DESIGN AND DEVELOPMENT OF ARTIFICIAL ZINC FINGER TRANSCRIPTION FACTORS AND ZINC FINGER NUCLEASES TO THE HTERT LOCUS

APPROVED BY SUPERVISORY COMMITTEE

Matthew Porteus, M.D., Ph.D.

Carole Mendelson, Ph.D.

Woodring Wright, Ph.D.

George DeMartino, Ph.D.

DEDICATION

Dedicated to my husband (Byron) and my two babies (Bruce and Dana) for their generous love, patience, and support. I would like to thank my parents for every encouragement they gave me during this time and all the help they have provided over the many years.

DESIGN AND DEVELOPMENT OF ARTIFICIAL ZINC FINGER TRANSCRIPTION FACTORS AND ZINC FINGER NUCLEASES TO THE HTERT LOCUS

by

KIMBERLY ANNE WILSON

DISSERTATION

Presented to the Faculty of the Graduate School of Biomedical Sciences

The University of Texas Southwestern Medical Center at Dallas

In Partial Fulfillment of the Requirements

For the Degree of

DOCTOR OF PHILOSOPHY

The University of Texas Southwestern Medical Center at Dallas

Dallas, Texas

August, 2010

DESIGN AND DEVELOPMENT OF ARTIFICIAL ZINC FINGER TRANSCRIPTION FACTORS AND ZINC FINGER NUCLEASES TO THE HTERT LOCUS

Kimberly Anne Wilson

The University of Texas Southwestern Medical Center at Dallas, 2010

Matthew H. Porteus, M.D., Ph.D.

The ability to direct *hTERT* expression through genetic control or tunable regulatory factors would advance our understanding of the transcriptional regulation of *hTERT*, and also potentially produce new strategies for addressing telomerase-associated disease. In this work, we describe the engineering of

artificial zinc finger transcription factors (ZFTFs) and zinc finger nucleases (ZFNs) to target sequences at the *hTERT* promoter.

We first explored expansions to the repertoire of sites that can be targeted by ZFNs and modifications of ZFN architecture to accommodate such sites. A ZFN is made of a zinc-finger DNA binding domain (ZFP) linked to the FokI nuclease domain by a short amino acid "inter-domain linker". The general sequence motif of a ZFN target is 5'-(ZFN site1)-(6 bp spacer)-(ZFN site2)-3' and each half-site is 5'-GNNGNNGNN-3'. Variations of this motif come in the forms of variable spacer lengths, extra basepairs in-between triplets, and the inclusion of non-GNN triplets. To explore these types of target sites, we created ZFN variants that contained different inter-domain linkers, lengthened inter-finger linkers, and DNA binding domains created through hybridizing the modular assembly and OPEN methodologies. We show that through altering ZFN architecture, target sites with 5-7-bp spacers and those with ANN, CNN, and TNN triplets can be efficiently recognized and cut by ZFNs. We then generated new ZFPs to five ZFN target sites with 5- or 6-bp spacers in the *hTERT* locus based on those findings and made ZFTFs by linking the ZFPs to the VP16 transcriptional activation domain. We were able to identify several active ZFTFs that demonstrate a dose-dependent response. The same ZFPs were also converted into ZFNs and screened in combinatorial pairs in cell-based single-strand annealing assays and gene

targeting assays. These screening strategies have pinpointed several ZFN pairs that may be useful in genomic editing of the *hTERT* locus.

Our findings provide guidelines for modifying ZFP architecture to a wider array of potential target sites for use in developing ZFTFs and ZFNs at the *hTERT* promoter, which may be applicable towards inheritable, telomerase-based diseases and answering basic science questions about *hTERT* transcriptional regulation.

TABLE OF CONTENTS

PRIOR PUBLICATIONS

Coque L, Neogi P, Pislariu C, **Wilson KA**, Catalano C, Avadhani M, Sherrier DJ, Dickstein R. Transcription of ENOD8 in Medicago truncatula nodules directs ENOD8 esterase to developing and mature symbiosomes. Mol Plant Microbe Interact. 2008 Apr;21(4):404-10.

Wilson KA, Chateau ML, Porteus MH. Design and Development of Artificial Zinc Finger Transcription Factors and Zinc Finger Nucleases to the hTERT Locus. PLoS One. Manuscript in Preparation.

Wilson KA, McEwen AE, Pruett-Miller SM, Zhang J, Kildebeck EJ, Porteus MH. Exploring the Target Site Structure and Zinc Finger Nuclease Architecture for Efficient Gene Targeting in Mammalian Cells. Nucleic Acids Research. Manuscript in Preparation.

LIST OF FIGURES

LIST OF TABLES

LIST OF DEFINITIONS

aa – amino acid (as a unit)

B2H – bacterial-2-hybrid

bp – basepair (as a unit)

CAD – Caspase activated DNase

CAT – Chloramphenicol acetyltransferase

CtIP – CtBP interacting protein

DD – Obligate heterodimer form of the Fn with amino acids D483 And D487

DR – wild type form of the Fn with amino acids D483 And R487

DSB – DNA Double-Stranded Break

EE - Obligate heterodimer form of the Fn with amino acids Q486E And E490

EL - Obligate heterodimer form of the Fn with amino acids Q486E And I499L

Fn – Nuclease domain of the FokI endonuclease

F1 – Finger 1

F2 – Finger 2

F3 – Finger 3

GFP – Green fluorescent protein

hEPO – Human erythropoietin

hBAX – Human BCL2-associated X protein

hCCK2R – Human cholecystokinin 2 receptor

hCCR5 – Human chemokine receptor 5

hIL2Rγ – Human interleukin 2 receptor γ

hMDR1 – Human multi-drug resistance gene

hPTHR1 – Human parathyroid hormone receptor 1

hTR – Human telomerase RNA

HR – Homologous recombination

hTERT – Human telomerase reverse transcriptase

hVEGF-A – Human vascular endothelial growth factor A

kb – Kilobase (as a unit)

KK - Obligate heterodimer form of the Fn with amino acids E490K And I538K

KRAB – Kruppel-associated box

mdx – Dystrophin deficient mouse

MRN – Mre, Rad50 and Nbs1 complex

NHEJ – Non-homologous end joining

OPEN – Oligomerized pool engineering

ObhetFn – Obligate heterodimer FokI nuclease domain

QE - Wild type form of the Fn with amino acids Q486 And E490

QK - Obligate heterodimer form of the Fn with amino acids Q486 And E490K

RR - Obligate heterodimer form of the Fn with amino acids D483R And R487

SDSA – Synthesis dependent strand annealing

SSA – Single strand annealing

TRAP – Telomerase repeat amplification protocol

wtFn – Wild type form of the FokI nuclease domain

y – Drosophila *yellow* gene

ZFN – Zinc finger nuclease

ZFP – Zinc finger DNA binding domain protein

ZFTF – Zinc finger transcription factor

CHAPTER 1: INTRODUCTION AND REVIEW OF THE LITERATURE

1.1 INTRODUCTION

The simplest linear progression of the Central Dogma describes the transcription of DNA sequence into RNA sequence before the translation of genetic information into protein synthesis (Crick 1970). To extend the Central Dogma into a model for the genetic basis of disease, the synthesized protein goes on further to produce cellular changes that lead to disruption of tissue function resulting in human disease. As a therapeutic strategy, gene therapy is typically designed to address mutations within a gene that result in altered protein activity (Porteus 2006). Sometimes, diseases do not arise from mutations in a transcribed region, but instead disease comes from the inappropriate expression of wild-type genes (notably cancer) (Hanahan and Weinberg 2000) (Figure 1.1). Therefore, the repertoire of gene therapies for regaining control over the genetic basis of disease will need to include other strategies to address more than one aspect of the Central Dogma.

One such strategy would be to achieve genetic control by editing the sequence or nucleotide content of the target locus. Alterations would include single point mutations, transgene insertions, or mutagenesis to change the content

Figure 1.1

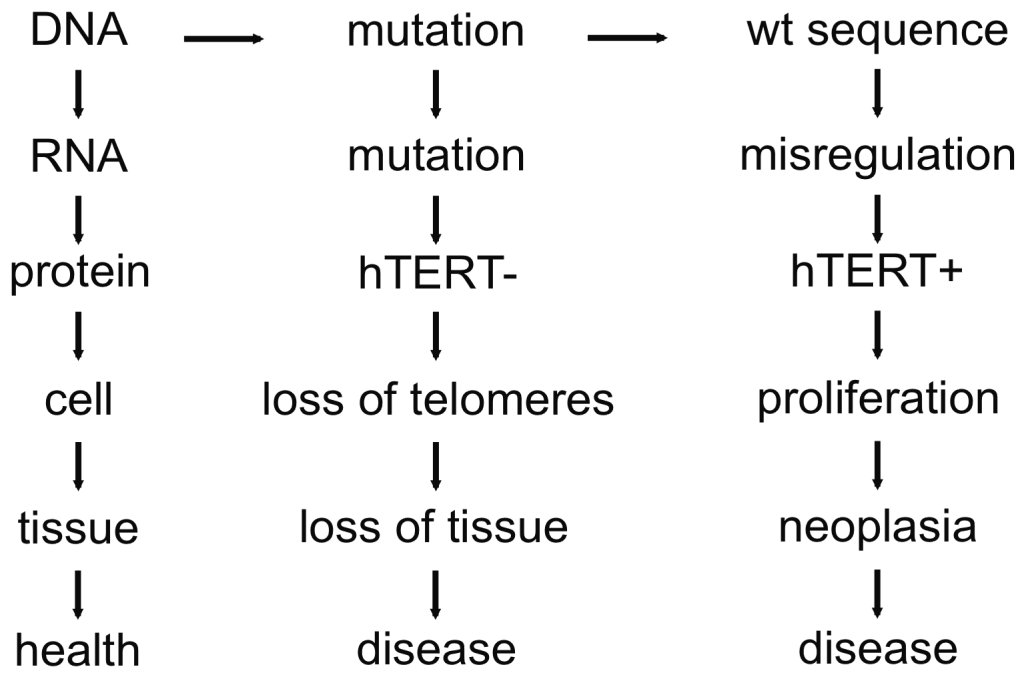

Figure 1.1: The Genetic Basis of hTERT-Associated Disease.

Mutations in hTERT and the misregulation of expression of wild type hTERT sequences can result in pathological disease as diagrammed in parallel with the Central Dogma.

of the sequence that would alleviate the disease state. Other strategies to direct regulation of transcription by exogenous adjustable factors could be used to control inappropriate expression (Figure 1.2). Emerging technologies that use the ZFP are quickly becoming the solution to implementing the types of strategies described above. By linking a ZFP to an effector domain, a new transcription factor can be developed to transcriptionally regulate the desired target sequences (Sera 2009). Linking a nuclease domain to a ZFP will create a method to site-specifically direct a DNA double-stranded break (DSB) to stimulate DNA repair pathways that can be recruited to alter the nucleotide content of a genomic locus (Durai et al. 2005).

The human telomerase reverse transcriptase gene (*hTERT*) is one such gene where mutations can cause inheritable diseases (e.g. dyskeratosis congenita) and misregulation of expression can enhance the disease progression (e.g. neoplasias) (Armanios M. et al. 2005, Hanahan and Weinberg 2000) (Figure 1.1). In this work, I will describe the design and development of ZFPs for use as ZFTFs and ZFNs to target sequences found in the promoter and exon 1 of the *hTERT* locus (Figure 1.2). The data presented in this dissertation will provide useful guidelines for future research into these emerging technologies that could not only answer basic science questions about *hTERT* regulation but also address telomerase-associated disease.

Figure 1.2

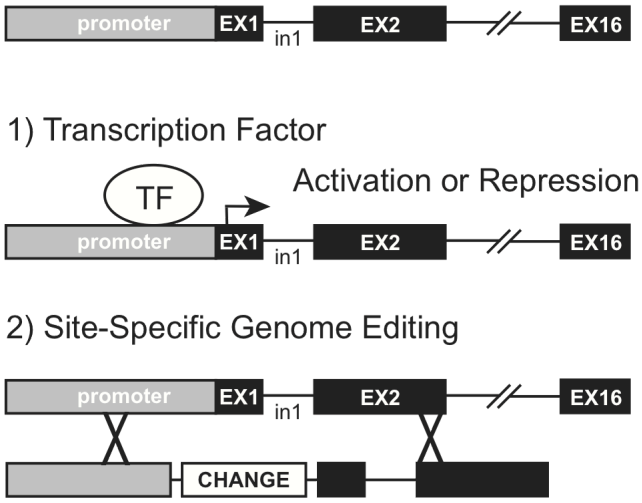

Figure 1.2: Two Strategies For Gaining Genetic Control of the hTERT promoter.

1) Engineering novel transcription factors to regulate hTERT expression. 2) Editing the endogenous locus in a site-specific manner to introduce new regulatory elements.

1.2 ZINC FINGER DNA BINDING DOMAIN

The zinc finger DNA binding domain is a well-conserved protein fold that has achieved rapid expansion with the evolution of increasingly complex species through gene duplication and it is predicted that three percent of genes in the human genome encode for this type of functional domain (Bateman et al. 2002). The history of the zinc finger DNA binding domains used in this work began with the characterization of the *Xenopus* transcription factor TFIIIA. Work done by several groups in the 1980s had lead to the discoveries that the TFIIIA protein recognized a 50 basepair (bp) sequence, contained zinc coordinated by cysteine and histidine pairs, and consisted of repeated sequences of amino acids which led to several hypotheses regarding the modular nature of the putative "zinc finger" that could mediate nucleotide recognition (Hanas et al. 1983, Miller J. et al. 1985, Pelham and Brown 1980, Picard and Wegnez 1979).

The exact nature of a ZFP was revealed by the resolved structure of the human transcription factor Zif268 as published by the laboratory of Dr. Carl Pabo (Pavletich and Pabo 1991). The Zif268 structure revealed a simple motif composed of a common $\beta\beta\alpha$ folding structure stabilized by the chelation of a zinc atom between coordinated pairs of cysteine and histidine residues (known as the C_2H_2 motif). The initial structure data suggested that zinc fingers could be independent and modular in nature as well, but further resolution of the Zif268 DNA binding domain found that context-dependent binding subtly governs the

overall affinity of a ZFP. In short, the binding strength of an individual zinc finger binding affected neighboring fingers around it (Elrod-Erickson et al. 1996). The amino acids composing the α-helix of the zinc finger motif probes into the major groove of the target DNA, while the displayed amino acid residues participate in hydrogen bonding with the target nucleotides. Every zinc finger motif recognizes a nucleotide triplet and the amino acids in positions -1, 1, 3, and 6 of the α-helix make the major recognition contacts while the amino acid in position 2 makes a minor contact (Pavletich and Pabo 1991). By studying the crystal structure of the Zif268 and using mutational analysis, it was found that an arginine in position 6 to recognize a guanine could make a significant contribution to binding and specificity through the formation of two hydrogen bonds. Thus, when selecting target sites and developing ZFPs, GNN recognition triplets are favored (Elrod-Erickson et al. 1998) (Figure 1.3).

Another important architectural element in the ZFP is the inter-finger linker that joins the individual zinc fingers together. This linker is defined as those amino acids (aa) that occupy the positions between the singular zinc finger α-helices immediately proceeding the histidine pair, but preceding the cysteine pair of the C_2 H_2 motif (...FingerA-HXXXH- (linker)-FQCXXXC...). This inter-finger linker is commonly 5 aa in length and contains the highly conserved TG(E/Q)KP consensus sequence. The 5-aa inter-finger linkers allow the ZFP to conform to the helical nature of DNA (Moore et al. 2001a). In some cases longer inter-finger

Figure 1.3

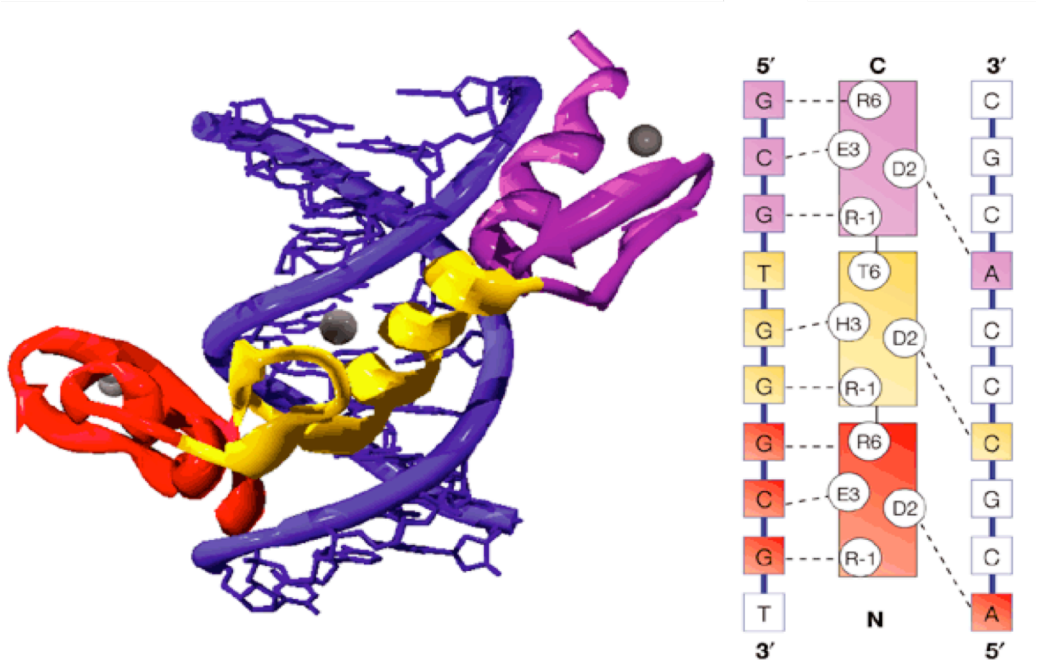

Figure 1.3: Structure of a Zinc Finger DNA Binding Domain and Schematic of Nucleotide Recognition.

A depiction of a 3-fingered zinc finger DNA binding domain as bound to a 9 bp recognition site. The amino acides mediating nucleotide recognition are schematized. Figure adapted from (Jamieson et al. 2003).

linkers are necessary to overcome periodicity in the target site helix or adapt to discontinuous recognition sequences. Moore et al. (2001) demonstrated that ZFPs with extended inter-finger linkers between 2-finger arrays in a 6-fingered ZFP could bind target sites with 1-bp insertions between the 2-finger array subsites with high affinity (Moore et al. 2001b). Work done by the Sugiura group has repeatedly highlighted the role of the inter-finger linker in the mode of DNA binding and target site selectivity (Imanishi et al. 2005, Nomura and Sugiura 2003, Yan et al. 2007). Therefore, the selection of amino acid length and content of the inter-finger linker is an important consideration in designing ZFPs to a desired target sequence.

1.3 ENGINEERING ZINC FINGER DNA BINDING DOMAINS

In the studies to be described in later chapters, 3-fingered ZFP platforms are used and these ZFPs recognize a 9 bp recognition site which are generally developed for 5'-GNNGNNGNN-3' sequence motifs (Elrod-Erickson et al. 1998). After the selection of the target site, work can begin on engineering the framework ZFP to recognize the nucleotide sequence. There are currently two popular methods by which protein engineers may develop ZFP domains: modular assembly, and Oligomerized Pool ENgineering (OPEN) protocols (Figure 1.4).

Figure 1.4

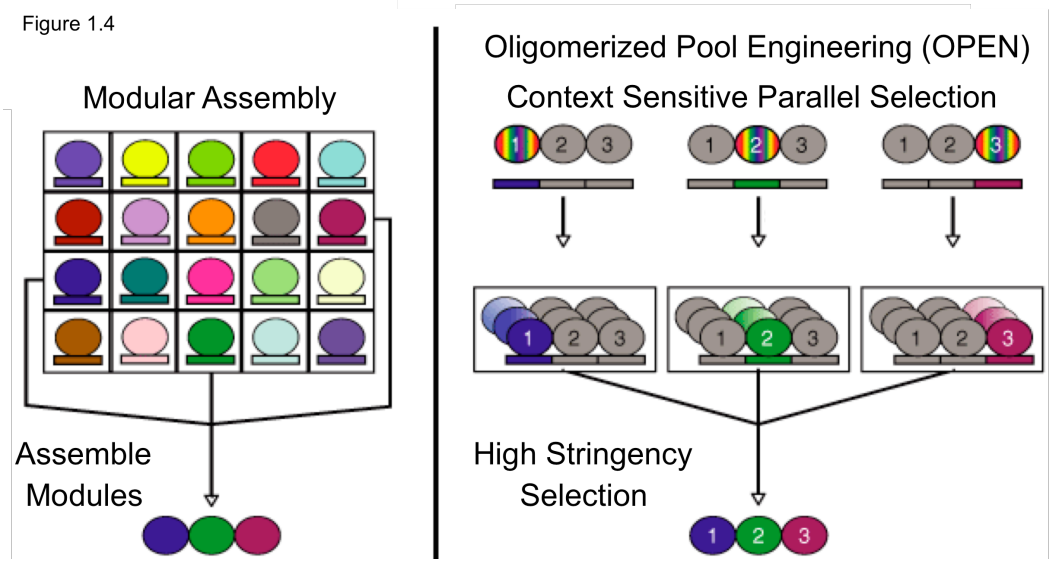

Figure 1.4: Comparison of Modular Assembly to Oligomerized Pool Engineering (OPEN).

Modular Assembly is a method of building ZFPs by consulting a list of individual zinc fingers derived by phage display for each recognition triplet and fusing individual fingers together as needed for the desired target site. This method rapidly produces novel ZFPs, but without the consideration of context-dependent binding. The OPEN platform utilizes master randomized libraries at each finger position and through two stages of B2H selections will produce ZFPs developed with context-dependent binding. Figure adapted from (Pruett-Miller et al. 2008).

1.3.1 Modular Assembly

Modular assembly is a method in which individual zinc fingers that are known to recognize a particular 3-bp target subsites (recognition triplets) are simply joined together to assemble a full ZFP (Kim et al. 2010). These individual zinc fingers were largely developed by work done with phage display. Recombinant phage displaying zinc fingers or ZFP domains on the capsid are selected for by binding column-bound oligonucleotides with increasingly stringent conditions (Rebar and Pabo 1994). Liu et al. (2002) first reported a table of recommended amino acid sequences (empirically derived from phage display) for the α–helix in the zinc finger position to recognize a GNN triplet within the 9-bp target half-site (Liu Q. et al. 2002). The Barbas laboratory and others have also contributed to the modular assembly methodology by producing similar tables for most GNN, ANN, CNN, and TNN triplets (Bae et al. 2003, Beerli and Barbas 2002, Dreier et al. 2005). Specifically designed oligonucleotides encoding for these recommended amino acids creates modules used in a fusion PCR strategy will assemble a full, 3-fingered ZFP in modular fashion (Gonzalez et al. 2010). This method allows for the engineering of a ZFP to a broader range of target sites, but the diversity of ZFPs to any single target site is limited due to few module fingers available for any given recognition triplet. Unfortunately, the convenience of modularity from this method does not take context-dependent binding into consideration. It is hypothesized that this is the reason why the majority of ZFPs

made by modular assembly fail to produce biological activity, although several modular assembly-based ZFPs have been demonstrated to have efficient activities in a variety of chimeric protein forms and model systems (Kim et al. 2010, Ramirez et al. 2008). Thus, the advantages of using this method continue to be debated. Currently all published module fingers that recognize 16 GNN, 15 ANN, 15 CNN, and 4 TNN triplets are available as a kit from Addgene.

1.3.2 Oligomerized Pool Engineering

The other popular method for developing the ZFP domains that bind to novel target sites requires two-stage bacterial-2-hybrid (B2H) selection assays used in the OPEN protocols (Hurt et al. 2003, Maeder et al. 2009, Maeder et al. 2008). Libraries that randomize the amino acids in the recognition α-helix are built within a 3-fingered Zif268 scaffold in which the specific binding of a randomized library ZFP to the cognate target site promotes survival of bacteria under very stringent conditions. This system selects for strong, context-dependent binding in all three zinc fingers and produces several candidate ZFPs for further use at the target site (as opposed to modular assembly methods which do not) (Hurt et al. 2003). Despite the laborious process of generating ZFPs by B2H methods, the advantage to using this system appears to be the higher rates of success in achieving biological activity with such ZFPs (Maeder et al. 2008). However, the OPEN protocols cannot provide as much target site triplet coverage as modular assembly. Currently there are 16 GNN and 7 TNN triplets available

for the Finger 1 (F1) position, 16 GNN and 5 TNN triplets available for the Finger 2 (F2) position, and 16 GNN and 6 TNN triplets available for the Finger 3 (F3) position in the OPEN finger libraries (also available from Addgene).

After the construction of a novel ZFP, options for choosing the amino acids that govern nucleotide recognition can further be explored using "recognition codes" (Wolfe et al. 1999). This method of rational design to site-specifically select amino acids at crucial positions is based on data from resolved crystal structure, phage display, and known interactions between nucleotides and amino acids. In this manner, DNA-binding properties of a ZFP may be improved by an informed amino acid choice at the -1, 2, 3, or 6 positions in the α-helix. Other methods to engineer ZFPs include phage display, ribosomal display, yeast-1-hybrid assays, bacterial-1-hybrid assays, and *in vitro* compartmentalization (Durai et al. 2006, Ihara et al. 2006, Rebar and Pabo 1994, Sepp and Choo 2005).

1.3.3 Summary

Taken together, each method of engineering a ZFP for use at a particular target site will have advantages and disadvantages. With natural domains, a characterized ZFP can be used for a particular target site, but the available natural domains are limited to those previously discovered and characterized. Mutational analysis and rational design can also be undertaken to attempt expanding the target site selectivity of any ZFP, but these efforts are generally inefficient (Wolfe et al. 1999). Assembling ZFPs with modular fingers allow for expedient

construction of a novel ZFP from zinc fingers demonstrated to have affinity and specificity for the chosen target triplet (Gonzalez et al. 2010). However, modular assembly-based ZFPs have shown to have high failure rates (Ramirez et al. 2008). Using the OPEN protocols to engineer B2H-based ZFPs is a labor-intensive process, but can more consistently produce usable ZFPs to a limited set of target triplets (Maeder et al. 2009). Therefore, the method by which a ZFP is developed for a desired target site should be chosen with serious consideration (Pruett-Miller et al. 2008).

1.4 ZINC FINGER TRANSCRIPTION FACTORS

1.4.1 Effector Domains

After the selection of a target site and the development of a ZFP to that target site, a functional domain can then be linked to the ZFP to create a biologically active artificial protein with site-specific activity. In the case of ZFTFs, an effector domain fused to a ZFP builds a novel protein with the potential to regulate gene expression. These effector domains recruit the necessary cellular machinery to the cognate target site as directed by the ZFP in order to transcriptionally activate or repress the target gene (Sera 2009). The most commonly used transcriptional activation domain is the VP16 domain from the herpes simplex virus (Triezenberg et al. 1988). Other activating effectors include domains derived from p65, heat shock proteins, and β-catenin (Pollock et al.

2000, Ruben et al. 1991, Tachikawa et al. 2004). The Kruppel-associated box (KRAB) domain is the most effective and frequently chosen repressor for fusion with ZFP (Urrutia 2003). In addition to these relatively simple transcriptional effector domains, small molecule- and ligand-binding domains have also been fused to ZFPs to act as gene switches. The use of these ligand binding domains such as nuclear hormone receptors have produced novel transcription factors that can modulate gene expression only during the time in which the small molecule is present, providing greater precision in directing the target gene response (Dent et al. 2007, Magnenat et al. 2008, Snowden et al. 2003, Xu et al. 2001). Due to the use of the VP16 effector domain in making ZFTFs described in later chapters, I will narrow the scope of this literature summary to the history of VP16-based ZFTFs in mammalian cells (Figure 1.5). Broader accounts of ZFTFs with other effector domains have been reviewed elsewhere (Klug 2010a, b, Sera 2009).

Figure 1.5

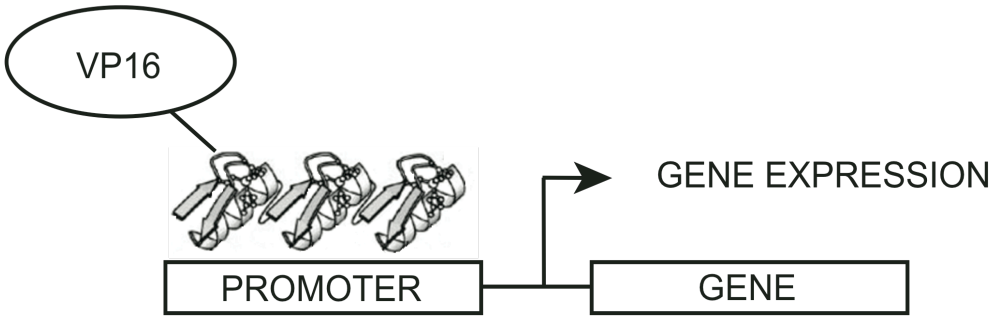

Figure 1.5: VP16-Based Artificial Zinc Finger Transcription Factor.

The fusion of the VP16 transcriptional activation domain to a ZFP results in a novel transcription factor that can up-regulate gene expression.

1.4.2 VP16-Based Zinc Finger Transcription Factors

The first report of a VP16-based ZFTF came in 1994 when Choo et al. (1994) developed a 3-fingered ZFP by phage display to target the *BCR-ABL* fusion point (Choo et al. 1994). They fused the ZFP to the VP16 effector domain in a transient reporter assay to demonstrate target site specificity. The ZFP was then used alone to block transcription of the *BCR-ABL* gene. In a similar fashion, Bartsevich and Juliano in 2000 published a VP16-based ZFTF that could regulate the expression of the human multi-drug resistance gene promoter (*hMDR1*), which transports drugs out of many types of cancer cells (Bartsevich and Juliano 2000). A 5-fingered ZFP was engineered by a yeast-1-hybrid strategy, and in transcriptional activator form the ZFTF could elicit a strong response from a luciferase reporter. However, the ZFP was converted into a repressor by fusion with the KRAB domain, which could reduce endogenous *hMDR1* expression by 95%. In the same year, work done by the Wolffe group of Sangamo Biosciences (Sangamo) led to the development of 3-fingered ZFTFs by their own proprietary methods (now available from Sigma-Aldrich) to the human erythropoietin (*hEPO*) promoter and they were able to achieve up-regulation of *hEPO* transcription at the endogenous locus (Zhang et al. 2000). They also discovered that the binding of these VP16-based artificial transcription factors to their target site in the *hEPO* promoter altered the chromatin state by inducing a remodeling of the region into a more open status. The Wolffe group has also constructed 3-

fingered ZFPs to the human vascular endothelial growth factor A (*hVEGF-A*) promoter before linking them to the VP16 effector domain. These ZFTFs could stimulate expression of *hVEGF-A* by ~10 fold (Liu P. Q. et al. 2001). This study was the first demonstration that multiple ZFTFs could be used in combination at the same promoter to achieve an additive effect on the expression of the target gene. In 2002, Sangamo also reported developing 3- and 6-fingered VP16-based ZFTFs to the mouse *VEGF-A* promoter sequences (Rebar et al. 2002). As a part of an *in vivo* mouse model, a virally transduced mVEGF-A ZFTF was able to induce *mVEGF-A* expression by almost three fold under hypoxic conditions. Then in 2003, the Juliano group built another set of 5-fingered ZFTFs by yeast-1-hybrid selected ZFPs linked to the VP16 effector domain for the human BCL2-associated protein X (*hbax*) promoter in order to re-upregulate hBAX expression in neoplasmic cells that are p53 deficient and reinstate the apoptotic pathways (Falke et al. 2003). Unfortunately, the efficacy of this strategy was ambiguous due to the ZFTFs reducing the cell viability of in the Saos-2 cell line but not U2OS cells, both of which were derived from osteogenic sarcomas. Recently, VP16-based ZFTFs have taken a leap forward as an emerging technology for treating disease. The Passananti and Corbi laboratories have used modular assembly to construct a 4-fingered ZFTF to the human utrophin promoter, an alternative protein to replace dystrophin in Duchenne muscular dystrophy (Di Certo et al. 2010, Mattei et al. 2007). These groups generated a transgenic mouse line that expressed the

utrophin ZFTF under a muscle-specific promoter before crossing into a dystrophin deficient (*mdx*) mouse, an animal model for muscular dystrophy. The double transgenic mouse exhibited improved muscle function and ameliorated the physiopathology of the muscle fibers normally seen in the *mdx* phenotype. This work provides the valuable proof-of-concept evidence that artificial ZFTFs may be useful in the future for therapeutic applications.

