OPTIMIZING ZINC FINGER NUCLEASES FOR USE IN MAMMALIAN CELLS

APPROVED BY SUPERVISORY COMMITTEE

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

Malu Tansey, Ph.D

Melanie Cobb, Ph.D.

David Chen, Ph.D.

DEDICATION

To my parents, Mike and Sheena Pruett, for your
many years of continuing love and support.

OPTIMIZING ZINC FINGER NUCLEASES FOR USE IN MAMMALIAN CELLS

by

SHONDRA M. PRUETT-MILLER

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

December, 2008

ACKNOWLEDGEMENTS

The successful completion of one's thesis work is a major milestone in any graduate student's life. Several people have been instrumental in my education, mentoring, and life during this time. I'd like to thank my mentor, Matthew Porteus, for his constant encouragement and mentoring. As one of Matt's first graduate students, I feel blessed to have seen Matt grow from an exceptional scientist to an exceptional scientist and mentor. I have learned so much more than just science from Matt and consider my time in his lab priceless both for my career and in my life.

I'd like to thank my thesis committee, Melanie Cobb, Malu Tansey, and David Chen for all of their advice and input into this project. I especially want to thank Melanie Cobb for her personal attention to my education. She has constantly encouraged me and persuaded me that I can do anything including excellent science. Both Melanie and Malu have been great role models as both are extremely talented female scientists that manage not only their labs but also their families with a great deal of grace.

Several members of the Porteus lab have been a huge part of my education as well. I want to thank Patrick Connelly for the many late night and weekend discussions, and the many trips to "get coffee." Brian Ellis, Shaina Porter, Tina

Dann, Joshua Checketts, and Charles Dann (he's not a Portei, but we consider him family) have been instrumental in my sanity during my graduate school years. I'd also like to thank all the members of the lab, past and present, for all of their hard work and helpful discussions during the past 5 years.

Several individuals outside of the UTSW walls have also contributed to my thesis work. I'd like to thank Keith Joung and the members of his laboratory for teaching me how to perform the B2H selection strategy for designing ZFNs and allowing me to do so in their lab space.

There is more to life than science, even during the graduate school years. To me the most important part of life is family and friends. I'd like to acknowledge all of the many friends that I have made during my time at Southwestern, and let you know that I really couldn't have done it without you. To my parents, I can't even begin to list all of the things that you have done for me, but I can say thank you. Thank you for all the things that I know about and especially for all of the things that I don't even have a clue about or overlook. To my loving husband, Ryan, thank you for being patient with me and for spending way too much time in lab when I told you that it would only take "5 minutes." You have my heart.

OPTIMIZING ZINC FINGER NUCLEASES FOR USE IN MAMMALIAN CELLS

SHONDRA M. PRUETT-MILLER

The University of Texas Southwestern Medical Center at Dallas, 2008

Supervising Professor: Matthew H. Porteus, M.D., Ph.D.

Homologous recombination is a well-established technique that has been used to manipulate the genomes of multiple model organisms with great precision and is therefore being explored as a potential way of performing gene therapy. However, the spontaneous rate of homologous recombination in human cells is too low (10^{-6}) to be therapeutically useful. The most powerful way of stimulating homologous recombination is by introducing a double strand break within the target locus. Zinc Finger Nucleases (ZFNs) are designed proteins that fuse a zinc

finger DNA binding domain to the nuclease domain from the *FokI* restriction endonuclease and have been used to induce double strand breaks at precise sequences. Although ZFNs have been successfully used to stimulate gene targeting at specific loci, several issues remain. First, a generalized optimal design strategy for making effective and safe ZFNs has yet to be established. Second, a systematic evaluation method needs to be established in which novel ZFNs are evaluated for both functionality and safety. We compare the gene targeting efficiencies and cytotoxicity of ZFNs made by the two established design strategies: modular assembly and a bacterial 2-hybrid selection strategy. We have found that ZFNs made by the bacterial 2-hybrid strategy are both more efficient at stimulating gene targeting and less toxic than ZFNs made by modular assembly. We have also found that ZFNs made via the bacterial 2-hybrid strategy are more efficient at gene targeting using a GFP reporter assay and show less cytotoxicity than previously published 4-finger proteins. We also present a generalized strategy for systematically evaluating new ZFNs, which includes a bacterial β-galactosidase transcription assay, a mammalian gene targeting assay, a mammalian flow cytometry based survival assay, and a 53BP1 foci formation assay. These assays provide a standard for future ZFN design and evaluation, particularly those that may be destined for therapeutic use. Because the issue of toxicity is such and important one, we have also developed possible strategies to reduce toxicity of ZFNs. We will discuss two strategies for regulating ZFN

protein expression using small molecules. We show that by regulating protein expression to create ZFNs with shortened half-lives, we can maintain high rates of ZFN mediated gene targeting while reducing ZFN toxicity.

TABLE OF CONTENTS

PRIOR PUBLICATIONS

Porteus, MH, Connelly, JP, **Pruett, SM**, (2006). A look to the future directions in gene therapy research for monogenic diseases. ***PLOS Genetics.*** *1285-1292.*

Pruett-Miller, SM, Connelly JP, Maeder, ML, Joung, KL, Porteus, MH. (2008). Optimizing Zinc Finger Nucleases for Gene Targeting in Mammalian Cells. ***Molecular Therapy***. 707-717.

Maeder, Morgan L., Thibodeau-Beganny, S, Osiak, A, Wright, DA, Anthony, RM,Eichtinger, M, Jiang, T, Foley, JE, Winfrey, RF, Townsend, JA, Unger-Wallace, E, Sander, JD, Muller-Lerch, Felix, Pearlberg, J, Gobel, C, Dassie, JP, **Pruett-Miller, SM**, Porteus, MH, Sgroi, DC, Iafrate, AJ, Dobbs, D, McCray, PB, Cathomen, T, Voytas, DF, Joung, JK. (2008). Rapid "open-source" engineering of customized zinc-finger nucleases for highly efficient gene modification. ***Molecular Cell.*** 31(2):294-301.

Pruett-Miller, SM, Reading, DW, Porter, SN, Porteus, MH. (2008). Attenuation of Zinc Finger Nuclease Toxicity by Small Molecule Regulation. *PLoS Genetics.* (submitted).

Wilson, KA, McEwen, AE. **Pruett-Miller, SM**, Porteus, MH. (2008). Expanding the Repertoire of Potential Zinc Finger Nuclease Sites and Determining the Zinc Finger Nuclease Architecture To Target Such Sites. *Nulceic Acids Research*. (submitted).

McEwen, AE, **Pruett-Miller**, **SM**, Wilson, KA, Porteus, MH. (2008). *In Vitro* Analysis of Zinc Finger Proteins Selected by a Bacterial 2-Hybrid Design Strategy. (manuscript in preparation).

Pruett-Miller, SM, Checketts, J, Porteus, MH. (2008). Comparison of the DNA-binding Specificities and Affinities of Zinc Finger Proteins Designed by a Bacterial 2-hybrid Strategy. (manuscript in preparation).

LIST OF FIGURES

LIST OF TABLES

LIST OF DEFINITIONS

ZFN-Zinc Finger Nuclease

ZFP-Zinc Finger Protein

AA-Amino Acid

DSB-Double Strand Break

BP-Base Pair

HR-Homologous Recombination

NHEJ-Non-Homologous End-Joining

MA-Modular Assembly

B2H-Bacterial 2-Hybrid

GFP-Green Fluorescent Protein

DD-Destabilization Domain

Ub-Ubiquitin

SCID-Severe Combined Immunodeficiency

IDLV-Integrase-Defective Lentiviral

HIV-Human Immunodeficiency Virus

MOI-Multiplicity of Infection

CHAPTER I: INTRODUCTION AND REVIEW OF LITERATURE

Gene Targeting to Treat Human Disease

Human diseases are caused by a combination of genetics and environmental factors. Certain diseases, however, are caused almost entirely by host genetics. Some such diseases can be attributed to a mutation in a single gene. Examples of these so-called monogenic diseases are Sickle Cell Anemia, Cystic Fibrosis, Huntington's Disease, and Hemophilia to name a few. The realization that so many devastating human diseases were caused by genetic mutations and the advent of recombinant DNA technology laid the groundwork for gene therapy. Gene therapy is based on the idea that one can manipulate the nucleic acid content of diseased cells and thereby cure genetic diseases. It is a strategy that traditionally uses DNA to treat diseases caused by defective genes.

There are two traditional approaches for gene therapy: gene addition and gene correction. The gene addition approach involves augmenting the deficient or mutant protein function by adding back a normal copy of the mutant gene randomly into the genome, typically using a viral vector (Ott, Schmidt et al. 2006). This approach has been used in clinical trials and has faced some major complications with *in vivo* administration of viral vectors and insertional

oncogenesis (Cavazzana-Calvo, Hacein-Bey et al. 2000; Aiuti, Slavin et al. 2002; Hacein-Bey-Abina, Von Kalle et al. 2003; Gaspar, Parsley et al. 2004). In contrast, the gene correction approach strives to correct the mutated gene at its endogenous locus. In theory, this approach will not only result in increased efficacy because the corrected gene will be under the control of its natural elements at its endogenous locus, but will also provide a better safety profile as there is no theoretical risk of insertional oncogenesis.

Our lab and others are interested in developing homologous recombination as a strategy for gene corrective gene therapy. Homologous recombination is a natural process that occurs in the cell for a variety of purposes including the repair of DNA double-strand breaks (DSBs) and is the most precise way to manipulate the genome (West, Chappell et al. 2000). In homologous recombination, the nucleic acid content of one piece of DNA is transferred to a homologous piece of DNA in efforts to repair a DSB. Mario Capecchi, Martin Evans, and Oliver Smithies were awarded the 2007 Nobel Prize in Medicine for their role in developing homologous recombination as a method to repress target genes in murine embryonic stem cells and thereby create knockout mice. One of the major limiting factors to using homologous recombination as a strategy for gene therapy, hence forth called gene targeting, is that the spontaneous rate in mammalian somatic cells is far too low (10^{-6}) to be useful therapeutically (Sedivy and Sharp 1989; Porteus and Baltimore 2003). In the 1990s, however, several

groups showed that the rate of gene targeting could be stimulated 1000-fold by creating a DSB using the I-SceI endonuclease (Rouet, Smih et al. 1994; Choulika, Perrin et al. 1995; Brenneman, Gimble et al. 1996; Taghian and Nickoloff 1997). Using optimized conditions and a reporter gene, Porteus and Baltimore demonstrated that the rate of DSB stimulated gene targeting could reach as high as 3-5% (Porteus and Baltimore 2003). Unfortunately, I-SceI has an 18 base pair recognition site that does not occur naturally in the human genome. It was therefore necessary to develop a method to create gene-specific DSBs.

Currently, there are several strategies to achieve site-specific DSBs using any of the following: modified triplex-forming oligonulceotides (Kuan and Glazer 2004), modified polyamides (Wurtz and Dervan 2000; Dervan and Edelson 2003), modified homing endonucleases (Belfort and Roberts 1997; Gimble 2000; Chevalier and Stoddard 2001; Chevalier, Kortemme et al. 2002) or zinc finger nucleases (ZFNs) (Porteus and Carroll 2005). We have chosen to pursue the ZFN method for creating site-specific DSBs.

Double Strand Break Repair

ZFNs are being pursued as reagents to induce DSBs at specific sequences within the genome in order to (1) correct mutant genes at their endogenous locus and (2) to generate targeted mutations. However, the introduction of DSBs needs

to be performed with great precision, as DSBs are the most dangerous type of DNA lesion. Delayed or aberrant DSB repair can lead to cell cycle arrest, cell death, translocations, and even tumorigenesis. Even though DSBs can be particularly dangerous for a cell and in turn the whole organism, DSBs occur daily during DNA replication, immune system development, and meiosis (Wyman and Kanaar 2006). To prevent harmful genomic instability, the cellular machinery has evolved several strategies to repair DSBs. The two major DSB repair pathways are non-homologous end-joining (NHEJ) and homologous recombination (HR) (Sonoda, Hochegger et al. 2006).

It has been shown by Maria Jasin's group and others that a DSB can stimulate homologous recombination by several thousand-fold (Rouet, Smih et al. 1994; Choulika, Perrin et al. 1995; Brenneman, Gimble et al. 1996; Taghian and Nickoloff 1997; Porteus and Baltimore 2003). If high enough rates of homologous recombination can be achieved, gene targeting could one day be used as a strategy to treat patients with monogenic diseases.

Site-specific DSBs can also lead to targeted mutagenesis by (1) being repaired by aberrant, non-homologous end-joining or (2) being repaired by precise homologous recombination such that an entire transgene could be integrated at the site of repair. This type of targeted gene addition has no theoretical risk of insertional oncogenesis as the site of DSB defines the site of integration.

Non-homologous End Joining

NHEJ is a DSB repair pathway that involves re-ligating break ends together without the use of a homologous template (Moore and Haber 1996). However, this pathway utilizes short homologous DNA sequences called microhomologies to direct repair. If break ends are compatible, the NHEJ pathway can repair the break accurately. If, however, the overhangs are not precisely compatible, NHEJ may lead to imprecise repair and loss or gain of nucleotides. Several groups are using ZFNs to create DSBs to stimulate this type of error-prone repair in order to knockout target genes. A number of proteins are essential for the repair of DSBs by NHEJ (KU70, KU80, DNA-PKcs, XRCC4, and LIGASE IV) (Burma, Chen et al. 2006).

Homologous Recombination

HR is another pathway by which DSBs can be repaired and requires a homologous segment of DNA to complete repair. After a DSB occurs, cellular machinery surveys the nucleic acid content for a homologous piece of DNA from which to repair the DSB. Normally, the cell uses a sister chromatid as a template for repair. However, the break may also be repaired from an extrachromosomal homologous segment of DNA, if supplied. Many groups are using ZFN-induced

DSBs to stimulate HR as a potential way to directly correct disease-causing mutations. Also, because ZFNs can create site-specific DSBs, ZFNs are being pursued as reagents to stimulate site-specific HR that would allow targeted integration of transgenes rather than random integration. A number of proteins are involved in HR (ATM, MRN complex MRE11/RAD50/NBS1, RAD51, RAD52, and RPA) (Wyman and Kanaar 2004)

Zinc Finger Nucleases: General Background and Architecture

Originally termed chimeric restriction enzymes, ZFNs were developed by Chandrasegaran and coworkers with the idea that novel sequences could be cleaved by fusing a nonsequence-specific cleavage domain to a different sequence-specific DNA binding domain (Kim and Chandrasegaran 1994). The FokI endonuclease is a Type IIS enzyme meaning that its DNA binding activity is separate from the DNA cleavage activity (Li, Wu et al. 1992). This modularity suggested that novel chimeric nucleases could be engineered by fusing the FokI cleavage domain to other DNA binding proteins. Using several different DNA binding motifs including (1) the *Drosophila melanogaster* homeobox domain (2) the yeast Gal4 DNA-binding domain and (3) the zinc finger-DNA binding domain, the Chandrasegaran group demonstrated that they could redirect the cleavage sites of the chimeras *in vitro* (Kim, Cha et al. 1996; Chandrasegaran and

Smith 1999; Kandavelou, Mani et al. 2005). The zinc finger nucleases are the most extensively studied group of chimeric nucleases because they contain a zinc finger-DNA binding domain that can be re-designed to recognize novel target sites, which allows the chimeras to be targeted to a wide range of sequences. ZFNs are composed of a zinc finger protein (ZFP) fused to a nuclease domain. Because all of the sequence specificity comes from the ZFP portion of the ZFN, it is incredibly important that the ZFP portion be designed to bind with high specificity for its target site. If the ZFP does not have a high enough specificity for its target sequence, undesired or "off-target" DSBs will result. Although most cells are quite good at repairing DSBs, too many DSBs can lead to cell cycle arrest, cell death, or even tumorigenesis.

ZFNs consist of N-terminal Cys2His2 zinc finger motifs fused with a linker to the cleavage domain of the FokI endonuclease. The Cys2His2 protein module was first discovered by Klug and colleagues in 1986 and has since been revealed as the most common motifs in the human genome, with over 700 proteins containing this motif (Diakun, Fairall et al. 1986; Notarangelo, Giliani et al. 2000; Lander, Linton et al. 2001; Venter, Adams et al. 2001; Jantz, Amann et al. 2004). Each zinc finger motif folds into a ßßα structure consisting of about 30 amino acids and is stabilized by the chelation of a zinc ion by the conserved Cys2His2 residues (Figure 1.1). The zinc finger motifs bind DNA by inserting the α-helix into the major grove of the DNA double helix (Pavletich and Pabo

1991). The α-helix is also called the recognition helix because key residues within the helix contribute most of the specific interactions of the zinc finger motif to its target site. Each zinc finger motif binds primarily to 3 basepairs (bp) of DNA (Pavletich and Pabo 1991). Relative to the start of the recognition helix, positions -1, +1, +2, +3, +4, +5, and +6 contribute most of the sequence-specific interactions with the target triplet (Shi and Berg 1995; Kim, Smith et al. 1998; Elrod-Erickson and Pabo 1999). Key residues -1, +3, and +6 each make direct contact with one of the nucleotides within the triplet DNA target on the sense strand, while the amino acid at position +2 has been shown to make minor contacts in the minor groove of the DNA on the antisense strand. Position +4 is a conserved leucine. The amino acids at position +1, and +5 were not originally thought to have major contributions to the DNA specificity, but recent data suggests that context-dependency and cooperativity within each recognition helix could play a major role in DNA recognition.

The 6 amino acids of the recognition helix (not including the conserved leucine) can be changed within the context of the remaining amino acids of the consensus backbone to generate ZFPs with different triplet sequence specificities (Desjarlais and Berg 1993; Wolfe, Greisman et al. 1999). In order to bind longer DNA sequences, several fingers can be linked together in tandem to form a multi-fingered zinc finger protein (ZFP). Zif268 is the most well studied, 3-finger protein and binds its 9-10 base pair (bp) site with picomolar affinity. The Zif268

crystal structure suggested that ZFP binding to DNA is modular (Pavletich and Pabo 1991). That is, each finger seemed to bind its cognate site independently of its neighboring fingers. This suggested that zinc finger proteins could be made to recognize new target sequences by swapping fingers.

Like the FokI restriction endonuclease, which requires dimerization of its nuclease domain in order to cleave DNA, ZFNs also require the nuclease domain to undergo dimerization in order to create a DSB (Bitinaite, Wah et al. 1998; Wah, Bitinaite et al. 1998; Smith, Bibikova et al. 2000; Mani, Smith et al. 2005). Therefore, the dimerization of two 3-fingered ZFN monomers (each of which recognizes a 9 bp target half-site) effectively gives an 18-bp full-site (Figure 1.2), which is long enough to specify a unique sequence within the 3 billion bp human genome.

The individual zinc finger motifs or "fingers" are numbered from N-terminus to C-terminus. For example, the N-terminal finger is called finger 1, while the C-terminal finger is called finger 3. Each ZFN binds in an anti-parallel fashion relative to the DNA target. That is, finger 3 binds at the 5' end of the DNA target, while finger 1 binds at the 3' end of the site. The FokI nuclease domain for each ZFN is attached via a variable linker to the last finger of a multi-fingered protein (Figure 1.1 and 1.2). Using a *Xenopus laevis* oocytes system, Bibikova and colleagues tested a number of other factors required for ZFN cleavage (Bibikova, Carroll et al. 2001). They were able to show using two 3-

finger ZFNs that most efficient cleavage could be achieved when the target half-sites are inversely oriented and separated by a six bp spacer (Figure 2). The spacer is defined as the DNA separating the inversely oriented ZFN half-sites. Our lab has furthered this dataset by showing that by varying the linker length between the nuclease domain and finger 3, different spacer lengths (5, 6, or 7 bp) can be used to obtain efficient cleavage (Wilson et al., manuscript in preparation).

Zinc Finger Nuclease Design

The major appeal of using ZFNs to create DSBs is that they can be modified to recognize novel targets. In theory, ZFNs can be made to essentially any target sequence. Currently, however, there is much debate about the best practical method by which to design ZFPs. Over the years, several design strategies have emerged. It should be noted that ZFP design has been driven mostly by those who are motivated by the possibility of developing artificial transcriptional activators and repressors to control gene expression.

When the crystal structure of Zif268 was published by Pavletich and Pabo in 1991, several critical features were observed that suggested that Cys2His2 zinc fingers might be an attractive framework for designing novel DNA-binding proteins. First, the simple modular structure of the Zif complex, where each finger binds its 3-bp target site independently of its neighboring fingers, led to the

prediction that by combining different zinc fingers with disparate triplet targets, one could change the overall binding specificity of the zinc finger protein. Second, the conserved pattern of the base contacts in tandemly linked fingers, where each nucleotide seemed to be contacted by a single amino acid (residues -1, 3, and 6), led to the prediction that by altering individual amino acids in the recognition helix, one could alter the specificity for an individual finger. These observations suggested that individual fingers with different target sites could be mixed and matched in order to create a protein with a novel target sequence. This is the intellectual theory behind the zinc finger design strategy known as "Modular Assembly" (MA).

After the Zif268 crystal structure was solved, many hoped that a ZFP-DNA recognition code could be created based on the idea that one amino acid contacts one nucleotide in the target sequence. Desjarlais and Berg developed a rule-based system to rationally alter the specificity of zinc fingers based on sequence and structural comparisons to Zif268 for the design of variants that could bind new DNA targets (Desjarlais and Berg 1992; Desjarlais and Berg 1992; Desjarlais and Berg 1992; Desjarlais and Berg 1993). Several others attempted similar rational design-based strategies but were met with only moderate success. That is, although several ZFPs were created with reasonable affinities and specificities, the general ability of these strategies to target a wide range of DNA sequences was limited (Nardelli, Gibson et al. 1991; Thukral,

Morrison et al. 1992; Kim and Berg 1995). Also, during these design studies, Desjarlais and Berg observed that swapping individual fingers with known specificities in a certain finger context to different finger positions in order to redirect protein binding led to significant context-dependent effects (Desjarlais and Berg 1993).

Although developing a simple recognition code of one amino acid to one nucleotide would have been a great advance had it ever materialized, ZFP-DNA binding is more complex than originally thought. Context-dependency and cooperitivity seem to play a major role in the binding of a ZFP to its target. That is, the manner in which one finger binds affects the other fingers. Likewise, the manner in which specific amino acids within the recognition helix bind to a target affects the ability of the other amino acids to bind. Some interactions are synergistic, while some interactions can be exclusionary. Since design strategies alone are not sufficient, selection strategies were developed in an attempt to re-design ZFPs to recognize new target sites. With this realization, several groups began developing phage display selection strategies. In general, the principle behind these phage display selections was to create randomized libraries of ZFPs, display them on the surface of phage, and determine which ZFPs bind to predetermined, immobilized target sequences. Rebar and Pabo were the first to publish their phage display strategy, which involved randomizing key amino acids (residues -1, 2, 3, and 6 as determined from the Zif268 crystal structure) in finger

1 of Zif268 and selecting for altered specificity using a DNA target with sequence changes in the 4bp subsite corresponding to finger 1 (Rebar and Pabo 1994). Using this strategy, they were able to produce Zif268 variants that had altered specificities at the finger 1 subsite. However, only two out of three of the attempted subsite sequences were successfully targeted demonstrating some limitations to this design strategy. Using similar selections by phage display based on the Zif268 framework, several other groups were successful at recovering zinc fingers with redirected specificities (Choo and Klug 1994; Choo and Klug 1994; Jamieson, Kim et al. 1994; Wu, Yang et al. 1995; Jamieson, Wang et al. 1996).

The next challenge was to design zinc finger proteins that could bind entirely novel sequences. Toward this goal, several novel phage display strategies have emerged: (1) parallel phage display selection, (2) sequential phage display selection, and (3) bipartite phage display selection.

Parallel phage selection is based on the assumption that zinc finger domains are functionally independent and can therefore be mixed and matched in any desired arrangement. The idea is that if the full zinc finger domain complement of the genetic code (64 domains for each DNA triplet) could be revealed, then any DNA sequence could be targeted by simply assembling the desired zinc fingers in the correct order (modular assembly). Currently, zinc finger domains specific for the 5'-GNN-3', 5'-ANN-3', and most of the 5'-CNN-

3' DNA target sequences have been developed (Segal, Dreier et al. 1999; Dreier, Segal et al. 2000; Dreier, Beerli et al. 2001; Liu, Xia et al. 2002; Dreier, Fuller et al. 2005). The Barbas group performs the parallel phage display selection by using two anchor fingers from a framework ZFP (Zif268 or a variant of Zif268) at finger 1 and finger 3. Then, a randomized library is used for finger 2 to select for individual fingers that bind with some specificity for the target triplet subsite. To make a novel multifinger ZFN, the Barbas group simply stitched the fingers that were selected for at the finger 2 position in the correct order. That is, fingers that were selected for at position two with the same anchor fingers are used at all finger positions to create a multifinger protein. Using this method, the Barbas group has been able to create ZFPs with up to six-fingers (Dreier, Beerli et al. 2001). In slight contrast, Liu and colleagues hypothesized that for some GNN triplets, different fingers may be needed at different finger positions. To address this concern, Liu et al performed phage display and used rational design to select for the 16, 5'-GNN-3' triplets at each of the three finger positions (Liu, Xia et al. 2002). This group was able to demonstrate that some of the GNN triplets showed significant positional dependence, and concluded that zinc fingers do not behave like completely independent modules. Moreover, context-dependency with respect to the other two fingers and to the binding target can all influence the affinity and specificity of ZFPs.

While the parallel approach allows one to construct ZFNs with relative ease by simply going to a table of pre-selected ZFNs and using PCR to assemble fingers to create a 3-fingered protein, it has several limitations. First, ZFPs designed using these pre-selected modules do not always have the desired sequence specificity or affinity. Second, a number of these known zinc finger motifs bind degenerate sites almost as well as their cognate sites and therefore require further optimization and empirical design to create zinc finger proteins with desired binding characteristics. Third, this strategy does not take cooperativity or context-dependency effects between fingers or within the recognition helix into account. The DNA subsites of neighboring fingers can overlap in such a way that the fingers often bind with synergy. These synergistic interactions in certain frameworks can cause interference between adjacent fingers. Most parallel selection strategies select for fingers that bind at the finger 2 position in the context of the Zif268 finger 1 and 3 anchor fingers which both recognize 5'-GCG-3' subsites. The Zif268 crystal structure revealed that an aspartate in position 2 of finger 2 contacts the binding site of finger 1 at position -1, and thus confirmed the importance of cross-strand interactions. This is one reason why the 5'-GNNGNNGNN-3' family of DNA targets have produced ZFPs with the best binding characteristics using this type of modular assembly design.

In contrast to parallel phage display selection, Greisman and Pabo developed a modified version of the original phage display experiments in which

they sequentially select one finger at a time in the context of the other zinc fingers in order to make a 3-fingered protein (Greisman and Pabo 1997). The overall idea behind this strategy is that in the first round of selection finger 1 is selected for in the context of two anchor fingers from the Zif268 protein. In the second round of selection, finger 2 is selected for in the context of the newly selected finger 1 clone and one anchor finger from Zif268. Finally in the last round of selection, finger 3 is selected for in the context of the newly identified finger 1 and finger 2 clones. This strategy has subsequently been modified to include a fourth round of selection in which finger 1 was reselected in its final context (Wolfe, Greisman et al. 1999). This strategy attempts to take both context dependency and cooperativity of neighboring fingers into account. However, because each round of selection produces only one zinc finger motif at each finger position to carry on to the next round of selection, the best finger modules in the context of their neighboring fingers may be lost. This strategy also requires that several randomized libraries be constructed for every ZFP that is to be designed. This is a much more labor and time intensive strategy than the parallel selection strategy and requires considerable expertise.

Another phage display strategy called the bipartite selection strategy utilizes two pre-fabricated zinc finger libraries in each of which 1½ fingers of the three-fingered Zif268 protein are partially randomized at the key amino acid residues that make contact with the DNA (Isalan and Choo 2001; Isalan, Klug et

al. 2001). This strategy of dividing the three-finger protein into two blocks allows one to optimize finger/finger interfaces as they will occur within the randomized blocks rather than between the modules. The two randomized libraries are divided up so that the N-terminal part of the ZFP is in one library and binds the 5 bp on the 3'-end of the target site, while the C-terminal part of the ZFP is in the other library and binds the 5 bp on the 5'-end of the DNA target. Parallel selections are carried out using the two libraries to create pools of ZFP "halves" that will bind their respective target half-site. Then another round of selection is performed in which selections from the randomized libraries are recombined and further rounds of selection are performed using the full-length target sequence. This strategy still requires several rounds of selection and technical expertise. The libraries generated by this strategy are held in the intellectual property vaults of Sangamo Biosciences, a small biotech company in California that purchased the libraries in 2000 from Genedaq, which was founded by Sir Aaron Klug (Scott 2005).

More recently, a bacterial based system has been developed in which zinc finger binding to its target is required for cell growth and survival (Joung, Ramm et al. 2000; Hurt, Thibodeau et al. 2003). This strategy was used to design the ZFNs used in most of this thesis and is described in more detail in Chapter 2 and Appendix A. Briefly, this strategy uses a bacterial 2-hybrid system to design ZFPs that bind to their target sequence with some affinity. This strategy involves

using two stages of selection in order to (1) create pools of each individual finger of a 3-finger protein that bind their target subsite with some affinity in the context of two anchor fingers and (2) after randomly recombining the three pools selected for in stage one, select for three-finger ZFPs that bind the full-length target with a relatively high affinity (Figure 2.1B).

Currently, the debate continues over which design strategy produces ZFNs that are both highly active and extremely specific. Several web-based zinc finger design software programs are now available to help experts and laypersons alike at the following URLs: http://www.zincfingers.org/software-tools.htm and http://www.zincfingertools.org (Mandell and Barbas 2006). However, the validity of ZFNs produced using this software has not been entirely validated. Also, a zinc finger consortium has been established by ZFP scientists and will soon make a B2H-designed finger pool library available to the academic community. The consortium website can be accessed at http://www.zincfingers.org.

Using Zinc Finger Nucleases in Model Systems

ZFNs have now been successfully designed and used to find and cleave their chromosomal targets in *Xenopus laevis* oocytes (Bibikova, Carroll et al. 2001), *Drosophila melanogaster* (Bibikova, Golic et al. 2002; Bibikova, Beumer

et al. 2003), plants (Lloyd, Plaisier et al. 2005; Wright, Townsend et al. 2005), *Caenorhabditis elegans* (Morton, Davis et al. 2006), *Danio rerio* (Doyon, McCammon et al. 2008; Meng, Noyes et al. 2008), CHO cells (Santiago, Chan et al. 2008), and in human cells (Urnov, Miller et al. 2005); (Porteus and Baltimore 2003; Alwin, Gere et al. 2005). In each of these systems, ZFNs have also been shown to induce HR and/or NHEJ at the site of cleavage to repair the DSB.

The first proof-of-principle experiment to show that ZFNs could stimulate HR through targeted cleavage was performed in frog oocytes using an extra-chromosomal target (Bibikova, Carroll et al. 2001). These experiments were performed using microinjections of different DNA substrates and the three-finger ZFN (ΔQQR-FN) directly into frog eggs. Bibikova et al. showed that the ΔQQR-ZFN could find and cleave its target *in vivo*. They were also able to show that the most efficient cleavage occurred when the binding site was an inverted repeat in the tail-to-tail orientation of the 9 bp target separated by a 6-bp spacer, and when the zinc finger and nuclease domains were fused using no intentional linker. Additionally, head-to-head orientation of the target site or a single copy of the target site proved to be ineffective at being a substrate for cutting by ZFNs. Although this study showed that ZFNs could find and cleave their target *in vivo*, two major questions remained: (1) can ZFNs be modified to recognize novel sequences, especially targets within endogenous genes and (2) can ZFNs show site-specific cleavage of a chromosomal target in cells?

Both of these questions were answered when Dana Carroll's lab successfully designed ZFNs and targeted the easily scorable *yellow* gene on the X-chromosome of the fruit fly (Bibikova, Golic et al. 2002). The ZFNs were designed using one of the 5'-GNN-3' tables of pre-selected finger modules. Two ZFNs were designed by stitching together three finger modules that bound a pure 5'-GNNGNNGNN-3' site on each of the strands separated by a 6bp spacer (giving a 5'-NNCNNCNNCNNNNNNGNNGNNGNN-3' full-length target site). The ZFNs were cloned under the control of a heat shock promoter. When the ZFNs were expressed from integrated transgenes during larval development by heat shock, mutations occurred specifically in the *yellow* gene as observed by the presence of yellow patches on a normally dark posterior abdomen. Sequencing of the mutants revealed that all of the mutations were small deletions and/or insertions at the ZFN cleavage site, which is indicative of the error-prone NHEJ DSB repair pathway. As expected, these mutations were also only observed in the males because females contain two copies of the X-chromosome. Therefore, female cellular machinery can use the second copy of the chromosome as a template off of which to precisely repair the DSB via HR. The male germ line transmission was very low. This was the first time that targeted cleavage by ZFNs was shown to facilitate the generation of directed mutations. Several groups are currently investigating this concept to knockout genes, especially in hard to manipulate cell types (Doyon, McCammon et al. 2008; Meng, Noyes et al.

2008; Santiago, Chan et al. 2008). Bibikova et al. then went on to show that when using the same larval heat shock system and when performed in the presence of a marked donor DNA segment to serve as a template for HR, that targeted gene replacement by HR could occur in both males and females (Bibikova, Beumer et al. 2003).

Subsequent experiments in the Carroll lab targeted two more fly genes, *ry* and *bw,* both involved in the pathway leading to red eye pigment (Beumer, Bhattacharyya et al. 2006). The *bw*-directed ZFNs produced very low rates of targeted mutagenesis, but the *ry* ZFNs produced much higher rates. When using an extrchromosomal linear mutant donor, more than 90% of the heat shock induced parents gave at least one mutant offspring. However, elevated toxicity and thus lethality was observed due to excessive cleavage when ZFNs were expressed at high levels.

More recently, the Carroll laboratory has focused some of its efforts on performing gene targeting experiments in *C. elegans*, which is a model organism in which knocking out genes has proven to be difficult. Morton and colleagues showed using both synthetic extrachromosomal sequence and an endogenous genomic site on the X-chromosome that the corresponding 3-finger ZFNs were very effective at stimulating gene targeting in worms (Morton, Davis et al. 2006). In the first experiments, the previously characterized ΔQQR-Fn ZFNs were used and placed under a heat shock promoter. The extrachromosomal target contained

the two inverted ΔQQR-Fn binding sites separated by a unique MluI restriction site. Both the ΔQQR-Fn plasmid and the target plasmid were introduced by co-injection to generate an extrachromosomal array. After larval heat shock to induce ZFN expression, PCR analysis and subsequent MluI digestion revealed that over 25% of the targets examined contained targeted mutations.

Morton and colleagues continued on to show that an endogenous sequence on the X-chromosome could be targeted using newly designed ZFNs. In the absence of a donor DNA template, NHEJ was used to repair the ZFN-induced DSB and led to targeted mutations. Unfortunately, ZFN expression has not been achieved in the germline of *C. elegans* and is thus not currently useful in creating knockouts in this organism.

Like in *C. elegans,* methods to create targeted mutations in plants are not readily available. In 2005, however, the genome of *Arabidopsis thaliana* was successfully modified using ZFNs (Lloyd, Plaisier et al. 2005). Similar to other strategies used in different model organisms, Lloyd and colleagues introduced the target site consisting of two inverted binding sites for the ΔQQR-FN, this time separated with an EcoRI site, into the *A. thaliana* genome. After heat shock induction of the QQR-ZFN, which was placed under the control of a heat-shock promoter, PCR analysis and subsequent EcoRI digestion revealed that some of the heat-shock seedlings had lost the EcoRI site. These results showed that ZFNs

could induce DSBs in plants that could be repaired by NHEJ and lead to targeted mutagenesis.

Not long after, Daniel Voytas' group was able to show, using tobacco plants, that ZFNs could mediate gene targeting via HR in plants (Wright, Townsend et al. 2005). In these experiments, Wright and colleagues were able to restore the function of a defective GUS:NPTII (β-glucuronidase/neomycin phosphotransferase fusion) reporter gene that had been integrated chromosomally into different transgenic tobacco plant lines. This reporter gene was made non-functional by the 600 bp deletion and the insertion of the 9bp Zif268 binding site in the inverted orientation and separated by 6 bp. The donor DNA template was a 5kb fragment of the GUS:NPTII gene that included the 600 bp missing from the non-functional target gene. If gene targeting via HR occurred, cells would express the corrected GUS:NPTII fusion, making then resistant to kanamycin and causing them to turn blue when incubated in the appropriate substrates. Using this system, Wright and colleagues were able to observe targeted integration and thus gene restoration in about 10% of the transformed protoplasts that had been exposed to both the donor DNA template and ZFN DNA.

During the last year, additional studies in model organisms have shown that ZFNs can induce DSBs in zebrafish and Chinese Hamster Ovary (CHO) cells and can be repaired via NHEJ resulting in targeted mutagenesis and gene knockout (Doyon, McCammon et al. 2008; Meng, Noyes et al. 2008; Santiago,

Chan et al. 2008). Although the use of ZFNs in model organisms has led to great strides in the understanding ZFN cleavage and subsequent repair, the use of ZFNs in mammalian cells is essential to developing ZFN-stimulated gene modification.

Using Zinc Finger Nucleases in Human Somatic Cells

Porteus and Baltimore were the first to demonstrate that ZFNs could stimulate gene targeting via HR in mammalian somatic cells (Porteus and Baltimore 2003). They developed a GFP gene targeting system that allows for easy detection of gene targeting using flow cytometry to assay for GFP+ (corrected) cells. Briefly, a stable HEK293 cell line was created in which a mutated form of GFP was randomly integrated into the human genome. The GFP gene was mutated by the insertion of a stop codon, a ZFN full-site, and the 18bp recognition sequence for the I-SceI endonuclease, which served as a control. The ZFN full-site consisted of the 9 bp target sites for Zif268 and QQR-Fn in a tail-to-tail inverted orientation and separated by a 6 bp spacer. Cells were transfected with a donor DNA template (repair substrate), which contained the homology needed to correct the integrated mutant GFP, and the plasmids encoding the Zif268 and QQR-Fn nucleases. The repair substrate contained an N-terminal truncation such that the transfection of the repair substrate alone could not produce GFP+ cells. After introduction of ZFNs and repair substrate, gene

targeting was measured using flow cytometry to assay for GFP+ cells and rates of about 3-5% were observed. Using the same system, Porteus was able to show that novel, modular assembly designed ZFNs could be made to target an endogenous site within the GFP reporter gene (Porteus 2006). As in other studies, the ZFNs used in these experiments exhibited elevated toxicity when expressed at high levels.

Using Zinc Finger Nucleases for Target Gene Addition in the Human Genome

More recently, a group from Sangamo BioSciences was able to successfully use ZFNs to generate targeted gene addition within the human genome (Moehle, Rock et al. 2007). In this study, cells were exposed to a previously characterized pair of 4-fingered ZFNs (discussed below) targeting the IL2Rγ (SCID) gene and a repair substrate containing up to 8000bp flanked by locus-specific homology arms. Remarkably, even in the absence of selection, the entire 8kb sequence was accurately reconstructed at the targeted endogenous locus at a frequency of 6%. This was the first proof-of-principle experiment demonstrating the possibility for targeted gene addition within the human genome.

Using Zinc Finger Nucleases for Targeted Gene Knockout in Mammalian Somatic Cells

Santiago and colleagues were the first to show targeted gene knockout in mammalian cells using ZFNs to create DSBs that were repaired by error-prone NHEJ (Santiago, Chan et al. 2008). In this study, ZFNs were designed to target the dihydrofolate reductase (DHFR) gene in the Chinese hamster ovary (CHO) cell line. DHFR is essential in the biosythesis of purines, thymidines, and glycine, and therefore, DHFR-/- cells are unable to grow unless culture medium is supplemented with essential metabolites (Yamane-Ohnuki, Kinoshita et al. 2004). This group chose to target the DHFR gene because it is a widely used selectable marker in CHO cells, and because DHFR-/- cell lines already exist to which they could functionally compare newly derived knockout cell lines.

After transfection with only a pair of 4-fingered DHFR-targeted ZFNs, Santiago et al. were able to generate targeted gene knockouts in CHO cells. Furthermore, they generated three new, genetically distinct DHFR-/- CHO cell lines. These cell lines were distinct because the breaks were repaired by NHEJ, which resulted in different mutations during the repair of the DSB. ZFN-induced biallelic gene disruption was observed at frequencies greater the 1%, further demonstrating the utility of ZFN induced DSBs and repair via NHEJ as a method to rapidly modify mammalian genomes.

Using Zinc Finger Nucleases to Target an Endogenous Human Gene: SCID gene

The first proof-of-principle experiments to successfully design and target ZFNs to cleave a disease-causing endogenous gene (IL2Rγ) in human cells were performed by Urnov and colleagues (Urnov, Miller et al. 2005). Mutations in the IL2Rγ gene lead to X-linked severe combined immunodeficiency (SCID). In this study, two optimized 4-finger proteins, which bind a 24bp full-site target, were designed using a proprietary archive of pre-selected 2-finger cassettes. The repair substrate used in these studies contained a portion of the IL2Rγ gene with a silent point mutation that creates a new BsrBI restriction site. If ZFN-induced HR occurred between the repair substrate and its chromosomal target, the BsrBI site would be incorporated at the cognate site, which could be assayed by restriction digest. Single clones were isolated and assayed for the insertion of the BsrBI site, and about 13% of the clones had been modified at a single allele, while about 6% of the clones had been modified at both alleles. These studies were performed in both K562 cells and in primary human CD4+ T lymphocytes, and similar results were obtained in both cell lines.

Using ZFNs to Disrupt an Endogenous Human Gene: CCR5

A naturally occurring, homozygous Δ32 deletion in the CCR5 seven-transmembrane chemokine receptor, which is the major co-receptor for HIV entry, has been shown to confer resistance to HIV-1 infection (Alkhatib, Combadiere et al. 1996; Deng, Liu et al. 1996). Perez et al. generated CCR5 targeted ZFNs with the idea that ZFN-mediated disruption could occur via aberrant NHEJ repair and lead to an HIV-resistant phenotype (Perez, Wang et al. 2008). This group transiently expressed the CCR5 targeted ZFNs in a pool of primary human CD4+ T-cells and were able to permanently disrupt about 50% of the CCR5 alleles generating the HIV-resistant genotype. The group continued on to test the stability and heritability of the ZFN-mediated disruption of the CCR5 gene *in vivo* using the NOG model of HIV infection (Watanabe, Terashima et al. 2007). Mice engrafted with ZFN-modified CD4+ T cells had lower viral loads and higher CD4+ T-cell counts than mock treated mice. The authors suggest a possible approach for the treatment of HIV-1 infection as follows: (1) treat CD4+ T cells with CCR5 targeted ZFNs (2) expand the treated CD4+ cells *ex vivo* (3) introduce expanded CD4+ cells into HIV/AIDS patients to reconstitute immune functions by maintaining a population of HIV-resistant CD4+ T cells.

ZFN Delivery to Human Primary Cells including ES Cells

In a majority of the experiments described in this chapter, ZFNs were introduced into cells via standard cell culture transfection methods or nucleofection. However, some of the cell types that have relevant targets for treatment of human disease, including stem cells, are not readily transfected by these methods. Codelivery of two ZFNs and the repair substrate in such relevant primary cells has been a major hurdle for using ZFNs to stimulate gene correction or targeted integration. Lombardo et al. addressed this issue by using an integrase-defective lentiviral (IDLV) vector delivery method (Lombardo, Genovese et al. 2007). Because of previous complications using integrating viral vectors, including random integration leading to insertional oncogenesis, non-integrating vectors are being pursued. Previous studies have shown that HIV vectors that are packaged with a D64V integrase mutant are able to infect cells and deliver vector DNA to the nuclease but are completely deficient in integrase-mediated intergration (Naldini, Blomer et al. 1996; Vargas, Gusella et al. 2004; Nightingale, Hollis et al. 2006; Philippe, Sarkis et al. 2006; Yanez-Munoz, Balaggan et al. 2006). Using the IDLV method, Lombardo and colleagues were able to show that most cell types including human primary hematopoietic stem cells and human ES cells were amenable to this type of delivery as shown by the presence of ZFN-mediated gene modification and targeted transgene addition.

The ZFNs used in this study were designed by Sangamo's proprietary library of two finger cassettes and empirical design and targeted the SCID gene and the CCR5 co-receptor (Urnov, Miller et al. 2005; Perez, Wang et al. 2008).

Zinc Finger Nuclease Toxicity

Several studies have shown toxicity associated with ZFNs (Porteus and Baltimore 2003; Beumer, Bhattacharyya et al. 2006; Porteus 2006). In fact, in whole organisms such as flies and zebrafish, high levels of ZFN expression has led to abnormal developmental mutations or "monsters" (Beumer, Bhattacharyya et al. 2006; Doyon, McCammon et al. 2008; Meng, Noyes et al. 2008). Some evidence suggests (described in Chapters 2 and 3) and it is hypothesized that this toxicity is due to off target DSBs.

Most cells are generally good at repairing DSBs as this process is essential to maintain genomic integrity. However, if too many DSBs occur within a single cell, this type of DNA lesion can be lethal or even lead translocations or tumorogenesis. Because two, 3-finger ZFNs should bind an 18bp site that statistically should occur at most one time in the 3 billion bases of the human genome, off-target breaks must be occurring because of a lack of specificity for their target site. In this thesis, we compare the gene targeting efficiencies and cytotoxicity of ZFNs made by the two established design strategies: modular

assembly and a bacterial 2-hybrid selection strategy. We have found that ZFNs made by the bacterial 2-hybrid strategy are both more efficient at stimulating gene targeting and less toxic than ZFNs made by modular assembly that target the same site. We also discuss strategies to reduce ZFN toxicity. We have found that small molecule regulation of ZFN protein expression is an effective way to reduce cytotoxicity without compromising targeting efficiency.

The major limiting factor to using ZFNs in any type of genome modification is making highly specific ZFNs. Whether the strategy involves using homologous recombination for gene targeting or it is for creating targeted mutagenesis, ZFNs need to be specific for their target in order to be used safely. The ZFP portion of the ZFN makes up the specific DNA-binding domain of the protein. Thus, it is of the utmost importance that ZFPs have excellent specificity for their target sites. The work presented here demonstrates that by using the B2H design strategy to design the ZFP portion, we were able to create highly active ZFNs with reduced toxicity. We also observed that toxicity decreased with a reduction in the amount of ZFN-DNA transfected. We hypothesized that if we could control ZFN protein expression, we would be able to reduce toxicity. That is, if a cell is saturated with ZFN, more off-target breaks will occur. In contrast, if less ZFN is present and the protein is highly specific, then fewer off-target breaks will occur. We have shown that by controlling ZFN protein expression using small molecules, we can reduce toxicity without losing activity. Fine-tuning ZFN

protein expression may be particularly beneficial to using ZFN mediated genome modification in a wide variety of cell types, including human stem cells.

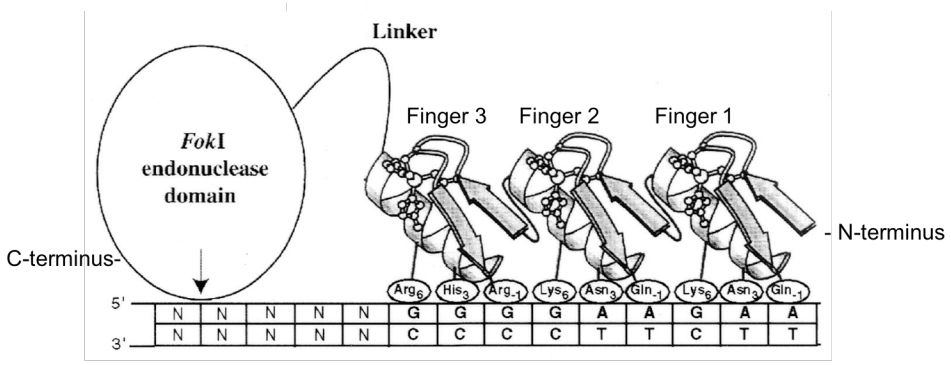


Figure 1.1. Schematic diagram of a zinc finger nuclease (ZFN).

The N-terminal sequence-specific DNA binding domain is composed of multiple, in this case 3, zinc finger binding motifs linked in tandem. Each zinc finger motif or "finger" is made up of about 30 amino acids that fold into a ββα structure. The folding of each finger is facilitated by the chelation of a zinc ion by a conserved Cys2His2 motif. The residues that provide the primary specificity-determining interactions with each triplet subsite are at positions -1, 3, and 6 relative to the start of the α-helix or recognition helix. The C-terminal nonsequence-specific DNA cleavage domain is from the FokI endonuclease and is fused to the last finger of the multifinger protein, in this case finger 3. A variable linker separates the last finger from the nuclease domain. The specific DNA-ZFN interaction illustrated above is for the QQR-Fn. Figure modified from (Smith, Bibikova et al. 2000).

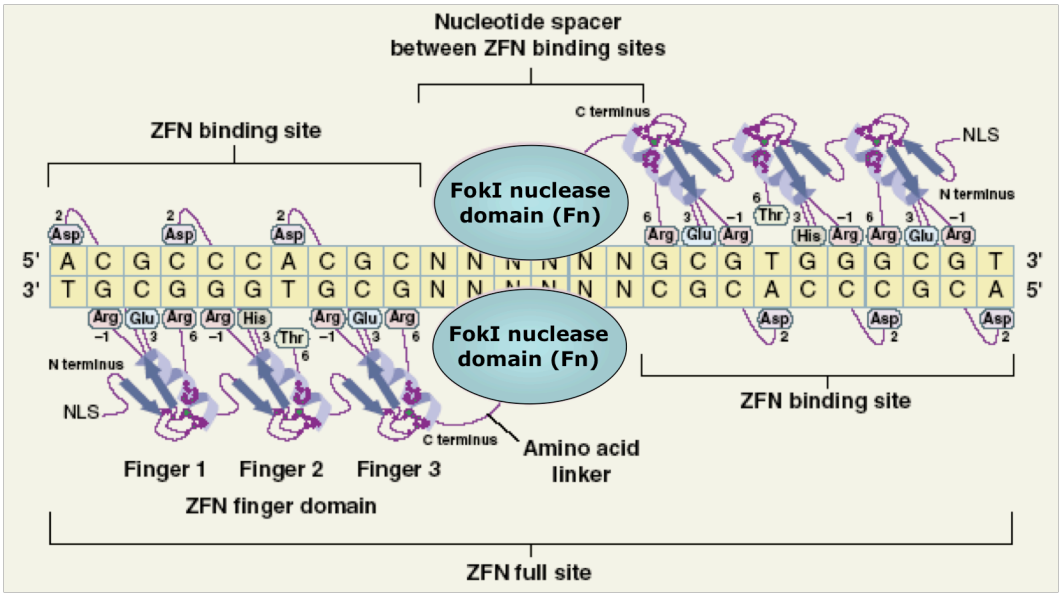


Figure 1.2. ZFN homodimer binding to DNA.

ZFNs bind their target half-sites as monomers, and cleave upon dimerization of two FokI cleavage domains. Shown above are two 3-fingered ZFNs binding their cognate target half-sites. Each half-site is made up of a 9-10 bp target site, giving an 18-20 bp ZFN full site. Dimerization of their FokI cleavage domain leads to cleavage of the nucleotides (spacer) between the two half-sites. The identity of these spacer nucleotides does not seem to be important, but it has been shown that most efficient cleavage is achieved when the half-sites are inverted in the tail-to-tail orientation separated by a 6bp spacer. The residues that provide the primary specificity-determining interactions with each triplet subsite are at shown above at positions -1, 3, and 6 relative to the start of the α-helix or recognition helix. The specific DNA-ZFN interaction illustrated above is for the Zif268. Figure modified from (Jantz, Amann et al. 2004; Porteus and Carroll 2005).

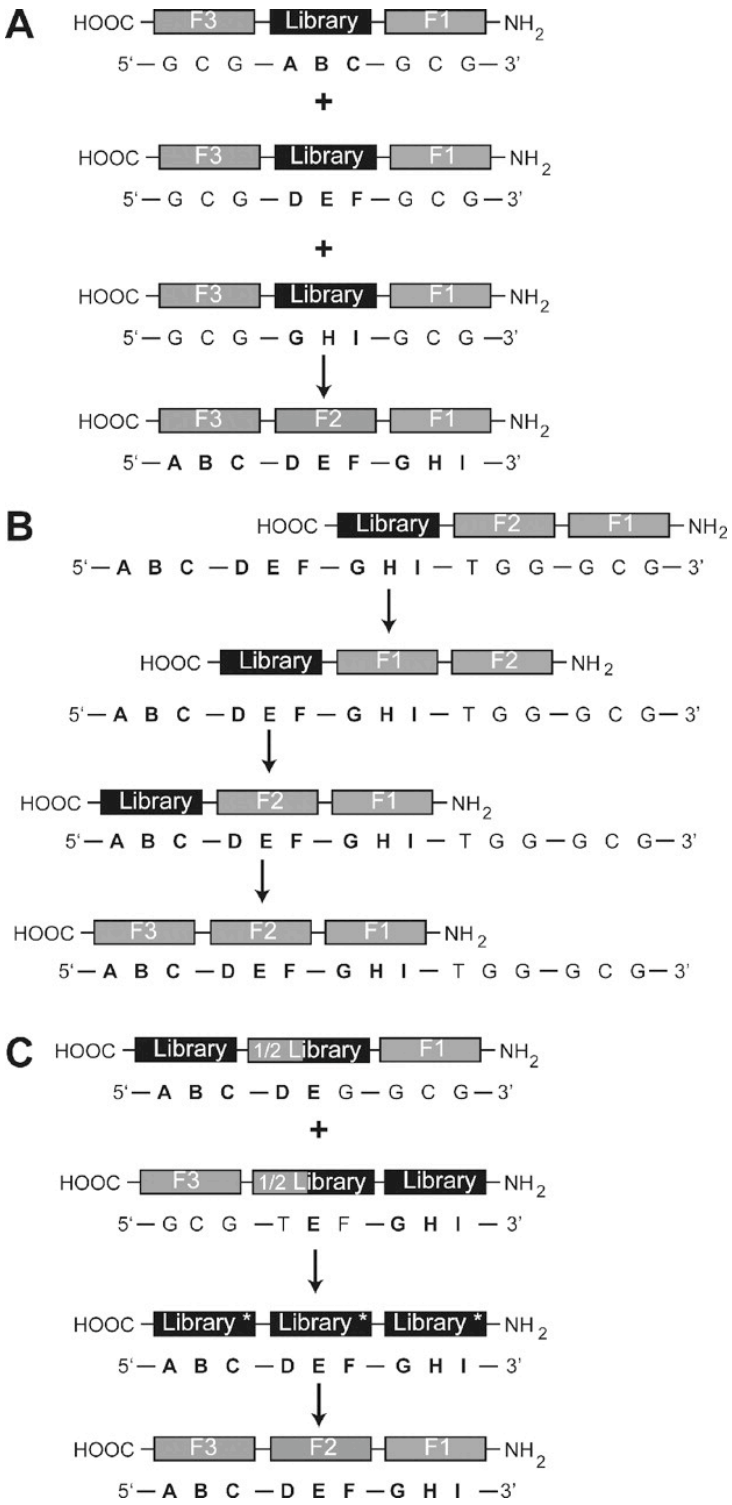

A
HOOC F3 Library F1 NH2
5'— G C G — A B C — G C G— 3'
+
HOOC F3 Library F1 NH2
5'— G C G — D E F — G C G— 3'
+
HOOC F3 Library F1 NH2
5'— G C G — G H I — G C G —3'
HOOC F3 F2 F1 NH2
5'— A B C — D E F — G H I —3'
B
HOOC Library F2 F1 NH2
5'— A B C — D E F — G H I — T G G — G C G— 3'
HOOC Library F1 F2 NH2
5'— A B C — D E F — G H I — T G G — G C G— 3'
HOOC Library F2 F1 NH2
5'— A B C — D E F — G H I — T G G — G C G— 3'
HOOC F3 F2 F1 NH2
5'— A B C — D E F — G H I — T G G — G C G— 3'
C
HOOC Library 1/2 Library F1 NH2
5'— A B C — D E G — G C G— 3'
+
HOOC F3 1/2 Library Library NH2
5'— G C G — T E F — G H I — 3'
HOOC Library * Library * Library * NH2
5'— A B C — D E F — G H I — 3'
HOOC F3 F2 F1 NH2
5'— A B C — D E F — G H I —3'

Figure 1.3. Phage display based strategies for designing novel ZFPs that bind a desired DNA sequence.

(A) Parallel phage display selection is based on the concept that each zinc finger motif is functionally independent from its neighboring fingers. Therefore, individual fingers are selected for in the context of two anchor fingers at finger positions 1 and 3. After 3 independent phage display selections to produce fingers that bind each of the target subsites for a 3-fingered protein, fingers are stitched together in the correct order. (B) Sequential phage display selection attempts to take into account cooperativity between neighboring fingers by sequentially selecting for each finger. First, finger 1 is selected for in the context of two anchor fingers. Then finger 2 is selected for in the context of the newly selected finger 1 and the remaining anchor finger. Next, finger 3 is selected for in the context of the newly selected finger 1 and finger 2 modules. Although not depicted in this figure, scientists have modified this strategy to include an addition round of selections in which they reselect finger 1 in the context of the newly selected finger 2 and finger 3 modules. (C) The bipartite phage display selection involves using two pre-generated ZFP libraries that have each been randomized at the key amino acids. Each library is randomized for half of the ZFP. That is, one library is randomized for the part of the ZFP that binds the 5, 5' nucleotides in the target sequence, while the other library is randomized in the 5, 3' nucleotides in the target sequence. Selections for each half of the ZFP are carried out in parallel, and then the resulting selections are recombined and subjected to an additional round of phage display selection using the full-length target sequence. Figure from (Beerli and Barbas 2002; Durai, Mani et al. 2005).

CHAPTER II: OPTIMIZING ZINC FINGER NUCELASES FOR USE IN GENE TARGETING IN MAMMALIAN CELLS

Abstract

Homologous recombination is a technique used to perform precise genomic manipulations, which makes it potentially ideal for gene therapy. The spontaneous rate of homologous recombination in human cells has been too low to be used experimentally or therapeutically, but inducing a DNA double-strand break (DSB) in the target gene can stimulate this rate. Zinc Finger Nucleases (ZFNs) are synthetic fusion proteins that can induce DSBs at specific sequences of DNA and stimulate gene targeting. Although ZFNs have shown success, several issues remain. First, an optimal generalized method of making effective and safe ZFNs needs to be determined. Second, a systematic method of evaluating efficiency and safety of ZFNs is needed. We compare the gene targeting efficiencies and cytotoxicity of ZFNs targeted to a single site made using modular assembly to those targeted to the same site made using a bacterial two-hybrid (B2H) selection-based method. We found that a ZFN pair made using the B2H strategy is more efficient at stimulating gene targeting and less toxic than a pair made using modular assembly. We demonstrate that a pair of 3-finger B2H

ZFNs is as efficient at stimulating gene targeting and display similar or reduced rates of toxicity as ZFNs with more fingers.

Introduction

Homologous recombination is the most precise way to manipulate the genome and is a natural mechanism that cells use for a variety of processes including the repair of DNA double-strand breaks (DSBs) (Wyman, Ristic et al. 2004). The spontaneous rate of gene targeting - the replacement of an endogenous DNA fragment with an exogenous fragment by homologous recombination - is too low in somatic mammalian cells (10^{-6}) to be of general use although it is used extensively in other experimental systems (Orr-Weaver, Szostak et al. 1981; Porteus and Baltimore 2003; Yamazoe, Sonoda et al. 2004; Capecchi 2005). There are several ways of increasing the rate of gene targeting, including the use of recombinant AAV (Hirata and Russell 2000). The most powerful way of stimulating gene targeting is by introducing a DSB into the target locus, which increases the rate by three to four orders of magnitude (Choulika, Perrin et al. 1995; Smih, Rouet et al. 1995; Brenneman, Gimble et al. 1996; Sargent, Brenneman et al. 1997). These results were originally obtained using the *I-SceI* (Sce) homing endonuclease, and investigators are studying how to re-target such homing endonucleases to endogenous gene-specific sequences (Belfort and Roberts 1997; Stoddard 2005).

Zinc Finger Nucleases (ZFNs) are showing promise as reagents that can create gene-specific DSBs (Durai, Mani et al. 2005). ZFNs are artificial proteins that fuse a specific zinc finger DNA binding domain with a non-specific endonuclease domain from the *FokI* restriction enzyme (Chandrasegaran and Smith 1999). ZFNs can create specific DSBs *in vitro* and DSBs that stimulate gene targeting or homologous recombination in *Xenopus* oocytes, *Drosophila melanogaster*, and plants (Kim and Chandrasegaran 1994; Smith, Bibikova et al. 2000; Bibikova, Carroll et al. 2001; Bibikova, Beumer et al. 2003; Lloyd, Plaisier et al. 2005). In mammalian cells, model ZFNs stimulate gene targeting by several-thousand fold using a GFP reporter system(Porteus and Baltimore 2003). In addition, ZFNs have been designed to recognize an endogenous gene (IL2RG) and stimulate gene targeting at one allele of the endogenous IL2RG locus at rates of 11% and both alleles in 6.5% of cells (Urnov, Miller et al. 2005). Several questions, however, remain: (1) what is the optimal way to make the zinc finger portion of the ZFN, (2) is there an optimal number of zinc fingers for a ZFN, and (3) what is an efficient way of measuring off-target toxicity.

There are over 700 proteins in the mammalian genome that contain Cys2His2 zinc finger motifs. The best studied is the 3-finger transcription factor Zif268, which binds its 10 base pair target site with picomolar affinity (Pavletich and Pabo 1991; Shi and Berg 1995; Elrod-Erickson and Pabo 1999). Each finger of Zif268 consists of ~30 amino acids that fold into a $\beta\beta\alpha$ structure and makes major contacts with a 3-

4 base pair site (Pavletich and Pabo 1991). These contacts are mediated through specific residues in the "recognition helix" of the zinc finger domain. The crystal structure suggested that zinc finger binding of DNA is modular; each finger seemed to bind its cognate site independently of its neighboring fingers. This suggested that zinc finger proteins could be made to recognize new target sequences by mixing and matching individual fingers with known recognition sites ("modular assembly"). New transcription factors and ZFNs have been made by modular assembly (Beerli, Dreier et al. 2000; Bibikova, Carroll et al. 2001; Liu, Rebar et al. 2001; Ren, Collingwood et al. 2002; Bae, Kwon et al. 2003; Bibikova, Beumer et al. 2003; Porteus and Baltimore 2003; Porteus 2006). One limitation of modular assembly is that it does not take into account the potential interactions between fingers in DNA binding interactions which are important in determining affinity and specificity (Greisman and Pabo 1997; Wolfe, Nekludova et al. 2000).

Currently, there are only two published papers describing the use of ZFNs to natural target sites in mammalian cells (Urnov, Miller et al. 2005; Porteus 2006). These papers use two different methods to make the zinc finger domain. In Urnov et al. (2005), a proprietary archive of two-finger cassettes was used to assemble new zinc finger proteins, and proprietary software was used to optimize these proteins (Urnov, Miller et al. 2005). The technology used to generate these proteins is not available to the academic community and the process can be expensive both in time and money. In Porteus (2006), the modular assembly approach was successfully

used to make a pair of ZFNs to the GFP gene (Porteus 2006). However, in Alwin et al. (2005), ZFNs made by modular assembly had only a 50% success rate and they were never able to target a natural site (Alwin, Gere et al. 2005). Clearly, there are deficiencies in the published methods to make ZFNs as they apply to mammalian cells. While most published ZFNs contain 3 fingers, the most striking published result was obtained using 4-finger ZFNs (Urnov, Miller et al. 2005). An important issue therefore is to determine if 3-finger proteins are fundamentally less well behaved than proteins with higher numbers of fingers.

In this work, we demonstrate that a pair of 3-finger ZFNs made using a strategy that takes context dependency into account are both more active and less toxic than those made by modular assembly; that these 3-finger ZFNs are at least as active and perhaps less toxic than a pair of ZFNs with 4-fingers; and establish quantitative measures for ZFN toxicity.

Results

Experimental work has shown several ways of making a zinc finger protein to a new target sequence. Modular assembly involves linking individual fingers with known binding sites together from published datasets of individual fingers for GNN, CNN, and ANN sites (Beerli, Segal et al. 1998; Dreier, Beerli et al. 2001; Beerli and Barbas 2002; Liu, Xia et al. 2002; Dreier, Fuller et al. 2005). In modular assembly, individual fingers from these datasets are fused together to

create a new zinc finger protein (Figure 2.1A). We have successfully made ZFNs to target the GFP gene using the modular assembly approach (Porteus 2006) (Table 2.1). These modular assembly GFP ZFNs were less active than the gold-standard I-SceI nuclease and showed significant cytotoxicity. Modular assembly does not take the potential inter-finger interactions into account. Alternative methods, in which combinations of fingers are subjected to iterative rounds of selection, are capable of producing zinc finger proteins with high specificity and affinity. While phage display has been used successfully for this purpose, another method of design is a bacterial 2-hybrid (B2H) selection system (Greisman and Pabo 1997; Wolfe, Greisman et al. 1999; Joung, Ramm et al. 2000; Hurt, Thibodeau et al. 2003) (Figure 2.1B). We used the B2H system to make zinc finger proteins, and subsequently ZFNs, to target the identical sites in GFP that we had previously targeted using ZFNs made by modular assembly (Table 2.1 and 2.2). The recognition helices for the modular assembly and B2H-ZFNs are shown in Table 2.2.

We used a previously described quantitative version of the B2H system (in which activity is measured by the transcriptional increase in β-galactosidase activity) to compare the ZFPs made by these two design methods (Wright 2006) (Table 2.2). As expected, we found that all of the zinc finger proteins identified by B2H selections had excellent activity. In contrast, only one of the two modular assembly proteins had significant activity.

We next compared the activity of the ZFNs made by modular assembly and B2H selection using a GFP gene targeting assay (Porteus and Baltimore 2003) (Figure 2.2A). Figure 2.3A shows that the B2H proteins are ~5 fold more active than the modular assembly proteins. When we performed a time-course assay, we found that the B2H proteins showed no evidence of toxicity while the modular assembly proteins showed toxicity (Figure 2.3B).

In these experiments, the modular assembly and B2H-ZFNs were made using a previously published nuclease domain that contains two point mutations (Porteus and Baltimore 2003; Porteus 2006). We remade the B2H ZFNs with the wild-type nuclease and found a further 5-fold increase in gene targeting activity (Figure 2.3C).

Because the issue of cytotoxicity is critical in further development of ZFN technology, we developed another quantitative assay to measure this effect. This assay is a flow-cytometry based assay that we call the "Toxicity Assay" (Figure 2.2B). In this assay, the percent of cells transfected with a potentially toxic nuclease is compared to the percent of cells transfected with Sce, a nuclease with minimal toxicity. A lower percent of cells with the nuclease is a sign of greater toxicity. Using this assay we found that the modular assembly ZFNs had significant toxicity even with the mutated nuclease (Figure 2.3E). We found that the B2H-GFP-ZFNs, which did not display toxicity in the gene targeting time course assay, showed cytotoxicity in the Toxicity Assay, and this toxicity was

increased with the wild-type nuclease domain (Figures 2.3D and 2.3E). These results show that the Toxicity Assay is a more sensitive measure of cytotoxicity than the gene-targeting time course assay.

We next determined if re-engineering the nuclease could offset the cytotoxic effects of the B2H-ZFNs while maintaining high rates of gene targeting. Cytotoxicity may be caused by homodimerization of ZFNs at off-target sites thereby causing off-target breaks. The dimerization interface of the nuclease is mediated by a pair of aspartate-arginine salt bridges (Bitinaite, Wah et al. 1998). Recently, Miller et al. and Szczepek et al., (2007) have demonstrated that modifying the nuclease domain so that the nuclease could only cut as an "obligate heterodimer" can reduce off-target effects including the DD, RR nuclease pair described here (Miller, Holmes et al. 2007; Szczepek, Brondani et al. 2007). Using slightly different modifications of the nuclease domain to make an obligate heterodimer pair (B2H-GFP-DD and B2H-GFP-RR), we find that cytotoxicity and off-target effects can be reduced but at the cost of reduced gene targeting efficiency (Figures 2.3D, 2.3E, and 2.3F). Thus, this way of engineering the nuclease domain to reduce cytotoxicity is probably not better than the way reported by Miller et al. and Szczepek et al (Miller, Holmes et al. 2007; Szczepek, Brondani et al. 2007).

We previously found that the rate of gene targeting increases as the amount of Sce or modular assembly GFP-ZFNs increases (Porteus and Baltimore

2003; Porteus 2006). To examine this relationship for the B2H-GFP-ZFNs, we performed a similar titration experiment (Figure 2.3G). We found that at lower amounts (1-10 ng) of B2H-GFP-ZFNs transfected, the rate of gene targeting increased with increasing amounts of nuclease. In contrast, at higher amounts (20-200 ng), the rate of gene targeting decreased as the amount of B2H-GFP-ZFNs was increased, thereby creating an inverse U-shaped curve (Figure 2.3G). We hypothesize that initially more ZFN creates more on-target DSBs and targeting, but further increases only create more off-target DSBs leading to cytotoxicity.

In our experimental design with the GFP gene targeting assay, it is possible that re-cutting could inactivate the corrected GFP gene. Specifically, when the GFP gene is targeted and corrected, it still contains the target site for the GFP-ZFNs and could be re-cut. If re-cutting occurs, mutagenic non-homologous end-joining could inactivate a gene that had initially been corrected by gene targeting. To determine if re-cutting might occur, we used a repair template ("protected repair substrate") with nucleotide changes that prevent re-cutting by the GFP-ZFNs but do not change the amino acid sequence (see figure 2.3 legend) of the GFP gene as had been done in ZFN studies done in Drosophila (Bibikova, Beumer et al. 2003; Beumer, Bhattacharyya et al. 2006). Using the protected repair substrate, low amounts of B2H-GFP-ZFNs transfected did not change the number of GFP positive cells (Figure 2.3H). In contrast, at higher amounts, there

was a 2-fold increase in the number of GFP positive cells (Figure 2.3H). These results suggest that with highly active ZFNs, re-cutting and subsequent mutagenic repair can reduce the final number of targeted events. In summary, we hypothesize that introduction of larger amounts of active nuclease decreases targeting events through two mechanisms: (1) cytotoxicity caused by off-target DSBs and (2) re-cutting of the correct target gene followed by mutagenic repair. By decreasing the amount of ZFN transfected one can mitigate both of these effects.

From the quantitative B2H β-galactosidase assay, we determined that all of the B2H-selected proteins were active while only one of two modular assembly proteins was active. We hypothesized that an inactive protein would be "dominant" in determining the overall gene targeting activity of the pair. To test this, we paired the B2H-GFP-ZFNs with the MA-GFP-ZFNs (Figure 2.3I). When the active MA-GFP-ZFN (GFP2-MA*) was paired with its corresponding B2H-GFP-ZFN partner (GFP1.4-B2H*), it induced gene targeting as well as the two B2H-GFP-ZFNs (Figure 2.3I). However, when the less active MA-GFP-ZFN protein (GFP1-S-MA*) was paired with its corresponding B2H-GFP-ZFN protein (GFP2-B2H*), it continued to have a low rate of gene targeting (Figure 2.3I). Therefore, a poor ZFN can have a dominant effect and high rates of gene

targeting can be obtained only if both proteins show good activity in the quantitative B2H β-galactosidase assay.

Analysis of ZFN Quality Based on the Number of Zinc Fingers

The best previously published ZFNs contained 4-fingers and an important consideration is whether these ZFNs were better because they had 4 rather than 3 fingers(Urnov, Miller et al. 2005). We compared our 3-finger B2H-GFP-ZFNs with previously published 4-finger proteins (SCID-ZFNs) using the GFP gene targeting and the Toxicity Assay (Figure 2.2A describes how we test the SCID-ZFNs using the GFP reporter assay).

We performed a titration of the SCID-ZFNs in the GFP gene targeting assay. Just as with the B2H-GFP-ZFNs, the SCID-ZFNs displayed an inverse U-shaped curve with maximal gene targeting activity between 10 and 20 nanograms (Figure 2.4A). The SCID-ZFN and B2H-GFP-ZFN titration experiments show that each pair of ZFNs has an optimal range for maximal gene targeting activity. The 3-finger B2H-GFP-ZFNs performed better than the 4-finger SCID-ZFNs in this assay (Figure 2.4B). In both the gene targeting time course assay and the Toxicity Assay, the 3-fingered B2H-GFP-ZFNs exhibited less cytotoxicity than the 4-fingered SCID-ZFNs (Figure 2.4C and 2.4D).

To further evaluate the impact of how finger number affects the activity and cytotoxicity of a ZFN, we compared the 3-finger B2H-GFP-ZFNs with a 6-

finger ZFN (Chk2). The 6-finger Chk2-ZFN was derived from Tan *et al.,* (2003) who reported that their 6-finger transcriptional repressor designed to bind to the Chk2 promoter was highly specific (Porteus, Cathomen et al. 2003). Unlike the 3 or 4-finger proteins, we did not identify an inverse U-shaped curve for gene targeting activity (Figure 2.5A). Instead, increasing amounts of Chk2-ZFN caused increased targeting until a plateau was reached. We found that the 3-finger proteins were as good as the 6-finger proteins in the gene targeting and gene targeting time course assays (Figure 2.5B and 2.5C). The single 6-finger protein showed less cytotoxicity than the pair of 3-finger ZFNs in the Toxicity Assay (Figure 2.5D).

In the experiments described thus far, we compared ZFNs by titrating the amount of ZFN transfected to determine the amount that gave maximal targeting activity. We also compared ZFN activity by normalizing to equal protein expression (Figure 2.6). We found that at equal expression levels the B2H-GFP-ZFNs were 5-10 fold more active than both the SCID and Chk2-ZFNs but also more toxic. We found, however, that the "B2H-DD/RR" constructs were 4-fold more active than the SCID-ZFNs and less toxic (Figure 2.6).

Analysis of Mechanism of Cytotoxicity

Given the importance of ZFN cytotoxicity, we began further studies exploring the mechanism. We hypothesized that cytotoxicity would increase as the amount of ZFN increased. In figure 2.7A, we performed a titration

experiment using the Toxicity Assay. For the 3, 4 and 6-finger proteins, cytotoxicity did increase as the amount of ZFN transfected increased. The obligate heterodimer "B2H-DD/RR" pair did not show this effect (Figure 2.7A). We also performed the Toxicity Assay in murine embryonic stem cells and primary human foreskin fibroblasts, and obtained similar results where the 3-finger B2H-ZFNs were less toxic than the 4-finger proteins and displayed a relatively equivalent level of toxicity to the 6-finger nucleases (Figures 2.7B, 2.7C).

Evidence has suggested that the cytotoxicity of ZFNs is from the creation of off-target DSBs (Beumer, Bhattacharyya et al. 2006). We used a foci formation assay to visualize DSBs. When a DSB occurs, it activates a signaling cascade including the phosphorylation of H2AX and the recruitment of various proteins to the site of the DSB (including 53BP1). By staining for 53BP1 foci, a quantitative measure of the number of DSBs at a given moment can be made (Schultz, Chehab et al. 2000). We transfected increasing amounts of the various ZFNs into primary human foreskin fibroblasts, stained for 53BP1 foci, and quantitated the number of foci in transfected cells as marked by their green fluorescence. Figure 2.7D shows representative cells after staining for 53BP1 foci. We first analyzed the data by determining the average number of foci per transfected cell. In this assay, untransfected cells or cells transfected with GFP alone had an average of less than 1 focus per cell. As further controls, cells

transfected with Sce had approximately 1 focus per cell, while cells transfected with CAD had approximately 8 foci per cell transfected (Figure 2.7E). As in the Toxicity Assay, we found that increasing amounts of ZFN transfected caused an increase in the number of 53BP1 foci per cell for all the ZFNs except for the "B2H-DD/RR" pair. Most striking is that the 4-finger ZFNs show many more foci per cell than other ZFNs.

We next quantitated the number of cells with a given number of foci and plotted this data in figure 2.7F. This analysis showed several important subtleties. Increasing the amount of ZFN, increased the number of cells with higher number of foci. In figure 2.7A and 2.7E, higher amounts of the "B2H-DD/RR" pair did not show increased cytotoxicity or foci. In figure 2.7F, however, we can see that higher amounts do cause an increase in the fraction of cells with 2-5 foci per cell transfected. We did not find cells with over 6 foci under any condition with the "B2H-DD/RR" pair and did not observe cytotoxicity in the Toxicity Assay with this pair (Figure 2.7A and Figure 2.7F). These data suggest that toxicity in the Toxicity Assay correlates with the number of cells having greater than 6 or more foci. In summary, the 53BP1 foci assay correlates extra DSBs with increased cytotoxicity and provides further evidence that one mechanism of cytotoxicity is through off-target DSBs. Moreover this assay provides an additional quantitative measure of the quality of a pair of ZFNs.

Discussion

In this study of a single site, we describe important steps towards efficiently assessing the activity and specificity of ZFNs. We found that by using a selection strategy (B2H) that takes inter-finger interactions into account, we can generate ZFNs that are both more active and less cytotoxic (Hurt, Thibodeau et al. 2003). These proof-of principle experiments suggest that 3-finger proteins made by the B2H system might be the optimal platform for ZFNs. Finally, using a quantitative B2H β-galactosidase assay, a mammalian GFP gene targeting reporter assay, a mammalian flow-cytometry based Toxicity Assay (Toxicity Assay), and a 53BP1 foci formation assay, we have developed a generalized method to systematically evaluate ZFNs. These assays or variations thereof should provide a standard for future ZFN development.

One of the important aspects to the development of ZFN technology is to establish an optimal design strategy for the zinc finger DNA binding domain. Modular assembly has been used to successfully make zinc finger transcription factors and ZFNs and has the advantage that it can be performed in any laboratory capable of standard molecular biology manipulations (Beerli, Segal et al. 1998; Beerli, Dreier et al. 2000; Liu, Rebar et al. 2001; Beerli and Barbas 2002; