic defects by gene addition rather than by HR-mediated gene correction. Integration usually occurs in a non-targeted manner, often resulting in sub-optimal expression, inactivation, or even loss of the construct. Site-specific gene correction by HR would be preferable to ensure appropriate expression and function.

Several approaches are under study to enhance the frequency of targeted HR for eventual application to gene therapy. One involves the use of single-stranded viral vectors, such as the adeno-associated virus (AAV), that may undergo homology-directed integration at increased frequencies [3, 4]. However, these vectors can still mediate significant levels of non-homologous integration events [3, 4]. Alternatively, studies have demonstrated that the production of site-specific chromosomal damage can substantially increase the frequency of homologous recombination by exogenous DNAs

[5]. In mammalian cells, evidence suggests that the search for homology is not necessarily the rate-limiting step for gene targeting [5]. Rather, DNA damage, and in particular strand breaks, may be the key step to sensitize the target site for HR. Work using rare-cutting endonucleases such as I-*Sce*I has shown that site-directed double-strand breaks (DSBs) can stimulate HR by 1,000-fold or more [6-11]. However, I-*Sce*I is not applicable to gene therapy because its use requires the prior introduction of its 18-bp recognition site into the genome.

To overcome this limitation, one approach has been the development of artificial nucleases with engineered binding domains as tools to create strand breaks at specified sites. Recent studies have demonstrated that expression of custom-designed zinc finger nucleases (ZFNs) in Drosophila embryos or in mammalian cells in culture can be effective at promoting site-directed recombination [12-15]. However, the off-target effects of such ZFNs have not been well-studied.

Another method to create site-specific DNA damage is based on the use of triplex-forming oligonucleotides (TFOs). Triplex DNA can be formed when oligonucleotides bind as third strands in a sequence-specific manner within the major groove of the double helix at polypurine/polypyrimidine stretches. In one motif, a pyrimidine-rich oligonucleotide binds in a direction parallel (in terms of 5' to 3' direction) to the purine strand in the duplex [16-19]. In this so-called pyrimidine or parallel-motif triplex, T binds to A:T base pairs (bp) and C binds to G:C bp [16]. In the alternative purine motif, favored at G:C bp-rich target sites, a purine-rich strand binds anti-parallel to

the purine strand in the Watson-Crick duplex, with A binding to A:T bp and G binding to G:C bp [20-24].

Initial studies testing the use of TFOs to induce recombination focused on the coupling of the TFOs to DNA-reactive conjugates, such as psoralen, to deliver site-specific DNA adducts that might provoke recombination [25]. However, subsequent work demonstrated that triple helix formation itself, without any associated covalent DNA damage, can stimulate recombination via recruitment of the nucleotide excision repair (NER) pathway in episomal shuttle vector targets in human cells or in plasmid substrates in cell-free extracts [26, 27]. Studies also demonstrated that there is no requirement for covalent linkage of the TFO and the donor molecule, and in the case of episomal targets in mammalian cells, there can be substantial distances of several hundred bp between the TFO binding site (TBS) and the recombination target [28]. In other work, it was shown that intramolecular recombination between tandem repeat genes at a chromosomal locus can also be induced by TFOs, but intermolecular recombination was not tested [29, 30].

Here, we report the ability of TFOs to stimulate intermolecular recombination at a single-copy chromosomal locus in mammalian cells. CHO-derived cell lines carrying single copies of *Firefly luciferase* (*Fluc*) genes inactivated by stop codon mutations were co-transfected with TFOs designed to form either parallel or anti-parallel triplexes at sites upstream of the *Fluc* coding region in combination with single-stranded donor DNAs (ssDNAs) of lengths between 30 and 69 nucleotides (nt) carrying wild-type *Fluc*

sequences. The mixture of the TFOs with the donor DNAs was 5- to 9-fold more effective at mediating *Fluc* gene correction and restoring *luciferase* activity than was transfection of the donor DNAs alone. Correction frequencies of up to 0.1% were obtained in a TFO sequence-, target site-, and dose-dependent manner. The results establish the utility of site-directed triplex formation as a tool to boost gene targeting by donor DNAs at chromosomal sites.

Materials and Methods

Oligonucleotides

Oligonucleotides were synthesized by The Midland Certified Reagent Company Inc. (Midland, Texas) and purified by RP-HPLC. Oligonucleotide sequences are as follows: AG30 (5' AGGAAG GGGGGG GTGGTG GGGGAG GGGGAG 3'); 12AG30 (5' GGAGGA GTGGAG GGGAGT GAGGGG GGGGGG 3'); s51 (5' TGGTAAAGCC ACCATGGAAG ACGCCAAAAA CATAAAGAAA GGCCCGGCGCC 3'); as30 5' (CTTTCTTTAT GTTTTTGGCG TCTTCCATGG 3'); as51 (5' GCCGGGCCTT TCTTTATGTT TTTGGCGTCT TCCATGGTGG CTTTACCAAGC 3'); as69 (5' CCAGCGGATA GAGTGGCGCC GGGCCTTTCT TTATGTTTTT GGCGTCTTCC ATGGTGGCTT TACCAAGCT 3'). AG30 and 12AG30 were 3'-end-protected with a C7 amine group. *Fluc* donor molecules including as30, as51, s51, and as69 were 5'- and 3'-end-protected with three phosphorothioate internucleoside linkages at each end. Donors with three locked nucleic acid (LNA) residues on each end LNA30, LNA51, and

LNA69 (of the same sequences as as30, as51, and as69, respectively) were synthesized by Eurogentec (Seraing, Belgium). Amidate18 was obtained from Transgenomic as described [31]. The TC18 sequence of amidate18 was 5'-TTTTCTTTTT TCTTTTCT-3' and consisted of phosphoramidate internucleoside linkages (3' bridging nitrogen) between all nucleosides. The TFO, LNA18, modified with LNAs at every other residue was also synthesized by Eurogentec and was of the same sequence as amidate18. N, N-dimethylaminopropylphosphoramidate (αDMAP and βDMAP) and diethylethylenediamine (αDEED) TFOs of the TC18 sequence with these chemistries between all nucleosides were kindly synthesized by Francoise Debart and Jean-Jacques Vasseur. (R)-1-O-[4-(1-pyrenylethynyl)phenylmethyl]glycerol, twisted intercalating nucleic acid (TINA) TFOs TINA849 and TINA855, synthesized by Vyacheslav Filichev, had the following substitutions 5'ttttPctttt ttctPtttctP3' and 5'UUUUPCUUUUP UUCUPUUUPCUa3', respectively, where P is a TINA, a is a C-7 amino group, C is 5-methyl-2'-deoxycytidine (5meC), and U is 5-(1-propynyl)-2'-deoxyuridine (pdU) (and c and t are unmodified nucleobases). Camptothecin (CPT) (topoisomerase I inhibitor) conjugated TFOs were synthesized by Paola Arimondo. The CPT TFOs all have the sequence 5'CUCUCUCUCU UUUUUU3' (where U is pdU and C is 5meC), with a four carbon linker conjugated to the 3'-end of the TFO and connected to either 10-carboxy camptothecin (TUCL4CPT), two phenol rings (TUCL4Ph2), or linker alone (TUCL4NH2). Psoralen was conjugated to the 5'-end of TUCL4CPT (psoTUCL4CPT) and TUCL4Ph2 (psoTUCL4Ph2). The topoisomerase II inhibitors have a $(CH_2CH_2O)_6$

linker conjugated to the same sequence as above (TUC-C18-ICP165) or to a 20-mer (5'CUCUCUCUCUUUUUUUCUCU3', TUC20-C18-ICP165).

Plasmid vectors

Plasmids containing the *Fluc* gene were created by subcloning of *Fluc*+ from Promega's pGL3-Basic Vector (Madison, Wisconsin) into pcDNA5/FRT (Invitrogen, La Jolla, California) as described [28, 31]. Briefly, an insert containing the duplex AG30 target or the duplex TC18 target was introduced upstream of the *Fluc* ATG start site. Nonsense mutations were created via site-directed mutagenesis at bp 13 relative to the *Fluc* start codon. Lack of luciferase activity in all mutated *Fluc* vectors was confirmed using the Promega Luciferase Assay (see below).

Cell lines

CHO-Flp cell lines were obtained from Invitrogen (Carlsbad, California) and maintained according to the manufacturers' instructions. Cells were transfected with a 1:10 ratio of the appropriate *Fluc* plasmid to pOG44 (Flp recombinase expression vector) to generate the following cell lines: CHO-RWT (AG30 purine TFO binding site, wild-type *Fluc*), CHO-R13 (AG30 purine TFO binding site, nonsense mutation at bp 13 of *Fluc*), CHO-YWT (TC18 pyrimidine TFO binding site, wild-type *Fluc*), CHO-Y13 (TC18 pyrimidine TFO binding site, nonsense mutation at bp 13 of *Fluc*), and CHO-YScr (A:T bp-rich site that has 9 mismatches relative to the TC18 site, with the *Fluc* nonsense mutation at bp 13). CHO-YScr and CHO-YScr202 have the TBS, 5'CUCUCUCUCU UUUUUU3' and the *Fluc* nonsense mutation at bp 13 and bp 202 from the start site,

respectively. Plasmid transfections were carried out using GenePORTER2 Transfection Reagent (Gene Therapy Systems, San Diego, California). Cells were placed in nonselective media for 48 hours and then placed into hygromycin selection (600 μg/mL); colonies were allowed to form and single-copy integration at the expected site was confirmed via Southern blot [32].

Establishment of relative light unit standards and determination of assay sensitivity

Using pairs of matched cell lines mixed in fixed ratios of wild-type to mutant *Fluc* cells, a standard curve was established. Cultures of matched single-copy *Fluc* wild-type and *Fluc* mutant cells for each target site set were grown to sub-confluence, detached by trypsinization, resuspended in medium and counted to establish cell density (cells/mL). The pairs were mixed as follows: CHO-R13 with CHO-RWT cells and CHO-Y13 with CHO-YWT cells. The suspensions of the mutant cells (either CHO-R13 or CHO-Y13) were adjusted to 1×10^5 cells per mL, and aliquots of 50,000 mutant cells were seeded into individual wells (4 cm^2/well). Wild-type *Fluc* cells (CHO-RWT or CHO-YWT) were adjusted to 5×10^4 cells per mL, serially diluted, and then added to the wells containing the mutant cells. A range of samples was generated to contain varying proportions of wild-type cells mixed with the *Fluc* mutant cells (6 replicates per data point). All cell cultures were allowed to grow for 48 hours and then the mixed cell populations were harvested and lysed. The cell lysates were analyzed for luciferase activity using Promega Luciferase Assay Reagent (Promega, Madison, Wisconsin) according to manufacturer's instructions. Relative light units (RLUs) were normalized to the quantity of protein per

tested sample. These standards were used as a benchmark for the RLU data. A plot of RLU/protein versus percentage of wild-type cells was generated for each experiment, and this was used to normalize the experimental data and allow the assignment of recombination frequencies to the experimental RLU values.

Chromosomal gene targeting protocol and luciferase assay

CHO-R13 or CHO-Y13 cells were transfected with the single-stranded donor DNA oligonucleotides (ssDNAs) and with selected TFOs (AG30, amidate18, or control oligonucleotides) using Geneporter 2 according to manufacture's instructions. Briefly, 24 hours prior to transfection, 12-well plates (4 cm^2 per well) were seeded with 5×10^4 cells/well. Quantities of TFOs and donor DNAs transfected were as indicated for each experiment; in all cases transfections were done in triplicate. Forty-eight hours following transfection, cells were harvested for lysate preparation and analysis of *Fluc* gene correction. Luciferase activity in the lysates was measured as described above. Data are presented as fold over donor alone for each experiment and as an absolute percent recombination. The recombination frequency was calculated based on normalization to the standard curves. For the DEED and DMAP TFOs, cells were electroporated with 2 μM TFO (Bio-Rad gene pulser; settings: 25 μF, 250W, 280 V; in 0.4 cm electrode gap cuvette) with oligonucleotide on the first day and a fixed amount of donor was introduced 24 hours later with GenePORTER2. As a positive control amidate18 was also electroporated at a concentration of 0.5 μM. Cells were lysed and measured for induced recombination 48 hours after donor transfection as above.

Results

Design of reporter cell lines containing a single-copy luciferase gene target

In order to test TFO-induced gene correction in an assay representative of a potential gene therapy target, we designed a series of CHO-derived cell lines each with a single chromosomal copy of a mutated version of the *Fluc* reporter gene. In each line, the *Fluc* locus was engineered to also contain a polypurine/polypyrimidine site amenable to triplex formation situated upstream of the coding region and within 71 bp of the mutation to be corrected. These cell lines were also engineered such that the *Fluc* constructs were inserted at the same chromosomal position, an outcome that was facilitated by use of a parental CHO cell line with a pre-existing Flp recombinase target (FRT) integrated at a single chromosomal site.

For one pair of cell lines, an expression cassette containing the *luciferase* gene downstream of the AG30 TFO binding site (at position -53 bp relative to the start ATG of the *Fluc* coding region) was sub-cloned into an FRT-containing plasmid vector. A matching plasmid with a nonsense mutation at bp 13 relative to the ATG start site (site-to-mutation distance of 66 bp) was derived by site-directed mutagenesis. These plasmids were separately used to transfect the FRT-containing CHO-Flp cells in conjunction with transient expression of the Flp recombinase. The desired cell lines were selected and screened as described in the methods section for site-specific integration of the target constructs. The resulting cell lines, CHO-RWT and CHO-R13, each were determined to

contain a single-copy of the luciferase gene at the standardized integration site, with either the wild-type *Fluc* or the *Fluc13* nonsense mutant, respectively.

Similar steps were performed to generate the CHO-YWT and CHO-Y13 cell lines containing either the wild-type *Fluc* or the *Fluc13* mutant downstream of an A:T bp-rich, pyrimidine-motif TBS for the TFO, amidate18. In addition, a third mutant cell line, CHO-YScr, otherwise identical to CHO-Y13 but with a scrambled version of the TC18 site, was established to serve as a target site control. This cell line contains an A:T bp-rich target which has 9 mismatches relative to the canonical TC18 target sequence. The site-to-mutation distance in both CHO-Y13 and CHO-YScr is 71 bp. The cell line CHO-YScr202 has the same TBS as CHO-YScr but the site to mutation distance is 260 bp.

As with the CHO-RWT and CHO-R13 cells, the CHO-Y cell set was confirmed to have the *Fluc*-containing constructs integrated into the CHO cell genome at the pre-existing FRT site, as determined by Southern blot analysis (not shown), thereby facilitating comparisons among the cell lines. Background luciferase enzyme activity was found to be negligible in lysates from all three of the cell lines containing the *Fluc13* genes with the bp 13 stop codon mutation and from the one cell line containing the *Fluc202* mutation.

Design of anti-parallel and parallel motif TFOs

AG30, a G-rich 30-mer TFO, has been previously shown to bind as a third strand with high affinity in the anti-parallel triplex motif at the G:C bp-rich duplex target present in the promoter regions of CHO-RWT and CHO-R13 (Fig. 1A). The binding

characteristics of AG30 for this duplex target have been reported [33]. 12AG30 is a G-rich 30-mer oligonucleotide containing the same base composition as AG30 but having 12 mismatches in the third-strand binding code. AG30 and 12AG30 consist of natural nucleobases and natural phosphodiester backbones except for end-capping with a propyl amine at the 3'-end to inhibit nuclease degradation. Amidate18 is a N3'->P5' phosphoramidate modified, T-rich TFO that binds in the parallel triplex motif to the polypurine strand of the duplex target present in the promoter region of CHO-YWT and CHO-Y13 (Fig. 1A). It consists of natural nucleobases but contains a fully substituted phosphoramidate backbone in which the 3' bridging oxygen is replaced with nitrogen. This modification has been shown to enhance third strand binding in the pyrimidine motif under physiologic conditions [31]. The binding characteristics of amidate18 for this target site have been described previously [31].

The luciferase reporter gene assay is highly reproducible and sensitive

For each experiment, a standard curve was established to correlate luciferase activity with the proportion of *Fluc* wild-type cells among *Fluc* mutant cells in mixed cell populations. Cell lines containing the wild-type copy of *Fluc* and either the AG30 or TC18 TFO binding site were mixed in defined ratios from 1:10,000 (0.01%) to 1:3 (33%) with matched cells containing the same TFO binding site but with a mutated version of the *Fluc* gene. Samples consisting of all *Fluc* positive (100%) and all *Fluc* mutant cells (0%) were also included. At each increment of wild-type cell percentage, cell lysates were tested for luciferase activity, and RLUs/protein values were determined.

As shown, there was a highly consistent, linear relationship between RLUs/protein and percentage of wild-type cells (Fig. 1C). In addition, this linear relationship was found to be essentially identical for both pairs of cell lines; this result is consistent with the fact that all cell lines were constructed to contain a single-copy of the same basic *luciferase* expression cassette in the same genomic locus. The linearity of the relationship between *luciferase* activity and the proportion of *luciferase* wild-type cells in mixed cell populations and the consistency between the two sets of cell lines provided the basis for us to compare data between experiments and to assign recombination frequencies based on *luciferase* enzyme activity in cell lysates of treated cells. In addition, the assay was found to be highly sensitive, being capable of detecting proportional *Fluc* activity (and therefore recombination events) in just 1 in 10,000 cells.

Parallel and anti-parallel motif TFOs induce sequence-specific recombination between a genomic target and a single-stranded donor DNA

TFO-induced recombination was tested in both the CHO-R13 and CHO-Y13 cell lines. The cells were transfected with either the single-stranded donor DNAs alone or donor DNAs plus specific, mismatched, or unrelated control oligonucleotides. Two days after transfection, correction of the *luciferase* mutation was determined by assaying *luciferase* activity in lysates of the treated cells. As shown (Fig. 2A), AG30 plus donor induced recombination 5.7-fold over donor alone (0.05% recombination) in the CHO-R13 cell line containing the matched G:C bp-rich AG30 binding site upstream of the *luciferase* coding region. However, AG30 had no effect above that of the donor alone in

cells containing the A:T bp-rich binding sites (CHO-Y13 or CHO-YScr) at the same relative position in the *luciferase* cassette. Specificity was also demonstrated by the lack of effect of 12AG30, an oligonucleotide containing a scrambled version of the AG30 sequence (creating 12 mismatches to the target sequence in CHO-R13 cells), on donor DNA recombination in the CHO-R13 cells (Fig. 2A).

Similarly, amidate18 induced recombination at a frequency 6.1-fold over donor alone (0.07% recombination) in the CHO-Y13 cells containing the matched A:T bp-rich binding site for the amidate18 TFO sequence (Fig. 2B) but not in the CHO-YScr cells (containing an A:T bp-rich binding site with 9 mismatches versus amidate18). As further evidence of specificity, amidate18 had no effect on recombination in the CHO-R13 cells with the G:C-rich binding site (Fig. 2B). Additionally, we also tested LNA18, an oligonucleotide with the same nucleotide sequence as amidate18 but containing modified 2'-O, 4'-C-methylene bridge ribose moieties. This TFO was previously shown to mediate triplex formation *in vitro* but to lack intracellular targeting activity [31]. LNA18 did not induce recombination above that of the donor alone in the CHO-Y13 cell line (Fig. 2B). These results show that, in two different CHO cell lines, TFOs can induce intermolecular recombination at a single-copy genomic locus in a site-specific manner. In addition, the induced recombination was mediated by both purine and pyrimidine TFOs designed to form anti-parallel and parallel triplexes, respectively. The specificity of the reaction was demonstrated not only by the lack of stimulation by control oligonucleotides but also by the inability of the AG30 and amidate18 TFOs to induce recombination at *luciferase* gene

loci with mismatched target sites.

TFO induction of recombination is dose-dependent

Using a fixed amount of donor DNA per transfection sample of 0.2 μg/well (~25 nM), the amount of TFO in the transfection mix was incrementally increased from 0.1 μg/well to 0.6 μg/well. In both the CHO-R13 and CHO-Y13 *Fluc* mutant cell lines, there was a TFO dose-dependent increase in recombination. In the CHO-R13 cells, at the lowest AG30 dose (0.1 μg/well), the fold induction was 3.8, while at the highest dose (0.6 μg/well), the fold induction over donor alone was 8.5, corresponding to recombination frequencies of 0.03% and 0.07%, respectively (Fig 3A). In the CHO-Y13 cells, at the lowest amidate18 dose (0.1 μg/well), the fold induction of recombination was 4.7, while at the highest dose (0.6 μg/well), the fold induction was 9.4, correlating to 0.06% and 0.11% recombination, respectively (Fig. 3B).

To exclude the possibility that the dose-dependent effects observed could be attributable simply to differences in the total quantities of nucleic acids transfected rather than the TFO dose, TFO-induced recombination was tested in the setting of a constant amount of total nucleic acid transfected per sample (Figs. 3C and 3D). In this experiment, the sum total amounts of donor DNA and TFO in the transfection mix were held constant at 0.6 μg/well; the quantities of the TFOs and donor DNAs were then varied in inverse proportion to each other (TFO/donor pairs of 0.0/0.6 μg, 0.2/0.4 μg, 0.4/0.2 μg, and 0.6/0.0 μg). In the CHO-R13 cells, increased recombination was seen

with higher amounts of AG30 (Fig. 3C), even in samples with lower amounts of donor DNA (except for the sample with no donor DNA); for example, in the samples containing 0.2 μg TFO/0.4 μg donor, there was a 2.1-fold induction; in contrast, samples containing 0.4 μg TFO/0.2 μg donor yielded a 4.1-fold induction. Similarly, with the CHO-Y13 cells, induced recombination was 3.2-fold (over donor alone) when amidate18 dose was 0.2 μg/well and 4.4-fold (over donor alone) when amidate18 dose was 0.4 μg/well (Fig. 3D). These results further demonstrate the TFO-dependent stimulation of recombination at the *Fluc* chromosomal loci in the two cells lines.

Donor DNA orientation and length

Previous studies focusing on recombination mediated by ssDNAs alone (without the use of TFOs) have observed different levels of activity for DNAs homologous to either the transcribed or the non-transcribed strand of the target. However, from target to target, no consistent preference has emerged [34, 35]. We tested whether there was any preference for donor DNA orientation in our assay in the context of TFO-induced recombination. The CHO-R13 cells were transfected with 0.2 μg/well of either sense (matching the non-transcribed strand) or antisense (homologous to the transcribed strand) single-stranded 51-mer donor DNAs, with or without 0.2 μg/well of AG30 (Fig. 4A). AG30 was found to boost recombination over both the sense and anti-sense donors alone, but the results with the anti-sense donor were superior. Similar results were seen with the CHO-Y13 cells and the amidate18 TFO in the comparison of sense and antisense donors (Fig. 4B). Hence, both anti-parallel and parallel third-strand binding induce

recombination at the *Fluc13* target to a greater degree when co-transfected with antisense versus sense donor DNAs. Antisense donors were used in all the experiments shown in Figures 2 and 3.

To further examine donor DNA characteristics, we compared TFO-induced recombination in conjunction with antisense donors of varying length: as30 (30-mer), as51 (51-mer), and as69 (69-mer). When co-transfected with amidate18, as30, as51, and as69 mediated recombination at levels 4.8-, 6.1-, and 3.7-fold over antisense donor alone, respectively (Fig. 5A). Hence, the short 30-mer donor was found to be active, but length greater than 51 nt did not necessarily yield increased activity. Additionally, we examined the effect of end-modification of the antisense donors with locked nucleic acids in place of phosphorothioates at the first and last three positions of the ssDNAs. The LNA-modified donors were less effective than the phosphorothioate-modified donors (Figures 5B and 5C).

Chemically modified TFOs

Other chemically modified TFOs in addition to amidate18 were examined as well. These modifications included the modified base: 5-(1-propynyl)-2'-deoxyuridine (pdU), the modified sugars: 2'-O, 4'-C-methylene bridged or locked nucleic acid (LNA), and 2'-O-(2-aminoethyl)-ribose (2'-AE), and the modified backbones: N, N-dimethylamino-propylphosphoramidate (αDMAP and βDMAP), N, N-diethylethylenediamine (αDEED), and (R)-1-O-[4-(1-pyrenylethynyl)phenylmethyl]glycerol, twisted intercalating nucleic acid (TINA). Chemical modifications and substitutions on the TFOs for the LNA, pdU,

2'-AE, DEED, and DMAP are shown in Appendix 1. The 2'-AE, pdU, and TINA849 showed minimal activity in CHO-Y13 cells compared to amidate18 (Fig. 6A). Several TINA oligonucleotides were tested but the best induction of recombination was 2.9-fold over donor for TINA849. The βDMAP oligonucleotide showed low levels of induced recombination, 3.1-fold over donor alone (Fig. 6B). The αDMAP oligonucleotide showed slightly higher levels of induced recombination, 4.0-fold over donor alone (Fig. 6B). The αDEED oligonucleotide showed 5.7-fold induction of recombination over donor (Fig. 6B). Amidate18 showed a 7.7-fold induction of recombination when electroporated in parallel with the DMAP and DEED TFOs (Fig. 6B). Percentages of induction of recombination for the βDMAP, αDMAP, αDEED, and amidate18 were 0.037%, 0.049%, 0.068%, and 0.092%, respectively.

The TINA oligonucleotides were also tested for induced recombination in the episomal targeting assay in COS-7 cells as outlined in Chapter 2. TINA849 did not show much effect but TINA855 showed an induced recombination frequency of 0.2% (compared to 0.37% for amidate18). Finally, the TINA oligonucleotides were tested for gene repression in cells containing the wild-type chromosomal target (CHO-YWT cells). In the *Fluc* cell lines used here, the *luciferase* gene is driven by a CMV promoter with the TFO binding site just upstream of the transcription start site. Thereby assaying for the level of *luciferase* expression in the wild-type lines allows for detection of inhibition of transcription due to TFO binding. The TINA oligonucleotides showed no gene repression effect.

In addition to the chemically modified TFOs mentioned above, TFOs conjugated to camptothecin (CPT) were assessed for their ability to induce recombination or suppress transcription. Using the cell line CHO-YScr, which contains the target for the CPT TFOs and a mutation 13 bp from the *Fluc* start codon and the cell line CHO-YScr202 (containing a mutation 202 bp from the *Fluc* start codon), the CPT TFOs were analyzed. Recombination between an antisense donor and the chromosomal target was induced with the addition of the camptothecin oligonucleotide (TUCL4CPT) at both distances in a dose responsive manner (Figures 7A and Fig. 7B). At the 0.2 μg/well TFO dose, recombination was induced 2.4-fold over donor alone in CHO-YScr cells (Fig. 7A) and 1.3-fold in CHO-Scr202 cells (Fig. 7B), corresponding to 0.03 and 0.02% induced recombination, respectively. A greater oligonucleotide concentration was required for a similar level of recombination to that seen at the 13 bp distance to be seen at the 202 bp distance.

Additionally, these oligonucleotides were tested for repression of wild-type *luciferase* activity. Repression observed with the camptothecin oligonucleotide (TUCL4CPT) was similar to that seen with the TUCL4NH2 and TUCL4NPh2 controls (Fig. 7C). Minimal repression was seen with the topoisomerase II inhibitor conjugated oligonucleotides (TUC-C18-ICP165 and TUC20-C18-ICP165) or with the psoralen conjugated oligonucleotides (psoTUCL4PH2 and psoTUCL4CPT), with or without psoralen activation (data not shown).

Discussion

The work reported here demonstrates the ability of both purine and pyrimidine motif TFOs to induce intermolecular recombination by short single-stranded DNAs at single-copy genomic loci in mammalian cells, with induced recombination frequencies of up to 0.11%. The TFO-induced recombination was shown to be sequence-specific based on comparisons of both mismatched oligonucleotides and mismatched chromosomal target sites. A TFO dose-dependence was seen in two sets of experiments, including one in which the total amount of transfected nucleic acid was held constant. Comparisons of sense and antisense orientation donor DNAs indicated a preference for antisense donors regardless of the orientation of the triplex, and a comparison of donor DNA length showed that the 51-mer was more active than either the 30-mer or the 69-mer,

This work was facilitated by a sensitive *Fluc* gene assay capable of reporting recombination events at frequencies over a wide range, from below 0.01% to 100%. Based on a consistent linear relationship between luciferase activity in cell lysates and the percentage of wild-type cells in mixed cell populations (as established using two separate sets of matched cell lines), the assay allows for straightforward calculation of recombination frequencies from enzyme activity. The use of the CHO-Flp parental cell line for site-specific introduction of the *Fluc* gene cassettes also provided an additional ability to compare TFO-induced recombination in different sequence contexts (and with both purine and pyrimidine TFOs) but still at the same chromosomal locus in otherwise isogenic cells.

The similar results in the two cell lines with two different types of TFOs provide further evidence for the fundamental ability of triple helix formation to create DNA helical distortions that can induce recombinagenic repair events in cells. Of note, there were no differences between parallel and anti-parallel binding TFOs vis à vis their ability to stimulate recombination. This is in contrast to the distinct effects seen when we compare donors homologous to the sense versus antisense strands of the target duplex. Hence, it is likely the extent of helical distortion created by the triplex, not its polarity, that confers the ability to stimulate recombination.

Compared to the amidate-modified TFO, LNA, 2'AE, and pdU-modified TFOs did not induce significant recombination in the genomic target. The pdU result was surprising as it induced recombination in the episomal target. The bulky base may be less effective in the single-copy chromosomal target compared to the episomal target. Several new chemistries not explored in Chapter 2 were also tested in this genomic target. The DEED analogue used in Chapter 2 showed poor binding likely due to the chirality of the phosphorus atom of the phosphoramidate internucleoside linkages [36, 37]. The use of a non-natural α anomery instead was proposed to offset this problem. Unlike their β analogues, and despite the chirality of the phosphorus atom, these molecules hybridize tightly to their DNA targets in an anti-parallel fashion at neutral pH and without $MgCl_2$ [38]. Between the αDEED TFO, the βDMAP TFO, and the αDMAP TFO, the αDEED TFO showed the greatest induction of recombination but that effect was still significantly less than that of amidate18. This may have been due to decreased binding to this target

site or due to decreased delivery. Twisted Intercalating Nucleic Acids (TINA) linkages as part of the backbone structure of the TFO [39] were also tested and none of the TINA oligonucleotides induced recombination significantly or repressed wild-type expression. The cellular effect of the TINA TFOs has not yet been determined. The TINA TFOs may influence intracellular binding or influence repair/recombination/damage recognition pathways. Binding of the DMAP and DEED TFOs and the TINA TFOs to the duplex used in Chapter 2 was not tested.

The CPT TFOs used in this work showed inhibition of an episomal target in HeLa cells [40, 41]. CPT TFOs can recruit topoisomerase I in a sequence-specific manner and induce stable cleavage complexes in cell nuclei [42]. CPT TFOs can inhibit transcription of an episomal target only when a cleavage site is present near the TFO end [42]. Topoisomerase II inhibitor conjugated TFOs form stable triplexes *in vitro* [43], but no *in vivo* activity of these complexes has been detected. The CPT-modified TFO induced recombination at low levels in cell lines containing the *Fluc* target mutated at two different distances from the TBS. No induction of recombination was seen with the other CPT TFOs. Repression of wild-type *Fluc* observed with the camptothecin oligonucleotide was similar to that seen with the controls.

It was thought that the intercalation of the camptothecin conjugation would create a structure that slows turnover or inhibits repair or recombination. Thus, CPT TFOs might create a camptothecin conjugate that blocks topoisomerase I which leads to a persistent strand break and therefore a higher likelihood of recombination. In fact, we

found the opposite; there was a substantial reduction of recombination, suggesting that a structure was created that the cell could not resolve. Perhaps merging the amidate modification with either the TINA or the CPT modifications would augment the effectiveness of induction of recombination in this or other genomic targets.

The ability of triplex formation to induce repair has been shown to depend on the nucleotide excision repair (NER) pathway. In a study using human cell-free extracts, TFO-induced intermolecular recombination was blocked by depletion of the NER damage recognition factor, xeroderma pigmentosum group A (XPA) protein, and was restored by supplementation of the depleted extracts with purified XPA protein [44]. Similarly, TFO-induced recombination in an episomal target was absent in human mutant cells deficient in XPA but was restored by complementation of the mutant cells by expression of XPA cDNA [45]. Hence, TFOs exert an effect on recombination by engaging the cellular repair machinery to generate recombinagenic intermediates. Initial studies demonstrated that DNA strand breaks are produced in the repair of triplex structures, but the exact pattern of repair-related incisions has not yet been mapped [5]. In a sense, the action of TFOs through cellular DNA repair pathways is conceptually similar to the action of exogenous small interfering RNAs through the cellular RNA processing pathways.

Recent reports of the ability of ZFNs to induce recombination further demonstrate the concept that the generation of site-specific DNA damage is recombinagenic. ZFNs and TFOs represent related approaches, in that both are major groove binding molecules

that can create site-specific DNA damage to stimulate recombination. ZFNs do this via the cleavage activity of the fused endonuclease domain. TFOs do this by forming triple helices, which are recognized by repair factors, leading to repair-associated strand breaks and thereby activating the target site for recombination. However, TFOs are chemically synthesized and obviate the need for expression constructs. The general toxicity and safety profile of oligonucleotides has been extensively studied, and specific oligonucleotides have already been approved for therapeutic use in humans [46]. In contrast, there are data suggesting that ZFNs can be toxic to cells when highly expressed [14], presumably due to off-target nuclease activity that is not completely regulated by the zinc finger domain.

Although TFOs were shown here to induce intermolecular recombination at chromosomal sites by short single-stranded DNAs, they may also be effective in combination with other DNA molecules, including long single- or double-stranded DNAs and even viral vector genomes, such as AAV. In the case of AAV vectors, it has already been demonstrated that I-*SceI*-induced DSBs can stimulate increases in AAV-mediated recombination rates [47]. Hence, TFO technology may prove to be complementary to a number of other approaches that are being developed to mediate gene correction in mammalian cells.

Although TFOs have been shown to induce site-directed sequence changes at chromosomal loci in cells in culture and even in mice following systemic administration, improvements are still needed in the efficiency of gene targeting by TFOs. The

frequencies of up to 0.1% obtained in this work represent up to 9-fold better than what can be achieved by transfection of the donor DNA fragments alone under the same conditions. However, efforts to improve intracellular binding properties of TFOs and to promote intra-nuclear TFO delivery will be needed to achieve the desired gains in efficiency. Nonetheless, the work reported here constitutes an important next step and establishes the ability of TFOs to stimulate the site-directed recombination at chromosomal loci in mammalian cells.

The majority of this work is currently accepted for publication in the journal Molecular Therapy with M.P. Knauert and J.M. Kalish as co-first authors. The luciferase chromosomal target was designed by and the CHO-R13 and CHO-RWT cell lines were made by Melissa Knauert. The AG30 data was collected with the help of Denise Hegan. All other cell lines were made by and data collection was done by J.M. Kalish.

References

1. Capecchi, M.R., *Altering the genome by homologous recombination.* Science, 1989. **244**(4910): p. 1288-92.
2. Doetschman, T., N. Maeda, and O. Smithies, *Targeted mutation of the Hprt gene in mouse embryonic stem cells.* Proceedings of the National Academy of Sciences of the United States of America, 1988. **85**(22): p. 8583-7.
3. Hendrie, P.C., R.K. Hirata, and D.W. Russell, *Chromosomal integration and homologous gene targeting by replication-incompetent vectors based on the autonomous parvovirus minute virus of mice.* J Virol, 2003. **77**(24): p. 13136-45.
4. Chamberlain, J.R., et al., *Gene targeting in stem cells from individuals with osteogenesis imperfecta.* Science, 2004. **303**(5661): p. 1198-201.
5. Vasquez, K.M., et al., *Manipulating the mammalian genome by homologous recombination.* Proc Natl Acad Sci U S A, 2001. **98**(15): p. 8403-10.
6. Jasin, M., *Genetic manipulation of genomes with rare-cutting endonucleases.* Trends Genet, 1996. **12**(6): p. 224-8.

7. Waldman, A.S., *Targeted homologous recombination in mammalian cells.* Critical Reviews in Oncology-Hematology, 1992. **12**(1): p. 49-64.

8. Taghian, D.G. and J.A. Nickoloff, *Chromosomal double-strand breaks induce gene conversion at high frequency in mammalian cells.* Mol. Cell. Biol., 1997. **17**: p. 6386-6393.

9. Donoho, G., M. Jasin, and P. Berg, *Analysis of gene targeting and intrachromosomal homologous recombination stimulated by genomic double-strand breaks in mouse embryonic stem cells.* Mol Cell Biol, 1998. **18**(7): p. 4070-8.

10. Lukacsovich, T., D. Yang, and A.S. Waldman, *Repair of a specific double-strand break generated within a mammalian chromosome by yeast endonuclease I-SceI.* Nucleic Acids Res, 1994. **22**(25): p. 5649-57.

11. Datta, H.J., et al., *Triplex-induced recombination in human cell-free extracts: Dependence on XPA and HsRad51.* J Biol Chem, 2001. **27**: p. 18018-18023.

12. Porteus, M.H. and D. Baltimore, *Chimeric nucleases stimulate gene targeting in human cells.* Science, 2003. **300**(5620): p. 763.

13. Porteus, M.H., et al., *Efficient gene targeting mediated by adeno-associated virus and DNA double-strand breaks.* Mol Cell Biol, 2003. **23**(10): p. 3558-65.

14. Bibikova, M., et al., *Enhancing gene targeting with designed zinc finger nucleases.* Science, 2003. **300**(5620): p. 764.

15. Urnov, F.D., et al., *Highly efficient endogenous human gene correction using designed zinc-finger nucleases.* Nature, 2005. **435**(7042): p. 646-51.

16. Broitman, S.L., D.D. Im, and J.R. Fresco, *Formation of the triple-stranded polynucleotide helix, poly(A.A.U).* Proceedings of the National Academy of Sciences of the United States of America, 1987. **84**(15): p. 5120-4.

17. Le Doan, T., et al., *Sequence-specific recognition, photocrosslinking and cleavage of the DNA double helix by an oligo-[alpha]-thymidylate covalently linked to an azidoproflavine derivative.* Nucleic Acids Research, 1987. **15**(19): p. 7749-60.

18. Letai, A.G., et al., *Specificity in formation of triple-stranded nucleic acid helical complexes: studies with agarose-linked polyribonucleotide affinity columns.* Biochemistry, 1988. **27**(26): p. 9108-12.

19. Moser, H.E. and P.B. Dervan, *Sequence-specific cleavage of double helical DNA by triple helix formation.* Science, 1987. **238**(4827): p. 645-50.

20. Kohwi, Y. and T. Kohwi-Shigematsu, *Magnesium ion-dependent triple-helix structure formed by homopurine-homopyrimidine sequences in supercoiled plasmid DNA.* Proc Natl Acad Sci U S A, 1988. **85**(11): p. 3781-5.

21. Beal, P.A. and P.B. Dervan, *Second structural motif for recognition of DNA by oligonucleotide-directed triple-helix formation.* Science, 1991. **251**(4999): p. 1360-3.

22. Kohwi, Y., S.R. Malkhosyan, and T. Kohwi-Shigematsu, *Intramolecular dG.dG.dC triplex detected in Escherichia coli cells.* Journal of Molecular Biology, 1992. **223**(4): p. 817-22.

23. Pilch, D.S., C. Levenson, and R.H. Shafer, *Structure, stability, and thermodynamics of a short intermolecular purine-purine-pyrimidine triple helix.* Biochemistry, 1991. **30**(25): p. 6081-8.

24. Postel, E.H., et al., *Evidence that a triplex-forming oligodeoxyribonucleotide binds to the c-myc promoter in HeLa cells, thereby reducing c-myc mRNA levels.* Proceedings of the National Academy of Sciences of the United States of America, 1991. **88**(18): p. 8227-31.

25. Faruqi, A.F., et al., *Recombination induced by triple helix-targeted DNA damage in mammalian cells.* Mol. Cell. Biol., 1996. **16**: p. 6820-6828.

26. Wang, G., M.M. Seidman, and P.M. Glazer, *Mutagenesis in mammalian cells induced by triple helix formation and transcription-coupled repair.* Science, 1996. **271**(5250): p. 802-5.

27. Faruqi, A.F., et al., *Triple-helix formation induces recombination in mammalian cells via a nucleotide excision repair-dependent pathway.* Mol Cell Biol, 2000. **20**(3): p. 990-1000.

28. Knauert, M.P., et al., *Distance and affinity dependence of triplex-induced recombination.* Biochemistry, 2005. **44**(10): p. 3856-64.

29. Faruqi, A.F., et al., *Triple-helix formation induces recombination in mammalian cells via a nucleotide excision repair-dependent pathway.* Mol Cell Biol, 2000. **20**(3): p. 990-1000.

30. Luo, Z., et al., *High-frequency intrachromosomal gene conversion induced by triplex-forming oligonucleotides microinjected into mouse cells.* Proc Natl Acad Sci U S A, 2000. **97**(16): p. 9003-8.

31. Kalish, J.M., et al., *Triplex-induced recombination and repair in the pyrimidine motif.* Nucleic Acids Res, 2005. **33**(11): p. 3492-502.

32. Southern, E.M., *Detection of specific sequences among DNA fragments separated by gel electrophoresis.* J Mol Biol, 1975. **98**(3): p. 503-17.

33. Wang, G. and P.M. Glazer, *Altered repair of targeted psoralen photoadducts in the context of an oligonucleotide-mediated triple helix.* J Biol Chem, 1995. **270**(38): p. 22595-601.

34. Sorensen, C.B., et al., *Site-specific strand bias in gene correction using single-stranded oligonucleotides.* J Mol Med, 2005. **83**(1): p. 39-49.

35. Olsen, P.A., M. Randol, and S. Krauss, *Implications of cell cycle progression on functional sequence correction by short single-stranded DNA oligonucleotides.* Gene Ther, 2005.

36. Chaturvedi, S., T. Horn, and R.L. Letsinger, *Stabilization of triple-stranded oligonucleotide complexes: use of probes containing alternating phosphodiester*

and stereo-uniform cationic phosphoramidate linkages. Nucleic Acids Res, 1996. **24**(12): p. 2318-23.

37. Dagle, J.M. and D.L. Weeks, *Positively charged oligonucleotides overcome potassium-mediated inhibition of triplex DNA formation.* Nucleic Acids Res, 1996. **24**(11): p. 2143-9.

38. Michel, T., et al., *Highly stable DNA triplexes formed with cationic phosphoramidate pyrimidine alpha-oligonucleotides.* Chembiochem, 2005. **6**(7): p. 1254-62.

39. Filichev, V.V. and E.B. Pedersen, *Stable and selective formation of hoogsteen-type triplexes and duplexes using twisted intercalating nucleic acids (TINA) prepared via postsynthetic Sonogashira solid-phase coupling reactions.* J Am Chem Soc, 2005. **127**(42): p. 14849-58.

40. Arimondo, P.B., et al., *Targeting topoisomerase I cleavage to specific sequences of DNA by triple helix-forming oligonucleotide conjugates. A comparison between a rebeccamycin derivative and camptothecin.* C R Acad Sci III, 1999. **322**(9): p. 785-90.

41. Matteucci, M., et al., *Sequence-Sepcific Targeting of Duplex DNA Using a Camptothecin-Triple Helix Forming Oligonucleotide Conjugate and Topoisomerase I.* J. Am. Chem. Soc., 1997. **119**: p. 6939-6940.

42. Arimondo, P.B., et al., *Exploring the cellular activity of camptothecin-triple-helix-forming oligonucleotide conjugates.* Mol Cell Biol, 2006. **26**(1): p. 324-33.

43. Duca, M., et al., *Triple helix-forming oligonucleotides conjugated to new inhibitors of topoisomerase II: synthesis and binding properties.* Bioconjug Chem, 2005. **16**(4): p. 873-84.

44. Datta, H.J., et al., *Triplex-induced recombination in human cell-free extracts. dependence on xpa and hsrad51.* J Biol Chem, 2001. **276**(21): p. 18018-23.

45. Chan, P.P., et al., *Targeted correction of an episomal gene in mammalian cells by a short DNA fragment tethered to a triplex-forming oligonucleotide.* J Biol Chem, 1999. **274**(17): p. 11541-8.

46. Kahan, B.D., et al., *Phase I and phase II safety and efficacy trial of intercellular adhesion molecule-1 antisense oligodeoxynucleotide (ISIS 2302) for the prevention of acute allograft rejection.* Transplantation, 2004. **78**(6): p. 858-63.

47. Miller, D.G., L.M. Petek, and D.W. Russell, *Human gene targeting by adeno-associated virus vectors is enhanced by DNA double-strand breaks.* Mol Cell Biol, 2003. **23**(10): p. 3550-7.

Figures

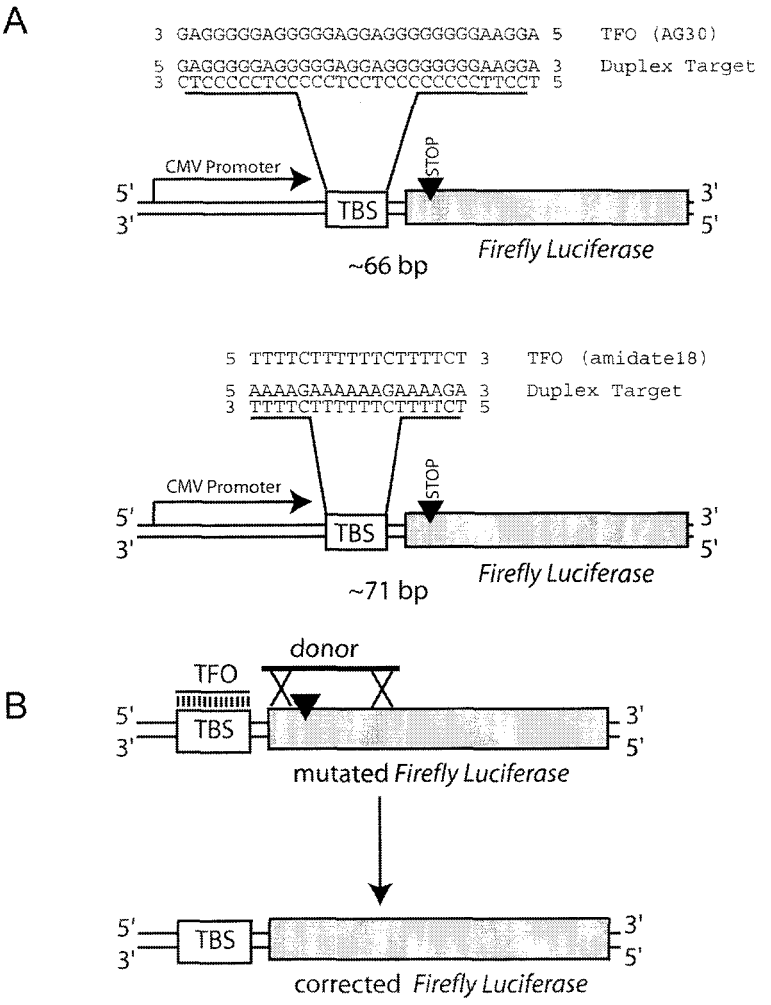


Figure 1A and B. Schematics of single-copy chromosomal targets. (A) Schematic of the luciferase reporter gene cassettes in the CHO-R13 and CHO-Y13 CHO cell lines used in this work. The sequences and positions of the TFO binding sites (TBS) relative to the stop codon mutations in the *Fluc* coding regions are shown with the distance from the start of the TBS to the mutation-to-be-corrected indicated below each construct. The mutant *Fluc* genes both contain a stop codon mutation 13 bp downstream from the start of the coding region, as indicated by the inverted triangle (▼). The TFOs, AG30 and amidate18, bind to their respective duplex target sites in the orientations shown; duplex TBS orientation is as indicated by the 5' and 3' labels. (B) Schematic of TFO-induced recombination and gene correction, with TFO binding to the TBS and alignment of single-stranded donor DNAs with the homologous target sequences, as indicated.

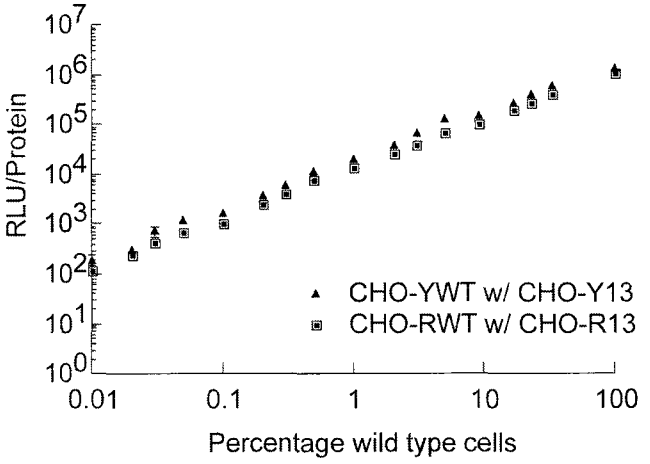


Figure 1C. Development of standard mixing curves. Graph of the relationship between the percent *Fluc* wild-type cells among *Fluc13* mutant cells in mixed cell populations and the *luciferase* activity in the lysates of the mixed populations as measured by RLUs per μg protein. Mixed cell populations were established with the indicated proportions of *Fluc* wild-type cells among *Fluc13* mutant cells to establish a standard curve. CHO-YWT w/ CHO-Y13 indicates the curve determined by mixing *Fluc* wild-type cells containing the *Fluc* target with the A-rich (TC18) TFO binding site and the corresponding *Fluc13* mutant cells. CHO-RWT w/ CHO-R13 refers to the curve determined by a similar analysis of mixed populations of *Fluc* wild-type cells containing the G-rich (AG30) TFO target site and the corresponding *Fluc13* mutant cells. Six replicates were performed per data point, and error bars indicate one standard error above and below the experimental mean.

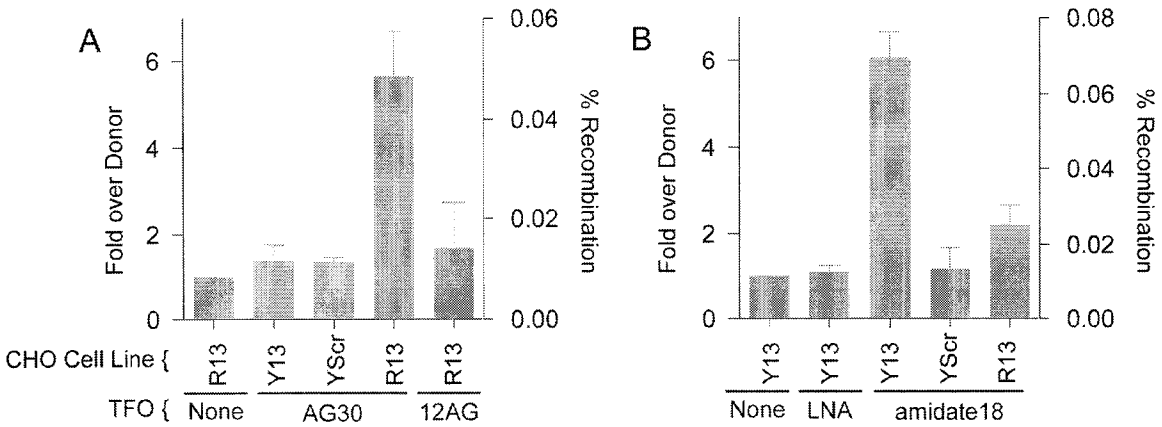


Figure 2. TFO-induced recombination at a single-copy chromosomal target. (A) Recombination in CHO-R13 cells induced by the matched sequence-specific TFO, AG30. Cells with the indicated TFO binding sites (CHO-R13: G-rich purine motif TBS matched to AG30; CHO-Y13: A-rich pyrimidine motif TBS matched to amidate18; CHO-YScr: A-rich pyrimidine motif TBS mismatched relative to amidate18) were transfected with either donor alone (no TFO), donor plus AG30 (specific for the G-rich target site) or donor plus 12AG30 (scrambled sequence G-rich oligomer mismatched at 12 bp versus the target in the CHO-R13 cells); cell lines are indicated by their suffix (R13, Y13 and YScr for CHO-R13, CHO-Y13, and CHO-YScr cells, respectively). Donor alone sample indicates CHO-R13 cells transfected with 0.2 μg/well of as51; all other samples received 0.2 μg as51 and 0.2 μg of the TFO or control oligonucleotide, as indicated. Gene correction to produce functional *luciferase* genes was detected by analysis of *luciferase* enzyme activity in the lysates of the treated cells. Results are presented as fold induction over donor alone (ssDNA 51-mer, as51) on the right axis and calculated percent recombination on the left axis. 12AG indicates 12AG30 as described in materials and methods; all other TFOs are indicated by their full designation as described in materials and methods. The frequencies of induced recombination were determined based on RLU/protein measurements and reference to the standard curves in Fig. 1C. At least six replicates were performed for each sample, and error bars indicate one standard error above the experimental mean. (B) Recombination in CHO-Y13 cells induced by the specific TFO, amidate18. Conditions and cell lines are as in (A); LNA indicates LNA18 as described in materials and methods.

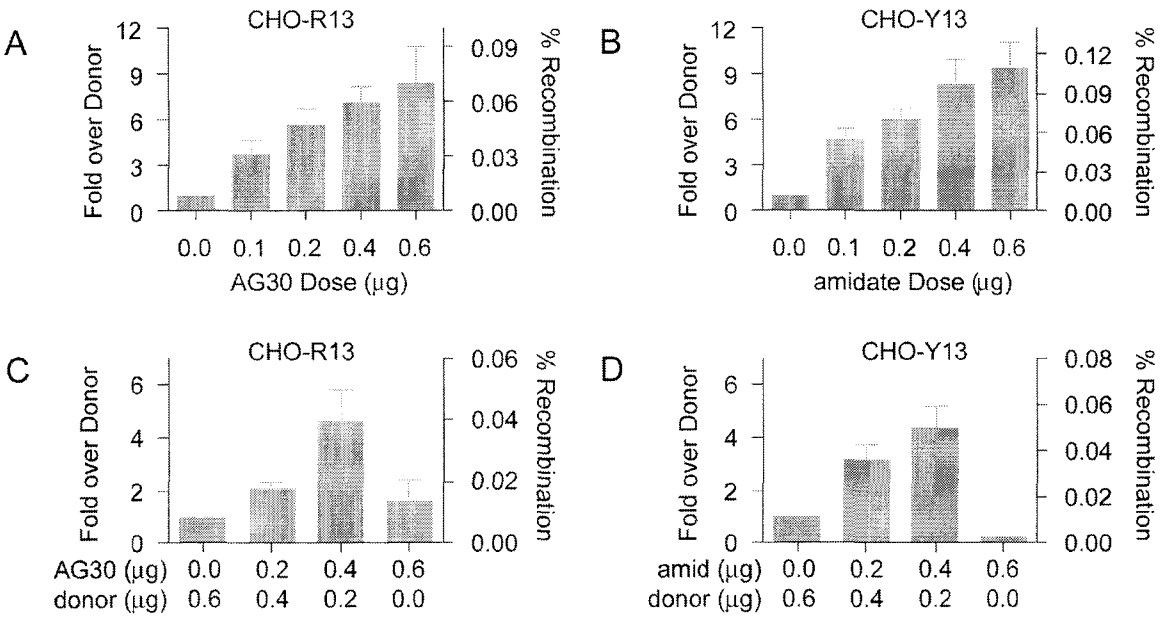


Figure 3. Dose dependence of TFO-induced recombination. (A) Dose dependence of induced recombination by the G-rich TFO, AG30. CHO-R13 cells were treated with 0.2 μg/well of as51 donor and increasing amounts of AG30, as indicated. Fold induction of recombination over donor alone samples is noted on the left axis and recombination frequency is presented on the right axis. The frequencies of induced recombination were determined based on RLU/protein measurements and reference to the standard curves in Fig. 1C; at least six replicates were performed for each sample, and error bars indicate one standard error above the experimental mean. (B) Dose-dependence of induced recombination by the T-rich TFO, amidate18. CHO-Y13 cells were treated with 0.2 μg/well of as51 donor and increasing amounts of amidate18 (amidate), as indicated. Results are presented as in (A). (C) Triplex-induced recombination in CHO-R13 cells as a function of TFO dose in samples transfected with a fixed total amount of nucleic acids (sum of TFO and donor DNAs). In each sample, CHO-R13 cells were treated with 0.6 μg total nucleic acid/well. Concentrations of as51 and AG30 were adjusted inversely as indicated. Results are presented as in (A). (D) Triplex-induced recombination in CHO-Y13 cells as a function of amidate18 (amid) dose in samples transfected with a fixed total amount of nucleic acids (sum of TFO and donor DNAs), as above (A).

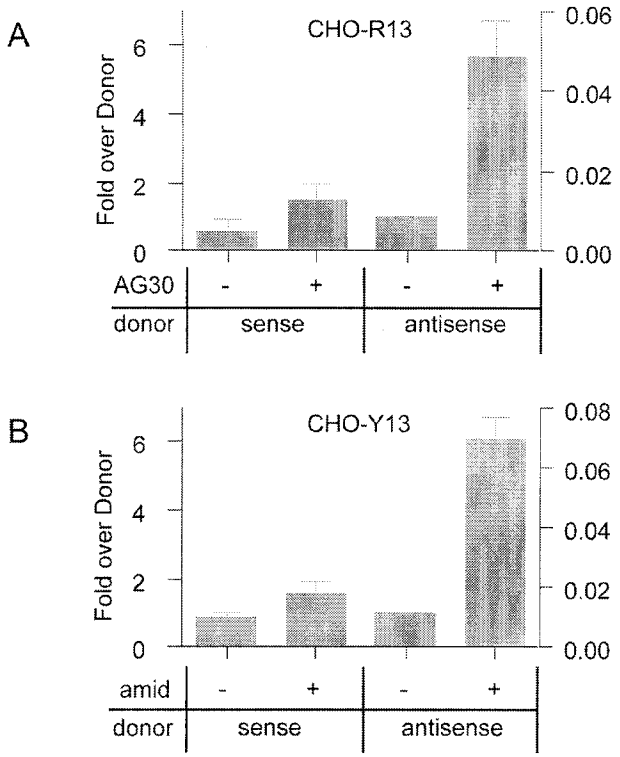


Figure 4. The effect of donor DNA orientation on recombination. (A) *Luciferase* gene correction by co-tranfection of either sense or antisense single-stranded donor DNAs and the TFO, AG30. CHO-R13 cell lines were treated with 0.2 μg of either sense or antisense 51-mer DNAs with (+) or without (-) 0.2 μg of AG30, as indicated. (B) *Luciferase* gene correction by co-tranfection of either sense or antisense single-stranded donor DNAs and the TFO, amidate18. CHO-Y13 cell lines were treated with 0.2 μg of either sense or antisense 51-mer donor DNAs with (+) or without (-) 0.2 μg of amidate18 (amid). All results (A and B) are presented as fold increase in recombination over as51 donor alone. The frequencies of induced recombination were determined based on RLU/protein measurements and reference to the standard curves in Fig. 1C; at least six replicates were performed for each sample, and error bars indicate one standard error above the experimental mean.

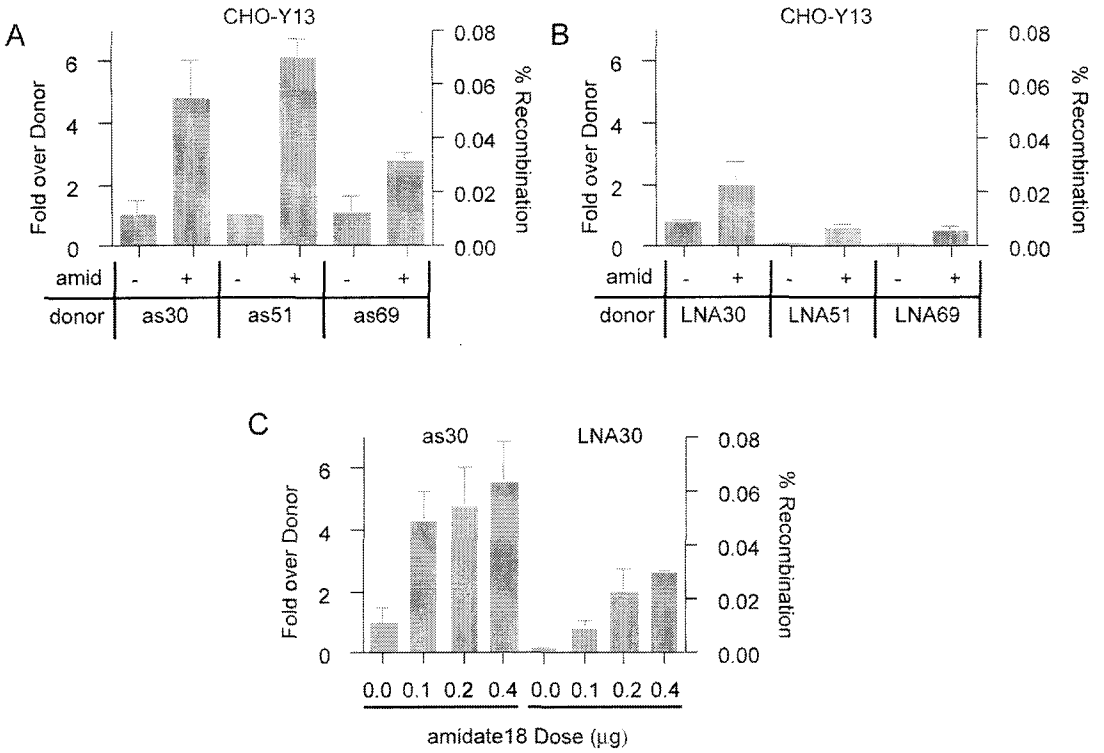


Figure 5. The effect of donor length and modification on recombination. (A) Influence of the length of the single-stranded donor DNAs. CHO-Y13 cells were treated with 25 nM of antisense, phosphorothioate end-modified donor DNAs with (+) or without (-) 0.2 µg of amidate18. Single-stranded donor DNAs tested were of length 30 nt, 51 nt and 69 nt. (B) Length comparison of LNA donors. CHO-Y13 cells were treated with 25 nM of antisense, LNA end-modified donor DNAs with (+) or without (-) 0.2 µg of amidate18. (C) LNA versus phosphorothioate donors. A fixed amount (25 nM) of donor (either as30 or LNA30 as labeled) was added to each well of CHO-Y13 cells. Doses of amidate18 as labeled were added. All results (A, B, and C) are presented as fold increase in recombination over as51 donor alone. The frequencies of induced recombination were determined based on RLU/protein measurements and reference to the standard curves in Fig. 1C; at least six replicates were performed for each sample, and error bars indicate one standard error above the experimental mean.

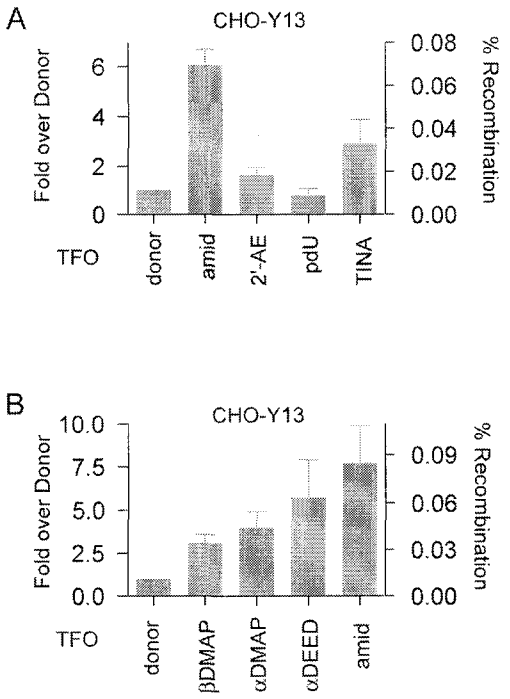


Figure 6. A comparison of chemically modified pyrimidine motif TFOs. (A) A number of chemically modified TFOs in the pyrimidine motif were compared. Amidate18 (amid), 2'-O-(2-aminoethyl)-ribose (2'-AE), 5-(1-propynyl)-2'-deoxyuridine (pdU), and (R)-1-O-[4-(1-pyrenylethynl)phenylmethyl]glycerol, Twisted Intercalating Nucleic Acid (TINA)-modified TFOs were compared in CHO-Y13 cells. TFOs are as described in materials and methods above or in Chapter 2. (B) A comparison of phosphoramidate modified-TFOs. N, N-dimethylaminopropyl-phosphoramidate (αDMAP and βDMAP) and N, N- diethylethylenediamine (αDEED) modified TFOs were compared with amidate18 for induction of recombination in CHO-Y13 cells, using electroporation for delivery of the TFO. as51 was used as the donor at 0.2 μg/well and TFOs were added at 2 μM, except for amidate18 which was added at 0.5 μM. TFOs are as described in materials and methods above. All results (A and B) are presented as fold increase in recombination over as51 donor alone. The frequencies of induced recombination were determined based on RLU/protein measurements and reference to the standard curves in Fig. 1C; at least six replicates were performed for each sample, and error bars indicate one standard error above the experimental mean.

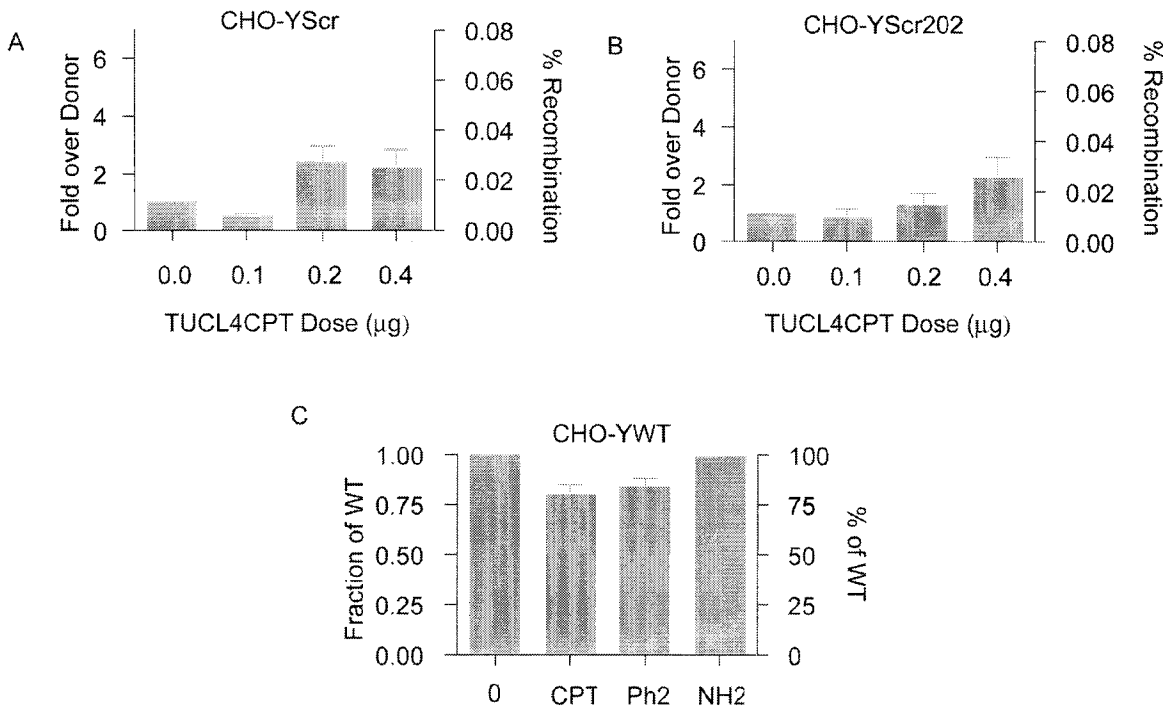


Figure 7. Camptothecin conjugated TFOs. (A) TUCL4CPT (CPT) induction of recombination in CHO-YScr cells. The dose effect of the CPT-conjugated TFO was demonstrated for cells with a mutation 13 bp from the ATG start site of *Fluc*. as51 at 0.2 µg/well was used as the donor. (B) TUCL4CPT induction of recombination in CHO-YScr202 cells. A 51-mer antisense ssDNA homologous to the wild-type *Fluc* sequence and flanking the 202 bp mutation was used at 0.2 µg/well. The dose effect of the CPT-conjugated TFO was demonstrated for cells with a mutation 202 bp from the ATG start site of *Fluc*. C. Repression of *Fluc* by CPT TFOs. The ability of CPT TFOs to repress wild-type Fluc was tested in CHO-YWT cells as compared to a phenyl-conjugated TFO, TUCL4Ph2 (Ph2) and an unconjugated TFO, TUCL4NH2 (NH2). TFOs were each added at 0.2 µg/well. Results (A and B) are presented as fold increase in recombination over as51 donor alone. The frequencies of induced recombination were determined based on RLU/protein measurements and reference to the standard curves in Fig. 1C; at least six replicates were performed for each sample, and error bars indicate one standard error above the experimental mean. Result C is presented as fraction and percent of wild-type *Fluc* activity as compared to untreated wild-type *Fluc* cells.

Chapter Four:
Repair of a Triplex Structure and Efforts Towards Understanding the Biochemical Mechanism

Introduction

Mammalian cells rely on the NER pathway to remove bulky DNA helix-distorting lesions (such as pyrimidine dimers) [1]. Unrepaired lesions lead to mutations or cell death; thereby NER serves (in conjunction with other repair pathways) to maintain genomic integrity. Deficiency in NER is responsible for the three rare autosomal recessive diseases, xeroderma pigmentosa (XP), Cockayne syndrome (CS), and the photosensitive form of trichothiodystropy (TTD) [2]. XP patients exhibit extreme ultraviolet radiation (UV) sensitivity and increased incidents of sunlight-induced cancers [2]. XP has been classified into 7 complementation groups, XPA through XPG [2]. Each group lacks or has a protein deficient in part of the NER complex.

NER involves damage recognition and incision, excision of the damage, repair resynthesis, and ligation [3] (Appendix 2). The NER reaction has been reconstituted *in vitro* [4-6] and the role of most of the factors involved is known, but the order of recognition and associated binding is still the subject of debate. Within the NER pathway there are two different sub-pathways based on the transcription status of the DNA containing the lesion, transcription coupled repair (TC-NER) and global genome repair (GGR) [3]. These pathways differ in their recognition but have a common incision process. Dual incisions are made 5-6 nucleotides 3' and 22-24 nucleotides 5' to thymine dimers in mammalian cell-free extracts with excision products 27-32 nucleotides long [7, 8]. The location of the incisions has been shown to vary with the type of lesion to be repaired [5]. XPG makes the 3' incision and XPF-ERCC1 makes the 5' incision [9]. In

repair of a cisplatin lesion, Evans et al. demonstrated that an open complex forms that is dependent on the core NER factors (all but XPF-ERCC1) and the incisions occur near the junctions of this open bubble with dsDNA [10]. Cross-link repair may require sequential excisions and repair on both strands and/or intra-strand recombination [11-13]. The location of the dual incisions in repair of a triplex is not known and would be essential in uncovering the mechanism of triplex repair.

Previous work in the Glazer lab demonstrated that triplex-induced mutagenesis and recombination is dependent on the nucleotide excision repair (NER) pathway [14-17]. Triplex-induced repair and recombination requires XPA, a NER recognition factor [16, 18, 19]. A 10 nucleotide TFO, conjugated to psoralen, and mono-adducted to duplex DNA produces excision products while a 30 nucleotide TFO does not produce excision products [20]. The gene modification activity of the 30-mer TFO, despite the lack of excision products suggests that multiple repair pathways may function in triplex repair. This idea is also consistent with the fact that the length of attachment of some TFOs (i.e. TFOs 30 nucleotides long) exceeds the canonical excision span of NER.

The mechanism of triplex repair is not currently understood and these experiments are designed to illuminate aspects of this mechanism. To understand the mechanism of triplex action and to elucidate the pathway by which triplexes promote repair and recombination, the incision step of the bulky adduct created by triplex binding needs to be mapped. The nature of these excision products and the location of the incisions have not been previously evaluated, and those are the goals of this work. In

order to ultimately develop an assay to locate the sites of incision made in repair of a TFO bound to duplex DNA, strides were taken to create substrates with altered helical structures to address this question.

In order to determine the biochemical nature of the excision of a triplex structure from DNA, a number of factors need to be considered. First, the repair pathways involved include at least NER, but given that NER alone cannot excise a 30-mer TFO bound to duplex DNA, other repair pathways are likely to be active as well. Second, in order to understand the biochemical mechanism of the incisions, the excision products themselves need to be visualized and deciphered. Finally, the complexity of these incisions and the alteration of incisions based on different altered helical structures by different TFOs need to be addressed. In the work presented in previous chapters, repair of pyrimidine motif triplexes was shown to be dependent on NER (as has been previously demonstrated for purine motif triplexes). Additionally, binding affinity alone is not sufficient for induction of repair and recombination. Chemical modifications that are more helically distorting are more inducing of repair and thereby more inducing of recombination. In attempting to capture excision products, a variety of techniques have been attempted as outlined here, and while some have been more productive than others, they demonstrate that there is at least one incision occurring in repair of 15-mer and 24-mer TFOs.

Materials and Methods

Oligonucleotides

pso-AG30, pso-5'AGGAAGGGGG GGGTGGTGGG GGAGGGGGAG3'; pso-New10, 5'pso-AGGAAGGGGG3'; pso-New14, 5'pso-AGGAAGGGGG AGGG3'; pso-New23, 5'pso-AGGAAGGGGG AGGGAGGGGA GGG3'; and pso-New29, 5'pso-AGGAAGGGGG AGGGAGGGGA GGGAGGGGG3' were synthesized by Oligos, Etc. psoNew11, psoNew15, psoNew24, and psoNew30 were created by adding [α-^{32}P]ddA (Amersham Biosciences) to the 3'-end of psoNew10, psoNew14, psoNew23, and psoNew29 as described below. IC1, 5'GGGACCTGAA CACGTACGAA ATTCGATATC CTCTAGCCAG ATCTGCGCCA GC3'; IC2, 5'TGAGGGGGAG GGGGTGGTGG GGGGGGAAGG ATC3'; IC3, 5'TTGGGCTGCA GCAGGTAGAC TCTAGAGGAT CACGAGCGAA CTCTAATGCA CC3'; IC4, 5'CAGATCTGGC TAGAGGATAT CGAATTTCGT ACGTGTTCAG GTC3'; IC5, 5'TGCAGCCCAA GATCCTTCCC CCCCCACCAC CCCCTCCCCC TCAGCTGGCG3'; IC6, 5'TTGGTGCATT AGAGTTCGCT CGTGATCCTC TAGAGTCTAC CTGC3'; New11, 5'AGGAAGGGGG A3'; New13, 5'AGGAAGGGGG AGG3'; New15, 5'AGGAAGGGGG AGGGA3'; New20, 5'AGGAAGGGGG AGGGAGGGGA3'; New24, 5'AGGAAGGGGG AGGGAGGGGA GGGA3'; New30, 5'AGGAAGGGGG AGGGAGGGGA GGGAG GGGGA3'; NEWPUR, 5'GGTACCGAAT TTCGGCCAGG GGGAGGGAGG GGAGGGAGGG GGAAGGATTC GAACCTT3'; and NEWPYR, 5'AAGGTTCGAA TCCTTCCCCC TCCCTCCCCT CCCTCCCCCT GGCCGAAATT

CGGTACC3' were synthesized by W.M. Keck Foundation Biotechnology Resource Laboratory (Yale University). New11, New15, New24, and New30 are of the same sequence as psoNew11, psoNew15, psoNew24, and psoNew30, described above without the psoralen on the 5'-end and the [α-^{32}P] ddA on the 3'-end.

CHO cell-free extracts

CHO cells were harvested from forty 15 cm dishes and centrifuged at 3000 rpm for 10 minutes at 4°C. Cells were washed twice with 10 times packed cell volume (PCV) of cold PBS, resuspended in 4 PCVs of 10 mM Tris-HCl pH 7.5, 1 mM EDTA, and 5 mM DTT, and incubated on ice for 20 minutes. Resuspended cells were homogenized and 4 PCVs of cold 50 mM Tris-HCl pH 7.5, 10 mM $MgCl_2$, 2 mM DTT, 25% sucrose, and 50% glycerol were added. 1 PCV of saturated ammonium sulfate (4.2 M, pH 7.2) was added drop wise and stirred for 30 minutes, following the last drop, at 4°C. The sample was then centrifuged at 45000 rpm in Ti70 rotor for 3 hours at 4°C. The supernatant was collected and the volume was measured, and 0.33 g of ammonium sulfate per mL of supernatant was added slowly. Then 100 µL 1 M NaOH/10 g ammonium sulfate was added and stirred for 30 minutes at 4°C. The sample was then centrifuged at 12000 rpm in a SS34 rotor for 30 minutes at 4°C. The pellet was resuspended in dialysis buffer (25 mM HEPES-KOH, pH 7.9, 100 mM KCl, 12 mM $MgCl_2$, 0.5 mM EDTA, 2 mM DTT, 15% glycerol) using 1/20 pellet volume of the supernatant and dialyzed at 4°C for 3 hours followed by another 3 hours with fresh dialysis buffer. The dialysate was centrifuged at 14000 rpm in Eppendorf tubes in a

microcentrifuge for 20 minutes at 4°C. This protocol was kindly given to us by Tadayoshi Bessho. Aliquots of 100 μL were made and frozen in liquid nitrogen. Protein concentrations were measured by Bradford assay [21].

Linear substrates

pNEWAG and pNEWTC were constructed from pIND/lacZ (Invitrogen) by cloning into the BamHI restriction site following the annealing of ssDNAs, 5'GCAGAAAGCG GATCCTTCCC CCTCCCTCCC CTCCCTCCCC CTAGAGACTT GTTCG3' and 5'CGAACAAGCT TCTAGGGGGA GGGAGGGGAG GGAGGGGGAA GGATCCGAAG CTTTCTGC3' that were synthesized by W.M. Keck Foundation Biotechnology Resource Laboratory (Yale University). The distinction between pNEWAG and pNEWTC is the orientation of the TBS. A 195 bp fragment was cut out of these plasmids with KpnI and XmnI, gel purified, and when not used in conjunction with end-labeled TFOs these substrates were 5'-end-labeled with [γ-^{32}P]ATP (Amersham Biosciences). Following phenol-chloroform extraction, TFOs were bound to this substrate and following mono-adduct formation, subjected to the incision reaction, as described below. Alternatively, one of two 57mers was end-labeled and annealed with its unlabeled complement.

End-labeled TFOs

TFOs (pso-New10, pso-New14, psoNew23, and pso-New29) were designed with psoralen on the 5'-end so that [α-^{32}P]ddA could be added to the 3'-end to complete the oligonucleotides (pso-New11, pso-New15, pso-New24, and pso-New30). [α-^{32}P]ddA

was added via terminal transferase (New England Biolabs) per manufacturer instructions. The synthesized sequences were as shown above and the final sequences were pso-New11, 5'pso-AGGAAGGGGG A3'; pso-Nw15, 5'pso-AGGAAGGGGG AGGGA3'; and pso-New24, 5'pso-AGGAAGGGGG AGGGAGGGGA GGGA3'; pso-New30, 5'pso-AGGAAGGGGG AGGGAGGGGA GGGAGGGGGA3'. *In vitro* binding and mono-adduct formation were done as described below. Samples were concentrated in microcon100s (Millipore), incision was as described below, then 10 μL of 2.5x formamide dye was added, samples were boiled for 5 minutes and loaded onto 15% polyacrylamide gels.

Internally labeled substrates

A double stranded 137 bp fragment was also built from 6 ssDNAs (IC1-IC6 as described above). One of the six ssDNAs was 5'-end-labeled with T4 kinase and [γ-^{32}P]ATP prior to construction and the remaining 5 ssDNAs were kinased together. The ssDNAs were mixed, phenol-chloroform extracted, ethanol precipitated, and resuspended in DNA dilution buffer (Roche, Rapid ligation kit). The sample was heated to 65°C for 10 minutes and then cooled at 23°C, for 20 minutes, then ligation buffer and ligase were added. Samples were incubated at 23°C for 2 hours and then for 18 hours at 4°C. Gel purification was as described below followed by *in vitro* binding as below.

Plasmid substrates

psupFG1, a plasmid containing the AG30 TBS sites was used as a plasmid substrate. *In vitro* binding was done as described below with psupFG1 and pso-AG30

followed by mono-adduct formation as described below. After the incision reaction described below, samples were resuspended in water and restriction enzyme cuts 200 bp apart were made and the fragments were separated on a 12% polyacrylamide gel. DNA was transferred to a membrane and probed with oligonucleotides labeled with [α-^{32}P]CTP (Amersham Biosciences). The probes were designed to contain all or part of the AG30 TBS. An internally labeled plasmid was created by cutting pMM2 [22] with BamHI (a cut at the end of the TBS), end-labeled with [γ-^{32}P]ATP, religated, and then used as the plasmid substrate for TFO binding and incision. *In vitro* binding, mono-adduct formation, and the incision reaction were as below. Following the incision reaction, stop reaction, phenol/chloroform extraction, and ethanol precipitation, a restriction digest was done as above.

In vitro binding

For linear substrates, 5 μg substrate and 10 μM TFO were used. For plasmids, *in vitro* binding was with 3 μg plasmid and 10 μM TFO. In all cases, binding was performed in 10 mM Tris (pH 7.6), 10 mM $MgCl_2$, 1 mM spermine, and 50% glycerol at 37°C for 16 hours.

Third-strand binding assays

Third-strand binding was measured using gel mobility shift assays under native conditions. Two complementary 30-mers (NEWPUR, NEWPYR) containing the NEW TBS target were synthesized. Duplex DNA was prepared by mixing 1000 pmol of each 57-mer together with 50 mM NaCl, incubating at 85°C for 20 minutes, and cooling to

room temperature before end-labeling with T4 polynucleotide kinase (New England BioLabs) and [γ-^{32}P]ATP. Duplex was gel purified, electroeluted, and filtered by centricon3 (Millipore). A fixed concentration of duplex at 5 x 10^{-8} M was added to binding reactions with increasing concentrations of oligonucleotides in 20 µl containing 10 mM Tris (pH 7.6), 10 mM $MgCl_2$, and 1 mM spermine. Samples were incubated at 37°C for 16 hours and were loaded onto 15% polyacrylamide gels [acrylamide/bisacrylamide (19:1)] containing 17.8 mM Tris and 17.8 mM boric acid (pH 7.6) and 10 mM $MgCl_2$ and the gels were run in the same concentration of Tris-boric acid and $MgCl_2$ at 60 V for 6 hours at 23°C.

Mono-adduct formation

To establish optimum mono-adduct formation and minimum cross-link formation, a range of times from 10 seconds to 6 minutes at 320 nm were evaluated. TFOs were bound to the 57-mer duplex, UVA treated for the indicated time, and visualized on a 15% polyacrylamide denaturing gel. Ten seconds at 320 nm was used for most of this work.

Removal of unbound and unadducted TFO

To dissociate unadducted TFO following mono-adduct formation, 10 mM Tris (pH 8), 1 mM EDTA, and 100 mM KCl were added to the binding reaction and incubated for 2 hours at 60°C. Samples were cooled to 23°C and concentrated in microcon100s, rinsed with water, and then incision reactions were performed as below.

Gel purification of substrates with mono-adducted TFOs

Prior to incision, TFO mono-adducted substrate was resolved on 10% denaturing gels and the DNA of the correct size was electroeluted from the gel in 1x TBE for two hours at 15 mA, spun in a centricon30 (Millipore), concentrated in a microcon3 (Millipore), washed with water, and subjected to the incision reaction.

Incision assay

The incision reaction consisted of, 10x repair buffer (0.3 M HEPES-KOH (pH 7.9), 0.4 mM KCl, 32 mM $MgCl_2$, 2 mM DTT, 1 mM EDTA (pH 8)), 100 mM ATP, 10 mg/mL BSA, 250 ng DNA substrate, and 100 μg of CHO cell-free extracts. Incision reactions were incubated for 1 hour at 32°C. Reactions were stopped by adding 1 µL of 10% SDS and 1 µL of 20mg/mL proteinase K (Sigma) and incubating at 37°C for 15 minutes, followed by phenol-chloroform extraction and ethanol precipitation. Glycogen (Sigma) was added to aid in precipitation of small DNA fragments. Samples were resuspended in 2.5x formamide dye solution, analyzed on a 12% polyacrylamide gel run at 65W for 2 hours. If samples were not previously labeled, DNA was transferred to membranes as described above. Otherwise gels were dried and analyzed on a phosphorimager (Storm 860, Molecular Dynamics, Amersham Pharmacia Biotech). This protocol was given to us by Tadayoshi Bessho.

Size standards

To monitor unbound TFO and for smaller size standards New11, New13, New15, New20, New24, New30 were 5'-end-labeled via T4 kinase with [γ-^{32}P]ATP.

Results

Altered helical substrates

The initial substrate to be used is a 195 bp duplex (either NEWAG or NEWTC) with a 3'-end-labeled TFO with a 5'-psoralen bound and mono-adducted just 5' of the TFO binding site (TBS) on the purine strand of the duplex (Fig. 1A). Also explored in this work are 195 bp duplexes (NEWAG and NEWTC) that are 5'-end-labeled followed by TFO binding and mono-adduct formation (Fig. 1B). Lastly, internally labeled duplexes are formed through ligation of 6 ssDNAs, one of which is 5'-end-labeled prior to ligation (Fig. 1C). IC2 is labeled in Fig. 1Ci, IC3 in Fig. 1Cii, and IC6 in Fig.1Ciii.

Incision activity of CHO cell-free extracts

In order to ensure excision activity of the cell-free extracts and our ability to detect the excision products, a 6-4 photo-adduct containing an internal radiolabel (a kind gift from Tadayoshi Bessho) was used. Excision of this 6-4 photoproduct was performed with extracts we made and the canonical excision pattern was seen (Fig. 2). The excision products ranging from 24 to 27 nucleotides in length with the primary product being 27 nucleotides were seen only in the presence of active cell-free extracts and were absent in the mock reaction without extracts. The amount of excision products was as expected for this substrate.

In vitro binding of 3'-end-labeled TFOs

The initial substrate for incision was a linear 195 bp fragment in conjunction with a series of 3'-end-labeled TFOs in the purine motif. To 3'-end-label the TFO,

dideoxyadenine was the only available P^{32} labeled dideoxynucleotide, therefore the purine TFO sequence had to be redesigned to have an adenine in the appropriate places to allow for comparison of different length TFOs 3'-end-labeled with [α-^{32}P]ddA. Binding studies were done with these TFOs at pH 7.6 with 10 mM $MgCl_2$ and 1 mM spermine to a 57-mer duplex substrate made of NEWPUR/NEWPYR. The dissociation constant (K_d), the molar concentration of TFO where half of the duplex substrate is bound to the TFO, was determined for each TFO. New11 did not have a detectable K_d. New15, New20, New24, and New30 had K_d's of 10^{-9}, 10^{-8}, 10^{-8}, 10^{-9} M, respectively (Fig. 3). To create the appropriate TFO target for these TFOs, their TBS site was cloned into pIND/lacZ in both orientations. The linear substrates, NEWAG and NEWTC were cut out of this plasmid with XmnI and KpnI creating a 195 bp linear fragment. NEWAG had the purine strand of the TBS 5' to 3' and NEWTC had the pyrimidine strand 5' to 3' in the plasmid.

Mono-adduct formation for incision substrates

Psoralen TFOs were chosen because cross-linking or mono-adducting psoralen to the plasmid substrate provided a greater likelihood of the TFO being present during the incision reaction in cell-free extracts. Given our interest in the TFO being adducted to the strand to which the TFO is bound and not to both strands of the duplex, the efficiency of mono-adduct formation with a given substrate needed to be determined. While previous work in the Glazer lab showed 12 minutes at 447 nm was optimum for mono-adduct formation and 6 minutes at 320 nm was optimum for cross-link formation [23], with these substrates and the psoNew TFOs 10 seconds at 320 nm proved optimum for

mono-adduct formation (Fig. 4). These results are for the pso-New10 TFO and are representative of the other pso-New TFOs used in this work. Lanes 1, 3, 5, 7, and 9 contain labeled duplex alone without TFO. Lanes 2, 4, 6, 8, and 10 contain labeled duplex bound to TFO prior to exposure to UVA. The lanes without TFO did not show any mono-adduct or cross-link formation. Given that this is a denaturing gel, only mono-adducted or cross-linked structures, which are both larger than the duplex, will move more slowly through the gel matrix. A TFO mono-adducted to one duplex strand is larger than the strand alone and a TFO cross-linked to both strands of the duplex is even larger. The smallest bands (seen in all lanes) represent the duplex while the slightly larger bands (seen in Lanes 2, 4, 6, 8, and 10) indicate mono-adduct formation between one of the duplex strands and the TFO. The largest bands (seen in lanes 2, 4, 6, 8, and 10) indicate cross-link formation between the TFO and both strands of the duplex. With time of UVA exposure increasing, mono-adduct formation decreases and cross-link formation increases. The optimum mono-adduct formation and minimal cross-link formation occurs at 10 seconds of UVA exposure.

Incision with 3'-end-labeled TFOs on linear substrates

Using psoNew24 with NEWAG and NEWTC linear substrates (unlabeled) binding, mono-adduction, and incision are as above. In Fig. 5, Lanes 1 and 2 show psoNew24 TFO only. The banding represents the psoNew24 TFO alone with its dimer and possible G-quartet (as this is a G-rich TFO). Lane 3 shows NEWAG with psoNew24 bound, mono-adducted, and incubated with CHO cell-free extracts. This lane shows a

small amount of large intact substrate, a middle size product (G-quartet formation), a smaller prominent product (unbound TFO), and a smaller ladder of possible incision products. Lane 4 shows NEWAG with psoNew24 bound, mono-adducted, and incubated with all of the same conditions as Lane 3 except dialysis buffer alone was added in place of cell-free extracts. This lane shows more intact substrate, G-quartet formation, unbound TFO, and less possible incision products. Lanes 5-9 show end-labeled New11, New 15, New20, New24, and New30, respectively. Recall that since the TFO carried the label, only fragments adducted to the TFO or TFO complexes alone are seen. The resulting image indicated a major product between 19 and 23 nucleotides in length (Fig. 5, Lane 3). This approximate size was determined by comparison to psoNew24 alone (Lanes 1 and 2) and end-labeled New11, New 15, New20, New24, and New30 in Lanes 5-9.

The signal distinction between the amount of this product with and without extract was small (compare Fig. 5, Lanes 3 and 4) and confounded by the fact that the labeled TFO ran at about the same size. Therefore, removal of the unbound TFO was necessary to confirm that these products were indeed excision fragments and not unbound TFO. To this end following TFO binding and mono-adduct formation, TFO binding was disrupted and size filtration was used to remove excess, unbound, and unadducted TFO [22] (Fig. 6A). Labeled TFO allows for monitoring of the removal of unbound TFO with size separation columns. Size filtration removed most of the unbound psoNew24 TFO, but a ladder similar to that seen in Fig. 5, Lane 3 remained (Fig. 6B). Gel purification on

a 10% denaturing gel was also employed and the yield of large substrate alone (and no small ladder) is shown in Fig. 6C.

The products resulting from incision after size purification showed a different pattern from the pre-purification results (Fig. 7). Lanes 1 and 2 showed the flow-through from each size filtration wash. Lane 3 shows NEWTC and psoNew24 bound, mono-adducted, and incubated with CHO cell-free extracts. Lane 4 shows the same binding and adduction, but mock incubation. Lane 5 shows NEWAG and psoNew24 incubated with cell-free extracts. Lane 6 shows the same substrate with a mock incision reaction. Lane 7 shows psoNew24 alone. Lanes 8-11 show end-labeled New15, New 20, New24, and New30. Note that the lanes with substrate and TFO but without extract (Lanes 4 and 6) look similar to the flow through lanes (Lanes 1 and 2) for all small products. However, comparing the large bands in Lane 3 to Lane 4, and Lane 5 to Lane 6, there is a product slightly smaller than the initial substrate in both lanes containing extracts (see arrows). After only size filtration, there were some small fragments of excision size with extract present but without extracts the amount of labeled TFO was still overwhelming possible fragments (Fig. 7). The presence of small excision fragments was still confounded by remaining psoNew24 TFO. However, it appears that there are larger products (closer in size to the initial substrates) that are present only when cell-free extracts are present and absent in the mock incision reaction. Given that the expected products were small, this gel (Fig. 7) was not run far enough to significantly separate the larger products.

In order to clarify the nature of these products and to determine if the small fragments were excision products or unbound TFO, gel purification on a 10% denaturing gel was done with the TFO bound and mono-adducted substrate prior to the incision reaction. The purified substrate was electroeluted out of the gel and then the triplex structure was allowed to re-form prior to the incision reaction. In Fig. 8, Lane 1 shows NEWAG and psoNew15, bound, mono-adducted, and incubated with cell-free extracts. Lane 2 shows the same substrate and TFO, bound, and mono-adducted but mock incubated. Lane 3 shows NEWAG and psoNew24, bound, mono-adducted, and incubated with cell-free extracts. Lane 4 shows the same substrate and TFO, bound, and mono-adducted but mock incubated. Note in both Lanes 1 and 3 there are two products smaller than the 195 bp substrate and TFO (see arrows), these products are not present in Lanes 2 and 4 (with the mock incision). Lanes 5 and 6 show the substrates and TFO without the mock or actual incision reaction. Lanes 7-10 show end-labeled New30, New24, New20, and New15, respectively. The purified substrate used for incision in Fig. 8 is visualized in Fig. 6C.

Following this extra purification step, there was a trace amount of small product, but there was minimal difference of this product between the incision reaction and the mock reaction (Fig. 8, compare Lane 1 to 2 and Lane 3 to 4). The major difference between the incision reaction products and the mock reaction was seen in the production of two fragments smaller than the initial substrate but too small to be the excision fragments that would be expected if a dual incision reaction occurred and the TFO-

adducted fragment was released. These products suggest that a single incision was made on one side or the other of the TFO mono-adducted and bound to the duplex. Since the TFO was labeled, only the strand adducted to the TFO was detected and the products visualized in Fig. 8 contain the TFO and some part of the initial substrate.

Internally labeled substrates

In order to locate the site of this incision, an internally labeled substrate was designed. This internally labeled substrate was similar to that used in previous work studying incision of thymine dimers [7, 8]. This internally labeled substrate was built by ligating 6 ssDNAs together, one of which was 5'-end-labeled prior to ligation. The internally labeled duplex was designed to be 137 bp long. The construction of the duplex is shown in Fig. 1C. Any one or more of the ssDNAs can be 5'-end-labeled before ligation, yielding several possible locations of a radioactive label on the substrate. The construction of these duplexes requires a gel purification step to completely isolate ligated products and to remove partially ligated products (Fig. 9). Lane 1 shows the 195 bp NEWAG substrate for size comparison. Lane 2 shows radiolabeled IC2 and minimal partially or fully labeled product. Lane 3 shows radiolabeled IC3 with fully and partially ligated products and IC3 alone. Lane 4 shows radiolabeled IC6 with fully and partially labeled ligation products and IC6 alone. Lanes 5 and 6 show IC3 and IC6 alone, respectively. IC2 was smaller and not visualized on this gel. The IC fragments to construct this 137 bp duplex ranged in size from 33 nt to 52 nt and several partially ligated products can be seen in Fig. 9. Fragments of 85 bp and 94 bp represent partial

ligations while fragments of 44 bp and 52 bp represent unligated IC6 and IC3, respectively. Very little full size 137 bp substrate is present.

In attempts to increase the yield of the internally labeled substrate, ratios of the ssDNAs and the amount of labeled ssDNA were varied. Ligation time and the ssDNA that was labeled were also varied and ssDNAs were gel purified prior to ligation. The amount of labeled ssDNA was scaled up markedly to 10 μg in order to increase the amount of 137 bp product. This substrate was limited in both yield, as the final ligation product showed a very low yield, and time, as the radioactive signal was time sensitive. No data was collected with this substrate to date.

Incision with linear end-labeled substrates

All substrates were end-labeled with radioactivity and the binding of the TFO to the substrate and mono-adduct formation was verified visually on a denaturing gel. Therefore the TFO bound state of the substrate could be verified and the percent of bound versus unbound substrate calculated and unbound TFO could be removed if necessary. Linear substrate duplexes allowed detection of the fragment between the incision and the 5'-end of the duplex. The 5'-end-labeled duplex substrate was easy to create in large quantities by restriction enzyme digestion of the appropriate plasmid containing the TBS. With these duplex substrates it was difficult to follow the fragment with the TFO adducted to it (the potential excision fragment).

Incision with plasmid substrates

The plasmid substrate psupFG1 which contains a TBS for AG30 was used with pso-AG30 as the TFO. The plasmid substrate required an additional step of restriction digestion to cut out a fragment small enough to be visualized on a polyacrylamide gel (about 200 bp). Additionally, since the plasmid was not internally labeled, the products were run on a gel, transferred to a membrane, and then probe hybridization was attempted to analyze the products. Probes to the AG30 target site and surrounding sequence were used. Although size filtration was used to remove unbound TFO from the sample mix before the incision reaction was carried out, probes could also bind to the remaining unbound TFO, and differentiating between excess TFO and incision products was difficult. Results with this substrate and method were limited and difficult to interpret. Attempts to use an internally labeled plasmid and a 3'-labeled TFO with a plasmid did not show positive results, perhaps due to loss of products in the extra restriction enzyme step required prior to visualization on a gel.

Discussion

Determining the nature of the incision products required forming a specific triplex structure, incubating that structure with the cell-free extracts for the incision reaction to occur, and then visualizing the products. Each of these steps had several obstacles to be overcome in order to create and visualize triplex products. Based on previous work in the Glazer lab, incisions were thought to be made 3' and 5' of the TFO when bound to the

TBS in removal of a 10-mer TFO [14], potentially analogous to the incision pattern seen in excision of a thymine dimer [7, 8]. However, a 30-mer TFO (AG30) did not yield a similar excision product. This led us to speculate that while NER may still make incisions near the triplex lesion, the "wing-span" of the NER complex cannot straddle such a large lesion/helical distortion. However, to test whether or not single or double incisions do occur, several approaches were considered.

The most promising approach used was a 3'-end-labeled TFO on a linear substrate, as the fragment bound to TFO could be tracked. The incision products detected have to be attached to the TFO because only the TFO was labeled. The confounding signal caused by the unbound TFO could be removed by size filtration and gel purification. The large size of the products was surprising given that the TFOs used were a 15-mer and a 24-mer and not a 30-mer. This suggests a different mechanism of excision compared to repair of a 10-mer. This mechanism may be similar to that for the 30-mer or could represent an entirely different mechanism. Previous notions of incision of a 10-mer but not a 30-mer were based on repair of intermediates of a TFO adducted to a plasmid [25]. While we have shown that these products are smaller than the initial substrate and represent a fragment adducted to the TFO, the exact location of the incision and the nature of these products remains to be determined.

Labeling the TFO at the 3'-end (as there is psoralen on the 5'-end) allows for the fragment with the mono-adduction to the TFO to be tracked following incision. This fragment should be of a different mobility than the labeled TFO alone, and may have a

Y-like structure (if the mono-adduct is still intact). This approach also allows for the direct comparison of different TFO lengths of the same sequence. One drawback of this approach is that the nature of the excision products will require deciphering a y-shaped structure and will require synthesis of a complex ladder of y-shaped structures of varying lengths to determine the length of the excision fragment and the location of the incision.

Additional approaches such as the internally labeled substrate can be used to decipher these products. The internally labeled substrate offers the advantage of mobility of the label. And perhaps further increases in the scale of the ligation can increase the amount of the fully ligated product. End-labeling the NEWAG and the NEWTC substrates instead of labeling the TFOs will allow for determination of the other products present in the incision reaction. The end-label approach uses a 195 bp fragment that is easy to make but the potential, labeled, incision product is a larger fragment representing the distance from the end of the initial substrate to one incision. Thereby the excised fragment if there is one, will not be visualized directly. Running samples in parallel with the TFO labeled in one sample, the substrate labeled in a second sample, and both labeled in a third sample will help piece together all of the fragments present after incision.

A complementary approach which has not yet been explored is to monitor transcription around these incisions to perhaps detect the location at which the incision is made based on where the polymerase falls off the substrate. This method has been used to detect peptide nucleic acids bound to a plasmid substrate [24]. The concept being that a disruption in the template could block transcription, whereas in this case either a

disruption caused by the altered helical structure or incisions made in repair of this structure could disrupt transcription and the location of this disruption can be mapped.

Using a plasmid substrate had the advantage of not introducing radioactivity until the last step but the disadvantage of using small fragment probes which bound to unbound TFO as well. Given that the incisions were likely to be in close proximity to the TBS, based on data showing that mutation frequencies are highest in the TBS and specifically near the psoralen mono-adduct or cross-link site [26], probing for the incision products requires including some of the TBS in the probe and therefore most probes binding to the incision products would also bind to free TFO remaining in the sample. There was no means to visually check this substrate because only the probe was radioactively labeled, thereby only the fragments adhering to the probe would be visualized. This approach is appealing because a variety of probes can be used to map the resultant excision products. However, these probes are likely to pick up unbound TFO as well.

Future directions of this work require extensive purification of the 137 bp substrate with internal labels. Additionally, end-labeling the 195 bp NEWAG and NEWTC fragments in order to attempt to detect the same larger products seen in Fig. 5 when the substrate is labeled instead of the TFO is a priority. Additionally, a similar repetition with the 137 bp internally labeled fragment would help to further isolate the location of the incision.

Once the incision sites for the psoralen mono-adducted TFO pso-New24 is confirmed by using these other substrates, the other 3'-end-labeled pso-New TFOs should be used in conjugation with these substrates. A comparison of the incision patterns between the different TFOs will suggest the TFO length limitation for the fragment excision as compared to the single incision. Additionally, TFOs in the pyrimidine motif can be evaluated and unconjugated TFOs in both motifs can be compared. Finally, peptide nucleic acid (PNA)-modified TFOs can be evaluated and compared. In summary, these techniques allow the location of incisions in repair of triplexes to be unveiled and thereby an important step towards understanding the mechanism of triplex activity is made.

References

1. Friedberg, E.C., G.C. Walker, and W. Siede, *DNA Repair and Mutagenesis*. 1995, Washington, D.C.: ASM Press. p. 283.

2. Bootsma, D., et al., *Nucleotide Excision Repair Syndromes: Xeroderma Pigmentosa, Cockayne Syndrome, and Trichothiodystropy*, in *The Metabolic & Molecular Bases of Inherited Disease*, C.R. Schriver, et al., Editors. 2001, McGraw-Hill: New York. p. 677-703.

3. Balajee, A.S. and V.A. Bohr, *Genomic heterogeneity of nucleotide excision repair.* Gene, 2000. **250**(1-2): p. 15-30.

4. Aboussekhra, A., et al., *Mammalian DNA nucleotide excision repair reconstituted with purified protein components.* Cell, 1995. **80**(6): p. 859-68.

5. Moggs, J.G., et al., *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, 1996. **271**(12): p. 7177-86.

6. Mu, D., D.S. Hsu, and A. Sancar, *Reaction mechanism of human DNA repair excision nuclease.* J Biol Chem, 1996. **271**(14): p. 8285-94.

7. Huang, J.C., et al., *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, 1992. **89**(8): p. 3664-8.

8. Svoboda, D.L., et al., *DNA repair by eukaryotic nucleotide excision nuclease. Removal of thymine dimer and psoralen monoadduct by HeLa cell-free extract and of thymine dimer by Xenopus laevis oocytes.* J Biol Chem, 1993. **268**(3): p. 1931-6.

9. Matsunaga, T., et al., *Human DNA repair excision nuclease. Analysis of the roles of the subunits involved in dual incisions by using anti-XPG and anti-ERCC1 antibodies.* J Biol Chem, 1995. **270**(35): p. 20862-9.

10. Evans, E., et al., *Mechanism of open complex and dual incision formation by human nucleotide excision repair factors.* Embo J, 1997. **16**(21): p. 6559-73.

11. Sladek, F.M., A. Melian, and P. Howard-Flanders, *Incision by UvrABC excinuclease is a step in the path to mutagenesis by psoralen crosslinks in Escherichia coli.* Proceedings of the National Academy of Sciences of the United States of America, 1989. **86**(11): p. 3982-6.

12. Cole, R.S., *Repair of DNA containing interstrand crosslinks in Escherichia coli: sequential excision and recombination.* Proceedings of the National Academy of Sciences of the United States of America, 1973. **70**(4): p. 1064-8.

13. Cheng, S., A. Sancar, and J.E. Hearst, *RecA-dependent incision of psoralen-crosslinked DNA by (A)BC excinuclease.* Nucleic Acids Research, 1991. **19**(3): p. 657-63.

14. Wang, G., M.M. Seidman, and P.M. Glazer, *Mutagenesis in mammalian cells induced by triple helix formation and transcription-coupled repair.* Science, 1996. **271**: p. 802-805.

15. Faruqi, A.F., et al., *Triple-helix formation induces recombination in mammalian cells via a nucleotide excision repair-dependent pathway.* Mol Cell Biol, 2000. **20**(3): p. 990-1000.

16. Datta, H.J., et al., *Triplex-induced Recombination in Human Cell-free Extracts. Dependence on XPA and HsRad51.* J Biol Chem, 2001. **276**(21): p. 18018-23.

17. Vasquez, K.M., et al., *Human XPA and RPA DNA repair proteins participate in specific recognition of triplex-induced helical distortions.* Proc Natl Acad Sci U S A, 2002. **99**(9): p. 5848-53.

18. Faruqi, A.F., et al., *Triple-helix formation induces recombination in mammalian cells via a nucleotide excision repair-dependent pathway.* Mol Cell Biol, 2000. **20**(3): p. 990-1000.

19. Kalish, J.M., et al., *Triplex-induced recombination and repair in the pyrimidine motif.* Nucleic Acids Res, 2005. **33**(11): p. 3492-502.

20. Wang, G. and P.M. Glazer, *Altered repair of targeted psoralen photoadducts in the context of an oligonucleotide-mediated triple helix.* The Journal of Biological Chemistry, 1995. **270**(39): p. 22595-22601.

21. Bradford, M.M., *A rapid and sensitive method for the quantitation of microgram quantities of protein utilizing the principle of protein-dye binding.* Anal Biochem, 1976. **72**: p. 248-54.

22. Macris, M.A. and P.M. Glazer, *Transcription dependence of chromosomal gene targeting by triplex-forming oligonucleotides.* J Biol Chem, 2003. **278**(5): p. 3357-62.

23. Gasparro, F.P., et al., *Site-specific targeting of psoralen photoadducts with a triple helix-forming oligonucleotide: characterization of psoralen monoadduct and crosslink formation.* Nucleic Acids Res, 1994. **22**(14): p. 2845-52.

24. Kim, K.H., P.E. Nielsen, and P.M. Glazer, *Site-Specific Gene Modification by PNAs Conjugated to Psoralen.* Biochemistry, 2006. **45**(1): p. 314-323.

25. Wang, G. and P.M. Glazer, *Altered repair of targeted psoralen photoadducts in the context of an oligonucleotide-mediated triple helix.* J Biol Chem, 1995. **270**(38): p. 22595-601.

26. Wang, G., et al., *Targeted mutagenesis in mammalian cells mediated by intracellular triple helix formation.* Mol Cell Biol, 1995. **15**(3): p. 1759-68.

Figures

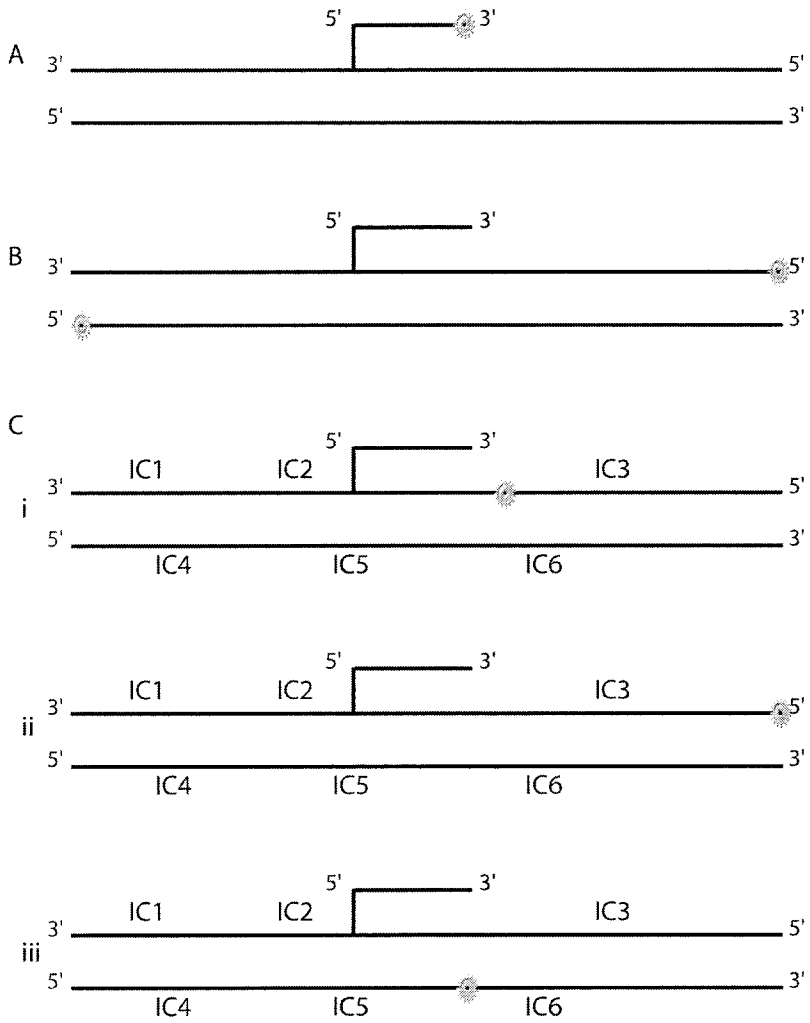


Figure 1. Substrates for incision assay. (A) 3'-end-labeled TFO with psoralen mono-adducted to the purine strand of the TBS on the 195 bp duplex. (B) 195 bp duplex (NEWAG and NEWTC) 5'-end-labeled with a psoNew TFO bound and mono-adducted. (C) Internally labeled substrates. i. IC2 fragment was 5'-end-labeled prior to ligation. ii. IC3 fragment was 5'-end-labeled prior to ligation. iii. IC6 fragment was 5'-end-labeled prior to ligation.

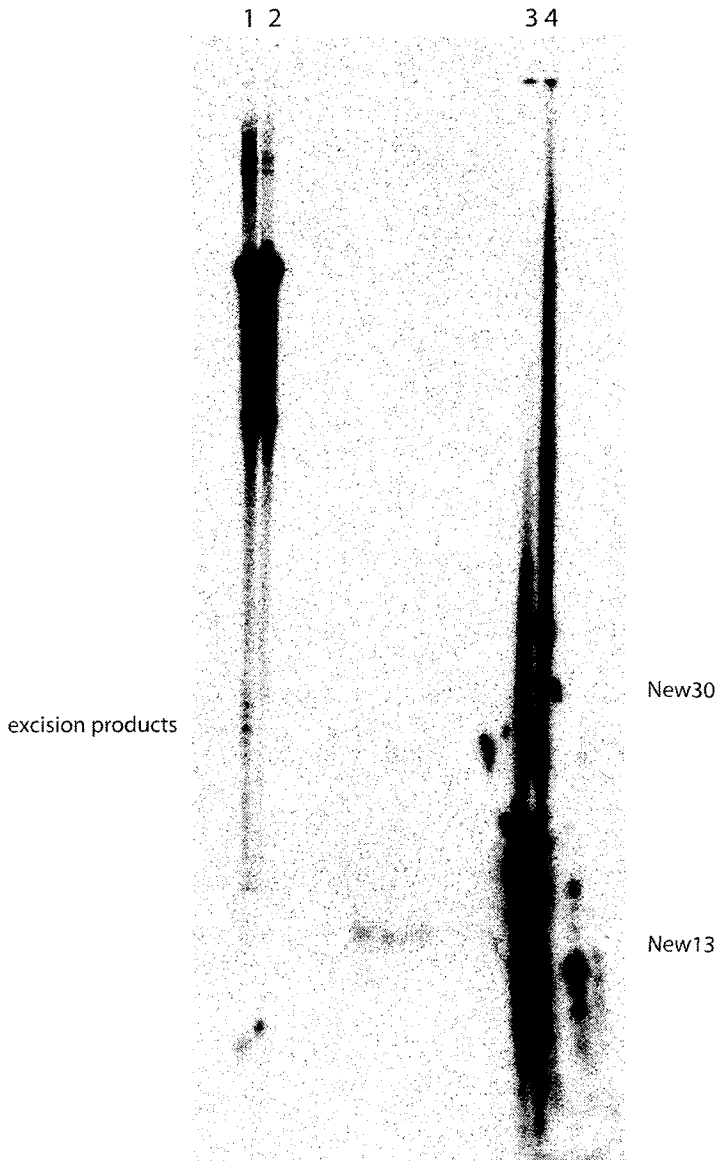


Figure 2. Verification of CHO cell-free extracts activity and visualization method. Lanes 1 and 2 contain a 6-4 photo-adduct in an internally labeled substrate. CHO cell-free extracts were added to Lane 1 and a mock incision reaction was added to Lane 2. Lane 3 contains an end-labeled New30 and Lane 4 contains end-labeled New13 as size markers.

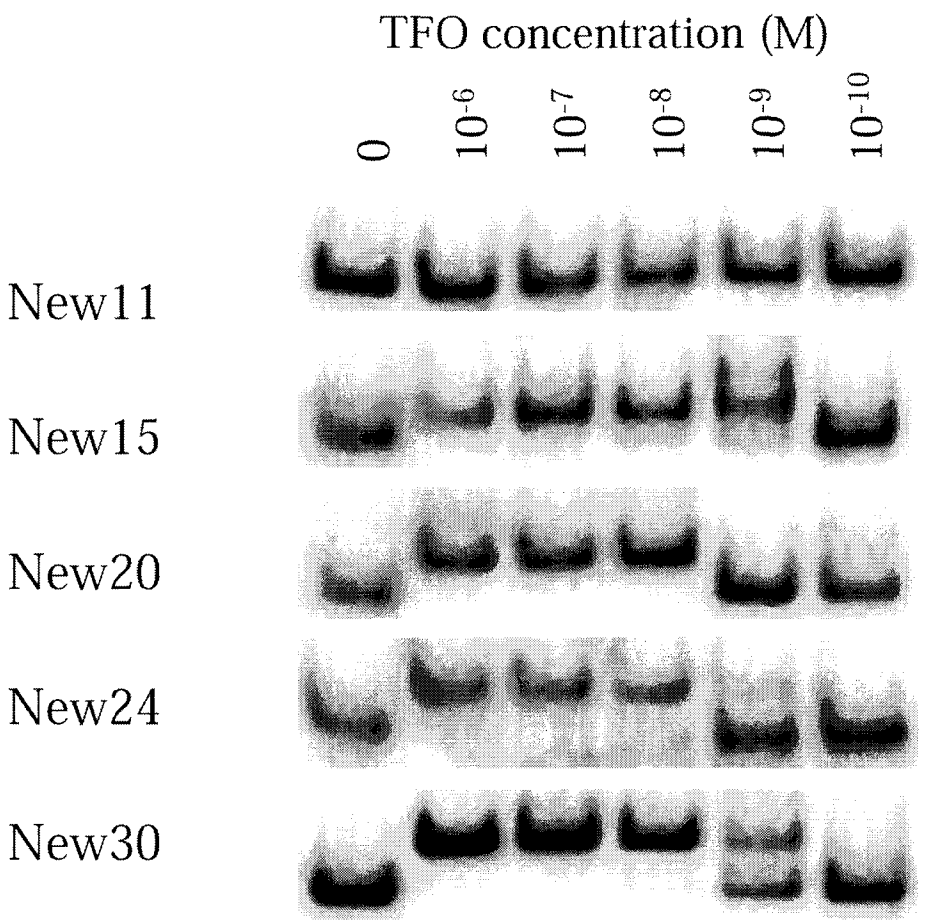


Figure 3. *In vitro* binding affinity of New TFOs. TFOs of different lengths were bound to a 57 bp end-labeled duplex and binding affinity was determined by the dissociation constant (K_d) which is the concentration of TFO where half of the duplex is bound to TFO and half is unbound, under the given set of conditions (10 mM Tris, pH 7.6, 10 mM $MgCl_2$, 1 mM spermine). The bands of reduced mobility relative to the duplex alone represent triplex formation.

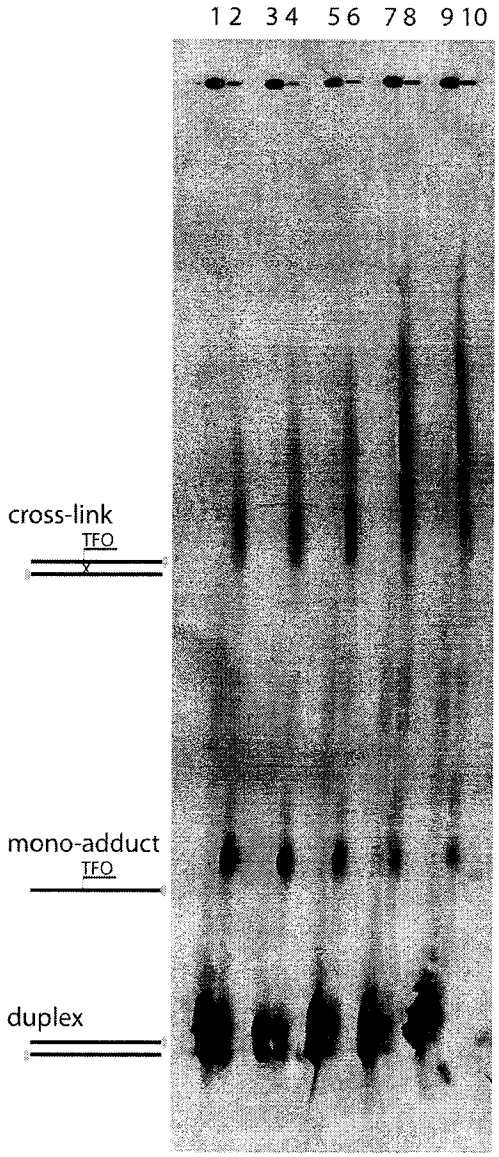


Figure 4. Mono-adduct versus cross-link formation with New TFOs. Using the same 57 bp duplex used in Fig. 3, the time of UVA exposure required to maximize mono-adduct formation was determined using psoNew10. Lanes 1, 3, 5, 7, and 9 represent duplex alone without psoNew10. Lanes 2, 4, 6, 8, and 10 represent duplex with psoNew10 bound as described in Fig. 3. Lanes 1 and 2 were treated with UVA at 320 nm for 10 seconds. Lanes 3 and 4 were treated for 30 seconds. Lanes 5 and 6 were treated for 1 minute. Lanes 7 and 8 were treated for 3 minutes. Lanes 9 and 10 were treated for 6 minutes.

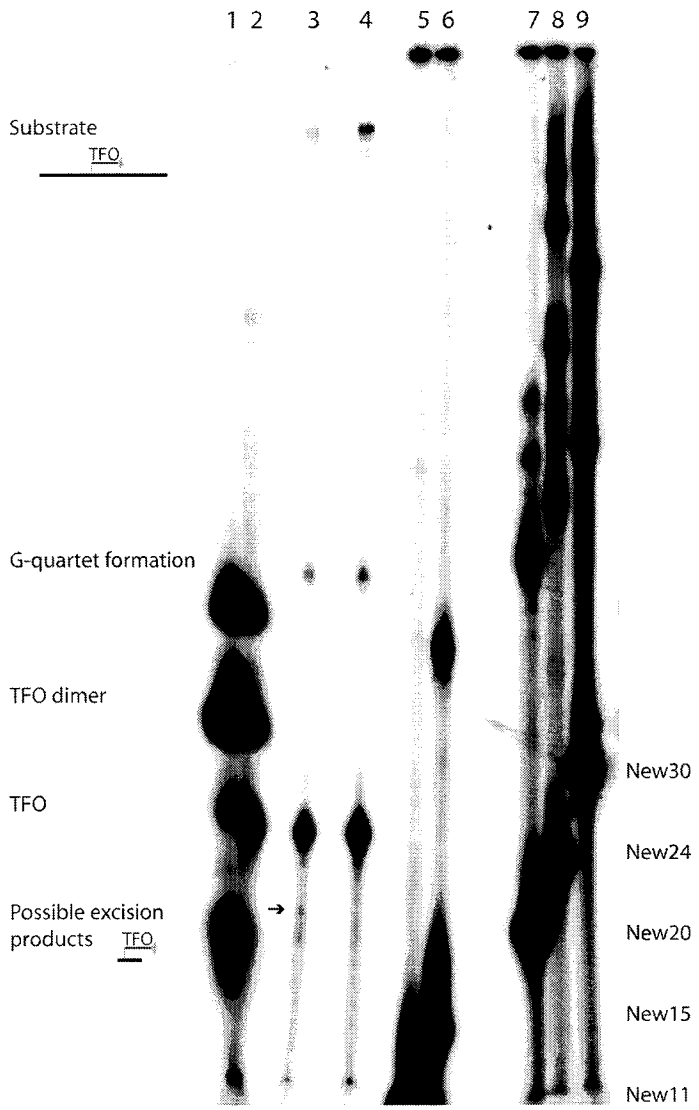


Figure 5. Incision of linear substrate adducted to a 3'-end-labeled TFO. Lanes 1 and 2 show 10 μM and 100 μM of psoNew24 alone. Lane 3 shows NEWAG and psoNew24, bound, mono-adducted, and incubated with CHO cell-free extracts. The arrow indicates excision products. Lane 4 shows the same binding, mono-adduction, and incubation in the same buffers as Lane 3, but no cell-free extracts were added. Lanes 5, 6, 7, 8, and 9 show end-labeled New11, New15, New20, New24, and New30, respectively.

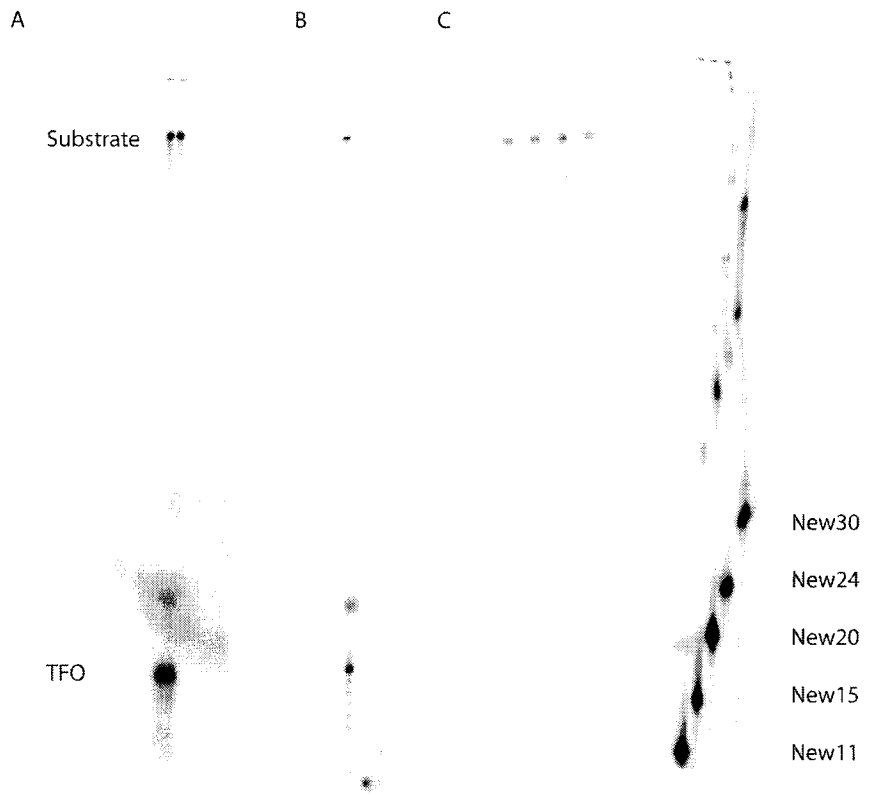


Figure 6. Purification of incision substrate. (A) A substrate containing the triplex binding of NEWAG (a 195 bp linear substrate) and psoNew24 followed by mono-adduct formation. (B) The substrate after disruption of triplex binding and size filtration to remove excess, unbound TFO. (C) Gel purification of the substrate.

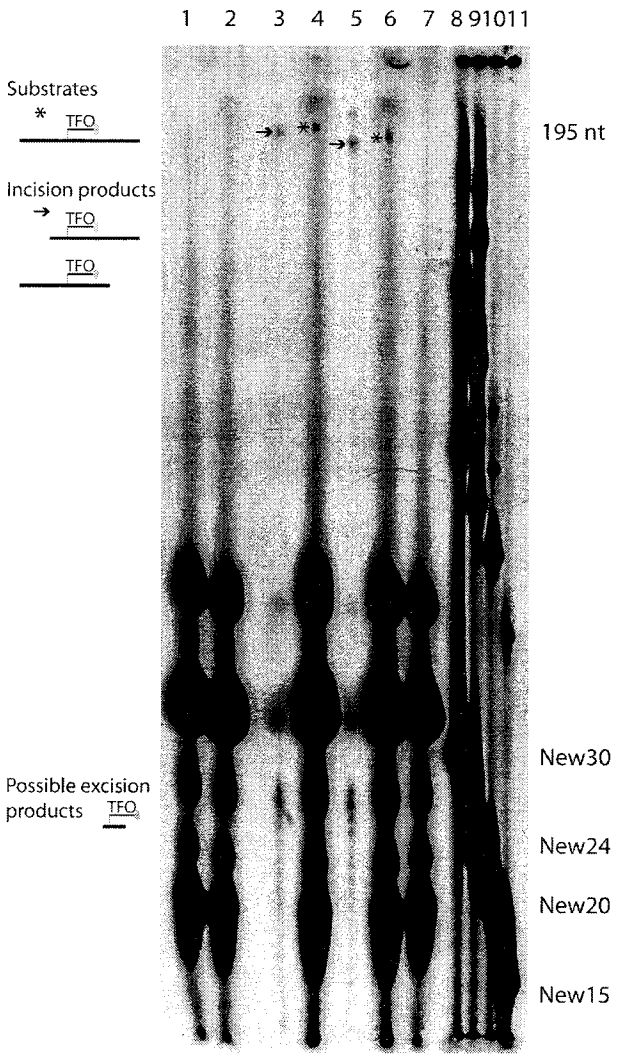


Figure 7. Incision of a substrate after size filtration. Lanes 1 and 2 show the flow-through as purification by size filtration proceeded. Lanes 3 and 4 show psoNew24 mono-adducted to NEWTC. CHO cell-free extracts were added to Lane 3 and a mock reaction was added to Lane 4. Lanes 5 and 6 show psoNew24 mono-adducted to NEWAG. CHO cell-free extracts were added to Lane 5 and a mock reaction was added to Lane 6. Lane 7 shows psoNew24 alone. Arrows indicate possible incision products and asterisks indicate substrate (NEWAG or NEWTC with psoNew24). Lanes 8-11 show end-labeled New30, New24, New20, and New15, respectively.

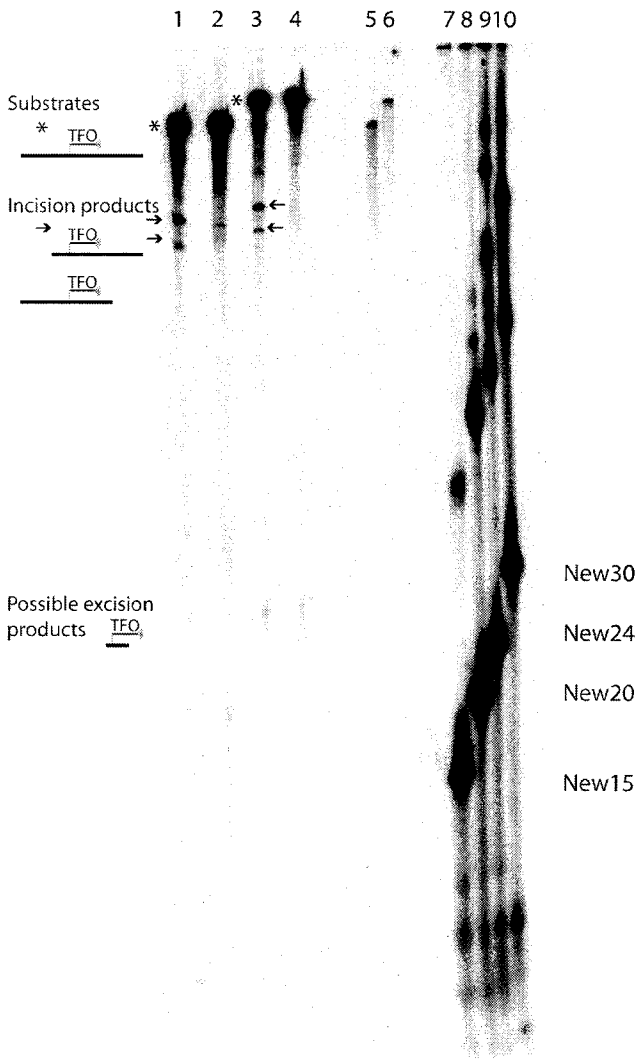


Figure 8. Incision of a purified, linear substrate adducted to psoralen TFOs. Lane 1 shows NEWAG and psoNew15, bound, mono-adducted, and incubated with cell-free extracts. Lane 2 had the same substrate but was mock incubated without extracts. Lane 3 shows NEWAG and psoNew24, bound, mono-adducted, and incubated with cell-free extracts. Lane 4 had the same substrate but was mock incubated without extracts. Lanes 5 and 6 show substrates NEWAG with psoNew15 and NEWAG with psoNew24 without extract incubation or mock incubation, respectively. Arrows indicate possible incision products and asterisks indicate substrate (NEWAG with psoNew15 or with psoNew24). Lanes 7-10 showed end-labeled New15, New20, New24, and New30, respectively.

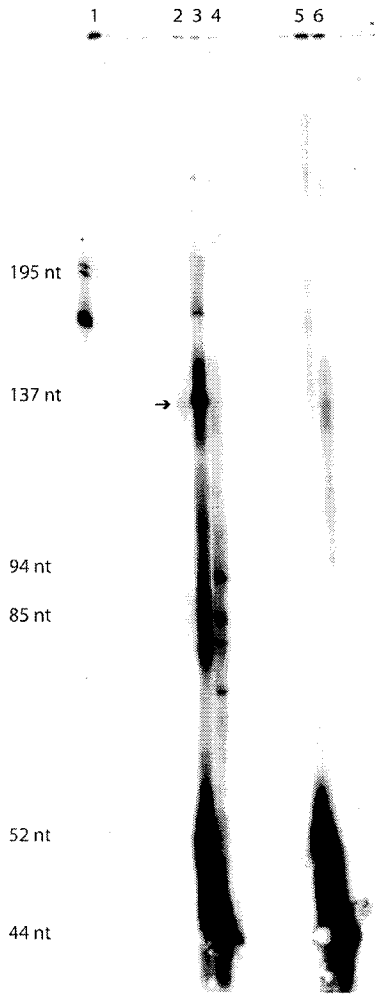


Figure 9. Construction of an internally labeled duplex. Lane 1 shows a 195 bp double stranded fragment. Lane 2 shows a ligation with IC2 labeled, fragment sizes within this lane include 137 bp, 85 bp, and 33 bp. Lane 3 shows a ligation with IC3 labeled, fragment sizes including 137 bp, 85 bp, and 52 bp are seen. Lane 4 shows a ligation with IC6 labeled, fragment sizes of 137 bp, 94 bp, and 44 bp are seen. Lane 5 shows the 52 bp fragment (IC3) and Lane 6 shows the 44 bp fragment (IC6) alone. 137 bp is the size of the fully ligated product (see arrow).

Chapter Five:
Conclusions and Future Directions

Gene targeting via homologous recombination (HR) offers a potential strategy for therapeutic correction of mutations in disease-related human genes. However, there is a need to improve the efficiency of site-specific recombination by transfected DNAs. Oligonucleotide-mediated triple helix formation has been shown to constitute a DNA lesion sufficient to provoke DNA repair and thereby stimulate recombination. Triplex-forming oligonucleotides (TFOs) bind to duplex DNA in a sequence-specific manner and are potential tools to achieve targeted gene modification. Initial studies demonstrated the ability of TFOs to deliver mutagenic agents in a DNA site-specific manner. Gene modification with TFOs includes induced recombination between a DNA target and a donor DNA molecule, a process that allows a TFO to exert an effect at a distance from the third-strand binding site. It was also found that TFOs in the purine motif can induce gene modification in chromosomal DNA via the effect of the triple helix itself but thus far, the effect of the pyrimidine motif TFO alone has not been demonstrated.

The work presented here expands the number of possible target sites for TFO-induced gene correction by demonstrating the effect of pyrimidine TFOs, independent of a DNA reactive conjugate, to induce recombination and repair. However, the ability of triplex-forming oligonucleotides to induce recombination between a target locus and a donor DNA was so far previously demonstrated only with multi-copy episomal targets in mammalian cells. This work demonstrates the activity of purine and pyrimidine motif TFOs in a single-copy chromosomal target. Previous work in the Glazer lab demonstrated that triplex-induced mutagenesis and recombination is dependent on the

nucleotide excision repair (NER) pathway [1-4]. The mechanism of incision of a triplex during repair is not currently understood and these experiments are designed to illuminate aspects of this mechanism.

We established that a number of chemically modified TFOs in the pyrimidine motif showed enhanced binding affinity. These included N3'->P5' phosphoramidate (amidate), 5-(1-propynyl)-2'-deoxyuridine (pdU), 2'-O, 4'-C-methylene bridged or locked nucleic acid (LNA), 2'-O-methyl-ribose (2'-O-Me), and 2'-O-(2-aminoethyl)-ribose (2'-AE)-modified TFOs. Other chemical modifications such as N, N-diethylethylenediamine (DEED), 5-methyl-2'-deoxycytidines (5meC), and morpholino (morph) which we expected to show enhanced binding affinity based on previous use of these modifications for antisense and other gene modification technology did not show such affinity.

Enhanced binding affinity was necessary but not sufficient to induce recombination in cells and repair in cell-free extracts as only the amidate and pdU-modified TFOs showed such activity. The amidate TFO showed similar induction of recombination to that demonstrated previously with the purine motif TFO, AG30, in an analogous reporter system. The amidate-induced recombination, like the AG30-induced recombination, was shown to be dependent on XPA. The amidate effect was not based on enhanced nuclease stability compared to the other chemically modified TFOs tested. These results show that amidate and pdU-modified TFOs can be used as reagents to stimulate site-specific gene targeting without the need for conjugation to DNA-reactive

molecules. The biological activity of the more DNA-like TFOs (amidate and pdU) demonstrated here and the lack of activity of the more RNA-like TFOs (LNA, 2'-O-Me, 2'-AE) in conjunction with NMR data by Asensio et al., which showed that triplexes with a DNA third strand are more helically distorting than those with a RNA third strand [5], suggests that greater helical distortion favors greater recognition by DNA repair mechanisms and therefore greater induction of recombination.

The efforts to test a series of chemical modifications on the third strand presented here led to the identification of the amidate modification as an inducer of recombination and repair based on the effect of the TFO alone. Demonstrating for the first time that an unconjugated TFO in the pyrimidine motif is biologically active and thereby, increasing the number of target sites possible for gene correction. These efforts provide a standardized screening method for testing novel chemical modifications as they are developed. Since the initial publication of the amidate work, we screened several new modifications but found that the amidate was still the most effective modification tested.

The biological activity of the amidate demonstrated above was in an episomal target in cells or in cell-free extracts. In order to demonstrate gene correction with TFOs as a model for gene therapy, a single-copy chromosomal target was designed using the *firefly luciferase* reporter gene. CHO cell lines containing triplex binding sites upstream of a *firefly luciferase* gene containing a stop codon, were made with purine and pyrimidine binding sites, respectively. The cell lines all had the same integration site for the TBS and reporter gene and cell lines with the wild-type *luciferase* gene were also

made. The sensitivity of the assay was determined to be less than 0.01%, where one wild-type cell could be detected among 10,000 mutant cells. Gene correction was achieved when TFOs were contransfected with single-stranded DNA donors at frequencies up to 0.1%. The TFO-induced correction was target site, sequence, and dose dependent.

The pyrimidine TFO, amidate18, and the purine TFO, AG30, showed similar levels of gene correction, demonstrating that correction was independent of the orientation of the TFO. Antisense DNA donors (homologous to the transcribed strand of the duplex) were more effective in inducing recombination in conjuction with TFO binding than sense donors. The 51-mer donor, end-modified with phosphorothioate linkages was most effective compared to a 30-mer and 69-mer donor and to LNA-modified donors. In comparison to other chemically modified TFOs in the pyrimidine motif, the amidate modified-TFO showed the highest percentage of induced recombination. In total, this work demonstrated the effectiveness of TFOs in targeting a single-copy chromosomal target and suggests the usefulness of TFOs as sequence-specific gene modification tools.

Finally, we developed an assay to detect incisions made in repair of a triplex. We demonstrated the excision activity of cell-free extracts and designed a triplex substrate to detect incisions. This substrate consisted of a 195 bp duplex and a TFO with a 5'-psoralen and a 3'-radiolabel. The TFO was bound to the duplex, mono-adducted, size filtered, and gel purified. The purified substrate was incubated with cell-free extracts or

mock-incubated without extracts. Using a 15-mer and a 24-mer TFO, incision products were detected in repair of a triplex. These products were smaller than the initial substrate and similar for both the 15-mer and the 24-mer TFO bound substrates. Smaller excision fragments were not detected. This work represents first steps towards understanding the nature of incisions made in repair of a triplex structure. Future work will further investigate the nature of the incisions detected here. Placing the radiolabel at different locations along the duplex DNA will allow for a more precise location of where the incisions were made.

Future directions of TFO technological development include the identification of novel chemical modifications to further enhance the biological activity of TFOs and the development of new techniques for TFO delivery to enhance the biological effectiveness of these reagents. Additionally, further steps towards identifying the mechanism of repair of a triplex such as the exact nature of the incisions and the nature of involvement of repair pathways besides NER need to be made. The interplay between the different repair pathways at work in repair of a triplex lesion is not understood and future directions of this work lie in deciphering these interactions.

In summary, our work suggests that high-affinity DNA binding ligands, such as TFOs, can be used to mediate site-specific genome modification. This capacity derives not only from the ability of TFOs to bind as third strands with sequence specificity but also from the ability of the resulting triple helices to provoke repair, leading to directed

mutagenesis, recombination, and potentially gene correction.

References

1. Wang, G., M.M. Seidman, and P.M. Glazer, *Mutagenesis in mammalian cells induced by triple helix formation and transcription-coupled repair.* Science, 1996. **271**: p. 802-805.

2. Faruqi, A.F., et al., *Triple-helix formation induces recombination in mammalian cells via a nucleotide excision repair-dependent pathway.* Mol Cell Biol, 2000. **20**(3): p. 990-1000.

3. Datta, H.J., et al., *Triplex-induced Recombination in Human Cell-free Extracts. Dependence on XPA and HsRad51.* J Biol Chem, 2001. **276**(21): p. 18018-23.

4. Vasquez, K.M., et al., *Human XPA and RPA DNA repair proteins participate in specific recognition of triplex-induced helical distortions.* Proc Natl Acad Sci U S A, 2002. **99**(9): p. 5848-53.

5. Asensio, J.L., Carr R., Brown, T., Lane, A.N., *Conformational and Thermodynamic Properties of Parallel Intramolecular Triple Helices Containing a DNA, RNA, or 2'-OMeDNA Third Strand.* J. Am. Chem. Soc., 1999. **121**: p. 11063-11070.