1.4.3 Summary and Future Directions

In summary, the use of VP16-based ZFTFs has significantly progressed over the past decade. The progress from simple transient reporters to animal models has been rapid and some ZFTFs have been incorporated into drug development strategies. Work done by Sangamo has led to the creation of 3-fingered ZFTFs to the regulatory sequences of the human cholecystokinin 2 receptor (*hCCKR2*) and human parathyroid hormone receptor 1 (*hPTHR1*) genes for the purpose of inducing expression of the protein of interest during small molecule library screens (Liu P. Q. et al. 2004, Liu P. Q. et al. 2005). Taken together, these studies have provided evidence of the efficacy of using VP16-based ZFTFs to stimulate gene expression, the ability to remodel "closed" chromatin, and potential for synergy when multiple ZFTFs are used in the same system. The versatility of these artificial transcription factors and continual improvements made in methodologies has made them an attractive option for addressing the genetic foundation of human disease.

1.5 ZINC FINGER NUCLEASES

1.5.1 Nuclease Domain of the FokI Endonuclease

ZFNs differ from ZFTFs in that instead of a VP16 effector domain, the ZFPs are fused to the nuclease domain of the FokI endonuclease (Fn) (Figure 1.6). The FokI endonuclease is a DNA cleaving enzyme derived from the prokaryote *F. okeankoites* and contains two functional domains. The N-terminus of the FokI protein contains a DNA binding domain that recognizes a 4-bp recognition site and the C-terminal end contains the nuclease domain (Wah et al. 1997). The actual site of the DSB created by FokI is 9 and 13 bp away from the recognition site regardless of the sequence in that location. This endonuclease achieves DNA cleavage in a bi-modal fashion where the DNA binding domain first binds to the target site before initiating enzymatic activity. In order to cleave DNA, the Fn must dimerize with another Fn, however, the dimerization event can occur without a DNA binding domain although with much less efficiency (Bitinaite et al. 1998, Hirsch et al. 1997, Wah et al. 1998). The dimerization event is mediated through the α-4 and α-5 helices of the Fn and mutations in these helices can inhibit nuclease activity. From biochemical assays and resolved crystal structure, it appears that the presence of the second FokI protein in the dimer (as bound to the DNA helix) makes significant contributions to the formation and stability of the Fn dimer/DNA complex (Bitinaite et al. 1998). Therefore, ZFNs are developed in pairs to allow the Fn domains to dimerize and

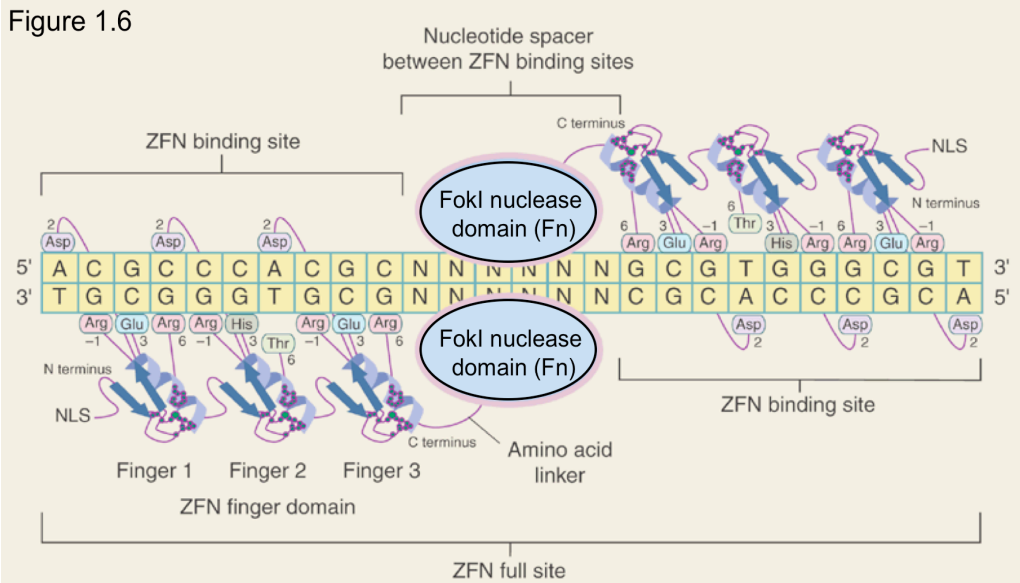

Figure 1.6: Schematic of A Pair of Artificial Zinc Finger Nucleases as Bound to Target Sequence.

The fusion of the FokI nuclease domain to a ZFP results in a novel nuclease that can be site-specifically engineered to deliver DSB to a desired target site. Figure adapted from Porteus and Carroll (2006).

cleave DNA in a site-specific manner. In recent years, obligate heterodimer forms of the Fn (obhetFn) domain have been designed to improve upon the specificity of Fn dimerization and reduce off-target cutting in ZFNs. When wild-type Fns (wtFns) homodimerize, hydrogen bonding occurs between D483 and R487 (DR), and Q486 and E490 (QE) on the α-4 helices of the dimer, as well as hydrophobic interactions between I499 and I538 between the α-4 and α-5 helices respectively (Bitinaite et al. 1998, Szczepek et al. 2007, Wah et al. 1998). The obhetFns studied in Chapter 2 use two double-mutant forms of the Fn in which one carries the E490K:I538K (KK) mutations and the other carries the Q486E:I499L mutations (EL) (Miller J. C. et al. 2007). In this rationally designed system, KK Fns may only dimerize with EL Fns. Electrostatic repulsion prevents KK/KK or EL/EL pairings or any wild type dimerization events. This has been shown to reduce toxicity of ZFN presence within a cell due to extraneous DSBs from off-target cutting, but the data is conflicting in regards to whether or not obhetFns improve ZFN activity (Pruett-Miller et al. 2008). Other forms of obhetFns have also been made. The wild-type DR form has been modified to create DD (D483:R487D) and RR (D483R:R487) heterodimers, as well as the QE form into EE (Q486E:E490) and QK (Q486:E490K) obhetFns (Szczepek et al. 2007). Interestingly, the obligate heterodimer strategy has also been extended to engineering of other meganucleases in order to expand the repertoire to

recognition sites beyond palindromic or quasi-palindromic sequences (Fajardo-Sanchez et al. 2008).

1.5.2 ZFN Target Sites

To meet the need for dimerization between the Fns of two ZFNs, the target site must be chosen to accommodate the necessary tail-to-tail orientation of the ZFNs and provide the physical space necessary for the Fn dimerization event (Porteus and Carroll 2005). Therefore, in a 3-fingered ZFN platform, the 9 bp recognition sites for the ZFPs used must be arranged in the general structure: 5'-(ZFN target site 1)-(spacer)-(ZFN target site 2)-3' where the spacer is a short stretch of nucleotides of any sequence to be the site of the DSB delivered by the dimerized Fns. Given the preferred 5'-GNNGNNGNN-3' sequence for a 3-fingered ZFP recognition site, the canonical full ZFN target site becomes 5'-NNCNNCNNC-(spacer)-GNNGNNGNN-3' in which each of the ZFP recognition sequences becomes a ZFN target half-site. This motif allows the proper tail-to-tail orientation for Fn dimerization and is long enough to be statistically unique sequence in the human genome.

1.5.3 Inter-Domain Linkers

The length of the spacer sequence that can be tolerated for efficient ZFN activity has been shown to be variable (Bibikova et al. 2001). The key to accommodating different spacer lengths is the length and amino acid content of

the inter-domain linker of the ZFNs used. This short amino acid (aa) linker connects the ZFP to the Fn domain after the terminal histidine pair of Finger 3 and but before the first amino acid of the Fn. The Carroll group was the first to broadly survey ZFNs with inter-domain linkers between 2-23 aa on extra-chromosomal reporter plasmids carrying the cognate target sites with spacer lengths of 4-20 bp in between the target half-sites in *Xenopus* oocytes (Bibikova et al. 2001). They found the highest ZFN activity to be in those with 5-aa linkers at target sites with spacer lengths of 6 bp, but also observed low activity at 5-, 7-, and 8-bp spacers. This result was further supported by the findings of Porteus and Baltimore (2003) that demonstrated the most efficient nuclease activity in human cells was seen in ZFNs with linkers of 5 aa at a stably-integrated reporter construct in which the target site had a 6-bp spacer, but also low activity at an 8-bp spacer (Porteus and Baltimore 2003). These studies led to the establishment of a 6-bp spacer being the canonical length of nucleotides found in a ZFN full site and a 5-aa inter-domain linker became standard in ZFN architecture. However, Sangamo then published a ZFN pair that could target an endogenous human gene with a 4-aa inter-domain linker to a target site spacer length of 5 bp (Urnov et al. 2005). In line with this linker length, the Cathomen laboratory also developed ZFNs with a different 4-aa linker to 6-bp spacer lengths in the target site in human cells at extra-chromosomal and stably integrated reporter constructs (Alwin et al. 2005). Then, Shimzu et al. (2009) reported the construction of ZFNs with a 6-aa

linker and in a limited survey of 4-8-bp spacer lengths in human cells, found the highest nuclease activity to be at a 6-bp spacer with minimal activity seen at spacer lengths of 4, 5, 7, and 8 bp at an extra-chromosomal reporter (Shimizu et al. 2009). Thus, there has been a growing body of evidence suggesting that the length and amino acid content of the inter-domain linker could be rationally designed to accommodate a range of ZFN target site spacer lengths. The Cathomen laboratory has since made another significant contribution to this ZFN engineering issue by completing another broad survey in multiple human cell lines at extra-chromosomal and stably-integrated reporters using ZFNs with inter-domain linkers of 0-20 aa to targets sites with spacer lengths of 4-18 bp (Handel et al. 2009). They found that several different linkers contributed to efficient nuclease activity at target sites with 5-, 6-, 7-, or 16-bp spacer lengths. Therefore, there appears to be a range of spacer lengths in target sites of 5-7 bp that can be effectively cleaved by ZFNs provided that a suitable inter-domain linker is installed in the intended ZFN. Should the findings of the Handel et al. (2009) human cell survey be recapitulated, then the repertoire of sites targetable by ZFNs has considerably expanded.

1.6 NUCLEASE-INDUCED DNA REPAIR

1.6.1 Introduction

Zinc finger nucleases are used to deliver DSBs in a site-specific manner in order to stimulate endogenous cellular repair pathways (Porteus and Baltimore 2003). The repair mechanisms can then be exploited to create desired changes in the target nucleotide sequence (Durai et al. 2005). When a ZFN-mediated DSB occurs, one of two main DNA repair pathways mend the damage. Non-homologous end joining (NHEJ) is one pathway that typically results in mutagenic repair. The broken ends of the DNA helix are ligated together wherever a few single-stranded nucleotides anneal (microhomology) and this may result in loss of genetic information (Perez et al. 2008). The other main pathway, homologous recombination (HR), uses a template of homologous sequence to repair the DSB and exogenously supplied repair templates can be supplied in order to incorporate new sequence content into a genomic locus (Urnov et al. 2005). Other DNA repair pathways exist (e.g. single strand annealing (SSA) and alternative NHEJ) but in general are variations of the two main pathways conceptually. Genomic modification by DNA repair is a process nearly as ancient as life itself and has an extensive background in the scientific literature. Therefore, I will discuss the mechanism and DSB-induced NHEJ and HR repair pathways before summarizing the background of ZFN-mediated genomic modification.

1.6.2 Mechanism of Non-Homologous End Joining

In contrast to the intricate choreography of HR, NHEJ pathways work to simply rejoin the broken ends of a DSB. NHEJ first came to light when a DNA repair mechanism beyond HR was suspected of being responsible for viral vector integration events in mammalian cells (Sambrook et al. 1980). In light many completed genome projects, NHEJ is also a highly conserved strategy for DSB repair across all levels of complex life. At the creation of a DSB, a dimerized Ku protein complex forms at the break site and serves as a tether to keep broken ends in close proximity. Then nucleases (Artemis, pol μ, and pol λ proteins in eukaryotes) are recruited to the Ku complex for resection. Specialized ligases (ligase IV) then directly rejoin the resected ends through the formation of microhomologous contacts between nucleotides found in the overhanging strands after resection. Yeast and vertebrate cells evolved to form heterogeneous Ku protein populations that form heterodimers (e.g. the Ku70/Ku80 complex). Other proteins such as regulatory kinases (DNA-PK, PNKP) and scaffold proteins (MRX) were also added for precise regulation of the machinery. Clean breaks that make ligatable overhangs can result in the perfect rejoining of the broken double-stranded helix. Extensively damaged DNA or the joining of different DSBs produce mutagenic couplings. Religating broken ends depends on compatible recognition contacts between a few nucleotides in the resected overhang and that is not sufficient to ensure precise religation that will preserve the original

sequence. Instead, short incompatible flaps of single-stranded nucleotide sequence can form at imprecise microhomologous joints. Nucleases simply trim the incompatible flaps and this results in the loss of genetic information. The mutagenic tendency of NHEJ pathways is exploited for clonal diversity in antibody production, but it also can contribute to chromosomal translocations (Lieber et al. 2006). Further details on the mechanism are specific to the organism employing NHEJ and reviewed elsewhere (Bowater and Doherty 2006, Daley et al. 2005, Lieber et al. 2004, Weterings and van Gent 2004).

1.6.3 Mechanism of Homologous Recombination

The ability to interchange sequences of genomic DNA is a highly conserved strategy used across all forms and manner of life. Inside a cell, the HR mechanism provides the means to site-specifically alter genetic sequences in a contained region while leaving the remaining genome undisturbed. At meiosis, homologous chromosomes synapse in prophase I to trade regions of DNA between the paternally- and maternally-inherited genetic contributions, thereby providing the necessary diversity (Szekvolgyi and Nicolas 2010). Most eukaryotic cells however, spend most, if not all, of the cell cycle in mitotic stages. In the mitotic mode, the standard use of HR is to repair DSBs while preserving the original sequence. In the event of a DNA lesion in a human cell, the broken ends are resected to generate single-stranded 3’ overhangs by the MRN complex (Mre, Rad50, and Nbs1), which are necessary in the search for homologous sequence

(Mimitou and Symington 2009). The 3' overhangs become coated with Rad51 to form nucleo-protein filaments and when the homolog is found, Rad51 will also catalyze the strand invasion into the homologous helix (Sung 1994). The mechanics of helical unwinding, heteroduplex formation, chromatin remodeling, and synapse formation are facilitated by Rad52, Rad54, and members of the XRCC protein family (Brenneman et al. 2002, Heyer et al. 2006, Liu P. Q. et al. 2004, Yamaguchi-Iwai et al. 1998). As a result, the single-stranded 3' overhangs from the resected lesion achieve strand displacement and begin DNA synthesis by pol η (McIlwraith et al. 2005). In the synthesis dependent strand annealing (SDSA) model, the concurrent replication of DNA as primed by the invading 3' overhangs forms Holliday junctions and the resolution of these junctions complete the repair of the resected DSB. Further detail on the HR and the SDSA model can be found in several excellent reviews (Adelman and Boulton 2010, Hinz 2010, Sung and Klein 2006).

1.6.4 Cell Cycle and DNA Repair

The HR and NHEJ pathways co-exist and compete for the repair of DSBs (Hartlerode and Scully 2009). In yeast, HR dominates over random integration (based on NHEJ) as the primary repair pathway (Orr-Weaver et al. 1981). However, vertebrate cells resort to NHEJ for the majority of DSBs routinely encountered and for the process of VDJ recombination. The prevailing hypothesis that explains the differential preference in eukaryotic cells is that choice between

the two pathways for DSB repair depends on the stage of the cell cycle. NHEJ is the predominant choice for DNA repair during the G1 phase of the cell cycle, but both NHEJ and HR are available during the S/G2 phase when sister chromatids are present (Branzei and Foiani 2008). A growing body of work has implicated the role of hCtIP (human CtBP interacting protein) in integrating cell cycle regulation with DNA repair in human cells. hCtIP is phosphorylated by CDK, which facilitates interactions with BRCA1, ATM, and the MRN complex to coordinate the DSB resection required for HR and cell cycle checkpoint activation (Chen et al. 2008, Li et al. 2000, Sartori et al. 2007). Thus, hCtIP provides the bridge across damage sensing, DNA repair, and cell cycle signal transduction pathways and directs the cellular decision between NHEJ and HR.

1.6.5 Nuclease-Induced Homologous Recombination

HR has been very useful experimentally. In the early years of applied molecular biology, it was found that yeast could be transformed and new strains engineered through the integration of exogenously supplied plasmids (Orr-Weaver et al. 1981). Hinnen et al. (1978) discovered that the integrations were largely a consequence of HR between the plasmid and the yeast genome (Hinnen et al. 1978). This study was further elaborated on by investigating the recombination rates between donor plasmids (circular and linearized) with and without homologous sequences flanking the intended sequence for integration, achieving recombination rates of 30% and 70-85% respectively (Orr-Weaver et

al. 1981). Encouraged by the very high rates of targeted gene conversion events vs. the very low random integration events found in yeast models, scientists have generated numerous engineered yeast strains as tools in experiments for eukaryotic investigations.

Unfortunately, site-specific HR occurs at a frustratingly inefficient and low rate in mammalian cells. This mechanism needs stimulation in order to be studied experimentally, or used therapeutically. The previously described studies done initially in the yeast models found that DSBs is the most powerful stimulant for inducing HR. The same was also found to be true in mammalian cells through the pioneering work of Maria Jasin's lab when they translated the use of homing endonucleases from yeast into mouse cells (Rouet et al. 1994a). They first used the I-SceI homing endonuclease to bring target DSBs to an extrachromosomal plasmid carrying its 18 bp recognition sequence inserted into the repeat-containing region of the chloramphenicol acetyltransferase (*CAT*) gene. Co-transfection of this mutated reporter plasmid and a I-SceI expressing plasmid into mouse NIH3T3 cells would lead to a DSB being made at the reporter plasmid by I-SceI at its recognition site and recombination events measured by the returned functionality of the *CAT* reporter gene when corrected by HR. They found that they could produce gene targeting frequencies of $4x10^{-4}$ events/total transfected cells whereas none were detected when a control plasmid was co-transfected. The Jasin lab furthered this paradigm by creating a NIH3T3 cell line with a stably

integrated *neo* gene mutated by the insertion of the I-SceI site (Rouet et al. 1994b). When this cell line was co-transfected with I-SceI expressing plasmids and a linear fragment of the normal *neo* gene sequence, they were able to measure gene targeting rates of 10% by observing the number of clones that gained G418 resistance after HR between the mutated *neo* gene and the linear fragment with the normal sequence. This is in contrast to the complete lack of converted clones after the co-transfection with a control fragment. Houlka et al. (1995) performed similar experiments in which they demonstrated the site-specific insertion of the *neo* resistance gene by HR into an integrated target construct bearing an I-SceI site after cleavage by the I-SceI homing endonuclease. They observed rates of gene targeting ranging from $1.8x10^{-4}$ to $4x10^{-4}$ but no targeted insertion in the absence of I-SceI expression (Houlka et al. 1995). As proof of principle, Gouble et al. (2006) took I-SceI induced HR in mammalian cells one step further by generating a transgenic mouse carrying a *LagoZ* gene repeat interrupted by a I-SceI recognition site. The mice were injected through the tail vein with adenovirus bearing an I-SceI expression cassette. Transduced cells would express the I-SceI homing endonuclease and the *LagoZ* gene function would then be rescued if HR occurred between repeats surrounding the targeted DSB at the cognate recognition site. X-gal staining of sectioned livers revealed that 1.3% of the hepatocyte population experienced recombination events (Gouble et al. 2006). Taken together, these studies indicate that in mammalian cells site-specific HR

can be stimulated by targeted DSBs both in *ex vivo* cell culture models and animal models, but not as vigorously as in yeast. These were, however, the first steps towards proving that using nuclease-induced HR to manipulate the sequence of a genomic target was a viable methodology.

1.6.6 Limited Use of Natural Meganucleases

The widely used I-SceI homing endonuclease belongs to a special class of enzymes called meganucleases. Meganucleases are broadly defined as endonucleases with specific DNA target sequences of 12 bp in length or longer. The membership of this enzyme class has grown extensive with many members being homing endonucleases (Chevalier B. S. and Stoddard 2001). The length of the cognate recognition site makes the sequence statistically infrequent, and thus meganucleases are considered ”rare cutters.” Otherwise, site-specific recombination events would be more difficult to regulate and the cell would be at high-risk for genomic instability. Unfortunately, it has become apparent that the catalog of known meganuclease recognition sequences will not be sufficient to target DSBs to many or all of the genes found in a complex genome. Therefore, in order to stimulate HR for the purpose of genomic sequence manipulation at a unique target site, meganucleases will need extensive engineering to accommodate any empirically chosen, novel target site. Typically, this process begins with a resolved crystal structure of a natural meganuclease. Then subsequent application of site-directed mutagenesis (either randomly or

rationally) at the amino acid positions determined to govern nucleotide sequence recognition produces artificial variants of the natural meganuclease scaffold (Smith et al. 2006). Success in producing effective artificial meganucleases and even artificial restriction enzymes has been greatly hindered by the paucity of resolved crystal structures and some structures reveal a discouraging level of complexity. There has been some success with using structure-based rational design in the I-CreI homing endonuclease and the use of computational methods in redesigning the I-MsoI protein (Chevalier B. et al. 2003, Sussman et al. 2004). A detailed account of the efforts put into engineering artificial homing endonucleases is reviewed elsewhere (Paques and Duchateau 2007).

1.6.7 ZFN-Induced Homologous Recombination

The use of ZFNs to stimulate HR-mediated gene targeting provides a technical solution to the naturally low rates of HR in mammalian cells and the difficulty in re-designing natural endonucleases. A progressive series of experiments by a number of scientific groups have provided evidence that ZFNs have a strong future in becoming the tool of choice for loci-specific genome manipulation through HR. Chandrasegaran along with Dana Carrol first used ZFNs to promote HR in a eukaryotic system by using ZFNs with natural ZFP domains (Kim et al. 1996, Kim et al. 1997, Smith et al. 1999, Smith et al. 2000). Purified ZFNs were co-injected with a plasmid carrying the cognate target sites into *Xenopus* oocytes. Self-recombination events were measured by changes in

electrophoretic mobility after restriction digest of the rescued plasmids. Their work uncovered recombination rates of 0-95% (Bibikova et al. 2001). Carroll with other members of his lab used a *Drosophila* larvae model that utilized a repair substrate to alter an endogenous genomic target. They developed a pair of ZFNs to a site within the *y* gene (*yellow*) and co-delivered the ZFNs to larvae with a linearized donor DNA carrying 8 kb of *y*-homologous sequence with a point mutation to introduce a novel XhoI restriction site. Recombination success was measured through the germ-line transmission of the recombined *y* locus to the F1 generation. The investigators reported that 14% of treated female larvae and 20% of treated male larvae produced mutant offspring. This landmark paper demonstrated that ZFN-induced HR could alter endogenous genomic sequences and that the permanent changes could engineer novel strains of model organisms. The first step in stimulating corrective HR by ZFNs in mammalian cells came through the work of Porteus and Baltimore (Porteus and Baltimore 2003). A green fluorescent protein (GFP) reporter construct, mutated through the insertion of a full ZFN target site and I-SceI recognition sequence, was stably integrated into HEK293 cells. Notably in these experiments, the ZFNs based on natural ZFP domains could not outperform the positive control, I-SceI. Porteus went on to create designed ZFNs using modular assembly to target half-sites found in the GFP gene as well as the human *β-globin* and the *IL2Rγ* (interleukin 2 receptor γ) genes in conjunction with ZFNs made with the natural Zif268 ZFP (Porteus

2006). By co-transfecting plasmids expressing the necessary pair of ZFNs and donor plasmids to act as a repair template, Porteus was able to efficiently alter or correct stably integrated reporter constructs similar in design to those used in Porteus and Baltimore. More importantly, it was recently demonstrated by Urnov et al. (2005) that using specifically developed chimeric proteins and a repair template; the endogenous sequence of a transcribed region could be altered (Urnov et al. 2005). They targeted the *hIL2Rγ* gene and altered single nucleotides to create new restriction endonuclease sites. By screening clones, Urnov et al. measured their efficiency of sequence changing at 12.2% for one allele and at 2.4% for both in asynchronous cultures. This landmark study has provided the basis for developing ZFNs to endogenous genomic targets responsible for human disease.

1.6.8 Other Forms of ZFN-Mediated Genome Modification

Beyond single-point or small-scale changes ZFN-mediated HR can also be used to insert large transgene insertions into a genomic locus. The Naldini group was the first to demonstrate that large transgene constructs could be site-specifically incorporated into the *IL2Rγ* locus by ZFN-induced HR (Lombardo et al. 2007). In their work, integrase-deficient lentiviral vectors were used as the repair template and successfully integrated a 3.6 kb GFP-expressing reporter construct into the *hIL2Rγ* locus in up to 6% of cells in the human K562 cell line. Orlando et al. (2010) has also reported transgene insertion at the hAAVS1 site

using exogenously supplied plasmid and linear repair templates with short arms (~750 bp) of homology at rates of 2-11% and 0.5-2.2% respectively (Orlando et al. 2010). In addition to HR, ZFN-mediated NHEJ in human cells has been developed to disrupt *hCCR5* gene (human chemokine receptor 5) by inducing error-prone repair within the translated regions. This strategy is currently being applied to primary human cell populations to limit the progression of HIV infection (Holt et al. 2010, Perez et al. 2008). Lee et al used NHEJ stimulated by ZFN-delivered DSBs to create chromosomal deletions at the CCR gene family clustered on chromosome 3 of up to 14.9 Mb (Lee et al. 2010). In summary, site-specific gene targeting, gene disruption, transgene insertion, and many other forms of genomic modification on all scales in many genetic model systems have been made possible by the imaginative use of ZFNs.

1.7 HUMAN TELOMERASE REVERSE TRANSCRIPTASE

1.7.1 hTERT and Telomerase

Cancer, aging, replicative senescence, and degenerative disease are all linked together by telomere status and its maintenance by the enzyme telomerase (Harley 2005). Telomeres lie at the ends of chromosomes and telomerase synthesizes a $(TTAGGG)_n$ hexanucleotide repeat extending for 4-14 kb, in human cells (de Lange et al. 1990). Telomeres function to prevent the ends of a chromosome from being recognized as a DNA break (Blackburn 1991). Every

cell division results in the erosion of the telomeres over time due to the inability of DNA polymerases to complete the last few hundred base pairs of a chromosome's end (Smogorzewska and de Lange 2004). Replicative senescence comes from the DNA damage checkpoint triggered by intolerably short telomere lengths, known as the Hayflick limit, prompting the cell towards apoptosis (Harley 2002).

The ribonucleo-protein holoenzyme of the human telomerase complex has two subunits: telomerase reverse transcriptase (hTERT) and telomerase RNA (hTR) (Feng et al. 1995, Meyerson et al. 1997). The hTERT component is a catalytic protein domain with reverse transcriptase activity and the other, hTR, is a single-stranded RNA component threaded through the active site of the hTERT protein (Lingner et al. 1997). In germ and stem cells, telomere length resists shortening by the continued activation of telomerase where telomere length maintenance enhances proliferation potential (Wright et al. 1996). In normal somatic cells, *hTR* remains actively expressed in all cell types and ages, while the constitutive repression of *hTERT* makes it the rate-limiting component in producing active levels of telomerase (Gunes et al. 2000, Harley 2002). Controlling active levels of telomerase by regulating hTERT depends primarily on controlling *hTERT* expression (Ramakrishnan et al. 1998).

1.7.2 Misregulation of hTERT Expression in Cancer Cells

Judicious control of *hTERT* expression balances genetic integrity, replicative senescence, and cellular aging against genomic instability, proliferative potential, and immortalization events leading to tumorigenesis. Inappropriate regulation of *hTERT* is also a hallmark in human cancer. Up to ~80% of human tumors will reactivate *hTERT* expression to maintain telomere length and promote immortalization (Hanahan and Weinberg 2000). The increase of active telomerase alone in a cell by experimental methods is not inherently tumorigenic. However, in the multiple-hit cancer model, the reactivation of telomerase is still potentially a very important step for a progressing cancer to take. The role of telomerase in the formation of neoplasias occurs in two stages. During "early crisis", the cell has reached its Hayflick limit and the p53- and $p16^{INK4A}$-dependent DNA damage/repair pathways have brought the cell to a senescent state. However, if the cell has managed to acquire mutations in both *p53* and $p16^{INK4A}$, it will continue to divide and function until the second stage, "genetic catastrophe". At this point, the cell has become so genetically unstable that more mutations have accumulated. If one of those mutations reactivates *hTERT* expression, the cell can rescue telomere function, thereby immortalizing the damaged cell (reviewed in Artandi and DePinho 2000). Conversely, Ozturk et al. (2006) has observed that spontaneous repression of hTERT expression in an

immortalized hepatocellular carcinoma-derived cell line will reinstate the senescence arrest program (Ozturk et al. 2006).

1.7.3 hTERT Promoter

The promoter of the *hTERT* gene extends to nearly 4 kb and contains many regulatory elements, both proven and putative. The first ~300 bp upstream of the transcriptional initiation codon comprises the core *hTERT* promoter (Horikawa et al. 1999, Takakura et al. 1999, Wick et al. 1999). This region has DNase sensitivity, suggesting a potentially open chromatin status (Wang and Zhu 2004). AP2 and Sp1 binding sites are the most abundant regulatory elements in the core promoter and they both regulate activation and repression (Kyo et al. 2000, Ma et al. 2003, Wooten and Ogretmen 2005). E2F boxes (E-boxes), when bound by E2F-1, are thought to result in repression of promoter activity (Crowe et al. 2001). However, the E-boxes in the *hTERT* promoter can be an activating or a repressing site, depending on the context of the cell (Horikawa et al. 2002, Won et al. 2004). Should Tax or the Upstream Stimulatory Factor occupy an E-box, *hTERT* promoter activity will decline (Chang et al. 2005, Gabet et al. 2003). If a c-Myc/Max complex controls the E-box, then the *hTERT* promoter will activate hTERT expression (Horikawa and Barrett 2003). Studies on regulatory elements in the *hTERT* promoter have depended on over-expression, transfected reporter plasmids, chromatin immunoprecipitation, and *in vitro* binding assays. These techniques have limited usefulness and may not accurately reflect what is

biologically relevant. Wang et al. (2003) demonstrated that luciferase reporter driven by the *hTERT* promoter sequence would produce activity in hTERT+ and hTERT- cell lines of the same lineage (Wang and Zhu 2003). This may be due to transfected plasmids not being subject to the chromatin status of the endogenous sequence. However, fusions between normal cells (hTERT-) and immortalized (hTERT+) cells result in the repression of *hTERT* expression and the application of a protein synthesis inhibitor (cycloheximide) can reactivate *hTERT* expression regardless of chromatin status (Wang and Zhu 2003, 2004). Their experiments point to labile proteins as a dominant factor in *hTERT* repression. Mutational analysis of the endogenous *hTERT* promoter sequence in a variety of normal and cancerous cell lines will be critical to definitively identifying the factors and sequence elements involved in the complex regulatory mechanisms of *hTERT* expression.

In addition to regulation by transcription factors, the *hTERT* promoter is subject to chromatin modifications. The core promoter sequences are within a large CpG island centered on the translational start codon (Wick et al. 1999). As with many genes, hypomethylation of the *hTERT* locus is correlated with transcriptional activity and hypermethylation with transcriptional repression (Shin et al. 2003). Zinn et al. (2007) demonstrated the *hTERT* promoter responds to demethylating agents with the re-upregulation of *hTERT* expression in hTERT- cell lines. Their analysis of the methylation patterns of the core promoter CpG

island revealed allelic heterogeneity in which only a few cells within hTERT+ cell lines can actually express *hTERT* (Zinn et al. 2007). Also, the heterochromatin status of the *hTERT* locus is a strong indicator of transcriptional activity. Several groups have made the observation that regulatory proteins binding the E-boxes and Sp1 sequences of the *hTERT* promoter (e.g. Sp1, Mad, and c-Myc) are mediating the increased *hTERT* expression after treatment with histone deacetylase inhibitors (Cong and Bacchetti 2000, Hou et al. 2002, Matuoka and Chen 2002, Takakura et al. 2001). In normal human somatic cells, the *hTERT* locus contains extensive tracts of heterochromatin. However, in hTERT+ cells, the core promoter becomes DNase hypersensitive, which indicates an "open" chromatin conformation amenable to *hTERT* expression. This suggests that chromatin modeling plays a part in the transcriptional regulation of this locus (Wang and Zhu 2003, 2004).

1.7.4 hTERT and Gene Therapy

Due to the position the telomerase enzyme holds within the mechanisms of cancer and cellular aging, hTERT can become the genetic basis of several diseases as a result of mutations, haploinsufficiencies, or deregulation of expression. Degenerative diseases, such as dyskeratosis congenita, aplastic anemia, idiopathic pulmonary fibrosis, and others have been linked to mutations in *hTERT* and exhibit decreased telomerase activity and shortened telomeres leading to progressive loss of the regenerative cell populations and fast-growing

tissues (Armanios M. et al. 2005, Armanios M. Y. et al. 2007, Tsakiri et al. 2007, Xin et al. 2007). Many types of transformed cells preferentially turn on the *hTERT* promoter to achieve immortalization. Therefore, the *hTERT* locus has become a candidate locus for gene therapy. Currently, in the field of gene therapy, *hTERT* gene and promoter have found many uses. Thus, several research groups have fused the *hTERT* promoter to an apoptosis-inducing gene product. Normal, healthy cells, which suppress *hTERT* expression, are spared. The *hTERT* promoter has also been utilized to drive *hTERT* siRNA in some applications (Shay et al. 2001). In addition, a growing body of work has pointed to the relief of *hTERT* repression producing potentially therapeutic results. Using gene therapy techniques, the upregulation of *hTERT* to reconstitute active telomerase levels in somatic cells has benefits. Fibroblasts, vascular smooth muscle cells, and cancer-specific cytotoxic-T cells are some examples of primary, normal human cells that have improved function, life span, and proliferation potential after *hTERT* transduction or transfection (Bodnar et al. 1998, Funk et al. 2000, Klinger et al. 2006, Murasawa et al. 2002, Verra et al. 2004, Yang et al. 2001). Unfortunately, an *hTERT* transfection is only a temporary solution, but the use of viral vectors for permanent *hTERT* control can incur unavoidable safety risks (Porteus et al. 2006). The development of a gene therapy technology to control *hTERT* expression without an integrative viral vector would be a step forward for the fields of telomerase biology and cancer.

1.8 SUMMARY OF RESEARCH GOALS

Considering the need to further characterize the sequence elements in the *hTERT* promoter and the benefits gained from purposefully controlling *hTERT* expression, the idea to apply ZFP-based chimeric proteins to telomerase biology is compelling. Therefore, I propose to develop VP16-based ZFTFs for the purpose of using them to regulate *hTERT* expression in human cells. Novel transcription factors to this locus could be used to study the regulatory mechanisms of transcriptional control in many types of human cells, normal and transformed. If these ZFTFs prove to be effective *hTERT* regulatory factors, then these ZFTFs could be pursued for therapeutic applications. As proof-of-concept, KRAB-based ZFTFs have been engineered to the *hTERT* promoter and were shown to repress *hTERT* expression leading to significant telomere shortening in HEK293 cells (Sohn et al. 2010).

After constructing the ZFPs, I also propose using them for the creation of ZFNs for nuclease-induced HR to edit the endogenous *hTERT* locus. Developing the means to change promoter sequence at the genomic levels will accomplish several important things. First, it will expand the repertoire of diseases addressed by gene therapy. Promoters can then become direct targets of gene therapy as opposed to an accessory in the methodology. Second, the complexity of promoter sequence and the variety of regulatory elements allows considerable freedom in the choice of repair templates, which improves the likelihood of success. In

addition, using ZFNs to stimulate HR through gene targeting at an untranscribed region will constitute a step forward in the gene therapy field. Lastly, the ability to selectively mutate sequence in the *hTERT* promoter will help settle conflicting *hTERT* expression data in the literature and specifically define the regulatory elements involved in control of *hTERT* transcription. In the following chapters, I will describe two major avenues of research into my proposed use of ZFP-based chimeric proteins.

First, there are several unresolved engineering issues involving ZFP architecture. A growing body of literature suggests that both the inter-finger linker of the ZFP and inter-domain linker of the ZFN make significant contributions to the performance of a ZFP-based chimeric protein based on the parameters of the target site (Bibikova et al. 2001, Handel et al. 2009). However, further exploration is required regarding the appropriate length and amino acid content of these linkers for use in human cells. In addition, in the engineering of ZFPs, due to the limitations on possible target sites presented by the OPEN platform and the high failure rates of modular assembly methods, a new strategy is necessary to not only expand the repertoire of potential target sites, but also generate ZFPs with high functionality (Maeder et al. 2008, Ramirez et al. 2008). Chapter 2 details the efforts made in surveying the activities of multiple linker ZFN variants on different types of target sites and the achievements made in creating a methodology that hybridizes both OPEN and modular assembly protocols.

Second, information generated by exploring design aspects of ZFP architecture was used to develop ZFPs that target sequences at the *hTERT* locus. The ZFPs generated were converted into VP16-based ZFTFs and ZFNs and screening strategies were used to find the most effective DNA binding domains. Chapter 3 provides the report on my findings and the data presented provides the necessary preliminary data for future research into applying these chimeric proteins to address telomerase-associated disease.

CHAPTER 2: EXPLORING THE TARGET SITE STRUCTURE AND ZINC FINGER NUCLEASE ARCHITECTURE FOR EFFICIENT GENE TARGETING IN MAMMALIAN CELLS

2.1 ABSTRACT

Recent studies have shown that ZFNs are powerful reagents in making site-specific genomic modifications. ZFNs have a generic structure of ZFP-(linker)-nuclease domain where the zinc finger DNA binding domain and nuclease domain are separated by an amino acid "linker". ZFNs cut genomic DNA at sites that have a generic structure (ZFNsite1)-(spacer)-(ZFNsite2) where the "spacer" separates the two ZFN binding sites. In this work, we compare the activity of ZFNs with different linkers on target sites with different spacer lengths. We found those ZFNs with linkers lengths of 2 or 4 aa efficiently cut at target sites with 5- or 6-bp spacers, and that those ZFNs with a 5-aa linker length efficiently cut target sites with 6- or 7-bp spacers. In addition, we demonstrate that the OPEN platform used for making 3-fingered ZFPs can be modified to incorporate modular assembly fingers (including those recognizing ANNs, CNNs, and TNNs) and we were able to generate ZFNs that efficiently cut cognate target sites. These findings establish guidelines for the identification of potential ZFN target sites and for the architecture of ZFNs to optimally target such sites.

2.2 INTRODUCTION

ZFNs are efficient tools to make site-specific modifications to genomic targets (Durai et al. 2005, Porteus and Baltimore 2003). These chimeric nucleases are engineered to recognize, bind, and cut specific DNA targets that have the general sequence 5'-(ZFNsite1)-(spacer)-(ZFNsite2)-3'. ZFNs create DSBs that are then repaired by the endogenous cellular repair machinery. If the DSB is repaired by a mutagenic non-homologous end-joining mechanism, then mutations consisting of small insertions and deletions can be created at the site of the specific ZFN induced break (Bibikova et al. 2003, Perez et al. 2008, Santiago et al. 2008). If the DSB is repaired using an exogenously provided template (donor DNA) by homologous recombination, then specific modifications to the genomic target can be created based on the donor sequence (gene targeting) (Porteus and Baltimore 2003, Urnov et al. 2005). Thus, ZFN mediated gene targeting can create small sequence changes or can be used to target transgene integration (Lombardo et al. 2007, Moehle et al. 2007, Porteus 2006).

The architecture of a ZFN contains three general parts: a polydactyl ZFP, a nuclease domain of the FokI restriction endonuclease (Fn), and short intervening amino acid linkers that connect the two domains (inter-domain linker) or the individual fingers in the ZFP (inter-finger linker). In order to cut DNA efficiently, the two Fn domains of a ZFN pair need to dimerize, which will occur when the two ZFNs bind to their cognate binding sites in the proper orientation (Smith et al.

2000). Once dimerized, ZFNs cut the DNA in the spacer region between the two ZFN binding sites (Figure 2.1).

Using a 3-fingered ZFP platform, we define the inter-domain linker to be the stretch of amino acids from the terminal histidine pair of the third zinc finger to the first amino acid of the nuclease domain (...Finger 3-HXXXH-(linker)-QLV...). A growing body of literature suggests that the inter-domain linker of a ZFN can be designed to accommodate a variety of spacer lengths between the two ZFN binding sites. Bibikova et al. (2001) showed that in *Xenopus* oocytes, optimal ZFN cutting was most efficient using a ZFN with a 5-aa inter-domain linker at a target site with a 6-bp spacer length, but inefficient cutting resulted at target sites with spacer lengths of 5, 7, or 8 bps using ZFNs with a wide array of inter-domain linkers (Bibikova et al. 2001). Consistent with Bibikova et al. (2001), Porteus and Baltimore (2003) found that ZFNs were more effective at stimulating gene targeting at sites with 6-bp rather than 8-bp spacer lengths (Porteus and Baltimore 2003). However, Urnov et al. (2005) showed that target sites with a 5-bp spacer could be efficiently targeted by ZFNs with a 4-aa inter-domain linker (Urnov et al. 2005). Alwin et al. (2005) also reported efficient targeting at a 6-bp spacer target site, but with a different 4-aa linker (Alwin et al. 2005). Finally, Handel et al. (2009) surveyed 11 inter-domain linker ZFN variants on targets with spacers ranging from 4-18 bp in mammalian cells and found that different inter-domain linkers could preferentially cleave target sites with 5-, 6-,

FIGURE 2.1

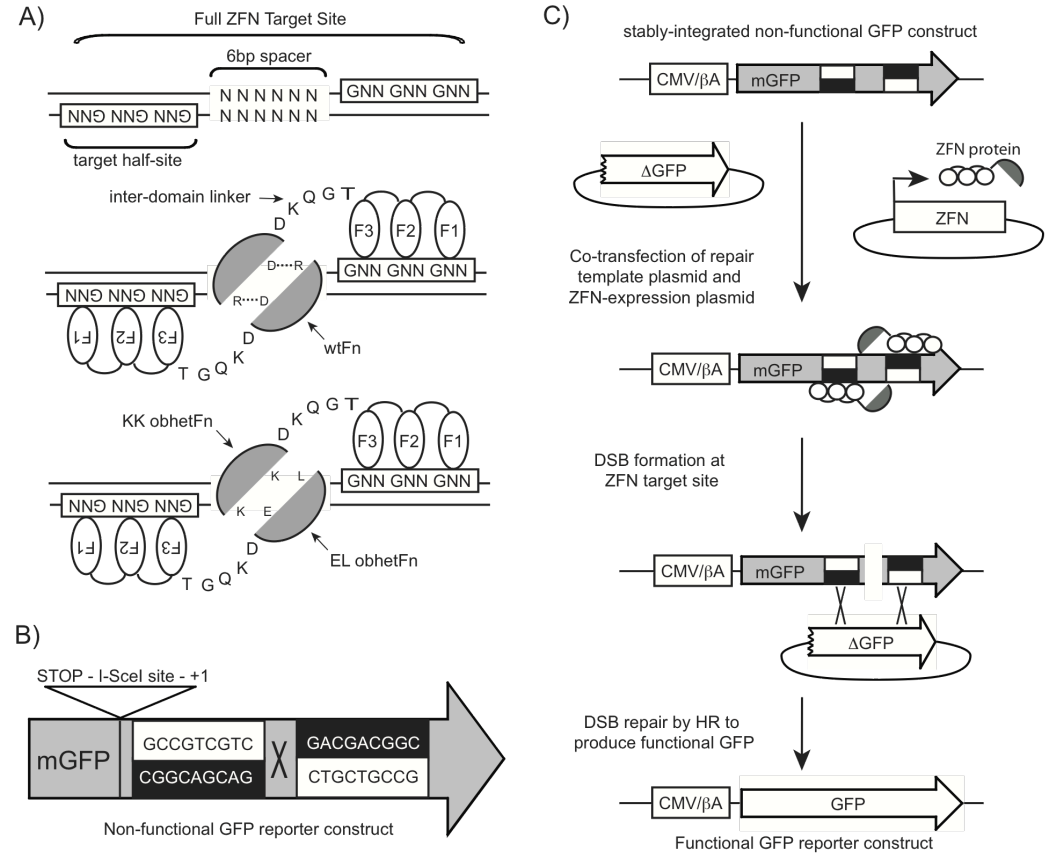

Figure 2.1: Schematic of ZFN Binding, GFP Reporter Constructs, and ZFN-mediated Gene Targeting.

A) Two ZFN target sites on opposing strands are separated by a short intervening sequence labeled as “spacer”. This diagram highlights the use of the TGQKD 5-aa linker that joins the three-fingered zinc finger DNA binding domain to the Fn domain. When a pair of ZFNs binds to their cognate binding sites, the nuclease domains can dimerize and cut DNA. The amino acids that mediate homodimerization between two wtFns by the formation of a salt bridge are depicted (Wah et al. 1998). The nuclease domain in the obligate heterodimer (obhetFn) has been modified to prevent homodimerization and is schematized to show that the KK nuclease can only dimerize with the EL nuclease and visa versa. B) Schematic of the target GFP gene construct mutated by several insertions (stop codon, I-SceI recognition site, and a frameshift) and the doubled GFP-ZFN2 half site (5’- GACGACGGC-3’, black boxes) to make a full ZFN target site where x = spacer length of 3, 4, 5, 6, or 7 bp. C) The mutated GFP reporter construct was used to generate monoclonal reporter cell lines. These cells were co-transfected with ZFN-expression plasmid(s) and a repair template plasmid (donor) bearing a truncated version of the GFP gene. The expressed ZFNs will bind to gene the GFP-ZFN2 target site in pairs to generate a DSB that stimulates homologous recombination between the reporter gene and the repair template plasmid. Repair of the DSB by homologous recombination will produce a functional GFP reporter gene.

7-, or 16-bp spacers (Handel et al. 2009). These studies suggest that the inter-domain linker may be an alterable component for target sites with variations in spacer length.

In addition to the inter-domain linker, we define the conceptually-related inter-finger linker to be those amino acids that occupy the positions between the singular zinc finger α-helices immediately proceeding the histidine pair but preceding the cysteine pair of the C_2 H_2 motif (...FingerA-HXXXH-(linker)-FQCXXXC...). Typically, this linker length is 5 aa, but in some polydactyl ZFP domains, inter-finger linkers may have to be lengthened to allow for conformation to the periodicity of the DNA helix or to allow for extra bps in a ZFP recognition site (Kim and Pabo 1998). Moore et al. (2001) tested 5-aa and 7-aa inter-finger linkers to join the zinc finger pairs in a 6-fingered ZFP for target sites that included 1 bp insertions between the 2-finger array subsites, and demonstrated that ZFPs with the modified inter-finger linker could bind such sites with high affinity (Moore et al. 2001b). In addition, a 6-aa inter-finger linker is used in the 4-fingered ZFNs reported by Urnov et al. (2005) to join pairs of two-finger arrays, and these produce high rates of gene targeting at the cognate target site (Urnov et al. 2005). Therefore, it appears that the inter-finger linker is also a candidate for customization to accommodate insertions within a binding site. However, it remains to be seen whether these studies can be applied to 3-fingered ZFN platforms.

Currently, three publically available methodologies exist for creating 3-fingered ZFNs. First, ZFNs can be made from naturally existing ZFPs to their native target sites. Second, ZFPs can be assembled in a modular fashion by connecting individual fingers of known specificity for the target subsite (Kim et al. 2010). Only 50 of the possible 64 triplets have modular fingers assigned and high failure rates can be expected in these ZFPs (Ramirez et al. 2008). Third, the OPEN method uses B2H selection strategies based on randomly combined libraries of fingers to develop 3-fingered ZFPs and has higher success rates in producing ZFNs with efficient activity. To date, 48 GNN and 18 TNN subsite libraries are available for public use in a 3-fingered ZFP platform (Hurt et al. 2003, Maeder et al. 2009, Maeder et al. 2008). While modular assembly and OPEN methodologies (materials available from Addgene) both allow the versatility of creating a new ZFP for a target sequence, neither can provide adequate coverage of 9-bp target sites not purely comprised of GNN triplets for a 3-fingered ZFP platform.

In this study, we survey a number of modifications in ZFN architecture to target sites that carry variations in the canonical sequence of two 5'-GNNGNNGNN-3' half-sites separated by a 6-bp spacer (Figure 2.1A). To explore a range of spacer lengths from 3-7 bp, we tested the on-target and off-target cutting activities of ZFN variants with four types of inter-domain linkers (2-5 aa) *in vitro* and in mammalian cell-based assays (Figure 2.1B-C). We found that

target sites with 5-, 6-, or 7-bp spacers could be cut by ZFNs with differential efficiencies based on the inter-domain linker used. Also, we sought to determine whether a ZFN based on a 3-fingered ZFP platform could continue to have efficient nuclease activity when the target half-site contains a 1 bp insertion between subsites (5'-GNNNGNNGNN-3' or 5'-GNNGNNNGNN-3'). In lengthening the endogenous 5-aa inter-finger linker to 6 aa between fingers corresponding to the insertion in the target site, we found that despite being able to develop inter-finger ZFN variants that could efficiently generate DSBs, these variants were not sufficiently specific for the intended target site. Finally, to address ZFN target half-sites that include non-GNN triplets, we describe a hybridized methodology that combines modular assembly with the OPEN method to produce functional ZFNs. All together, our work not only further defines the robustness of the 3-fingered ZFN platform, but in doing so, our data provides guidelines on how to modify ZFN architecture to variations in potential target sites.

2.3 RESULTS

2.3.1 Experimental Strategy For Testing GFP-ZFN2 Inter-Domain Linker Variants

We have previously reported a ZFN (GFP-ZFN2) that was designed to recognize a target site 5'-GACGACGGC-3' within the GFP gene (Pruett-Miller et

al. 2008). This ZFN was previously shown to have high activity and lower toxicity in prior work and binds to its target site with high affinity (unpublished data). We constructed a series of ZFNs from the GFP-ZFN2 with different inter-domain linkers: GS (2 aa), LRGS (4 aa), TGQKD (5 aa), and AAARA (5 aa). The LRGS, TGQKD, and AAARA linkers represent actual or modified linkers used in mammalian cells from the published literature (Alwin et al. 2005, Pruett-Miller et al. 2008, Urnov et al. 2005). We made the GS inter-domain linker to explore the effect of a shorter linker. To test these linker variants, we made a series of reporter constructs in which two GFP-ZFN2 binding sites were arranged as inverted repeats separated by spacers of 3, 4, 5, 6, or 7 bp (Figure 2.1B). Finally, we also made modifications of the nuclease domain originally designed to prevent homodimerization and previously demonstrated in literature to reduce toxicity (Miller J. C. et al. 2007, Pruett-Miller et al. 2008, Szczepek et al. 2007). We tested whether these nuclease modifications changed the activity of a ZFN on different spacer constructs.

2.3.2 In Vitro Nuclease Activity of the TGQKD Inter-Domain Linker GFP-ZFN2 Variant Protein

Using an *in vitro* cutting assay, we tested the purified TGQKD 5-aa inter-domain linker variant (unmodified GFP-ZFN2) protein for its ability to cut DNA *in vitro* on a substrate in which the spacer varied from 3-7 bp (Figure 2.2). We measure specific cutting by the ZFN-mediated cutting of a linear 3 kb fragment

FIGURE 2.2

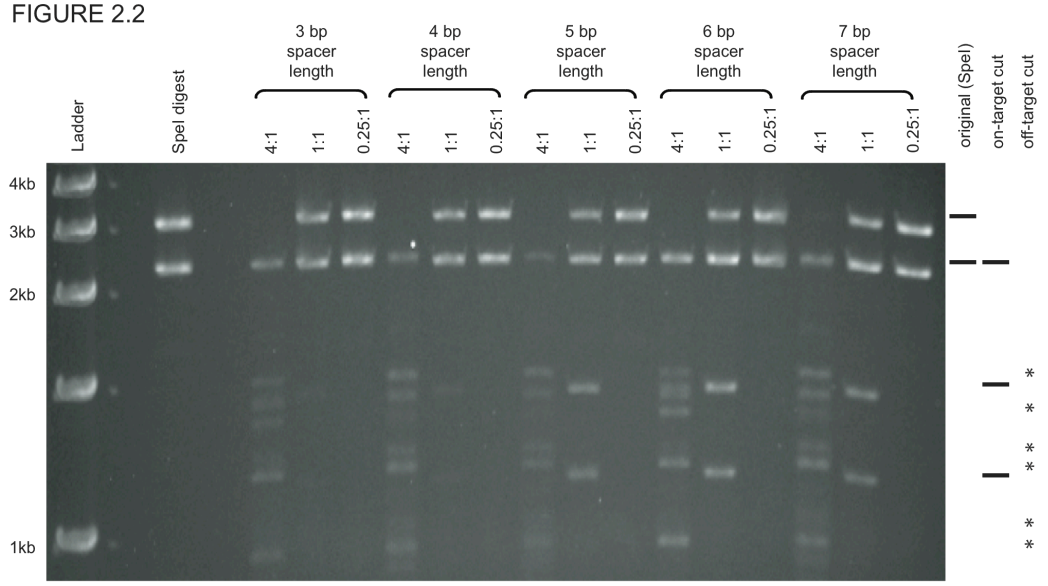

Figure 2.2: ZFN *in vitro* Cutting.

Purified GFP-ZFN2 (TGQKD inter-domain linker, wtFn) protein was added to 200 ng (0.05 picomoles) of the GFP target construct (Figure 1B) linearized by a SpeI digest (spacer length noted) in molar ratios of [(0.25 - 4) : 1] as [protein:DNA]. Specific cutting by the ZFN protein is demonstrated by the digest of the top 3.0 kb fragment into two fragments of ~1.8 kb and ~1.2 kb. Off-target cutting is observed through the degradation in either of the original two bands and appearance of products of different molecular weights. Bars on the right-hand side of the gel mark the locations of the two original SpeI fragments and the specific-cutting fragments, generated by on-target digestion with the ZFN.

into two fragments of ~1.8 kb and ~1.2 kb. We also measured off-target cutting by the disappearance of a 2.4 kb band or by the presence of other bands that are not ~1.8 kb or ~1.2 kb. In our results, we found significant off-target cutting on all the substrates at high concentrations (4:1 ratio of protein to DNA) of ZFN protein. Since the products of this digestion had identical weights, the ZFN cutting appears specific, albeit off-target (Figure 2.2). The off-target cutting *in vitro* is consistent with the off-target DSBs ZFNs make in cells (Alwin et al. 2005, Beumer et al. 2006, Bibikova et al. 2002, Cornu et al. 2008, Miller J. C. et al. 2007, Pruett-Miller et al. 2008). At low concentrations of ZFN protein (0.25:1 ratio of protein to DNA), we found that the ZFNs did not cut any of the substrates efficiently in these conditions. Finally, at intermediate amounts (1:1 ratio of protein to DNA), the ZFN cut the substrates with 5-, 6-, or 7-bp spacers better than they cut the substrates with 3 or 4 bp spacer lengths (Figure 2.2, as shown by brighter specific bands in the 1:1 ratio lane for the 5, 6, and 7 bp spacer constructs). These *in vitro* results, however, did not accurately predict which spacer lengths the nuclease would cut efficiently when integrated into the mammalian genome (see below). For this reason, we did not pursue *in vitro* characterization of the other linker variants. These results suggest that ZFN *in vitro* activity profiles cannot necessarily predict activity in cells. The possibility, however, of establishing *in vitro* conditions with predictive ability about ZFN-mediated target site cutting in mammalian cells remains.

2.3.3 Gene Targeting By GFP-ZFN2 Inter-Domain Linker Variants on Target Sites With Different Spacer Lengths

Since our *in vitro* studies demonstrated that an inter-domain linker variant ZFN had activity at both target and off-target sites, we next tested on-target activity (Figure 2.2). We created a set of stable cell lines that contained an integrated GFP reporter gene with the ZFN target sites separated by 3-7 bp, using GFP reporter constructs identical to those used as cutting substrates in the *in vitro* assays (Figure 2.1B and C). Thus, for each spacer length, we generated a different cell line (five total). The reporter in each of these cell lines contained a recognition site for I-SceI as an internal standard. By using I-SceI as an internal standard, we could compare the relative activities of the ZFN variants across different cell lines and control for local positional effects and chromatin status. Prior work has found that the activity of ZFNs can vary depending on the amount of ZFN plasmid transfected (Cornu et al. 2008, Pruett-Miller et al. 2008, Szczepek et al. 2007). We performed these experiments, therefore, at low (20 ng) and high (100 ng) amounts of ZFN-expressing plasmids.

We found no evidence of targeting in cells when the spacer was 3 or 4 bp in length no matter how much ZFN was transfected or what type of inter-domain linker was used (Figure 2.3A-B, 2.3D-E). In the reporter with the 5-bp spacer, the 2-aa and 4-aa variants gave the best gene targeting activity (Figure 2.3G-H). Figure 2.3H best demonstrates the improved targeting with these shorter linkers.

FIGURE 2.3

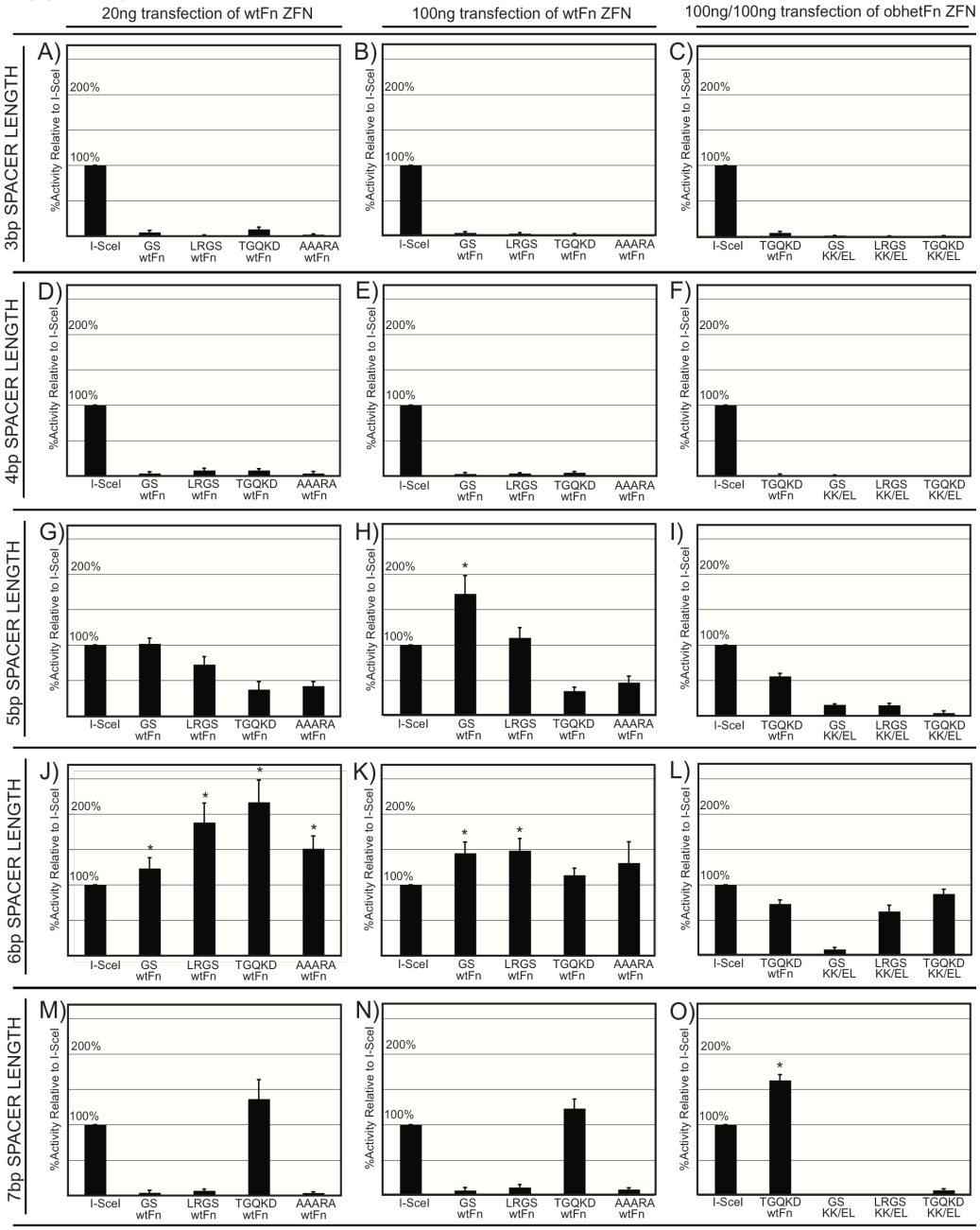

Figure 2.3: Gene Targeting Assays Using Inter-Domain Linker Variant ZFNs.

The data are presented as rates of gene targeting as normalized to a percentage of I-SceI generated events. Since the absolute rate of targeting varies between different cell lines, we use I-SceI as an internal standard, which allows comparison of activity of different ZFNs across different cell lines due to positional effects and chromatin status. Statistical Analysis: Asterisk indicates architectures that are statistically significantly better than the I-SceI positive control (* = $p<0.05$, Student's one-tailed t-test). We consider any ZFN architecture of linker and spacer that gives activity at least as good as the I-SceI standard to be highly functional, however. Experiments done in the same cell line to test the impact of a particular spacer length between the two GFP2 target half-sites are grouped in rows. Abbreviations: wtFn=wild-type nuclease domain; obhetFn=obligate heterodimer nuclease domain; KK/EL=pair of ZFNs with obhetFns.

We found that all inter-domain linker variants stimulated efficient gene targeting using the 6-bp spacer (Figure 2.3J-K). Finally, with the 7-bp target, we found that only the TGQKD 5-aa linker gave efficient targeting (Figure 2.3M-N). The cell based assay results with the TGQKD linker in which efficient targeting was achieved best on spacers with 6 or 7 bp is in contrast with the *in vitro* results where the TGQKD variant cut the 5, 6, or 7 bp constructs equally (Figure 2.2).

2.3.4 Gene Targeting Using Obligate Heterodimer GFP-ZFN2 Inter-Domain Linker Variants on Target Sites With Different Spacer Lengths

Prior studies have shown that ZFN toxicity can be decreased by making modifications to the nuclease domain to prevent homodimerization, called obligate heterodimer variants (obhetFn). In this study, we tested the modifications described in Miller et al. (2007) (Miller J. C. et al. 2007). In these variants, one ZFN contains the following changes E490K:I538K (KK) while the other contains the changes Q486E:I499L (EL), where the numbering reflects the amino acid position in the wtFn. We incorporated these changes into our inter-domain linker variants and tested them for targeting activity using the spacer reporter lines described above (Figure 2.1A-C). We found that just as with the wtFn, the obhetFn variants had no activity on the 3-bp or 4-bp spacer (Figure 2.3C and 2.3F), but also had no activity on the 7-bp spacer (Figure 2.3O). With the 5-bp spacer, the 2-aa and 4-aa inter-domain linker obhetFn pairs gave significantly less activity than the wtFn counterparts (compare Figure 2.3H to 2.3I). With the

TGQKD inter-domain linker, the obhetFn pair showed only 20% of the activity given by the TGQKD variant with the wtFn (Figure 2.3I). Overall, the TGQKD inter-domain linker variant showed the broadest activity (spacer lengths of 5, 6, and 7 bp), which is in contrast to the AAARA inter-domain linker variant which only efficiently targets 5-bp and 6bp spacers (Figure 2.3H, 2.3K, and 2.3N). With the 6-bp spacer, the obhetFn pair gave equal activity (Figure 2.3L) to the wtFn with the TGQKD linker but significantly less activity with the 2-aa and 4-aa inter-domain linker (compare Figure 2.3K with 2.3L).

2.3.5 Cell-Based Assays to Measure Toxicity of GFP-ZFN2 Inter-Domain Linker Variants

To determine if the inter-domain linker GFP-ZFN2 variants had different off-target effects in cells leading to extraneous DSBs that may result in cell death, we used two assays: 1) a cell-based survival assay and 2) a DSB foci formation assay (Pruett-Miller et al. 2008). In the cell-based survival assay, we demonstrated that at low amounts of ZFN transfected, the GFP-ZFN2 inter-domain linker variants did not show appreciable toxicity (Figure 2.4A). At the higher amount of ZFN transfected, we found that the TGQKD variant showed some increased toxicity and the LRGS variant showed the most (compare Figure 2.4A to 2.4B). Results from the foci formation assay show similar results; only the LRGS inter-domain linker variant with the wtFn had significantly ($p<0.05$) more cells with 6 or more foci than the non-toxic controls (Figure 2.4D).

FIGURE 2.4

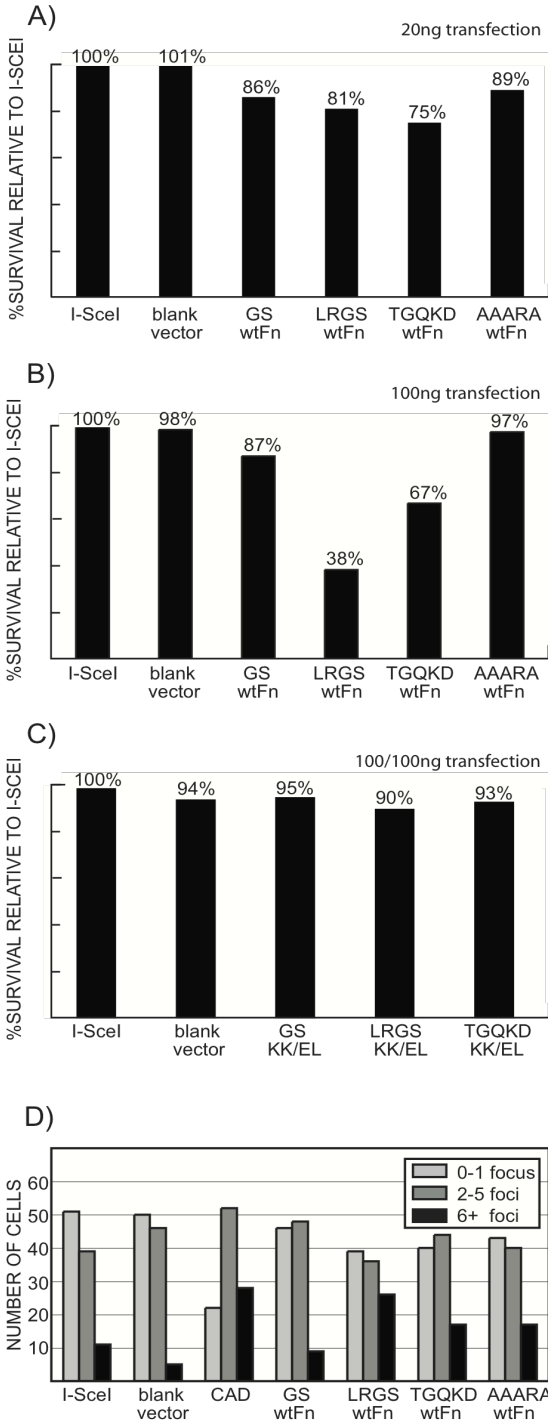
A)
20ng transfection
%SURVIVAL RELATIVE TO I-SCEI
100%
101%
86%
81%
75%
89%
I-SceI
blank vector
GS wtFn
LRGS wtFn
TGQKD wtFn
AAARA wtFn
B)
100ng transfection
%SURVIVAL RELATIVE TO I-SCEI
100%
98%
87%
38%
67%
97%
I-SceI
blank vector
GS wtFn
LRGS wtFn
TGQKD wtFn
AAARA wtFn
C)
100/100ng transfection
%SURVIVAL RELATIVE TO I-SCEI
100%
94%
95%
90%
93%
I-SceI
blank vector
GS KK/EL
LRGS KK/EL
TGQKD KK/EL
D)
NUMBER OF CELLS
60
50
40
30
20
10
0-1 focus
2-5 foci
6+ foci
I-SceI
blank vector
CAD
GS wtFn
LRGS wtFn
TGQKD wtFn
AAARA wtFn

Figure 2.4: Toxicity of Inter-Domain Linker Variant ZFNs.

The different ZFN variants were analyzed for toxicity using two different previously described assays: a cell survival assay and a double-strand break foci formation assay (Pruett-Miller et al. 2008). In the cell survival assay, a lower percent survival relative to I-SceI is a marker of greater toxicity. In the double-strand break foci formation assay, an increased number of cells with 53BP1 foci are a marker of greater toxicity. A) Cell survival after transfection of 20 ng of each ZFN with a wild-type nuclease domain. B) Cell survival after transfection of 100 ng of each ZFN with a wild-type nuclease domain. C) Cell survival after transfection of 100 ng of each ZFN with a modified nuclease domain to prevent homodimerization. D) Double-strand break foci formation assay in which double-strand breaks are identified by 53BP1 foci after immunostaining. The number of foci was counted in one hundred transfected cells for each condition. The cells were then grouped into 3 bins (0-1 foci, 2-5 foci, 6 or more foci). In prior work, we have found that a large number of cells with 6 or more foci correlate best with toxicity (Pruett-Miller et al. 2008). As negative controls, cells were transfected with either I-SceI alone or an empty expression vector ("blank vector"). As a positive control, cells were transfected with Caspase Activated DNAase (CAD). Almost 30% of the cells had greater than 6 foci when transfected with CAD. The 4 aa LRGS ZFN, 5 aa TGQKD ZFN, and the 5 aa AAARA ZFN all showed increases in DSB formation relative to the negative controls while the 2 aa GS ZFN did not. In only the 4aa LRGS ZFN, however, was the increase statistically significantly different than the negative controls (chi-square analysis with $p<0.05$).

Since toxicity increases as the amount of ZFN transfected increases, we determined the relative expression levels of the inter-domain linker ZFN variants (Figure 2.7A). We found that the LRGS-wtFn variant had higher levels of expression relative to all other variants and correlates to the higher levels of toxicity seen in Figure 2.4A, 2.4B, and 2.4D. In contrast, we also found that the LRGS-obhetFn variants did not express well, even at very high transfection amounts, which may account for the lower rates of gene targeting (Figure 2.3I, 2.3L, and 2.3O). These expression studies suggest that the toxicity of the LRGS inter-domain linker variant with the wtFn may be more the result of its expression level rather than any intrinsic property of the ZFN architecture itself (Figure 2.4A-C, Figure 2.7). Table 2.1 summarizes our findings for inter-domain linker ZFN variant activity.

	GS-wtFn	LRGS-wtFn	TGQKD-wtFn	AAARA-wtFn	GS-KK/EL Fns	LRGS-KK/EL Fns	TGQKD-KK/EL Fns
3 bp	-	-	-	-	-	-	-
4 bp	-	-	-	-	-	-	-
5 bp	++++	+++	+	+	+	+	-
6 bp	++++	++++	++++	++++	-	++	+++
7 bp	-	-	+++	-	-	-	-
Expression	+++	++++	+++	++++	++	+	++
Toxicity	Lower	Higher	Medium	Lower	Lower	Lower	Lower

Table 2.1: Qualitative Summary ZFN and Target Site Variants.

In the first five rows, ZFN linker variant gene targeting activity is evaluated relative to I-SceI (defined as +++) on targets with 3-7-bp spacers. The row labeled "expression" is the relative level of ZFN expression relative to the expression level of GFP-TGQKD-ZFN2 with the wild-type nuclease domain (defined as +++). The ZFNs with the modified nuclease domain are listed as pairs. The row labeled "Toxicity" states the amount of toxicity found with each ZFN as either "Lower," "Medium" or "Higher."

2.3.6 Experimental Strategy For Testing GFP-ZFN2 Inter-Finger Linker Variants

To investigate whether lengthening modifications to the inter-finger linker as seen in some 6- and 4-fingered ZFPs could also be applied to a 3-fingered ZFN platform to increase the target half-site repertoire from 9 bp to 10 bp, we chose to continue studying variations of the GFP-ZFN2 and target site (Kim and Pabo 1998, Moore et al. 2001b, Urnov et al. 2005). The GFP2 target half-site was mutated to include a single bp insertion between selected subsites that correspond to triplets recognized by individual fingers. While the similar mutation strategy in Moore et al. (2001) used only thymine for the inserted bp, we elected to insert guanine and adenine mutations as well to discern any potential for sequence selectivity in ZFNs modified for these new non-bound regions of the cognate target half-site (Moore et al. 2001b). Based on the inter-finger linker variant found in Urnov et al. (2005), we first modified GFP-ZFN2 canonical inter-finger linker TGEKP to TGSEKP or TGSQKD between F1-F2 or F2-F3 while maintaining the same recognition helices (**Figure 2.5**A) (Urnov et al. 2005). Then we used the OPEN method to generate new 3-fingered ZFP libraries to the normal GFP2 and four mutated sites. These libraries were made from the same single-finger archives used to generate the original ZFP for the GFP-ZFN2, and the inter-finger linker modifications (TGSEKP only) were installed during the PCR-based

FIGURE 2.5

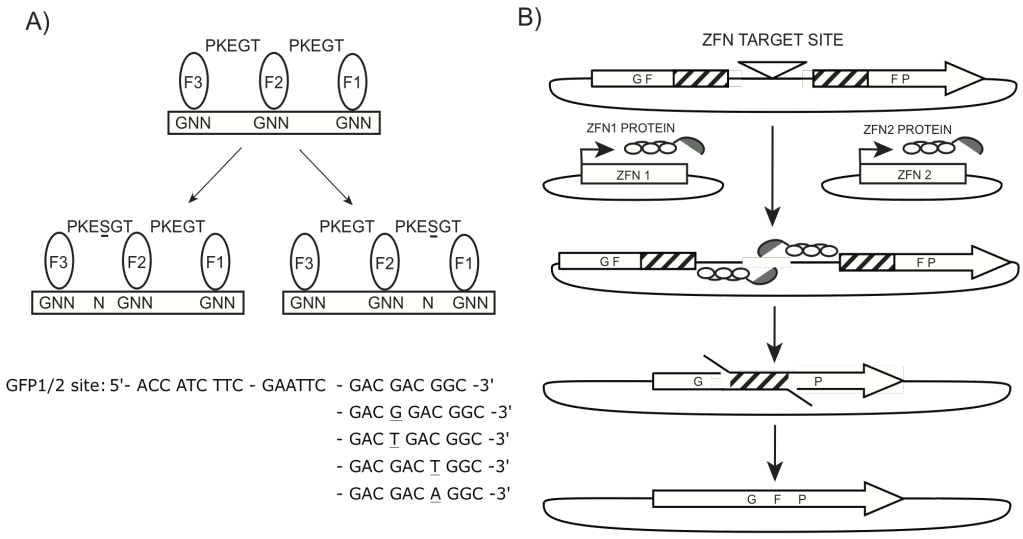

Figure 2.5: **Schematic of Inter-Finger Linker Strategy, GFP2 Target Half-Site mutations, and ZFN-Mediated Repair of an Extra-Chromosomal GFP Reporter Target by Single Strand Annealing.**

A) Our experimental model for adapting the 3-fingered platform to a 10 bp target site by inserting a single bp between the F1-F2 or F2-F3 target subsites. The diagram depicts a ZFP with TGEKP linkers that are then modified to include an extra serine residue (TGSEKP) to accommodate the insertion. Listed below are the four GFP2 target half-site mutations made as they are paired with the GFP1 target half-site. B) The target sites listed are inserted between two repeated regions of the GFP gene to create a GFP-based reporter plasmid. When co-transfected, the expressed GFP1 and GFP2 variant ZFNs cut the target site and the results DSB is repaired single-strand annealing repair mechanisms to produce a functional GFP gene.

ZFN	F3	Linker	F2	Linker	F1	Found In Selection For Site:
pGFP-ZFN2	EGGNLMR		DRSNLTR		APSKLDR	GAC GAC GGC
pM527-9a	EGGNLMR	TGSEKP	DRSNLTR		APSKLDR	none
pM527-10a	EGGNLMR	TGSQKD	DRSNLTR		APSKLDR	none
pM527-11c	EGGNLMR		DRSNLTR	TGSEKP	APSKLDR	none
pM527-12b	EGGNLMR		DRSNLTR	TGSQKD	APSKLDR	none
pKW878	DPSNLRR	TGSEKP	DRSNLTR		SPSKLIR	GAC GAC GGC
pKW877	DPSNLRR		DRSNLTR	TGSEKP	APSKLDR	GAC GAC A GGC
pKW848	DPSNLRR		DRSNLTR		APSKLTV	GAC GAC T GGC
pKW849	DPSNLRR		DRSNLTR		TKAHLDV	GAC GAC T GGC

Table 2.2: GFP-ZFN2 Inter-Finger Linker Variants.

The variant ZFNs used in this study are listed with the amino acid identities of the recognition helices of each zinc finger, which target site the ZFP was derived from in the B2H selections of the OPEN protocols, and in which position the inter-finger linker variant can be found. The inter-finger linker is TGEKP unless otherwise noted. The M527 series was made by modifying the existing ZFP found in GFP-ZFN2 and the KW8 series was derived by B2H-based OPEN protocols.

protocol for randomly recombining the single-finger archives (Maeder et al. 2009). The inter-finger linker mutations and aa identities of the recognition helices for the ZFNs made from modifying GFP-ZFN2 and OPEN-based ZFPs reported in this study are listed in **Table 2.2** (see also **Figure 2.7**B for expression analysis).

2.3.7 Repair of an Extra-Chromosomal GFP Reporter by SSA Using GFP-ZFN2 Inter-Finger Linker Variants

To assess the ability of a GFP-ZFN2 inter-finger linker variant to recognize and cut a target site with an insertion mutation compared to unmodified GFP-ZFN2 activity at the normal target site, we tested the ZFNs for nuclease activity in GFP-based extra-chromosomal SSA assays where the GFP2 (and insertion mutants) were paired with the GFP1 target half site and cloned into repeated sections of the GFP gene (Figure 2.5B). All versions of the GFP-ZFN2 were co-transfected with the GFP-ZFN1 and an SSA reporter plasmid carrying one of the five versions of the GFP1/2 site (listed in Figure 2.5A) in a combinatorial manner. All ZFN activity was normalized as a percentage of unmodified GFP-ZFN1 and GFP-ZFN2 on the normal GFP1/2 site as a positive control to provide context for how well an inter-finger variant ZFN could perform at a target site with the 1 bp insertion (Table 2.3).

ZFN	GAC GAC GGC	GAC G GAC GGC	GAC T GAC GGC	GAC GAC T GGC	GAC GAC A GGC
pGFP-ZFN2	100%	11 (± 3.9)	13.4 (± 7.2)	13.1 (± 0.7)	5.8 (± 3.4)
pM527-9a	44.1 (± 6.9)	15.2 (± 6.7)	9.9 (± 8.1)	8.4 (± 2.2)	3.5 (± 1.7)
pM527-10a	38.9 (± 11.6)	13.5 (± 4.9)	0.6 (± 0.6)	5.7 (± 1.1)	2.9 (± 2.9)
pM527-11c	19.4 (± 1.7)	5.2 (± 2.7)	2.9 (± 2.6)	13.7 (± 3.7)	2.37 (± 1.6)
pM527-12b	18.2 (± 4.7)	8.1 (± 5.4)	6.9 (± 5.7)	10.4 (± 5.4)	4.3 (± 4.0)
pKW878	77.9 (± 25.5)	26.2 (± 7.7)	147 (± 15.2)	44.4 (± 10.1)	0.9 (± 0.9)
pKW877	31.9 (± 5.2)	2.9 (± 2.3)	0.5 (± 0.5)	134.1 (± 32.2)	5.8 (± 3.2)
pKW848	56.7 (± 9.1)	0.9 (± 0.8)	0	264.7 (± 63.7)	24.6 (± 4.4)
pKW849	24.2 (± 2.6)	0	0	296 (± 94.9)	23.1 (± 5.0)

Table 2.3: Nuclease Activities of GFP-ZFN2 Inter-Finger Linker Variants.

The SSA reporter plasmids (20ng) were individually co-transfected in to HEK293 cells with 100ng of the GFP-ZFN1 and 100ng of a GFP-ZFN2 variant for a combinatorial strategy. The data are presented as rates of GFP repair as normalized to a percentage of the nuclease activity of the GFP-ZFN1 and GFP-ZFN2 pair on the normal GFP1/2 site (mean ± SEM). Refer to **Table 2.2** for information regarding individual ZFNs.

In our results, we found that the unmodified GFP-ZFN2 had low but measurable activity on all the mutated sites (~10%) in comparison to the normal site. These data points provide a glimpse into the specific types of sequences at which this ZFN may be delivering off-target DSBs (refer to TGQKD-wtFn data in **Figure 2.2** and 2.4). For the inter-finger linker variants where the canonical TGEKP linker was mutated to TGSEKP or TGSQKD at different inter-finger positions (M527 series in **Table 2.2** and **Table 2.3**), we broadly found that these lengthened inter-finger linkers reduced ZFN activity. With the M527-9a and -10a ZFNs, the F3-F2 inter-finger linker variants showed ~40% of the activity of the unmodified GFP-ZFN2 at the normal site but low activity on all other sites, even on the 5'-GAC<u>G</u>GACGGC-3' or 5'-GAC<u>T</u>GACGGC-3' GFP2 site mutants for which one could hypothesize these ZFNs to be a match for. Likewise, in the M527-11c and M527-12b ZFNs, the F2-F1 inter-finger linker variations showed less activity on the normal GFP2 site (~20%) and low activity on all other sites. In this series of ZFNs, the addition of a serine in the inter-finger linker reduced nuclease activity in these SSA assays, and that the position of the inter-finger linker variation made more of a difference in activity than the amino acid sequence.

Of the many ZFPs generated across all five GFP2-type sites in the OPEN protocol that we converted to ZFNs for testing (data not shown), only four

showed efficient nuclease activity and are reported in this study (KW8 series listed in**Table 2.2** and **Table 2.3**). The KW878 ZFN was found in the OPEN-based selection from interrogating the GFP2 F3-F2 inter-finger linker variant library on the normal GFP2 site. In our SSA assays, we found that this ZFN still maintains efficient nuclease activity for that normal site (~75%) but also demonstrated very high activity (~150%) for the 5'-GACTGACGGC-3' relative to the positive control. However, the KW878 ZFN has much less activity on the similar 5'-GACGGACGGC-3' site (~25%). The KW877 ZFN was found in the OPEN-based selection using the GFP2 F2-F1 inter-finger linker variant library on the 5'-GACGACAGGC-3'site. When tested in the SSA assays, this ZFN instead showed very high preferential activity for the 5'-GACGACTGGC-3' site. Interestingly, the KW848 and KW849 ZFNs were derived when the normal GFP2 library was interrogated on the 5'-GACGACTGGC-3' site and thus have no inter-finger linker variations. These ZFNs showed the most activity on the "selected for" site but also demonstrated measurable off-target nuclease activity at the normal GFP2 site. It is interesting to note that the highest nuclease activities for this system are consistent with a thymine insertion in the GFP2 target site (Moore et al. 2001b). In summary, based on these data, inter-finger linker variations may be tolerated in the 3-fingered ZFN platform, but we do not believe that this strategy can help broaden the repertoire of ZFN target sites from 9 bp to 10 bp without decreases in specificity. We are also unable to conclude that specifically

matching 6-aa inter-finger linkers to the positions the target site insertions in this platform can be systematically accomplished for three reasons: 1) the high failure rate of inter-finger linker ZFN variants made by either modifying an existing 3-finger ZFPs or those generated by OPEN protocols; 2) discrepancies between sites with high nuclease activity in the SSA assays when compared to the site selections of ZFP origin; and 3) significant levels of off-target cutting.

2.3.8 Hybridizing the Modular Assembly and OPEN Methodologies

Given the lack of full coverage for all triplet subsites publically available for modular assembly, and even more limited subsite coverage from OPEN methods, the preference for developing a 3-fingered ZFP to the 5'-GNNGNNGNN-3' target half-site motif remains. However, full ZFN target sites with two such half-site motifs are found too infrequently within the human genome to always have convenient proximity to a desired locus. Therefore, to address developing ZFPs to target sites containing non-GNN triplets, we propose incorporating modular assembly fingers into the OPEN protocols to not only broaden the range of potential target sites, but also increase the rates of success in the resulting ZFPs when module fingers are used (Kim et al. 2010, Maeder et al. 2009, Maeder et al. 2008, Ramirez et al. 2008). For this study, we identified four different full ZFN target sites where at least one of the target half-sites contains a non-GNN subsites. The sequences are given in Table 2.4. We then adapted the module fingers that recognize the non-GNN subsites by including them in the

PCR-based protocol for randomly recombining the single-finger archives to make the 3-finger cassettes used in full site selection (schematized in Figure 2.6A). The resulting ZFNs from this hybridized method of ZFP development are also listed in Table 2.4.

2.3.9 Repair of an Extra-Chromosomal GFP Reporter by SSA Using ZFNs Developed From Hybrid Methodologies

In this study, we tested the ZFNs made by the hybridized method for nuclease activity with SSA strategies very similar to those described in Figure 2.5B. Briefly, the four target sites were inserted into repeated sections of the GFP gene to create reporter plasmids that were co-transfected with the corresponding pair of ZFNs listed in Table 2.4. The activity of these ZFNs were normalized relative to the GFP1/2 positive control and reported as a percentage of that activity (Figure 2.6B). We found that when the JZ90A ZFN (ACG module at F2) is used with the JZ110 ZFN, this pairing can show ZFN activity on the F2-ACG site that is ~one-third of the GFP1/2 positive control. In comparison, the JZ154/JZ144 ZFN pair, which also contains an ANN module for F2, exhibits nuclease activity for the F2-AAC site that approaches the GFP1/2 standard (~85%). We also installed a module finger recognizing the F1-CAG for the JZ99C2 ZFN. When used as a pair with the JZ108 ZFN, the mean nuclease activity for the F1-CAG site exceeds the positive control (~125%), although this

FIGURE 2.6

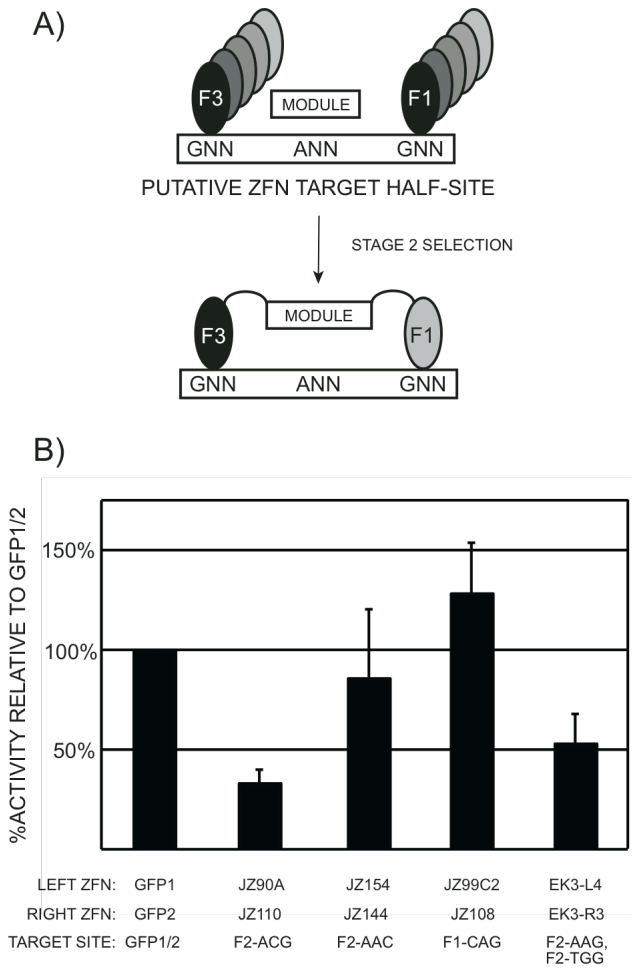

Figure 2.6: Experimental Strategy For Developing ZFPs Using a Hybridized Method and the Nuclease Activity of the Resulting ZFNs as Measured by Extra-Chromosomal Repair of a GFP-based Reporter Plasmid.

A) Module fingers were integrated into the B2H-based OPEN methods in the creation of 3-fingered libraries before the second stage of B2H selections for the full target half-site. In this manner, module fingers can be used to generate ZFPs made through the paradigm of context-dependent binding. B) In a strategy similar to that depicted in Figure 2.5B, SSA reporter plasmids were made by inserting the target sites listed in Table 2.4 between repeated regions of the GFP gene to create a GFP-based reporter plasmid. Each reporter plasmid (20ng) was co-transfected into HEK293 cells with 100ng of each ZFN-expressing plasmid in appropriate pairs. Extra-chromosomal repair of the resulting DSB by SSA mechanism produces a functional GFP gene. The data are presented as rates of gene targeting as normalized to a percentage of the nuclease activity of the GFP-ZFN1 and GFP-ZFN2 pair on the normal GFP1/2 site (Error = SEM).

Site Name	Full Site Sequence	ZFN	F3	F2	F1
GFP1/2					
	5'- ACC ATC TTC - gaattc - GAC GAC GGC -3'				
	LEFT ZFN: GNN GNN GNN	pGFP1	QHPNLTR	VAHNLTR	TRQKLGV
	RIGHT ZFN: GNN GNN GNN	pGFP2	EGGNLMR	DRSNLTR	APSKLDR
F2-ACG					
	5'- TAC CGT GTC - caagac - GGA GAC GAG -3'				
	LEFT ZFN: GNN ACG GNN	pJZ90A	DQGNLIR	RTDTLRD	RAAVLVR
	RIGHT ZFN: GNN GNN GNN	pJZ110	TNNVLNT	DRSNLTR	KHSNLTR
F1-CAG					
	5'- CTG CTC AAC - atcgcc - GTG GCT GAC -3'				
	LEFT ZFN: GNN GNN CAG	pJZ99C2	VNSSLGR	RDKNLTR	RADNLTE
	RIGHT ZFN: GNN GNN GNN	pJZ108	RNDALRR	LSQTLKR	DEANLRR
F2-AAC					
	5'- TCC CAC AGC - tcctg - GGC AAC GTG -3'				
	LEFT ZFN: GNN GNN GNN	pJZ154	VRNTLNR	RTEILRN	DNAHLAR
	RIGHT ZFN: GNN ACC GNN	pJZ144	KNVSLNN	DSGNLRV	RSTSLHR
F2-AAG,F2-TGG					
	5'- CTC CTT GCC - tagtct - GGA TGG GCA -3'				
	LEFT ZFN: GNN AAG GNN	pEK3-L4	KNASLGH	RKDNLKN	RMSNLDR
	RIGHT ZFN: GNN TGG GNN	pEK3-R3	QRTHLRV	RSDHLTT	DRSQLAR

Table 2.4: Sequences of Target Sites and ZFNs Used to Test a Hybridized Method of Developing ZFPs to Non-GNN Subsites.

Listed here are four ZFN target sites we identified where one or two subsites were not GNN triplets. By using modular assembly and the B2H-based OPEN method, we generated pairs of ZFNs to target each site. Listed are those used in this study along with the amino acid identity of the recognition helices for each zinc finger.

is not statistically significant ($p = 0.19$, Student's one-tailed t-test). In addition to studying ZFN pairs with only one modular finger, we used the F2-AAG, F2-TGG site as an opportunity to investigate the feasibility of using two modular fingers. Both the EK3-L4 and EK3-R3 ZFNs have F2 modules and as a pair show half as much nuclease activity for the F2-AAG, F2-TGG site relative to the activity of the GFP1/2 ZFNs. All together, these data support our proposal to develop ZFNs by OPEN-based methods using module fingers in the F1 and F2 positions, and also suggests that functional ZFNs can result from this hybridized methodology.

FIGURE 2.7

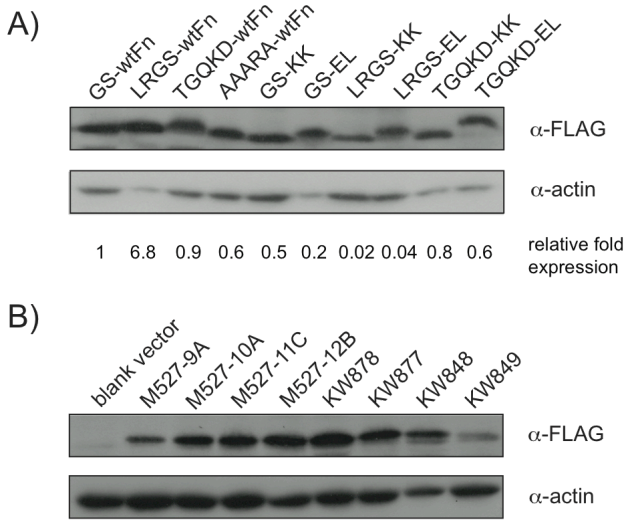

Figure 2.7: Expression Analysis of GFP-ZFN2 Linker Variants.

A) GFP-ZFN2 inter-domain linker variants used in this study were transiently transfected into HEK293 cells at 150 ng (lanes 1-5, 9-10) or 750 ng (lanes 6-8) and the resulting lysates were loaded in varying amounts until the α-FLAG signal became relatively equal by Western blot analysis. Fold expression differences were calculated using densitometric analysis by normalizing the α-FLAG signal to the corresponding α-actin loading control. The GFP-ZFN2 variants are labeled according to the amino acid content of the inter-domain linker and the Fn used. The KK variants all seemed to migrate slightly faster than the wild-type nuclease or EL variants for unexplained reasons. The low expression of the LRGS-KK and LRGS-EL ZFNs is a possible explanation for the lower gene targeting activity found in **Figure 2.3**. The other KK and EL variants with the GS and TGQKD inter-domain linkers, however, express relatively better. Thus expression level does not provide an explanation for why the obhetFn variants did not stimulate gene targeting as effectively as the wild-type nuclease domain variants. B) GFP-ZFN2 inter-finger linker variant ZFNs were also transiently transfected into HEK293 cells at 500 ng each and expression was assayed by Western blot analysis using an α-FLAG antibody (see **Table 2.2** for complete description of each ZFN). In each lane, blotting for actin was used as a loading control. In the M527 series of ZFNs, changes in the position and amino acid content of inter-finger linker do not result in highly variable ZFN expression. However, more variation in expression is seen in the KW8 ZFN series despite each ZFN demonstrating efficient nuclease activity in the SSA assays (**Table 2.3**).

2.4 DISCUSSION

Determining the range of sites that can be targeted by ZFNs is an important part of the development of ZFN-mediated genome modification. Therefore we surveyed a range of target site variations to investigate whether architectural elements of a ZFN can be adapted for efficient nuclease activity. In this study, we challenged variations such as spacer lengths, insertions in the target half-site, and the inclusion of non-GNN subsites with changes in ZFN inter-domain linkers, inter-finger linkers, and protocols for ZFP creation.

Previous work has demonstrated that full ZFN target sites with 5-, 6-, 7-, and 16-bp spacers could be targeted efficiently at a chromosomal locus (Alwin et al. 2005, Bibikova et al. 2002, Handel et al. 2009, Porteus 2006, Porteus and Baltimore 2003, Urnov et al. 2005). To target sites with these different spacer lengths, however, requires ZFNs with slightly different architectures. Our work not only validates the efficacy of targeting such sites, but also provides additional inter-domain linker solutions to differences in target site spacer length (Figure 2.3). More specifically, we have also found that sites with 3- or 4-bp spacers could not be targeted efficiently. In our work, target sites with 5-bp spacer lengths showed the most efficient ZFN-mediated gene targeting when the linker between the zinc finger DNA binding domain and the nuclease domain is 2 aa or 4 aa (Figure 2.3G-H). In contrast, our work demonstrates that ZFN variants with a 5-aa TGQKD inter-domain linker can efficiently target sites with a 7 bp spacer. In

comparing the AAARA and TGQKD 5-aa inter-domain linker constructs, we found the TGQKD linker had a broader range of activity but also had a higher degree of toxicity (Figure 2.3M-N, 2.4A-B, and 2.4D).

The evaluation of new ZFNs by an *in vitro* cutting assay could be a convenient way of assessing on-target and off-target effects (Figure 2.2). Unfortunately, we find that the ability of a ZFN to cut a target site *in vitro* does not always predict its ability to cut its target when embedded in the genome (Figure 2.3). This result is most clearly demonstrated by the TGQKD 5-aa linker variant which cuts a 5-bp target site as efficiently as a 6-bp site *in vitro* but has significantly less activity on the 5-bp target in cell-based assays (compare Figure 2.3G-H to 2.3J-K). Thus in evaluating a ZFN, *in vitro* assays alone are not sufficient.

Minimizing ZFN toxicity is also an important aspect of ZFN development (Alwin et al. 2005, Cornu et al. 2008, Miller J. C. et al. 2007, Pruett-Miller et al. 2008). In this work, we have confirmed that engineering modifications in the nuclease domain to produce obligate heterodimerization can reduce toxicity (Figure 2.4C-D). The data also shows that these modifications, however, can result in reduced nuclease activity. How these modifications reduce toxicity requires further investigation to elucidate the full mechanism. The elimination of homodimer formation should only result in a two-fold reduction in toxicity (preventing homodimer (ZFN1 with ZFN1 and ZFN2 with ZFN2) cutting at off-

target sites but not preventing off-target cutting by heterodimers (ZFN1 with ZFN2 and ZFN2 with ZFN1), but published studies have demonstrated a significantly greater decrease (Miller J. C. et al. 2007, Pruett-Miller et al. 2008, Szczepek et al. 2007). One hypothesis is that these modifications reduce toxicity not just by eliminating homodimerization but also by decreasing the affinity of the nuclease domains for each other, thereby requiring a more stable complex between the zinc finger DNA binding domain and its target binding site to form before cutting can occur. In addition, we did not find that the length or content of the linker affected toxicity. Instead, we found that the length and content of the linker and modifications of the nuclease domain could affect expression levels, and that the change in expression correlated best with toxicity (Figure 2.4 and Figure 2.7). This expression data also suggests the possibility of improving toxicity if ZFN expression could be controlled (Pruett-Miller et al. 2009).

In addition to variations in target site spacer lengths, we investigated the possibility of developing ZFNs to 9-bp target half-sites containing 1 bp insertions between subsites (10 bp total). We hypothesized that lengthening the inter-finger linker from 5 aa to 6 aa at the corresponding position of the insertion may allow a ZFN to conform to the new spatial requirements in the DNA helix for efficient nuclease activity (**Figure 2.5**). Unfortunately, our efforts resulted in very few functional ZFNs. Any inter-finger linker variant ZFN with efficient nuclease

activity demonstrated a preference for thymine insertions and significant amounts of off-target cutting (**Table 2.2** and **Table 2.3**). One possibility is that the TGSEKP or TGSQKD linkers may not provide an optimal solution to this problem whether in terms of length or amino acid content. Also, and more likely, the 3-fingered ZFP platform may not be robust enough to tolerate a lengthened inter-finger linker or target site insertion in the same manner that a 4-fingered or 6-fingered ZFP can (Kim and Pabo 1998, Moore et al. 2001b, Urnov et al. 2005). Based on these findings, we cannot recommend developing 3-fingered ZFNs to 10-bp target half-sites.

For those target half-sites sites that include non-GNN subsites, the options are limited for successfully developing a ZFP to such target half-sites using publically available protocols and materials. Modular assembly methods have 50 out of 64 bp triplets available, but rates of success are low (Ramirez et al. 2008). The OPEN method has higher rates of success, but only GNN subsites and some TNN subsites can be considered (Maeder et al. 2008). We sought to solve the problems of low success rates and low coverage by hybridizing these methods by installing module fingers for non-GNN subsites in the OPEN protocols (Figure 2.6A) (Maeder et al. 2009). We generated ZFNs by this hybridized methodology to four different target sites where a modular finger was used at the F1 or F2 position. Of the four ZFN pairs tested, we were successful in demonstrating nuclease function, although to varying degrees of efficiency on an extra-

chromosomal target (Table 2.4 and Figure 2.6B). Using this hybridized method, our data lead us to believe that module fingers can be reasonable candidate helices for the F1 or F2 positions, and presumably for the F3 position as well. This preliminary study provides encouraging data that endogenous loci may also be targeted with ZFNs developed with such hybrid ZFPs, and with the incorporation of module fingers into OPEN protocols, more target sites are eligible for consideration.

In summary, this work provides a guide to identify possible full ZFN target sites that vary from the canonical structure of two 5'-GNNGNNGNN-3' target half-sites separated by a 6-bp spacer and methods to adapt ZFN architecture to cut those sites. The probability of finding such canonical sites is 1 in 4096 bp and may not be in sufficient proximity to a desired locus for achieving high rates of gene conversion (Elliott et al. 1998). However, when target site criteria is expanded to include 5-, 6-, or 7-bp spacer and the possibility of using module fingers and OPEN protocols together, the probability of finding a candidate target site can dramatically increase to 1 in 17 bp (Table 2.5). Thus, the guidelines provided by the data in this study should be immediately useful to researchers attempting to develop ZFNs to perform efficient, site-specific genome modifications.

Target Site Conditions	Probability
"GNN" triplets with a 6 bp spacer	1 in 4096 bp
"GNN" triplets with a 5, 6, or 7 bp spacer	1 in 1382 bp
OPEN B2H triplets with a 6 bp spacer	1 in 608 bp
OPEN B2H triplets with a 5, 6, or 7 bp spacer	1 in 216 bp
OPEN B2H triplets with one module for F1 or F2 on ONE side with 5, 6, or 7 bp spacer	1 in 63 bp
OPEN B2H triplets with one module for F1 or F2 on BOTH sides with 5, 6, or 7 bp spacer	1 in 17 bp

Table 2.5: Target Site Criteria and Associated Probabilities.

Based on the target site guidelines provided by this study, we calculated the probability of finding such sites using Monte-Carlo based methods (99% confidence). We assumed 50% GC content and as inputs, used the modular assembly fingers and the B2H-based OPEN single finger pool archives (as available from Addgene). By allowing spacer lengths of 5, 6, or 7 bp in the target site and using hybridized methods to generate ZFPs, the probability of finding a full ZFN site in close proximity to a locus of interest increases dramatically relative to the probability of only considering canonical sites.

2.5 MATERIALS AND METHODS

2.5.1 GFP-ZFN2 Inter-Domain Linker Variants with the Wild-Type Nuclease Domain

The 3-fingered GFP2 zinc finger DNA binding domain (ZFP) was developed through the B2H protocols of the OPEN methodology described previously (Pruett-Miller et al. 2008). The DBD recognizes the 9-bp target half-site 5'-GACGACGGC-3' and the recognition helices for the three fingers are as follows: Finger 1: APSKLDR; Finger 2: DRSNLTR; Finger 3: EGGNLMR. Using standard molecular biology procedures, three variants of the original nuclease were made: one with a 2-aa linker (GS), a 4-aa linker (LRGS), and another 5-aa linker (AAARA) (Alwin et al. 2005, Urnov et al. 2005). All ZFN variants were cloned using the previously characterized GFP-ZFN2-B2H vector plasmid as a template, which already has the TGQKD inter-domain linker and wild-type FokI nuclease domain (wtFn) and referred to as GFP-ZFN2 in this study. All of the ZFNs in this paper were cloned into expression vectors with a CMV promoter.

2.5.2 GFP-ZFN2 Inter-Domain Linker Variants with the Obligate Heterodimer Nuclease Domain

The obhetFn domains were made by PCR mutagenesis using the GFP-ZFN2-B2H vector as a template. The "KK" variant contains the E490K and

I538K mutations and the "EL" variant contains the Q486E and I499L mutations where the numbering is with respect to the wild-type FokI enzyme (Miller J. C. et al. 2007). Six ZFNs with obhetFns (three KK/EL pairs) were made for the GS, LRGS, and TGQKD inter-domain linkers.

2.5.3 Generation of Reporter Cell Lines with Target Sites of Different Spacer Lengths

The ZFN linker variants were tested for targeting at sites with different spacer lengths using the previously reported GFP gene targeting assay (Durai et al. 2005, Porteus and Baltimore 2003). In this reporter, an inverted repeat of the GFP2 binding site was inserted into the middle of a mutated GFP gene and adjacent to the recognition site for I-SceI using standard molecular biology techniques (Figure 2.1B). Separate reporter plasmids in which the GFP-ZFN2 binding sites were separated by either 3, 4, 5, 6, or 7 bp were made and each reporter was used to generate a monoclonal HEK-293 cells line with a single copy of the reporter as previously described (Porteus 2006, Porteus and Baltimore 2003).

2.5.4 In Vitro Cutting Assay

The *in vitro* cutting assay was performed using a protocol to be described elsewhere (AEM and MHP, manuscript in preparation). Briefly, His-tagged GFP-ZFN2 (TGQKD inter-domain linker and wtFn) was purified using a Ni^{+2} metal

affinity column. We then linearized the five different target construct plasmids with the 3-7 bp spacer lengths using a SpeI digest and *in vitro* combined purified ZFN protein at molar ratios of 4:1-0.25:1 (protein:linearized plasmid, where (Bodnar et al.)=0.3nM). Cutting of the target constructs was evaluated by agarose gel electrophoresis.

2.5.5 GFP-Based Gene Targeting Assays

Gene targeting assays were done as described (Porteus and Baltimore 2003). ZFNs were transfected at 20 ng and 100 ng with 300 ng of repair template (donor) plasmid into HEK293 cells. All GFP-ZFN2 linker variants with KK or EL nuclease domains were transfected at 100 ng each with 300 ng of repair template plasmid. As an internal control for each experiment, separate wells were transfected with the repair template and GFP-ZFN2 (TGQKD inter-domain linker and wtFn) at 200 ng. Transfection efficiency was determined 48 hours post-transfection and gene targeting rates were measured at 72 hours post-transfection. Repair of the mutated GFP gene to restore function was measured by flow cytometry using a FACS Calibur (Becton-Dickerson, San Jose, CA, USA). Each of these experiments was done between 3-11 times. All gene targeting events were normalized to the transfection efficiency and normalization as a percentage of I-SceI activity. By normalizing to the internal I-SceI standard, we could compare the efficiency of ZFN variants across different cell lines. Significance

was measured using Student's one-tailed t-test where $p<0.05$ and error bars indicate standard error of the mean (SEM).

2.5.6 Flow Cytometry Assay For Cell-Survival

Cell survival assays were performed as previously described (Pruett-Miller et al. 2008). Briefly, HEK293 cells were transfected with 300 ng of a GFP expression plasmid along with 20 ng or 100 ng of GFP-ZFN2 variants and analyzed by flow cytometry for GFP expression at days 2 and 6 after transfection. Percent survival is calculated as ratio of ratios of percent GFP+ populations at day 2 and day 6: ((ZFN day 6 / ZFN day 2) / (I-SceI day 6 / I-SceI day 2)) x 100.

2.5.7 Foci Formation Assay and Immunofluorescence Staining

Primary cultures of human foreskin fibroblasts were maintained and transfected with nucleofection techniques as previously described (Amaxa Inc., Gaithersburg, MD, USA). Briefly, one million fibroblasts were co-nucleofected using 2 μg of GFP-expression plasmids and 2 μg of each ZFN variant-expressing plasmid. DSBs were highlighted 48 hours post nucleofection though incubation with rabbit α-p53BP1 primary antibody (Cell Signaling, Danvers, MA, USA) and goat α-rabbit Rhodamine Red-X secondary antibody (Invitrogen, Carlsbad, CA, USA). Cells were mounted onto slides using Vectashield with 4,6-diamidino-2-phenylindole (DAPI) (Vector Laboratories, Burlingame, CA). Foci within the GFP+ fibroblasts were manually counted in a blinded manner and binned

according to number of foci/nucleus. Bins were set at 0-1 focus as background or minimal toxicity, 2-5 foci as moderate toxicity, and 6+ foci as severe toxicity. Differences between ZFN variants were analyzed statistically by a chi-squared analysis with a significance threshold of $p<0.05$.

2.5.8 Western Blot Analysis

Expression of ZFNs was determined by Western blotting using an anti-FLAG antibody (Sigma, St. Louis, MO, USA) to the FLAG tag at the N-terminus of each ZFN using standard procedures. Briefly, each ZFN was transfected into HEK-293 cells and cell lysates were made 48 hours post-transfection. Densitometric analysis was performed using Image J software.

2.5.9 Generation of GFP2 Inter-Finger Linker ZFNs Variants Using the OPEN Method

First, using the GFP-ZFN2-B2H vector plasmid as a template, the canonical 5 aa TGEKP inter-finger linkers at either the Finger1-Finger2 (F1-F2) or Finger2-Finger3 (F2-F3) junctions were mutated to TGSEKP or TGSQKD, while preserving the original α-helices and TGQKD inter-domain linker to create four new ZFN variants using standard molecular biology techniques. New ZFPs were generated with B2H selections in the OPEN methodology using the same zinc finger pools used to generate the original ZFP of the GFP-ZFN2. The mutant TGSEKP inter-finger linker was incorporated into the three-fingered cassettes to

generate two new libraries, F1-TGSEKP-F2 and F2-TGSEKP-F3, for stage two B2H selections (Maeder et al. 2009). These libraries were then interrogated in five bacterial reporter strains bearing either the normal GFP2 target site or one of four GFP2 target site variants with a 1bp insertion between subsites: 5'-GACGGACGGC-3', 5'-GACTGACGGC-3', 5'-GACGACTGGC-3', and 5'-GACGACAGGC-3'. New ZFNs with TGQKD inter-domain linkers were cloned from the rescued ZFPs found in the surviving colonies.

2.5.10 Hybrid Method For Creating ZFPs Using Modular Fingers and the OPEN Protocols

We identified four new full ZFN target sites in which one or both of the half-sites have a non-GNN triplet. The names and sequences are as follows: F2-ACG: 5'-TACCGTGTC-ccagac-GGAGACGAG-3'; F1-CAG: 5'-CTGCTCAAC-atcgcc-GTGGCTGAC-3'; F2-AAC: 5'-TCCCACAGC-tcctg-GGCAACGTG-3'; and F2-AAG,F2-TGG: 5'-CTCCTTGCC-tagtct-GGATGGGCA-3'. Modular assembly fingers (Addgene) were used in the finger positions in the non-GNN target half-site. The modules were incorporated into the recombinant PCR reaction to generate three finger cassettes for making five hybrid libraries in addition to the three other conventional OPEN-based libraries at the stage two B2H selections (Maeder et al. 2009). These 3-fingered libraries were then interrogated in corresponding bacterial reporter strains bearing the appropriate target half-site. New ZFNs with TGQKD inter-domain linkers were cloned from

the rescued ZFPs found in the surviving colonies, with the exception of the ZFNs developed to target the F2-AAC site for which the GS inter-domain linker was used.

2.5.11 Generating GFP-based SSA Reporter Plasmids

Nine SSA reporter plasmids were made by inserting one full ZFN target site into repeated sequences within the GFP gene using standard molecular biology techniques. These include the normal GFP1/2 ZFN target site (5'-ACCATCTTC-gaattc-GACGACGGC-3') as a positive control, four GFP2 insertion target half-site variants which were paired with the GFP1 half-site (5'-GAAGATGGT-3') with a 6-bp spacer to create a new full target site, and the four ZFN sites identified to have non-GNN triplets listed above.

2.5.12 SSA Assays For ZFN Variants

Extra-chromosomal SSA assays were performed in HEK-293 cells by transfection similar to those described previously (Szczepek et al. 2007). Briefly, 20 ng of a SSA reporter plasmid (described above) with 100 ng of each of the plasmids expressing a single ZFN. Repair of the mutated GFP gene by the endogenous SSA repair mechanism to restore GFP function was measured by flow cytometry at 48 hours post transfection. The GFP-ZFN2 inter-finger linker variants were co-transfected with the GFP-ZFN1 which recognizes the GFP1 half-site and the recognition helices for the three fingers are as follows: Finger 1:

TRQKLGV; Finger 2: VAHNLTR; Finger 3: QHPNLTR (previously described as GFP1.4-B2H) (Pruett-Miller et al. 2008). The hybrid ZFP ZFNs were co-transfected as pairs corresponding to the chosen target site. As a positive control, GFP-ZFN2 and GFP-ZFN1 were transfected with an SSA reporter plasmid with the normal GFP1/2 ZFN target site. Activities of all ZFNs were normalized to this positive control and reported as a percentage that activity. Each of these experiments was done 3-4 times and error is reported as SEM.

2.6 ACKNOWLEDGEMENTS

We thank Dr. Sandeep Burma for his assistance and advice in developing the foci formation assay and Dr. Toni Cathomen for discussions with regards to the AAARA inter-domain linker.

This work was supported by National Institutes of Health grants R01 HL079295, The March of Dimes, a career development award from the Burroughs Welcome Fund, and funding through University of Texas Southwestern Medical Center by the State of Texas.

CHAPTER 3:DESIGN AND DEVELOPMENT OF ARTIFICIAL ZINC FINGER TRANSCRIPTION FACTORS AND ZINC FINGER NUCLEASES TO THE HTERT LOCUS

3.1 ABSTRACT

The *hTERT* promoter is a complex region of regulatory sequences that controls *hTERT* expression and thus, indirectly, the levels of the telomerase holoenzyme. The ability to direct *hTERT* expression through either genetic control or tunable regulatory factors would advance not only our understanding of the transcriptional regulation of *hTERT*, but also potentially produce new strategies for addressing telomerase-associated disease. In this work, we describe the engineering of artificial ZFTFs and ZFNs by modular assembly and OPEN protocols to target sequences within the *hTERT* promoter and exon 1. The ZFTFs were generated by linking the zinc finger DNA binding domain ZFP to the VP16 transcriptional activation domain and we were able to identify several active ZFTFs that demonstrate a dose-dependent response by screening a cell-based transcriptional reporter. The same ZFPs used to make ZFTFs were also converted into ZFNs and screened in combinatorial pairs in cell-based extra-chromosomal single-strand annealing assays and stably-integrated constructs in gene targeting assays. These screening strategies have pinpointed several ZFN pairs that may be useful in genomic editing of the *hTERT* locus. Our work provides an encouraging

foundation in further developing these artificial zinc finger proteins for use at the *hTERT* locus.

3.2 INTRODUCTION

Telomeres protect human chromosomes from genomic instability and replicative attrition that results as a consequence of their linear structure and the inability of DNA polymerases to completely replicate chromosomal ends (Smogorzewska and de Lange 2004). These specialized chromosomal end-caps are comprised of TTAGG repeats synthesized by the telomerase enzyme. Maintenance of telomere length by telomerase allows a cell to keep proliferative capacity while the absence of active telomerase permits telomere shortening with every cell cycle (Bodnar et al. 1998). This becomes as important mechanism behind the control of cellular aging, lifespan, proliferative potential, and eventual senescence. The human telomerase holoenzyme is a complex of two components: hTR, a single-stranded RNA template, and hTERT, a reverse transcriptase. As the limiting component, *hTERT* expression is constitutively repressed in normal somatic cell types. It is often through the re-upregulation of *hTERT* expression that transformed cells are able to achieve immortalization (Artandi and DePinho 2010). However, the re-introduction of *hTERT* expression alone is not sufficient for transformation (Harley 2002, Morales et al. 1999).

The promoter region controlling *hTERT* expression is densely composed of many regulatory elements. Many of these can be found within the first 300 bp,

which is considered to be the core promoter. Studies have suggested that transcriptional regulation of *hTERT* occurs in a cell context-dependent manner and the transcription factors controlling the promoter at different sites can vary between cell lines (Horikawa and Barrett 2003). Haploinsufficiencies and mutations in *hTERT* have also been implicated in a number of inheritable genetic diseases that result from increasingly shortened telomere lengths. These diseases (such as idiopathic pulmonary fibrosis, aplastic anemia, and dyskeratosis congenita) exhibit attrition of fast growing tissues, stem cell depletion, and anticipation in subsequent generations (Armanios M. et al. 2005, Armanios M. Y. et al. 2007, Tsakiri et al. 2007). It would therefore benefit many lines of research into the *hTERT* locus if the genomic sequence could be modified or the promoter put under directed control. Systematic site-specific mutational analysis of the endogenous *hTERT* locus or the use of a tunable regulatory factor at the *hTERT* promoter across many human cell lines would yield many potentially useful insights about the regulation of this crucial gene.

Artificial ZFTFs and ZFNs have shown promise in becoming solutions to such technical needs. Both chimeric proteins are based on the well-characterized zinc finger DNA binding protein domain (ZFP). The polydactyl ZFP is custom-engineered for the chosen target site and can be generated through combining individual zinc fingers in a modular fashion, using randomly recombined libraries of zinc fingers in the B2H-based OPEN platform, or a hybrid methodology that

incorporates elements of both (Kim et al. 2010, Maeder et al. 2009, Maeder et al. 2008). A novel, chimeric protein with site-specific activity is then made by attaching another functional domain to the ZFP.

ZFPs, when linked to an effector domain, become ZFTFs that will recruit the cellular machinery necessary to activate or repress transcription and are typically developed to sequences found in the promoter regions of the target gene. Due to the ability of a ZFTF to site-specifically regulate transcription in cell culture and in live animal models, these artificial proteins are becoming attractive new avenues of drug development (Sera 2009). Recently, Sohn et al. developed a set of four-fingered ZFTFs that target sequences within in the *hTERT* core promoter and demonstrated transcriptional repression of *hTERT* in HEK293 cells by linking the KRAB repressor domain to ZFPs made by modular assembly (Sohn et al. 2010). Several ZFTFs have also been successfully made for research to other target genes such as *hbax*, *hVEGF-A*, and utrophin, all of which have therapeutic potential (Beerli et al. 2000, Mattei et al. 2007, Snowden et al. 2003).

ZFNs are made by joining a ZFP to the nuclease domain of the FokI endonuclease (Fn) and are developed in pairs to target sites of the general structure: 5'-(ZFN target site 1)-spacer-(ZFN target site 2)-3'where the target half-sites commonly contain the 5'-GNNGNNGNN-3' sequence motif (Durai et al. 2005, Porteus and Baltimore 2003). The binding of the two ZFNs at the cognate target half-sites allows the Fns to dimerize in the spacer region and create

a DSB (Wah et al. 1998). If the DSB is repaired by HR and an exogenous repair donor is supplied, then the ZFN-targeted locus can be edited to introduce any number of modifications from single point mutations to the installment of large transgenes. Repair of the DSB through error-prone pathways (e.g. NHEJ) can also result in changes to the nucleotide content of the target site (Bibikova et al. 2003, Santiago et al. 2008). ZFNs have been used to perform genomic editing in human cells and are currently undergoing rapid development for the purpose of gene therapy (Cathomen and Joung 2008, Porteus et al. 2006).

In this study we report on the progress made towards engineering new ZFTFs for transcriptional activation and ZFNs that target the *hTERT* locus. We have identified five ZFN full sites (ten target half-sites total) within the *hTERT* promoter and exon 1. Multiple three-fingered ZFPs were generated for each target half-site using the OPEN platform strictly for eight target half-sites and a hybridized methodology that incorporated module fingers at the Finger 1 or Finger 2 positions in the OPEN protocols for the other two target half-sites. The resulting ZFPs were converted into ZFTFs or ZFNs by linking these DNA binding domains to either a VP16 transactivator domain or a Fn respectively. ZFTFs were screened for the ability to upregulate transcriptional activity through co-transfection with an episomal *hTERT* promoter-driven GFP reporter construct in HEK293 cells. We found that the ZFTFs could not only induce GFP expression in a dose-dependent manner, but also be used in combination. ZFN versions of many

of the same ZFPs used as ZFTFs were also screened in sets of combinatorial pairs for each of the target full sites through SSA assays first, and then in a chromosomally-integrated GFP gene targeting reporter assay for one set ZFNs developed to a site nearest the translational start codon of the hTERT gene. Our results present promising preliminary data towards new tools for not only studying the *hTERT* locus, but also developing new methods for approaching telomerase-associated genetic disease.

3.3 RESULTS

3.3.1 Identification of Target Sites in the hTERT Locus

In our search for suitable target sites, we chose to limit our focus to the *hTERT* core promoter and exon 1 sequences for several reasons. First, these sequences are a very GC-rich region of the human genome and expectations are reasonable that a ZFP can be developed to a target half-site if it follows the 5'-GNNGNNGNN-3' sequence motif (Maeder et al. 2009, Maeder et al. 2008). Second, this part of the *hTERT* locus is thought to be where the majority of the transcriptional control occurs and contains a high density of both known and putative regulatory sequences. Finally, for ZFN applications, higher rates of gene conversion are possible if a DSB is delivered in close proximity to the desired site of change (Elliott et al. 1998, Stark et al. 2004). Thus, if genomic editing to the

core promoter or 5' coding regions were to be accomplished, then it would be an advantage to engineer ZFNs to target sites within them.

We identified five full ZFN target sites (ten target half-sites total) either within or in proximity to the *hTERT* core promoter and exon 1 and have listed them in Table 3.1. All target half-sites in the chosen full ZFN target sites follow the 5'-GNNGNNGNN-3' sequence motif, except for the hTERT5SR target half-site, which contains a CNN triplet. Furthermore, the hTERT5 and hTERT5C target sites overlap as a frame shift of three basepairs. For the remainder of this report, we will in general narrow our scope to the hTERT5 and hTERT6 target sites. The hTERT5R and hTERT6L target half-sites have also been described as ZFP target sites 1853L in the *ntSuRB* gene and F-5R (partial) in the *hTERT* promoter, respectively (Sohn et al. 2010, Townsend et al. 2009). Both the hTERT5 and hTERT6 target sites are located in exon 1 of the *hTERT* locus with the hTERT6 site containing the translational start codon within the hTERT6R target half-site (Figure 3.1).

Full Site Name	Half-Site Name	Full Site Sequence	Start	End
hTERT5		5'-AGCCACTACcgcgaGGTGCTGCC-3'	46	68
	hTERT5L	5'-GTAGTGGCT-3'	46	54
	hTERT5R	5'-GGTGCTGCC-3'	60	68
hTERT5C		5'-CACTACCGCgaggtGCTGCCGCT-3'	49	71
	hTERT5CL	5'-GCGGTAGTG-3'	49	57
	hTERT5CR	5'-GCTGCCGCT-3'	63	71
hTERT5S		5'-CCCGCCCTCtcctcGCGGCGCGA-3'	-85	-63
	hTERT5SL	5'-GAGGGCGGG-3'	-85	-77
	hTERT5SR	5'-GCGGCGCGA-3'	-71	-63
hTERT6		5'-CCCGGCCACccccgcGATGCCGCG-3'	-16	8
	hTERT6L	5'-GTGGCCGGG-3'	-16	-8
	hTERT6R	5'-GATGCCGCG-3'	-1	8
hTERT6B		5'-GCCAGCGGCcaaaggGTCGCCGCA-3'	-507	-484
	hTERT6BL	5'-GCCGCTGGC-3'	-507	-499
	hTERT6BR	5'-GTCGCCGCA-3'	-492	-484

Table 3.1: hTERT Target Sites and Sequences.

Five ZFN full target sites with five or six basepair spacer lengths were identified in the *hTERT* promoter and exon 1. Sequences of the target left and right half-sites are provided in the 5'-3' orientation found on either DNA strand and position is given relative to the translational start codon.

FIGURE 3.1

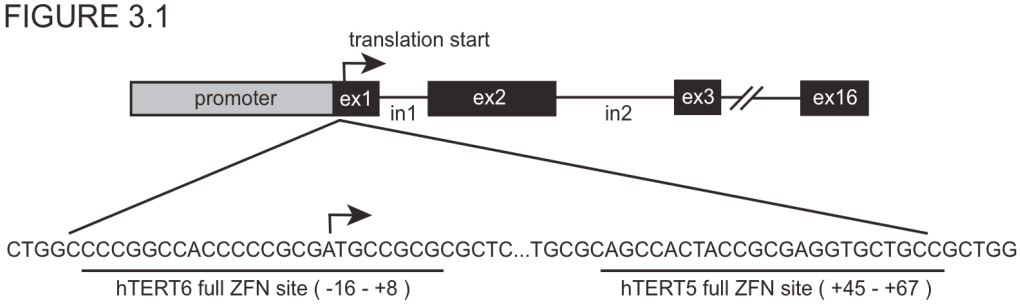

Figure 3.1: Approximate Location and Sequence of the hTERT5 and hTERT6 Sites in the *hTERT* Locus.

3.3.2 Engineering ZFPs by Modular Assembly and OPEN Methods

In order to create three-fingered ZFPs for the ten target half-sites, we utilized three different methods. In general, we used the B2H-based OPEN protocols to generate candidate DNA binding domains for eight of the target half-sites, including those for hTERT5 and hTERT6 full sites (Table 3.2 and Table 3.3). For the hTERT5SL and hTERT5SR half-sites, we installed a module finger (Addgene) in place of one of the zinc fingers within the three-fingered zinc finger libraries used in the OPEN selections (Table 3.3). In addition to those ZFPs generated through OPEN protocols for hTERT6L and hTERT6R target half-sites, we constructed ZFPs made purely by modular assembly (Table 3.2). Thus, we were able to generate multiple candidate ZFPs for each target half-site and within each set of ZFPs, there appears to be a degree of consensus for the amino acids found in the α-helices governing sequence recognition. However, there is also sufficient diversity between individual ZFPs to warrant using screening strategies. The ZFPs generated for the hTERT5R site closely resemble those made by the OPEN protocol for gene targeting at the *ntSuRB* gene, but none were identical (Fu et al. 2009, Townsend et al. 2009). Also, the module fingers used to make the hTERT6L modular assembly ZFP were different from the module fingers in the F-5R ZFP designed to target the same sequence (Sohn et al. 2010). Therefore, our results provide a unique set of custom-engineered, three-fingered ZFPs for the *hTERT* locus.

hTERT5L	**GTA**	**GTG**	**GCT**		
METHOD	**Finger 3**	**Finger 2**	**Finger 1**	**ZFTF**	**ZFN**
OPEN	QSPSLNR	RREVLMN	NAHQLKR	-	pKW741
OPEN	QSTSLQR	RHEVLVN	GRQALDR	-	pKW744
OPEN	QKVSLKR	RKQVLTN	QRQALDR	-	pKW745
OPEN	QIVSLRR	RIEILRN	NGQGLRR	-	pKW746
hTERT5R	**GGT**	**GCT**	**GCC**		
METHOD	**Finger 3**	**Finger 2**	**Finger 1**	**ZFTF**	**ZFN**
OPEN	QNHGLLR	LSQTLKR	DRRTLDR	pKW639	pKW663
OPEN	LMHQLNR	LSQTLKR	DGGTLHR	pKW640	pKW664
OPEN	QNHGLLR	LSQTLKR	DSSTLAR	-	pKW665
OPEN	LMHGLVR	LSQTLKR	DSSTLAR	pKW641	-
OPEN	HRHGLGR	LSQTLKR	DSTTLRR	pKW642	pKW666
OPEN	LMHGLVR	LSQTLKR	DRRTLDR	pKW643	-
OPEN	VMHHLAR	LSQTLKR	DRRTLDR	pKW644	-
OPEN	LMHGLVR	LSQTLKR	ERRGLDR	pKW645	-
OPEN	LMHGLVR	LSQTLKR	DGRGLER	pKW646	-
hTERT6L	**GTG**	**GCC**	**GGG**		
METHOD	**Finger 3**	**Finger 2**	**Finger 1**	**ZFTF**	**ZFN**
modular	RSDALTR	DRSDLTR	RSDHLSR	pKW308	-
OPEN	RADALTR	DSSVLRR	RREKLRI	pKW616	-
OPEN	RADSLTR	DRSVLVR	RQDHLAR	-	pKW632
OPEN	RRDALLR	DTSVLNR	KQDHLNR	pKW618	pKW633
OPEN	RRDLATR	DRSVLAR	KKDHLDR	pKW619	pKW634
OPEN	RADSLTR	DVSVLKR	ARDKLVR	pKW620	pKW635
OPEN	RVDALTR	DVSVLKR	KQDHLAR	pKW621	pKW637
OPEN	RRDALTR	DTSVLTR	RREQLGR	pKW622	pKW638
OPEN	RGDSLTR	DTSVLNR	RKEKLRI	pKW623	-
hTERT6R	**GAT**	**GCC**	**GCG**		
METHOD	**Finger 3**	**Finger 2**	**Finger 1**	**ZFTF**	**ZFN**
modular	QSSNLAR	DRSDLTR	RSDDLTR	pKW309	-
OPEN	VRSNLGR	DRSVLRR	RQHSLRR	pKW599	pKW609
OPEN	VTHNLNR	DTSVLNR	RRHNLDR	pKW598	pKW608
OPEN	LTHNLRR	DRTTLKR	TQHNLDR	pKW600	-
OPEN	LTHNLRR	DPTVLKR	TTHNLKR	pKW601	pKW611
OPEN	LTHNLRR	ESTTLHR	RQHSLRR	pKW602	-
OPEN	LSGNLRR	DRTTLHR	RTHSLAR	pKW603	pKW613
GFP1	**GAA**	**GAT**	**GGT**		
METHOD	**Finger 3**	**Finger 2**	**Finger 1**	**ZFTF**	**ZFN**
P-M 2008	QHPNLTR	VAHNLTR	TRQKLGV	-	GFP-ZFN1
GFP2	**GAC**	**GAC**	**GGC**		
METHOD	**Finger 3**	**Finger 2**	**Finger 1**	**ZFTF**	**ZFN**
P-M 2008	EGGNLMR	DRSNLTR	APSKLDR	-	GFP-ZFN2

Table 3.2: List of Novel Zinc Finger DNA-Binding Domains Generated in This Study for the hTERT5 and hTERT6 Target Sites.

Three-fingered zinc finger arrays were made for the hTERT5L, hTERT5R, hTERT6L, and hTERT6R target half-sites by either modular assembly or the OPEN platform. The amino acid identities of the α-helices used to mediate DNA binding for the hTERT sites and the GFP positive control sites are listed under the target site triplet recognized by the individual finger. These ZFPs were converted into ZFTFs and/or ZFNs and their corresponding designations are also given.

hTERT5CL	GCG	GTA	GTG	
METHOD	**Finger 3**	**Finger 2**	**Finger 1**	**ZFN**
OPEN	RLDMLAR	QSASLRR	RRHILHN	pKW705
OPEN	RKDGLTR	QRSSLRR	RKEVLSR	pKW706
OPEN	RLDMLAR	QSPSLKR	RRHILLN	pKW707
OPEN	RTDSLPR	QGVVLKN	RRHILGN	pKW708
OPEN	RLQGLRP	QRPVLVN	KRDLLRV	pKW709
OPEN	RGQGLQN	QKPILTN	RPSVLHV	pKW710
hTERT5CR	**GCT**	**GCC**	**GCT**	
METHOD	**Finger 3**	**Finger 2**	**Finger 1**	**ZFN**
OPEN	LRASLVR	ESTTLIR	QTADLRR	pKW711
OPEN	LGHTLNR	DSSVLRR	TKQGLQR	pKW712
OPEN	LRHDLQR	DTSVLTR	RKQVLQR	pKW713
OPEN	VSNTLTR	ESTTLVR	ARQGLNR	pKW714
OPEN	VGGSLNR	DRSVLGR	QGQGLVR	pKW715
hTERT5SL	**GAG**	**GGC**	**GGG**	
METHOD	**Finger 3**	**Finger 2 (modular)**	**Finger 1**	**ZFN**
hybrid	RADDLRR	DRSHLAR	RREKLRI	pKW717
hybrid	RRDNLNR	DRSHLAR	RAEKLRI	pKW718
hybrid	RRDNLNR	DRSHLAR	KQDHLNR	pKW719
hybrid	RKDGLTR	DRSHLAR	KGDHLLR	pKW720
hybrid	RSDGLTR	DRSHLAR	KKDHLDR	pKW721
hybrid	RADNLGR	DPGHLVR	KQDHLAR	pKW722
hTERT5SR	**GCG**	**GCG**	**CGA**	
METHOD	**Finger 3**	**Finger 2**	**Finger 1 (modular)**	**ZFN**
hybrid	RPDALPR	RRDTLRR	QLAHLKE	pKW723
hybrid	RLDMLAR	RLDILTR	QLAHLKE	pKW724
hybrid	RSDSLPR	RRDMLRR	QLAHLKE	pKW726
hybrid	REDSLPR	RRDTLRR	QLAHLKE	pKW728
hTERT6BL	**GCC**	**GCT**	**GGC**	
METHOD	**Finger 3**	**Finger 2**	**Finger 1**	**ZFN**
OPEN	DASNLPR	LSQTLNR	TKAKLAI	pKW729
OPEN	AEPLNTG	LSQTLSR	SNSKLAR	pKW730
OPEN	AEPLNTG	LSQTLRR	APSKLAR	pKW731
OPEN	DPSNLRR	LSQSLSR	SNSKLAR	pKW732
OPEN	DKSVLAR	LSQTLRR	APSKLMR	pKW734
hTERT6BR	**GTC**	**GCC**	**GCA**	
METHOD	**Finger 3**	**Finger 2**	**Finger 1**	**ZFN**
OPEN	DMTPLGR	ESTTLVR	STEQLSR	pKW735
OPEN	DHTPLTR	DTSVLTR	DGAQLTR	pKW736
OPEN	EKRGLRR	DRSVLRR	DTGQLAR	pKW737
OPEN	NVSGLRR	DSSVLRR	SREQLDR	pKW738
OPEN	DRSSLRR	DRSVLGR	DRAQLGR	pKW739
OPEN	AKARLLP	DRTTLKR	SPEQLAR	pKW740

Table 3.3: List of Novel Zinc Finger DNA-Binding Domains Generated in This Study for the hTERT5C, hTERT5S, and hTERT6B Target Sites.

Three-fingered zinc finger arrays were made for the hTERT5CL, hTERT5CR, hTERT5SL, hTERT5SR, hTERT6BL, and hTERT6BR target half-sites by either the OPEN platform or module fingers adapted into the OPEN protocols (hybrid). The amino acid identities of the α-helices used to mediate DNA binding for the hTERT sites are listed under the target site triplet recognized by the individual finger. These ZFPs were converted into ZFNs and their corresponding designations are also given.

3.3.3 hTERT ZFTFs Can Stimulate Transcription of a GFP Reporter Through the hTERT Promoter

Many of the ZFPs generated for the hTERT5R, hTERT6L, and hTERT6R target half-sites were converted into ZFTFs (23 total) by linking the VP16 transcriptional activation domain to the N-terminus of the candidate ZFP (Triezenberg et al. 1988). The ability of the resulting candidate ZFTF to up-regulate transcription in a site-specific manner was tested through the use of an extra-chromosomal GFP reporter plasmid. In this system, approximately 3 kb of the *hTERT* promoter and exon 1 were cloned upstream of a GFP reporter gene (pMC6). We hypothesized that the candidate ZFTF would bind to the cognate target site within the *hTERT* promoter-exon 1 sequence and recruit the transcriptional machinery necessary to drive GFP expression (Figure 3.2A). We then co-transfected the ZFTF-expressing plasmids and pMC6 into HEK293 cells and measured GFP expression by flow cytometry. Our results show that while the *hTERT* promoter-driven, extrachromosomal GFP reporter was a leaky system, in the presence of a ZFTF, the stimulation of %GFP+ cells over background could be as high as ~40-fold (Figure 3.2B).

In our screen of all ZFTFs made, we found that stimulation of GFP expression could vary significantly between individual ZFTFs (Figure 3.2C). Overall, the ZFTFs made with ZFPs from the OPEN platform performed better than those made purely by modular assembly, although that was not always the

FIGURE 3.2

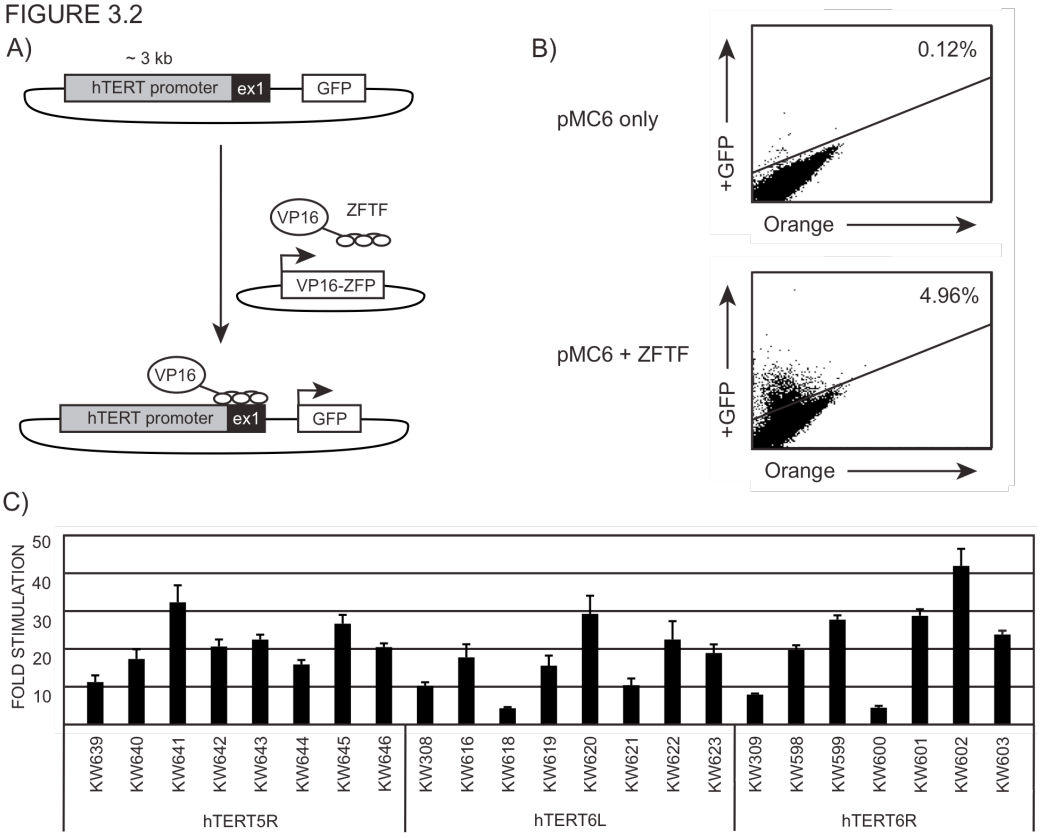

Figure 3.2: Screening hTERT ZFTFs For Transcriptional Activation Activity.

A) Schematic of the pMC6 transcriptional reporter, which contains 3 kb of the *hTERT* promoter and exon 1 sequence upstream of a GFP reporter gene. When co-transfected with a ZFTF expressing plasmid, the ZFTF will bind to the pMC6 reporter plasmid and recruit the cellular machinery necessary to express GFP. B) Flow cytometry plots demonstrating the increase in GFP+ cells after co-transfection of 100 ng of the MC6 reporter plasmid with 700 ng of plasmid expressing the KW602 ZFTF in HEK293 cells. C) All ZFTFs listed in Table 3.2 were screened for transcriptional activation activity. Our results identify which ZFPs efficiently recognize and bind the hTERT target half-sites as ZFTFs. Error is reported as the standard deviation.

case (compare KW308 and KW309 to KW618 and KW600 respectively) (Ramirez et al. 2008). The success or failure of a ZFTF to stimulate GFP expression on this extra-chromosomal reporter could be the result of multiple factors. First, variations in expression from each ZFTF may cause the differences in reporter stimulation. Second, in the case of the hTERT6R ZFTFs, the native translational start codon found within exon 1 had been mutated from ATG to GTG to enforce translation at the GFP reporter gene and this changes the hTERT6R target half-site from 5'-GATGCCGCG-3' to 5'-GGTGCCGCG-3'. While the mutation does not drastically alter the hTERT6R target half-site, it may generate a bias in the transcriptional activation activities of ZFTFs depending on the specificity of the ZFP used for the hTERT6R sequence. Finally, even though many of the ZFPs used to make ZFTFs conform to a general amino acid consensus and allow the candidate ZFP to survive the B2H-based OPEN protocol, the individual amino acid deviations may still subtly influence the DNA binding properties of the ZFP to yield the variations seen in our ZFTF screen (Table 3.2).

Of the 23 ZFTFs screened, we chose the best performing ZFTFs from each target half-site set (KW602, KW620, and KW641) and tested them for a dose-dependent response. We first co-transfected single ZFTFs in increasing amounts with the pMC6 reporter plasmid and saw that as more ZFTF was transfected, the more GFP expression we were able to observe (Figure 3.3). Then, we began transfecting the three ZFTFs in combination as pairs and finally all

FIGURE 3.3

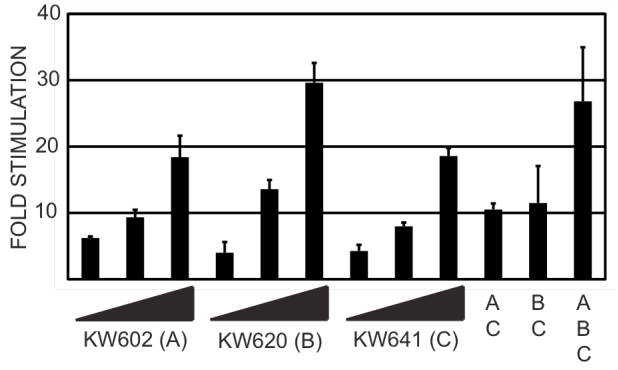

Figure 3.3: hTERT ZFTFs Exhibit a Dose-Dependent Response.

The best performing ZFTF for each target half-site were transfected in increasing amounts of ZFTF-expressing plasmid (200, 400, and 600 ng) with 100 ng of pMC6 into HEK293 cells. The KW602, KW620, and KW641 ZFTFs were also co-transfected in pairs or all three combined (200 ng each). The increase in transcriptional activity between pairs of ZFTFs and in combination is statistically significant (student's t-test, $p<0.05$). Error is reported as the standard deviation.

three together in equal amounts. Again, we observed an additive response from the pMC6 reporter as more ZFTF-expressing plasmid was co-transfected (Figure 3.3). These data further reinforce previous findings demonstrating that ZFTFs can elicit a tunable response from a target gene and suggests that this may be possible for the *hTERT* locus as well (Magnenat et al. 2008, Snowden et al. 2003).

3.3.4 hTERT ZFNs Can Initiate Repair of an Extra-Chromosomal GFP Reporter by SSA

To further test the ZFPs engineered to site-specifically target the hTERT locus, we converted a total of 39 ZFPs into ZFNs by installing a wild-type Fn to the C-terminus of the ZFP for all target sites (Table 3.2 and Table 3.3). Due to variations in activity seen in the ZFTFs and individual deviations from the consensus sequence in multiple ZFPs, we again designed a screening strategy to test pairing of ZFNs in a combinatorial manner for each target site. To determine the ability of a ZFN pair to recognize, bind, and cut the cognate target site, we constructed four GFP-based SSA reporter plasmids where a full hTERT ZFN target site would be inserted between two repeats within the GFP gene. The insertion also includes the full GFP1/2 site, which can be targeted by the GFP-ZFN1 and GFP-ZFN2 ZFN pair as a positive control and internal standard. In this assay, co-transfection of the GFP-based SSA reporter plasmid with ZFN expressing plasmids can result in the formation of a functional GFP gene after a DSB is delivered to the target sites that separate the repeated sequences and is

repaired by the cellular SSA repair machinery to restore function of the GFP reporter gene (Figure 3.4). We report hTERT ZFN activity as a relative percentage of the nuclease activity measured for the GFP-ZFN1/2 internal standard and positive control (Table 3.4-Table 3.8).

After surveying all five combinatorial sets of pairings, we found that the hTERT5 and hTERT6 ZFN sets were, broadly, the most active sets of ZFNs (Table 3.4-Table 3.5). For the ZFNs developed for the hTERT5 target site, the KW744/KW664 ZFN pairing exhibited equal nuclease activity to the GFP-ZFN1/2 positive control pair. Also, eight out of 16 possible pairings were shown to have more than 50% activity relative to the control (Table 3.4). The hTERT6 ZFN set produced the most active nucleases out of those tested. In this combinatorial screen, the KW633/KW609 hTERT6 ZFN pair was more active than the GFP-ZFN1/2 standard by three-fold. We also observed that all but six pairs (associated with the hTERT6R KW610 ZFN) demonstrated nuclease activity that was higher than the GFP-ZFN1/2 standard (Table 3.5). Therefore, we find these ZFN screens, based on SSA assays, to have identified highly active ZFN pairs targeting sequences found at the hTERT locus.

FIGURE 3.4

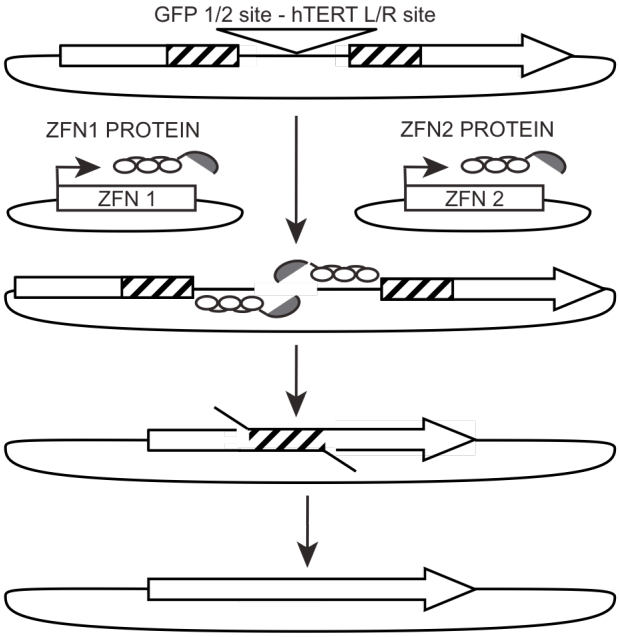

Figure 3.4: Diagram of the GFP-Based SSA Assay.

For the purpose of screening nuclease activities of various ZFN pairs, we inserted the GFP1/2 ZFN full site and a hTERT ZFN full site between two repeats within a GFP reporter gene. Co-transfection of this reporter plasmid (20 ng) with ZFN-expressing plasmids (100 ng each) into HEK293 cells results in the delivery of a DSB between the repeated GFP sequences which is then repaired by the endogenous SSA machinery to produce a functional GFP reporter gene.

	GFP-ZFN2	KW663	KW664	KW665	KW666
GFP-ZFN1	100%	-	-	-	-
KW741	-	56.8	61	62.9	63.9
KW744	-	74.5	103.5	77.7	66.6
KW745	-	37.1	32.2	13.3	23.7
KW746	-	32.8	34.3	24.9	29.8

Table 3.4: Combinatorial Screen of hTERT5 ZFN Pairs in the SSA Assay.

Four hTERT5L ZFNs and four hTERT5R ZFNs were paired in combinatorial fashion and co-transfected as ZFN-expressing plasmids (100 ng each) into HEK293 cells with an SSA reporter plasmid bearing the hTERT5 ZFN full site (20 ng, Figure 3.4). All hTERT ZFN nuclease activities were normalized as a relative percentage of the GFP-ZFN1/2 positive control pair (set to 100%).

	GFP-ZFN2	KW608	KW609	KW610	KW611	KW613
GFP-ZFN1	100%	-	-	-	-	-
KW632	-	105.9	193.8	58.8	166.2	121.7
KW633	-	280.6	297.9	34.6	254.6	208.4
KW634	-	227.6	202.7	11.8	183.2	141.5
KW635	-	239.3	171.1	46.2	231.3	194.5
KW637	-	280.9	238	14.7	238.1	222.6
KW638	-	236	279.6	22.2	280.1	225.8

Table 3.5: Combinatorial screen of hTERT6 ZFN Pairs in the SSA Assay.

Six hTERT6L ZFNs and five hTERT6R ZFNs were paired in combinatorial fashion and co-transfected as ZFN-expressing plasmids (100 ng each) into HEK293 cells with an SSA reporter plasmid bearing the hTERT6 ZFN full site (20 ng, **Figure 3.4**). All hTERT ZFN nuclease activities were normalized as a relative percentage of the GFP-ZFN1/2 positive control pair (set to 100%).

	GFP-ZFN2	KW711	KW712	KW713	KW714	KW715
GFP-ZFN1	100%	-	-	-	-	-
KW705	-	10.6	13.2	2.4	0	2.6
KW706	-	3.7	20.1	14.5	0	10.3
KW707	-	0	0.8	0	0	2.1
KW708	-	0	6.3	0	0	6.3
KW709	-	0	0	0	0	0.8
KW710	-	0	0	0	0	0

Table 3.6: Combinatorial screen of hTERT5C ZFN Pairs in the SSA Assay.

Six hTERT5CL ZFNs and five hTERT5CR ZFNs were paired in combinatorial fashion and co-transfected as ZFN-expressing plasmids (100 ng each) into HEK293 cells with an SSA reporter plasmid bearing the hTERT5CL ZFN full site (20 ng, Figure 3.4). All hTERT ZFN nuclease activities were normalized as a relative percentage of the GFP-ZFN1/2 positive control pair (set to 100%).

	GFP-ZFN2	KW723	KW724	KW726	KW728
GFP-ZFN1	100%	-	-	-	-
KW717	-	9.1	20.8	44.8	81.7
KW718	-	13.8	20.4	67.9	15.6
KW719	-	10.2	46	81.9	12.9
KW720	-	52.2	3.2	13.4	20.8
KW721	-	12.3	41.2	78.3	11.2

Table 3.7: Combinatorial screen of hTERT5S ZFN Pairs in the SSA Assay.

Five hTERT5SL ZFNs and four hTERT5SR ZFNs were paired in combinatorial fashion and co-transfected as ZFN-expressing plasmids (100 ng each) into HEK293 cells with an SSA reporter plasmid bearing the hTERT5S ZFN full site (20 ng, Figure 3.4). All hTERT ZFN nuclease activities were normalized as a relative percentage of the GFP-ZFN1/2 positive control pair (set to 100%).

	GFP-ZFN2	KW735	KW736	KW737	KW738	KW739	KW740
GFP-ZFN1	100%	-	-	-	-	-	-
KW729	-	0	0	0	0	0	0
KW730	-	0	0	0	5	0	0
KW731	-	0	8.1	0	19.6	0	8.1
KW732	-	19.9	22.9	0	19.6	0	0
KW734	-	0	0	0	0	0	0

Table 3.8: Combinatorial screen of hTERT6B ZFN Pairs in the SSA Assay.

Four hTERT6BL ZFNs and four hTERT6BR ZFNs were paired in combinatorial fashion and co-transfected as ZFN-expressing plasmids (100 ng each) into HEK293 cells with an SSA reporter plasmid bearing the hTERT5 ZFN full site (20 ng, Figure 3.4). All hTERT ZFN nuclease activities were normalized as a relative percentage of the GFP-ZFN1/2 positive control pair (set to 100%).

3.3.5 Gene Targeting by the hTERT6 ZFNs at a Stably-Integrated GFP Reporter Construct

Due to the high levels of nuclease activity seen in the SSA assays for the ZFNs developed to the hTERT6 site, we chose to further study this set of ZFNs by assaying for ZFN-stimulated homologous recombination at a genomic target. Based on a previously described system, we designed another GFP-based reporter construct that was rendered non-functional through the insertion of an hTERT full ZFN target site and a I-SceI homing endonuclease recognition site as an internal standard (Porteus and Baltimore 2003). This insertion does not disrupt the endogenous GFP1/2 target site that can be found in the GFP gene and the GFP-ZFN1/2 pair was used as a positive control in this assay (Figure 3.5). Briefly, we generated a monoclonal HEK293 cell line containing the stably integrated reporter construct and co-transfected the same amount of ZFN-expressing plasmids in pairs used in the SSA assay screen along with a repair donor plasmid. Function of the GFP gene was restored if the ZFNs cut their cognate target site and stimulated HR-mediated repair with the donor plasmid. Again, ZFN activity is normalized as a relative percentage of the activity measured for the internal standard, I-SceI (Table 3.9). In this assay, we found the GFP-ZFN1/2 positive control pair to have ~80% activity relative to I-SceI and is comparable with a previous reports of this ZFN pair in similar assays, thus verifying the robustness of this particular type of gene targeting reporter (Pruett-Miller et al. 2008).

FIGURE 3.5

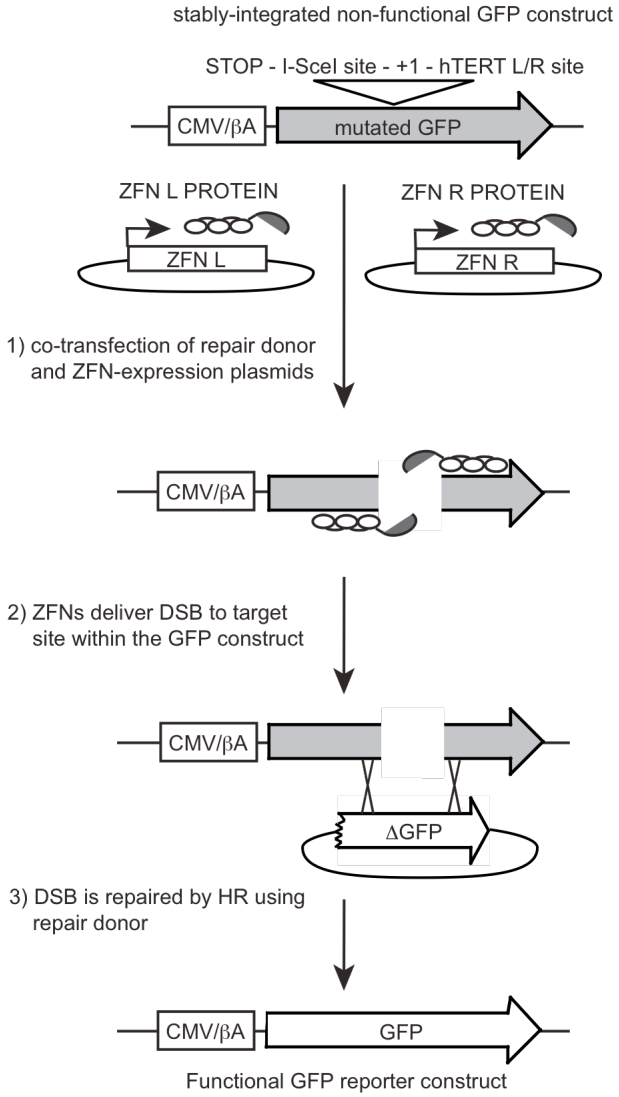

Figure 3.5: Schematic of the GFP-based Gene Targeting Reporter Cell Line.

Based on a previously described strategy, we rendered a GFP reporter gene non-functional through the insertion of a stop codon, an I-SceI recognition site, a frame shift nucleotide, and the hTERT6 ZFN full site. Using this construct, we generated a stably-integrated monoclonal HEK293 cell line in which we co-transfected pairs of ZFN-expressing plasmids and a repair donor plasmid. Nuclease activity of the ZFN pair is measured by the ZFN-mediated HR repair event between the cut GFP reporter and the donor plasmid producing a functional GFP reporter gene.

I-SceI	100%					
GFP-ZFN1/2	79.4 (± 18.7)					
	KW99	**KW608**	**KW609**	**KW610**	**KW611**	**KW613**
KW98	0	5.3 (± 0.6)	5.6 (± 4.6)	0	1.8 (± 0.6)	1.7 (± 1.7)
KW632	0	27.8 (± 15.8)	18.3 (± 16.6)	0	12.8 (± 6.2)	11.1 (± 4.4)
KW633	0	8.4 (± 2.9)	14.7 (± 6.5)	0	33.9 (± 12.1)	26.7 (± 9.3)
KW634	0.7 (± 0.6)	16.3 (± 8.4)	3.7 (± 1.9)	0	19.4 (± 5.8)	10.4 (± 3.2)
KW635	0	27.9 (± 14.8)	15.7 (± 2.2)	0	21.3 (± 1.9)	41.7 (± 16.9)
KW637	0	18.7 (± 4.8)	18.3 (± 10.5)	0	21.7 (± 7.9)	30.0 (± 14.9)
KW638	0	35.5 (± 18.8)	11.9 (± 4.1)	0	34.9 (± 14.6)	26.7 (± 11.1)

Table 3.9: Combinatorial Screen of hTERT6 ZFN Pairs in the Gene Targeting Assay.

Six hTERT6L ZFNs and five hTERT6R ZFNs were paired in combinatorial fashion and co-transfected as ZFN-expressing plasmids (100 ng each) along with a repair donor plasmid (300 ng) into a monoclonal HEK293 cell line bearing the GFP-based reporter construct described in Figure 3.5. All hTERT ZFN nuclease activities were normalized as a relative percentage of the nuclease activity measured for the I-SceI internal standard (set to 100%). Error is reported as SEM.

In repeating the hTERT6 ZFN combinatorial pairing screen, we found that all hTERT6 ZFN pairs stimulated less gene targeting than the GFP-ZFN1/2 positive control. The best performing pair, KW635/KW613, produced ~40% activity relative to I-SceI (Table 3.9). We again observed that, in general, the ZFNs made from ZFPs with purely module fingers showed low activity with this screening strategy when compared to nucleases from which the ZFPs were made by OPEN methodologies (Ramirez et al. 2008). The KW610 ZFN, made from the same ZFP used to make the KW600 ZFTF, produced no evidence of nuclease activity in this gene targeting assay even though we observed measurable activity in the SSA assay screens. Also, despite the observation all other hTERT6 ZFN pairs tested in the SSA assays show nuclease activity higher than the GFP-ZFN1/2 standard, many pairings show low activity in the GFP-based gene targeting assay. One reason may be that a stably-integrated, single copy target construct of the gene targeting assay is a more difficult target at which to stimulate DNA repair than an extra-chromosomal, potentially multi-copy target construct. Thus, by including the GFP-based gene targeting assay into our hTERT6 ZFN screens, we were able to narrow the list of candidate ZFN pairs that warrant further characterization and development for gene targeting at the endogenous *hTERT* locus.

3.4 DISCUSSION

Due to the role hTERT plays in telomere regulation, the *hTERT* locus has become an important target gene for research and drug development (Philippi et al. 2010). The ability to direct *hTERT* transcription in a dose-dependent manner or edit the nucleotide content of the locus could yield many potentially useful insights into the regulation of this gene and cellular aging. To this end, we have developed many ZFPs that recognize sequences within the *hTERT* promoter and exon 1 (Table 3.1 and Figure 3.1). We built the ZFPs using modular assembly methods (Addgene), the B2H-based OPEN platform, and a hybrid method that combines the two protocols (Maeder et al. 2009, Maeder et al. 2008). This resulted in several candidate ZFPs for every target half-site and many of them are consistent with a consensus in the amino acids that govern nucleotide recognition (Fu et al. 2009). We screened the ZFPs for the ability to recognize and direct activity to the cognate target site by converting them to ZFTFs or ZFNs (Table 3.2 and Table 3.3).

As ZFTFs, we were able to demonstrate that many of our ZFPs linked to the VP16 effector domain could stimulate the transcription of a hTERT promoter-exon1 driven GFP reporter gene (Figure 3.2). We were also able to demonstrate that ZFTFs developed for sites in the *hTERT* locus could elicit transcriptional activation in dose-dependent manner in which incremental increases in ZFTF produced an additive response from the GFP-based reporter (Figure 3.3). This is

consistent with the findings of others using ZFTFs at other gene targets (Magnenat et al. 2008, Snowden et al. 2003). Our ZFTF screen also identifies some well-performing candidate ZFTFs (e.g. KW602, KW620, and KW641) that may be pursued for future research in endogenous hTERT transcriptional regulation. In comparison to the ZFTFs made by modular assembly described in Sohn et al., which were demonstrated to be effective regulators of hTERT transcription, we found the KW308 and KW309 modular assembly ZFTFs to be less effective at regulating GFP transcription than the OPEN-generated ZFTFs (Sohn et al. 2010). The differences in efficacy for modular assembly ZFPs as ZFTFs may be due to the differences in the module fingers used to construct the ZFPs, the effector domains, and design of the reporter assays. Taking these two studies together, we still believe that our results provide an informed contribution to the repertoire of ZFTFs that could target the *hTERT* locus.

To continue testing the candidate ZFPs, but in nuclease form, we then screened sets of ZFN pairs in a combinatorial fashion for each hTERT target site identified using GFP-based SSA assays (all sites) and gene targeting assays (hTERT6 only) (Figure 3.4). ZFNs to the hTERT5 and hTERT6 site were the most active sets of nucleases in the SSA assays (Table 3.4-Table 3.8). Further screening of the hTERT6 ZFNs in our gene targeting assays have suggested which ZFN pairs may best stimulate DNA repair at a chromosomal target (Figure 3.5 and Table 3.9). Future development of these ZFNs for genomic editing will

require demonstrating efficacy for ZFN-mediated homologous recombination at the endogenous *hTERT* locus.

When assessing the performance of our candidate ZFPs across multiple screening assays, we found it difficult to reliably use the results of one assay to predict the results of another. According to our results, a low quality ZFP would yield low activity in many of the screens. For example, the hTERT6 ZFPs fully made by modular fingers stimulated some of the lowest levels of transcriptional activation as ZFTFs and nuclease activity in the gene targeting assay as ZFNs (Table 3.2, Figure 3.2C, and Table 3.9). A similar result was seen for the ZFP used to make the KW600 ZFTF and KW610 ZFN (Table 3.2, Figure 3.2C, Table 3.8-Table 3.9). However, we could not establish a threshold of ZFTF activity that could predict which ZFPs could produce highly efficient ZFNs, or vice versa. When comparing the relative ZFTF vs. ZFN activity of KW618 vs. KW633, KW640 vs. KW664, and KW647 vs. KW666, no extrapolation can be formed (Table 3.2, Figure 3.2C, and Table 3.4-Table 3.5 and Table 3.9). Even between the two assays used to screen hTERT6 ZFNs, where the same amount of ZFN-expressing plasmids were transfected into cells, different pairs of ZFNs exhibited the highest nuclease activity. The comparison of the KW633/KW609 ZFN pair in SSA assay and KW635/KW613 ZFN pair in the gene targeting assay particularly highlights this observation (Table 3.4-Table 3.5). Therefore, when developing novel ZFPs for use at the *hTERT* locus, a good strategy may be to not assume that

one ZFP can be used for all applications, but that ZFPs ought to be screened in assays that most closely match the intended application to identify the most appropriate ZFP for use, whether as a ZFTF or a ZFN.

In summary, the data generated by our screening strategies provide the necessary preliminary results to move forward with using these ZFTFs and ZFNs as potential tools to modulate *hTERT* expression and edit the genomic locus. However, many of these ZFPs do require further characterization, such as demonstrating effectiveness at eliciting changes at the endogenous locus. Therefore, we do strongly encourage the pursuit of using artificial ZFP-based proteins for not only researching *hTERT* regulation, but also as an emergent technology in alleviating telomerase-associated disease.

3.5 MATERIALS AND METHODS

3.5.1 Generating Zinc Finger DNA Binding Domains to Target Sites at the hTERT Locus

Five full ZFN target sites were identified in the *hTERT* core promoter and exon 1 and are listed in Table 3.1. We used the publically available OPEN platform to construct three-fingered ZFP arrays for each target half-site (Maeder et al. 2009, Maeder et al. 2008). For the hTERT5SL and hTERT5SR target half-sites, we generated ZFPs by including a module finger at the Finger 2 and Finger 1 positions (respectively) during the creation of the three-fingered cassettes used

in the stage B2H selections (Addgene). In addition to the OPEN-based ZFPs made for the hTERT6L and hTERT6R target half-sites, we constructed ZFPs by assembling module fingers at each zinc finger position (KW308 and KW309). For each target half-site, the OPEN protocols produced multiple candidate ZFPs, some of which were chosen for further screening (Table 3.2 and Table 3.3).

3.5.2 Cloning ZFTFs and ZFNs

To create ZFTFs, chosen candidate ZFPs made by the OPEN-based methods were cloned into the PRK5.AD-GFZ3 vector provided by Toni Cathomen to link the VP16 transcriptional activator domain to the N-terminus of the ZFP (Alwin et al. 2005). ZFNs were made by cloning many of the same ZFPs into the previously characterized GFP-ZFN2-B2H vector to link a wild-type Fn domain to the C-terminus of the ZFP (Pruett-Miller et al. 2008). All cloning was performed using standard molecular biology techniques.

3.5.3 Screening hTERT ZFTFs For Transcriptional Activation Activity

Approximately 3 kb of the *hTERT* promoter and exon 1 were subcloned upstream of a GFP reporter. In this plasmid construct (pMC6), the endogenous hTERT translational start codon was mutated from ATG to GTG, but the start codon for GFP was preserved. All ZFTFs were screened for transcriptional activation activity by co-transfecting 700 ng of ZFTF-expressing plasmid with 100 ng of pMC6 into HEK293 cells. GFP transcription was assayed by measuring

%GFP+ cells by flow cytometry (FACS Calibur, Becton-Dickerson, San Jose, CA, USA). pMC6 was co-transfected with a blank vector as a negative control. Then, one ZFTF from each target half-site group (three total) was co-transfected singularly in increasing amounts (200-600 ng) and in combination (200 ng each) with pMC6 to assay for any dose response. Fold stimulation of transcriptional activation was calculated as (experimental/background) and error is reported as standard deviation.

3.5.4 SSA Assays For Screening hTERT ZFN Pairs

Based on a previously described SSA strategy, reporters were cloned by inserting a hTERT full ZFN target site between repeated sequences within GFP gene (four total) (Szczepek et al. 2007). Each of these SSA reporters contains a single hTERT target site except for the hTERT5 reporter, which also carries the hTERT5C site due to the overlapping sequence shared (Table 3.1). The reporter constructs also include the GFP1/2 full ZFN target site (5'-ACCATCTTC-ttcaag-GACGACGGC-3') as a positive control and internal standard. These extrachromosomal SSA reporter plasmids were used to assay for on-target nuclease activity for sets of combinatorial pairs of ZFNs developed for each hTERT target site. Briefly, 100 ng of each ZFN-expressing plasmid and 20 ng of appropriate reporter plasmid were co-transfected into HEK293 cells and the GFP1/2 ZFN pair was used as a positive control. Repair of the GFP gene, after cutting by a pair of ZFNs, was measured by %GFP+ cells using flow cytometry.

The nuclease activities of all hTERT ZFN pairs tested were normalized to the activity of the positive control and are reported as a percentage of GFP1/2 ZFN activity.

3.5.5 Gene Targeting Assay For hTERT6 ZFNs

The hTERT6 ZFNS were tested for the ability to stimulate HR at a chromosomal target by using a previously described GFP gene targeting assay (Porteus and Baltimore 2003). This assay uses a GFP reporter that has been mutated by the insertion of a hTERT full ZFN target site and an I-SceI recognition site as a positive control and internal standard. Within the GFP gene, the GFP1/2 full ZFN target site can be found as well. This reporter construct was electroporated into HEK293 cells and a monoclonal cell line was derived. ZFN-expressing plasmids were co-transfected into the monoclonal cell line in pairs (100 ng each) for a combinatorial screen with 300 ng of repair donor plasmid. Nuclease-mediated gene targeting events were measured as %GFP+ cells by flow cytometry. The nuclease activities of all ZFN pairs tested were normalized to the I-SceI positive control and are reported as a relative percentage of that activity. Error is reported as standard error of the mean.

3.6 ACKNOWLEDGEMENTS

We would like to thank Sierra Sciences for cloning vectors containing the hTERT promoter and exon 1 sequences and Toni Cathomen for the PRK5.AD-

GFZ3 plasmid. Also, we are grateful to Tracy Chow for her helpful discussions regarding cell culture and telomerase biology.

This work was supported by National Institutes of Health grants R01 HL079295, The March of Dimes, a career development award from the Burroughs Welcome Fund, and funding through University of Texas Southwestern Medical Center by the State of Texas.

CHAPTER 4:CONCLUSIONS AND FUTURE DIRECTIONS

4.1 CONCLUSIONS

4.1.1 Inter-Domain Linkers

Given the diversity of sequence structure and GC content across different genes, the flexibility to choose appropriate ZFN full target sites is very important. Several groups have attempted to survey the affects of a variety of inter-domain linkers in ZFNs at target sites with a range of spacer lengths to elucidate how to best engineer a ZFN to the parameters of a particular target site (Bibikova et al. 2001, Handel et al. 2009, Shimizu et al. 2009). Unfortunately, it is difficult to do a comprehensive meta-analysis on the data generated by these groups because of the different inter-domain linkers surveyed and assays performed to measure ZFN activity. However, what is becoming clear is that spacer lengths of 5-7 bp can be easily targeted by ZFNs given the choice of inter-domain linker that accommodates the target site conditions. Target sites of longer spacer lengths may be targeted by ZFNs as well given the periodicity of the DNA helix. The contribution that our inter-domain linker survey gives to the discussion of how to alter ZFN architecture comes from the recapitulation of the work done by Handel et al. (2009) that spacer lengths of 5-7 bp can be targeted by ZFNs with relatively short inter-domain linkers in human cells (Handel et al. 2009). Also, new inter-domain linkers are characterized and presented as new options in engineering

ZFN architecture. Previously characterized inter-domain linkers, such as the 4 aa LRGS and 5 aa AAARA sequence, were also active and selective for target site spacer lengths as seen in previous studies (Alwin et al. 2005, Handel et al. 2009, Urnov et al. 2005). Therefore, this subset of data allows for the probability of finding a suitable ZFN full target site at 1 in 216 bp. The ability to find a target site with a 5-7-bp spacer should place a ZFN target site within a range of distance from the site of desired change in which it would be reasonable to expect efficient rates of gene targeting (Elliott et al. 1998).

4.1.2 Inter-Finger Linkers

When selecting target sites for ZFP development, particularly long target sites, there is the distinct possibility that a discontinuous target site could be found. Extra basepairs between target subsites (or recognition triplets) is common and it may be necessary to find a ZFP architectural solution to efficiently target such sites. Based on previous findings that ZFPs could be developed to target sites with 1-bp insertions between recognition triplets given the use of a lengthened inter-finger linker, we attempted to survey the activity inter-finger linker variant ZFNs at 10-bp target sites using a 3-fingered ZFP platform (Moore et al. 2001a, Moore et al. 2001b). The data were able to support previous findings that 3-fingered ZFP-based proteins can have activity at discontinuous target sites where the insertions are thymine, but not guanine, or adenine. We also observed

significant levels of non-specific activity in ZFNs engineered to these discontinuous target sites with lengthened inter-finger linkers.

4.1.3 Obligate Heterodimer Fns

Any ZFN architectural modification that can increase on-target cutting and decrease off-target cutting is a welcomed step forward in the ZFN field. The rational design of obhetFns creates a system where the Fn dimerization events are enforced and should occur only between two complementary Fn domains of two ZFNs bound to the intended target site. Work done by Miller et al. (2007) and Szczepek et al. (2007) has demonstrated that ZFNs with obhetFns show increased on-target cutting and decreased cellular toxicity due to off-target cutting (Miller J. C. et al. 2007, Szczepek et al. 2007). In our use of the KK/EL obhetFns, we did see reduced toxicity when compared to analogous ZFNs with the wild-type Fn, but we also observed reduced on-target cutting as well. Our lab has previously reported similar findings with the DD/RR obhetFn forms (Pruett-Miller et al. 2008). It is difficult to know whether the reduced toxicity of the KK/EL obhetFn ZFN variants is truly due to heterodimerization preventing off-target cutting, or simply due to reduced nuclease activity overall. The source of the discrepancy between our results and those in previously published literature may come from the different ZFPs used in the ZFNs tested. Bitinaite et al. (1998) determined through structure analysis and biochemical assays that the presence of the second FokI endonuclease bound to the target site stabilizes the Fn dimerization event

(Bitinaite et al. 1998). In the ZFN platform, two ZFNs made from two high quality ZFPs may be required in order to sufficiently stabilize the obhetFn dimerization event to produce the increased rates of gene targeting with decreased cellular toxicity. The methods necessary to test such a hypothesis are beyond the scope of those used in our studies.

4.1.4 Hybrid Methods For Generating ZFPs

It cannot be overemphasized that building a ZFP with high affinity and high specificity for the target site is one of the most important steps in successfully engineering a ZFP-based chimeric protein. The two most characterized methods for building ZFPs are modular assembly and the OPEN platform (Gonzalez et al. 2010, Maeder et al. 2009). In light of the high failure rates seen in purely modular ZFPs and the target subsite limits in the OPEN platform, a new engineering solution was needed to produce active ZFPs across a wider range of potential target sites. To this end, we hybridized the two methods by incorporating module fingers into the B2H protocols of the OPEN platform. We produced successful ZFNs made from hybrid ZFPs to four different sites. Our work provides the proof-of-concept evidence that elements of each method can be synthesized to overcome some of the disadvantages of using the original protocols. In combination with the possibility of targeting ZFN target sites with 5-7-bp spacers, the probability that a usable target site can be found has increased to 1 in 17 bp, essentially almost anywhere in a target locus.

4.1.5 Making hTERT ZFPs

Using B2H-based methods to engineer ZFPs can result in multiple candidates for future use in chimeric proteins (Maeder et al. 2009, Maeder et al. 2008). Our methods of using modular assembly, OPEN, and hybridized methodologies yielded several ZFPs for each 9-bp target half-site selected. The advantage of having options in the choice of ZFP becomes a technical challenge in determining which ZFP is best suited for the intended purpose. The candidate ZFPs from the B2H-based protocols showed a consensus sequence in the a-helices that mediate sequence recognition, especially in the -1, 1, 3, and 6 amino acid positions which suggests a particular set of amino acids are useful for recognizing the target site. The candidate ZFPs also showed subtle diversity, especially in positions 2 and 4 ,which mediate minor recognition contacts. The diversity seen in the output of the OPEN platform prompted us to enact screening strategies to determine which ZFPs, from a set of candidates, are most capable of targeting the cognate target site.

4.1.6 hTERT Zinc Finger Transcription Factors

We first fused most of the candidate ZFPs to the VP16 transcriptional activation domain and screened the resulting ZFTFs for the ability to upregulate the expression of a reporter gene on transiently transfected reporter plasmids. Our results showed high variability in the upregulation of the reporter gene between ZFTFs made by the same method to the same target site. This finding supports the

strategy of screening candidate ZFPs. The best ZFTFs from each set of candidates were then used in increasing amounts and in combination. We observed a dose-dependent response that is consistent with previously published ZFTFs (Magnenat et al. 2008, Snowden et al. 2003). Therefore, we identified ZFTFs that can efficiently upregulate gene expression through hTERT promoter sequences.

4.1.7 hTERT Zinc Finger Nucleases

Many of the candidate ZFPs were also fused to Fns and screened for ZFN activity on reporter constructs. We screened combinatorial pairs of ZFNs on extra-chromosomal and stably integrated repair constructs. The two ZFN screening strategies were able to identify high-performing pairs of ZFNs. However, the "best" identified ZFN pairs were different in each nuclease assay. The performance of a ZFP as a ZFTF or a ZFN could not be correlated or predicted across the screening strategies used. The results of our screens have highlighted the need for testing ZFP-based chimeric proteins with assays that best reflect the intended application.

4.2 FUTURE DIRECTIONS

As it always is in scientific investigation, one answered question invariably leads to more questions.

4.2.1 The Effects of Inter-Domain Linkers

Our survey of inter-domain linkers in ZFNs at target sites of variable spacer lengths has revealed that the length and amino acid content of the inter-domain linker can affect nuclease performance. What is not clear is how this short stretch of amino acids is mediating nuclease activity. Do the inter-domain linkers factor into the spatial issues of Fn dimerization and accommodate steric boundaries? Also, does amino acid content of the inter-domain linker affect the flexibility or rigidity of this architectural element? If so, how does this affect ZFN activity? Some aspects of these questions can be answered through a resolved crystal structure of ZFNs bound as a pair to a DNA sequence. Structural data could elucidate the spatial limits of ZFN-binding and provide insight into how inter-domain linkers influence Fn dimerization. The contributions of structure in inter-finger linkers to the entropy and binding energies of a ZFP or ZFTF have been investigated, but have not yet been investigated or applied to the inter-domain in a linker ZFN platform (Moore et al. 2001a).

4.2.2 The Effects of Inter-Finger Linkers

In addition to the biophysical aspects of linkers in a ZFP described above, more questions arise regarding the relationship between the amino acid length and content of the inter-finger linker and the sequence of the discontinuous target site. Previously published studies on this matter have put forth several types of inter-finger linkers to match the extra nucleotides found between recognition triplets

(Imanishi et al. 2005, Moore et al. 2001b, Nomura and Sugiura 2003, Yan et al. 2007). These studies report on the activity of ZFP-based proteins (usually ZFTFs) on target sites where the inserted basepairs were most often thymine. Our survey of inter-finger linker variant ZFNs on target sites with thymine, guanine, and adenine insertions found highly efficient ZFN activity only on thymine insertions and lower levels on non-specific activity on guanine and adenine insertions. This discrepancy prompts questions regarding the sequence preference of ZFPs engineered to discontinuous sequences. Again, structural data could inform this issue. More importantly, our methods need to be repeated at other target sites based on the 5'-GNNGNNGNN-3' target site motif and target sites that include CNN, ANN, and TNN triplets. Other inter-finger linkers described in other studies should also be included in these repeated surveys for cross-comparison in a 3-fingered ZFN platform.

4.2.3 Obligate Heterodimer Fns

The rational design of a obhetFn intends for the dimerization event to only occur when the two ZFNs of a pair are both bound to the cognate target site and unable to deliver off-target DSBs elsewhere (Miller J. C. et al. 2007, Szczepek et al. 2007). This should result in increased rates of on-target cutting and less off-target cutting. However, our results are in conflict with the outcomes of previous studies. One hypothesis is that the ZFPs used in our studies and in others are different and may have different DNA-binding properties. These differences may

be affecting the stability of the obhetFn dimerization event leading to discrepancies in ZFN activity. A more comprehensive survey of ZFNs in which the ZFPs used are well characterized in terms of specificity and affinity should be performed to better define how obhetFns can result in higher rates of on-target cutting in mammalian cells.

4.2.4 hTERT Zinc Finger Transcription Factors

We have developed ZFTFs that bind target sequences within the *hTERT* promoter and exon 1 and can induce the expression of a reporter gene on a transiently transfected reporter plasmid. While these accomplishments constitute progress towards the goal of therapeutic application, much more evidence of efficacy is needed. The next step would be to determine the direct effects of these ZFTFs on the *hTERT* locus in hTERT- cell lines. First, assays quantifying the stimulation of hTERT expression, such as qRT-PCR, would establish the immediate effects of applying ZFTFs to the *hTERT* promoter sequence. Next, we would need to demonstrate increases in active telomerase enzyme after the application of ZFTFs and telomerase repeat amplification protocol assays (TRAP) would be best suited for detecting any changes. Unfortunately, Western blotting is not helpful in this line of research due to the lack of highly specific antibody and the very little amount of hTERT protein necessary to enact major cellular changes (Wu et al. 2006). Thus, making the visualization or quantification of hTERT protein expression difficult and TRAP assays provide a broader dynamic range

within which to measure changes in telomerase activity, even if the assay does not directly test for the hTERT protein. Finally, the changes in cellular proliferation potential, function, and lifespan associated with hTERT ZFTF exposure would need to be characterized. Should these ZFTFs demonstrate the ability to up-regulate *hTERT* expression and consequently produce cellular changes, then these ZFNs can be considered for therapeutic applications. For example, the short-term up-regulation of hTERT expression to increase proliferative potential and cellular function would be useful in tissue engineering where cells explanted from patients would need to undergo many cell divisions to rebuild the needed tissue while avoiding accelerated cellular aging. The ZFTFs presented in our study show promise and we encourage any effort made to further characterize the efficacy of their use.

4.2.5 hTERT Zinc Finger Nucleases

Our screening strategies have also identified highly active ZFN pairs to more than one target site in the *hTERT* promoter and exon 1 sequences in assays testing nuclease activity in extra-chromosomal and stably integrated reporter constructs. In the future, ZFN-induced genomic modification will need to be tested at the endogenous *hTERT* locus. In general, changes made in the genomic sequence can be detected by Southern blot strategies and in some cases Cel1 assays. The *hTERT* locus is strongly regulated by chromatin status and methylation, and it is not known whether these conditions will effect ZFN activity

(Wang and Zhu 2003, 2004, Zinn et al. 2007). Thus, the use of histone deacetylase inhibitors and demethylating agents may be needed to enhance ZFN performance by making target sites more accessible. In addition to the technical considerations, the repertoire of potential modifications to the *hTERT* locus is extensive. ZFN-mediated NHEJ could be used for created mutagenic repair events to study the site-specific disruption of the *hTERT* expression. Given the appropriate repair donor, ZFN-mediated HR could site-specifically introduce single nucleotide changes to characterize regulatory elements or correct disease-causing mutations. Eliminating or introducing new regulatory sequences into the *hTERT* promoter is also possible. In theory, the entire *hTERT* promoter could also be exchanged for new promoters responsive to exogenous control. More importantly, any of these strategies could be used in gene therapy applications to treat many kinds of telomerase-associated disease.

4.2.6 Summary

The data presented in this thesis addresses many issues of ZFP design and implementation across the entire process of selecting a target site to demonstrating the activity of the chimeric protein. Our results provide informed recommendations for the types of sequences that are eligible for ZFP-targeting, which methods may be needed for generating ZFPs, and how to further customize ZFP architecture for the intended target site. This information was then used in the development of ZFP-based chimeric proteins at the hTERT locus where novel

artificial ZFTFs and ZFNs were generated and screened for activity at hTERT target sites. Taken together, our data offers guidelines for the engineering of ZFPs to other target sites as well as supporting groundwork for developing these emerging technologies for use at the hTERT locus in the future.

REFERENCES

Adelman CA, Boulton SJ. 2010. Metabolism of postsynaptic recombination intermediates. FEBS Lett. Epub ahead of print.

Alwin S, Gere MB, Guhl E, Effertz K, Barbas CF, 3rd, Segal DJ, Weitzman MD, Cathomen T. 2005. Custom zinc-finger nucleases for use in human cells. Mol Ther 12: 610-617.

Armanios M, et al. 2005. Haploinsufficiency of telomerase reverse transcriptase leads to anticipation in autosomal dominant dyskeratosis congenita. Proc Natl Acad Sci U S A 102: 15960-15964.

Armanios MY, et al. 2007. Telomerase mutations in families with idiopathic pulmonary fibrosis. N Engl J Med 356: 1317-1326.

Artandi SE, DePinho RA. 2000. Mice without telomerase: what can they teach us about human cancer? Nat Med 6: 852-855.

—. 2010. Telomeres and telomerase in cancer. Carcinogenesis 31: 9-18.

Bae KH, et al. 2003. Human zinc fingers as building blocks in the construction of artificial transcription factors. Nat Biotechnol 21: 275-280.

Bartsevich VV, Juliano RL. 2000. Regulation of the MDR1 gene by transcriptional repressors selected using peptide combinatorial libraries. Mol Pharmacol 58: 1-10.

Bateman A, Birney E, Cerruti L, Durbin R, Etwiller L, Eddy SR, Griffiths-Jones S, Howe KL, Marshall M, Sonnhammer EL. 2002. The Pfam protein families database. Nucleic Acids Res 30: 276-280.

Beerli RR, Barbas CF, 3rd. 2002. Engineering polydactyl zinc-finger transcription factors. Nat Biotechnol 20: 135-141.

Beerli RR, Dreier B, Barbas CF, 3rd. 2000. Positive and negative regulation of endogenous genes by designed transcription factors. Proc Natl Acad Sci U S A 97: 1495-1500.

Beumer K, Bhattacharyya G, Bibikova M, Trautman JK, Carroll D. 2006. Efficient gene targeting in Drosophila with zinc-finger nucleases. Genetics 172: 2391-2403.

Bibikova M, Golic M, Golic KG, Carroll D. 2002. Targeted chromosomal cleavage and mutagenesis in Drosophila using zinc-finger nucleases. Genetics 161: 1169-1175.

Bibikova M, Beumer K, Trautman JK, Carroll D. 2003. Enhancing gene targeting with designed zinc finger nucleases. Science 300: 764.

Bibikova M, Carroll D, Segal DJ, Trautman JK, Smith J, Kim YG, Chandrasegaran S. 2001. Stimulation of homologous recombination through targeted cleavage by chimeric nucleases. Mol Cell Biol 21: 289-297.

Bitinaite J, Wah DA, Aggarwal AK, Schildkraut I. 1998. FokI dimerization is required for DNA cleavage. Proc Natl Acad Sci U S A 95: 10570-10575.

Blackburn EH. 1991. Structure and function of telomeres. Nature 350: 569-573.

Bodnar AG, Ouellette M, Frolkis M, Holt SE, Chiu CP, Morin GB, Harley CB, Shay JW, Lichtsteiner S, Wright WE. 1998. Extension of life-span by introduction of telomerase into normal human cells. Science 279: 349-352.

Bowater R, Doherty AJ. 2006. Making ends meet: repairing breaks in bacterial DNA by non-homologous end-joining. PLoS Genet 2: e8.

Branzei D, Foiani M. 2008. Regulation of DNA repair throughout the cell cycle. Nat Rev Mol Cell Biol 9: 297-308.

Brenneman MA, Wagener BM, Miller CA, Allen C, Nickoloff JA. 2002. XRCC3 controls the fidelity of homologous recombination: roles for XRCC3 in late stages of recombination. Mol Cell 10: 387-395.

Cathomen T, Joung JK. 2008. Zinc-finger nucleases: the next generation emerges. Mol Ther 16: 1200-1207.

Chang JT, Yang HT, Wang TC, Cheng AJ. 2005. Upstream stimulatory factor (USF) as a transcriptional suppressor of human telomerase reverse transcriptase (hTERT) in oral cancer cells. Mol Carcinog 44: 183-192.

Chen L, Nievera CJ, Lee AY, Wu X. 2008. Cell cycle-dependent complex formation of BRCA1.CtIP.MRN is important for DNA double-strand break repair. J Biol Chem 283: 7713-7720.

Chevalier B, Turmel M, Lemieux C, Monnat RJ, Jr., Stoddard BL. 2003. Flexible DNA target site recognition by divergent homing endonuclease isoschizomers I-CreI and I-MsoI. J Mol Biol 329: 253-269.

Chevalier BS, Stoddard BL. 2001. Homing endonucleases: structural and functional insight into the catalysts of intron/intein mobility. Nucleic Acids Res 29: 3757-3774.

Choo Y, Sanchez-Garcia I, Klug A. 1994. In vivo repression by a site-specific DNA-binding protein designed against an oncogenic sequence. Nature 372: 642-645.

Cong YS, Bacchetti S. 2000. Histone deacetylation is involved in the transcriptional repression of hTERT in normal human cells. J Biol Chem 275: 35665-35668.

Cornu TI, Thibodeau-Beganny S, Guhl E, Alwin S, Eichtinger M, Joung JK, Cathomen T. 2008. DNA-binding specificity is a major determinant of the activity and toxicity of zinc-finger nucleases. Mol Ther 16: 352-358.

Crick F. 1970. Central dogma of molecular biology. Nature 227: 561-563.

Crowe DL, Nguyen DC, Tsang KJ, Kyo S. 2001. E2F-1 represses transcription of the human telomerase reverse transcriptase gene. Nucleic Acids Res 29: 2789-2794.

Daley JM, Palmbos PL, Wu D, Wilson TE. 2005. Nonhomologous end joining in yeast. Annu Rev Genet 39: 431-451.

de Lange T, Shiue L, Myers RM, Cox DR, Naylor SL, Killery AM, Varmus HE. 1990. Structure and variability of human chromosome ends. Mol Cell Biol 10: 518-527.

Dent CL, Lau G, Drake EA, Yoon A, Case CC, Gregory PD. 2007. Regulation of endogenous gene expression using small molecule-controlled engineered zinc-finger protein transcription factors. Gene Ther 14: 1362-1369.

Di Certo MG, et al. 2010. The artificial gene Jazz, a transcriptional regulator of utrophin, corrects the dystrophic pathology in mdx mice. Hum Mol Genet 19: 752-760.

Dreier B, Fuller RP, Segal DJ, Lund CV, Blancafort P, Huber A, Koksch B, Barbas CF, 3rd. 2005. Development of zinc finger domains for recognition of the 5'-CNN-3' family DNA sequences and their use in the construction of artificial transcription factors. J Biol Chem 280: 35588-35597.

Durai S, Bosley A, Abulencia AB, Chandrasegaran S, Ostermeier M. 2006. A bacterial one-hybrid selection system for interrogating zinc finger-DNA interactions. Comb Chem High Throughput Screen 9: 301-311.

Durai S, Mani M, Kandavelou K, Wu J, Porteus MH, Chandrasegaran S. 2005. Zinc finger nucleases: custom-designed molecular scissors for genome engineering of plant and mammalian cells. Nucleic Acids Res 33: 5978-5990.

Elliott B, Richardson C, Winderbaum J, Nickoloff JA, Jasin M. 1998. Gene conversion tracts from double-strand break repair in mammalian cells. Mol Cell Biol 18: 93-101.

Elrod-Erickson M, Benson TE, Pabo CO. 1998. High-resolution structures of variant Zif268-DNA complexes: implications for understanding zinc finger-DNA recognition. Structure 6: 451-464.

Elrod-Erickson M, Rould MA, Nekludova L, Pabo CO. 1996. Zif268 protein-DNA complex refined at 1.6 A: a model system for understanding zinc finger-DNA interactions. Structure 4: 1171-1180.

Fajardo-Sanchez E, Stricher F, Paques F, Isalan M, Serrano L. 2008. Computer design of obligate heterodimer meganucleases allows efficient cutting of custom DNA sequences. Nucleic Acids Res 36: 2163-2173.

Falke D, Fisher M, Ye D, Juliano RL. 2003. Design of artificial transcription factors to selectively regulate the pro-apoptotic bax gene. Nucleic Acids Res 31: e10.

Feng J, et al. 1995. The RNA component of human telomerase. Science 269: 1236-1241.

Fu F, Sander JD, Maeder M, Thibodeau-Beganny S, Joung JK, Dobbs D, Miller L, Voytas DF. 2009. Zinc Finger Database (ZiFDB): a repository for information

on C2H2 zinc fingers and engineered zinc-finger arrays. Nucleic Acids Res 37: D279-283.

Funk WD, Wang CK, Shelton DN, Harley CB, Pagon GD, Hoeffler WK. 2000. Telomerase expression restores dermal integrity to in vitro-aged fibroblasts in a reconstituted skin model. Exp Cell Res 258: 270-278.

Gabet AS, Mortreux F, Charneau P, Riou P, Duc-Dodon M, Wu Y, Jeang KT, Wattel E. 2003. Inactivation of hTERT transcription by Tax. Oncogene 22: 3734-3741.

Gonzalez B, Schwimmer LJ, Fuller RP, Ye Y, Asawapornmongkol L, Barbas CF, 3rd. 2010. Modular system for the construction of zinc-finger libraries and proteins. Nat Protoc 5: 791-810.

Gouble A, et al. 2006. Efficient in toto targeted recombination in mouse liver by meganuclease-induced double-strand break. J Gene Med 8: 616-622.

Gunes C, Lichtsteiner S, Vasserot AP, Englert C. 2000. Expression of the hTERT gene is regulated at the level of transcriptional initiation and repressed by Mad1. Cancer Res 60: 2116-2121.

Hanahan D, Weinberg RA. 2000. The hallmarks of cancer. Cell 100: 57-70.

Hanas JS, Hazuda DJ, Bogenhagen DF, Wu FY, Wu CW. 1983. Xenopus transcription factor A requires zinc for binding to the 5 S RNA gene. J Biol Chem 258: 14120-14125.

Handel EM, Alwin S, Cathomen T. 2009. Expanding or restricting the target site repertoire of zinc-finger nucleases: the inter-domain linker as a major determinant of target site selectivity. Mol Ther 17: 104-111.

Harley CB. 2002. Telomerase is not an oncogene. Oncogene 21: 494-502.

—. 2005. Telomerase therapeutics for degenerative diseases. Curr Mol Med 5: 205-211.

Hartlerode AJ, Scully R. 2009. Mechanisms of double-strand break repair in somatic mammalian cells. Biochem J 423: 157-168.

Heyer WD, Li X, Rolfsmeier M, Zhang XP. 2006. Rad54: the Swiss Army knife of homologous recombination? Nucleic Acids Res 34: 4115-4125.

Hinnen A, Hicks JB, Fink GR. 1978. Transformation of yeast. Proc Natl Acad Sci U S A 75: 1929-1933.

Hinz JM. 2010. Role of homologous recombination in DNA interstrand crosslink repair. Environ Mol Mutagen 51: 582-603.

Hirsch JA, Wah DA, Dorner LF, Schildkraut I, Aggarwal AK. 1997. Crystallization and preliminary X-ray analysis of restriction endonuclease FokI bound to DNA. FEBS Lett 403: 136-138.

Holt N, et al. 2010. Human hematopoietic stem/progenitor cells modified by zinc-finger nucleases targeted to CCR5 control HIV-1 in vivo. Nat Biotechnol.

Horikawa I, Barrett JC. 2003. Transcriptional regulation of the telomerase hTERT gene as a target for cellular and viral oncogenic mechanisms. Carcinogenesis 24: 1167-1176.

Horikawa I, Cable PL, Afshari C, Barrett JC. 1999. Cloning and characterization of the promoter region of human telomerase reverse transcriptase gene. Cancer Res 59: 826-830.

Horikawa I, Cable PL, Mazur SJ, Appella E, Afshari CA, Barrett JC. 2002. Downstream E-box-mediated regulation of the human telomerase reverse transcriptase (hTERT) gene transcription: evidence for an endogenous mechanism of transcriptional repression. Mol Biol Cell 13: 2585-2597.

Hou M, et al. 2002. The histone deacetylase inhibitor trichostatin A derepresses the telomerase reverse transcriptase (hTERT) gene in human cells. Exp Cell Res 274: 25-34.

Hurt JA, Thibodeau SA, Hirsh AS, Pabo CO, Joung JK. 2003. Highly specific zinc finger proteins obtained by directed domain shuffling and cell-based selection. Proc Natl Acad Sci U S A 100: 12271-12276.

Ihara H, Mie M, Funabashi H, Takahashi F, Sawasaki T, Endo Y, Kobatake E. 2006. In vitro selection of zinc finger DNA-binding proteins through ribosome display. Biochem Biophys Res Commun 345: 1149-1154.

Imanishi M, Yan W, Morisaki T, Sugiura Y. 2005. An artificial six-zinc finger peptide with polyarginine linker: selective binding to the discontinuous DNA sequences. Biochem Biophys Res Commun 333: 167-173.

Kim JS, Pabo CO. 1998. Getting a handhold on DNA: design of poly-zinc finger proteins with femtomolar dissociation constants. Proc Natl Acad Sci U S A 95: 2812-2817.

Kim JS, Lee HJ, Carroll D. 2010. Genome editing with modularly assembled zinc-finger nucleases. Nat Methods 7: 91; author reply 91-92.

Kim YG, Cha J, Chandrasegaran S. 1996. Hybrid restriction enzymes: zinc finger fusions to Fok I cleavage domain. Proc Natl Acad Sci U S A 93: 1156-1160.

Kim YG, Shi Y, Berg JM, Chandrasegaran S. 1997. Site-specific cleavage of DNA-RNA hybrids by zinc finger/FokI cleavage domain fusions. Gene 203: 43-49.

Klinger RY, Blum JL, Hearn B, Lebow B, Niklason LE. 2006. Relevance and safety of telomerase for human tissue engineering. Proc Natl Acad Sci U S A 103: 2500-2505.

Klug A. 2010a. The discovery of zinc fingers and their applications in gene regulation and genome manipulation. Annu Rev Biochem 79: 213-231.

—. 2010b. The discovery of zinc fingers and their development for practical applications in gene regulation and genome manipulation. Q Rev Biophys 43: 1-21.

Kyo S, Takakura M, Taira T, Kanaya T, Itoh H, Yutsudo M, Ariga H, Inoue M. 2000. Sp1 cooperates with c-Myc to activate transcription of the human telomerase reverse transcriptase gene (hTERT). Nucleic Acids Res 28: 669-677.

Lee HJ, Kim E, Kim JS. 2010. Targeted chromosomal deletions in human cells using zinc finger nucleases. Genome Res 20: 81-89.

Li S, Ting NS, Zheng L, Chen PL, Ziv Y, Shiloh Y, Lee EY, Lee WH. 2000. Functional link of BRCA1 and ataxia telangiectasia gene product in DNA damage response. Nature 406: 210-215.

Lieber MR, Yu K, Raghavan SC. 2006. Roles of nonhomologous DNA end joining, V(D)J recombination, and class switch recombination in chromosomal translocations. DNA Repair (Amst) 5: 1234-1245.

Lieber MR, Ma Y, Pannicke U, Schwarz K. 2004. The mechanism of vertebrate nonhomologous DNA end joining and its role in V(D)J recombination. DNA Repair (Amst) 3: 817-826.

Lingner J, Hughes TR, Shevchenko A, Mann M, Lundblad V, Cech TR. 1997. Reverse transcriptase motifs in the catalytic subunit of telomerase. Science 276: 561-567.

Liu PQ, et al. 2004. Cell lines for drug discovery: elevating target-protein levels using engineered transcription factors. J Biomol Screen 9: 44-51.

—. 2005. Isogenic human cell lines for drug discovery: regulation of target gene expression by engineered zinc-finger protein transcription factors. J Biomol Screen 10: 304-313.

—. 2001. Regulation of an endogenous locus using a panel of designed zinc finger proteins targeted to accessible chromatin regions. Activation of vascular endothelial growth factor A. J Biol Chem 276: 11323-11334.

Liu Q, Xia Z, Zhong X, Case CC. 2002. Validated zinc finger protein designs for all 16 GNN DNA triplet targets. J Biol Chem 277: 3850-3856.

Lombardo A, et al. 2007. Gene editing in human stem cells using zinc finger nucleases and integrase-defective lentiviral vector delivery. Nat Biotechnol 25: 1298-1306.

Ma H, Urquidi V, Wong J, Kleeman J, Goodison S. 2003. Telomerase reverse transcriptase promoter regulation during myogenic differentiation of human RD rhabdomyosarcoma cells. Mol Cancer Res 1: 739-746.

Maeder ML, Thibodeau-Beganny S, Sander JD, Voytas DF, Joung JK. 2009. Oligomerized pool engineering (OPEN): an 'open-source' protocol for making customized zinc-finger arrays. Nat Protoc 4: 1471-1501.

Maeder ML, et al. 2008. Rapid "open-source" engineering of customized zinc-finger nucleases for highly efficient gene modification. Mol Cell 31: 294-301.

Magnenat L, Schwimmer LJ, Barbas CF, 3rd. 2008. Drug-inducible and simultaneous regulation of endogenous genes by single-chain nuclear receptor-based zinc-finger transcription factor gene switches. Gene Ther 15: 1223-1232.

Mattei E, et al. 2007. Utrophin up-regulation by an artificial transcription factor in transgenic mice. PLoS One 2: e774.

Matuoka K, Chen KY. 2002. Telomerase positive human diploid fibroblasts are resistant to replicative senescence but not premature senescence induced by chemical reagents. Biogerontology 3: 365-372.

McIlwraith MJ, Vaisman A, Liu Y, Fanning E, Woodgate R, West SC. 2005. Human DNA polymerase eta promotes DNA synthesis from strand invasion intermediates of homologous recombination. Mol Cell 20: 783-792.

Meyerson M, et al. 1997. hEST2, the putative human telomerase catalytic subunit gene, is up-regulated in tumor cells and during immortalization. Cell 90: 785-795.

Miller J, McLachlan AD, Klug A. 1985. Repetitive zinc-binding domains in the protein transcription factor IIIA from Xenopus oocytes. EMBO J 4: 1609-1614.

Miller JC, et al. 2007. An improved zinc-finger nuclease architecture for highly specific genome editing. Nat Biotechnol 25: 778-785.

Mimitou EP, Symington LS. 2009. DNA end resection: many nucleases make light work. DNA Repair (Amst) 8: 983-995.

Moehle EA, Rock JM, Lee YL, Jouvenot Y, DeKelver RC, Gregory PD, Urnov FD, Holmes MC. 2007. Targeted gene addition into a specified location in the human genome using designed zinc finger nucleases. Proc Natl Acad Sci U S A 104: 3055-3060.

Moore M, Choo Y, Klug A. 2001a. Design of polyzinc finger peptides with structured linkers. Proc Natl Acad Sci U S A 98: 1432-1436.

Moore M, Klug A, Choo Y. 2001b. Improved DNA binding specificity from polyzinc finger peptides by using strings of two-finger units. Proc Natl Acad Sci U S A 98: 1437-1441.

Morales CP, Holt SE, Ouellette M, Kaur KJ, Yan Y, Wilson KS, White MA, Wright WE, Shay JW. 1999. Absence of cancer-associated changes in human fibroblasts immortalized with telomerase. Nat Genet 21: 115-118.

Murasawa S, Llevadot J, Silver M, Isner JM, Losordo DW, Asahara T. 2002. Constitutive human telomerase reverse transcriptase expression enhances regenerative properties of endothelial progenitor cells. Circulation 106: 1133-1139.

Nomura W, Sugiura Y. 2003. Effects of length and position of an extended linker on sequence-selective DNA recognition of zinc finger peptides. Biochemistry 42: 14805-14813.

Orlando SJ, et al. 2010. Zinc-finger nuclease-driven targeted integration into mammalian genomes using donors with limited chromosomal homology. Nucleic Acids Res.

Orr-Weaver TL, Szostak JW, Rothstein RJ. 1981. Yeast transformation: a model system for the study of recombination. Proc Natl Acad Sci U S A 78: 6354-6358.

Ozturk N, et al. 2006. Reprogramming of replicative senescence in hepatocellular carcinoma-derived cells. Proc Natl Acad Sci U S A 103: 2178-2183.

Paques F, Duchateau P. 2007. Meganucleases and DNA double-strand break-induced recombination: perspectives for gene therapy. Curr Gene Ther 7: 49-66.

Pavletich NP, Pabo CO. 1991. Zinc finger-DNA recognition: crystal structure of a Zif268-DNA complex at 2.1 A. Science 252: 809-817.

Pelham HR, Brown DD. 1980. A specific transcription factor that can bind either the 5S RNA gene or 5S RNA. Proc Natl Acad Sci U S A 77: 4170-4174.

Perez EE, et al. 2008. Establishment of HIV-1 resistance in CD4+ T cells by genome editing using zinc-finger nucleases. Nat Biotechnol 26: 808-816.

Philippi C, Loretz B, Schaefer UF, Lehr CM. 2010. Telomerase as an emerging target to fight cancer - Opportunities and challenges for nanomedicine. J Control Release.

Picard B, Wegnez M. 1979. Isolation of a 7S particle from Xenopus laevis oocytes: a 5S RNA-protein complex. Proc Natl Acad Sci U S A 76: 241-245.

Pollock R, Issner R, Zoller K, Natesan S, Rivera VM, Clackson T. 2000. Delivery of a stringent dimerizer-regulated gene expression system in a single retroviral vector. Proc Natl Acad Sci U S A 97: 13221-13226.

Porteus MH. 2006. Mammalian gene targeting with designed zinc finger nucleases. Mol Ther 13: 438-446.

Porteus MH, Baltimore D. 2003. Chimeric nucleases stimulate gene targeting in human cells. Science 300: 763.

Porteus MH, Carroll D. 2005. Gene targeting using zinc finger nucleases. Nat Biotechnol 23: 967-973.

Porteus MH, Connelly JP, Pruett SM. 2006. A look to future directions in gene therapy research for monogenic diseases. PLoS Genet 2: e133.

Pruett-Miller SM, Reading DW, Porter SN, Porteus MH. 2009. Attenuation of zinc finger nuclease toxicity by small-molecule regulation of protein levels. PLoS Genet 5: e1000376.

Pruett-Miller SM, Connelly JP, Maeder ML, Joung JK, Porteus MH. 2008. Comparison of zinc finger nucleases for use in gene targeting in mammalian cells. Mol Ther 16: 707-717.

Ramakrishnan S, Eppenberger U, Mueller H, Shinkai Y, Narayanan R. 1998. Expression profile of the putative catalytic subunit of the telomerase gene. Cancer Res 58: 622-625.

Ramirez CL, et al. 2008. Unexpected failure rates for modular assembly of engineered zinc fingers. Nat Methods 5: 374-375.

Rebar EJ, Pabo CO. 1994. Zinc finger phage: affinity selection of fingers with new DNA-binding specificities. Science 263: 671-673.

Rebar EJ, et al. 2002. Induction of angiogenesis in a mouse model using engineered transcription factors. Nat Med 8: 1427-1432.

Rouet P, Smih F, Jasin M. 1994a. Expression of a site-specific endonuclease stimulates homologous recombination in mammalian cells. Proc Natl Acad Sci U S A 91: 6064-6068.

—. 1994b. Introduction of double-strand breaks into the genome of mouse cells by expression of a rare-cutting endonuclease. Mol Cell Biol 14: 8096-8106.

Ruben SM, Dillon PJ, Schreck R, Henkel T, Chen CH, Maher M, Baeuerle PA, Rosen CA. 1991. Isolation of a rel-related human cDNA that potentially encodes the 65-kD subunit of NF-kappa B. Science 251: 1490-1493.

Sambrook J, Botchan M, Hu SL, Mitchison T, Stringer J. 1980. Integration of viral DNA sequences in cells transformed by adenovirus 2 or SV40. Proc R Soc Lond B Biol Sci 210: 423-435.

Santiago Y, et al. 2008. Targeted gene knockout in mammalian cells by using engineered zinc-finger nucleases. Proc Natl Acad Sci U S A 105: 5809-5814.

Sartori AA, Lukas C, Coates J, Mistrik M, Fu S, Bartek J, Baer R, Lukas J, Jackson SP. 2007. Human CtIP promotes DNA end resection. Nature 450: 509-514.

Sepp A, Choo Y. 2005. Cell-free selection of zinc finger DNA-binding proteins using in vitro compartmentalization. J Mol Biol 354: 212-219.

Sera T. 2009. Zinc-finger-based artificial transcription factors and their applications. Adv Drug Deliv Rev 61: 513-526.

Shay JW, Zou Y, Hiyama E, Wright WE. 2001. Telomerase and cancer. Hum Mol Genet 10: 677-685.

Shimizu Y, Bhakta MS, Segal DJ. 2009. Restricted spacer tolerance of a zinc finger nuclease with a six amino acid linker. Bioorg Med Chem Lett 19: 3970-3972.

Shin KH, Kang MK, Dicterow E, Park NH. 2003. Hypermethylation of the hTERT promoter inhibits the expression of telomerase activity in normal oral fibroblasts and senescent normal oral keratinocytes. Br J Cancer 89: 1473-1478.

Smith J, Berg JM, Chandrasegaran S. 1999. A detailed study of the substrate specificity of a chimeric restriction enzyme. Nucleic Acids Res 27: 674-681.

Smith J, Bibikova M, Whitby FG, Reddy AR, Chandrasegaran S, Carroll D. 2000. Requirements for double-strand cleavage by chimeric restriction enzymes with zinc finger DNA-recognition domains. Nucleic Acids Res 28: 3361-3369.

Smith J, et al. 2006. A combinatorial approach to create artificial homing endonucleases cleaving chosen sequences. Nucleic Acids Res 34: e149.

Smogorzewska A, de Lange T. 2004. Regulation of telomerase by telomeric proteins. Annu Rev Biochem 73: 177-208.

Snowden AW, et al. 2003. Repression of vascular endothelial growth factor A in glioblastoma cells using engineered zinc finger transcription factors. Cancer Res 63: 8968-8976.

Sohn JH, Yeh BI, Choi JW, Yoon J, Namkung J, Park KK, Kim HW. 2010. Repression of human telomerase reverse transcriptase using artificial zinc finger transcription factors. Mol Cancer Res 8: 246-253.

Stark JM, Pierce AJ, Oh J, Pastink A, Jasin M. 2004. Genetic steps of mammalian homologous repair with distinct mutagenic consequences. Mol Cell Biol 24: 9305-9316.

Sung P. 1994. Catalysis of ATP-dependent homologous DNA pairing and strand exchange by yeast RAD51 protein. Science 265: 1241-1243.

Sung P, Klein H. 2006. Mechanism of homologous recombination: mediators and helicases take on regulatory functions. Nat Rev Mol Cell Biol 7: 739-750.

Sussman D, Chadsey M, Fauce S, Engel A, Bruett A, Monnat R, Jr., Stoddard BL, Seligman LM. 2004. Isolation and characterization of new homing endonuclease specificities at individual target site positions. J Mol Biol 342: 31-41.

Szczepek M, Brondani V, Buchel J, Serrano L, Segal DJ, Cathomen T. 2007. Structure-based redesign of the dimerization interface reduces the toxicity of zinc-finger nucleases. Nat Biotechnol 25: 786-793.

Szekvolgyi L, Nicolas A. 2010. From meiosis to postmeiotic events: homologous recombination is obligatory but flexible. FEBS J 277: 571-589.

Tachikawa K, Schroder O, Frey G, Briggs SP, Sera T. 2004. Regulation of the endogenous VEGF-A gene by exogenous designed regulatory proteins. Proc Natl Acad Sci U S A 101: 15225-15230.

Takakura M, Kyo S, Kanaya T, Hirano H, Takeda J, Yutsudo M, Inoue M. 1999. Cloning of human telomerase catalytic subunit (hTERT) gene promoter and identification of proximal core promoter sequences essential for transcriptional activation in immortalized and cancer cells. Cancer Res 59: 551-557.

Takakura M, Kyo S, Sowa Y, Wang Z, Yatabe N, Maida Y, Tanaka M, Inoue M. 2001. Telomerase activation by histone deacetylase inhibitor in normal cells. Nucleic Acids Res 29: 3006-3011.

Townsend JA, Wright DA, Winfrey RJ, Fu F, Maeder ML, Joung JK, Voytas DF. 2009. High-frequency modification of plant genes using engineered zinc-finger nucleases. Nature 459: 442-445.

Triezenberg SJ, Kingsbury RC, McKnight SL. 1988. Functional dissection of VP16, the trans-activator of herpes simplex virus immediate early gene expression. Genes Dev 2: 718-729.

Tsakiri KD, Cronkhite JT, Kuan PJ, Xing C, Raghu G, Weissler JC, Rosenblatt RL, Shay JW, Garcia CK. 2007. Adult-onset pulmonary fibrosis caused by mutations in telomerase. Proc Natl Acad Sci U S A 104: 7552-7557.

Urnov FD, Miller JC, Lee YL, Beausejour CM, Rock JM, Augustus S, Jamieson AC, Porteus MH, Gregory PD, Holmes MC. 2005. Highly efficient endogenous human gene correction using designed zinc-finger nucleases. Nature 435: 646-651.

Urrutia R. 2003. KRAB-containing zinc-finger repressor proteins. Genome Biol 4: 231.

Verra NC, et al. 2004. Human telomerase reverse transcriptase-transduced human cytotoxic T cells suppress the growth of human melanoma in immunodeficient mice. Cancer Res 64: 2153-2161.

Wah DA, Bitinaite J, Schildkraut I, Aggarwal AK. 1998. Structure of FokI has implications for DNA cleavage. Proc Natl Acad Sci U S A 95: 10564-10569.

Wah DA, Hirsch JA, Dorner LF, Schildkraut I, Aggarwal AK. 1997. Structure of the multimodular endonuclease FokI bound to DNA. Nature 388: 97-100.

Wang S, Zhu J. 2003. Evidence for a relief of repression mechanism for activation of the human telomerase reverse transcriptase promoter. J Biol Chem 278: 18842-18850.

—. 2004. The hTERT gene is embedded in a nuclease-resistant chromatin domain. J Biol Chem 279: 55401-55410.

Weterings E, van Gent DC. 2004. The mechanism of non-homologous end-joining: a synopsis of synapsis. DNA Repair (Amst) 3: 1425-1435.

Wick M, Zubov D, Hagen G. 1999. Genomic organization and promoter characterization of the gene encoding the human telomerase reverse transcriptase (hTERT). Gene 232: 97-106.

Wolfe SA, Greisman HA, Ramm EI, Pabo CO. 1999. Analysis of zinc fingers optimized via phage display: evaluating the utility of a recognition code. J Mol Biol 285: 1917-1934.

Won J, Chang S, Oh S, Kim TK. 2004. Small-molecule-based identification of dynamic assembly of E2F-pocket protein-histone deacetylase complex for

telomerase regulation in human cells. Proc Natl Acad Sci U S A 101: 11328-11333.

Wooten LG, Ogretmen B. 2005. Sp1/Sp3-dependent regulation of human telomerase reverse transcriptase promoter activity by the bioactive sphingolipid ceramide. J Biol Chem 280: 28867-28876.

Wright WE, Piatyszek MA, Rainey WE, Byrd W, Shay JW. 1996. Telomerase activity in human germline and embryonic tissues and cells. Dev Genet 18: 173-179.

Wu YL, et al. 2006. Immunodetection of human telomerase reverse-transcriptase (hTERT) re-appraised: nucleolin and telomerase cross paths. J Cell Sci 119: 2797-2806.

Xin ZT, Beauchamp AD, Calado RT, Bradford JW, Regal JA, Shenoy A, Liang Y, Lansdorp PM, Young NS, Ly H. 2007. Functional characterization of natural telomerase mutations found in patients with hematologic disorders. Blood 109: 524-532.

Xu L, Zerby D, Huang Y, Ji H, Nyanguile OF, de los Angeles JE, Kadan MJ. 2001. A versatile framework for the design of ligand-dependent, transgene-specific transcription factors. Mol Ther 3: 262-273.

Yamaguchi-Iwai Y, Sonoda E, Buerstedde JM, Bezzubova O, Morrison C, Takata M, Shinohara A, Takeda S. 1998. Homologous recombination, but not DNA repair, is reduced in vertebrate cells deficient in RAD52. Mol Cell Biol 18: 6430-6435.

Yan W, Imanishi M, Futaki S, Sugiura Y. 2007. Alpha-helical linker of an artificial 6-zinc finger peptide contributes to selective DNA binding to a discontinuous recognition sequence. Biochemistry 46: 8517-8524.

Yang J, Nagavarapu U, Relloma K, Sjaastad MD, Moss WC, Passaniti A, Herron GS. 2001. Telomerized human microvasculature is functional in vivo. Nat Biotechnol 19: 219-224.

Zhang L, Spratt SK, Liu Q, Johnstone B, Qi H, Raschke EE, Jamieson AC, Rebar EJ, Wolffe AP, Case CC. 2000. Synthetic zinc finger transcription factor action at an endogenous chromosomal site. Activation of the human erythropoietin gene. J Biol Chem 275: 33850-33860.

Zinn RL, Pruitt K, Eguchi S, Baylin SB, Herman JG. 2007. hTERT is expressed in cancer cell lines despite promoter DNA methylation by preservation of unmethylated DNA and active chromatin around the transcription start site. Cancer Res 67: 194-201.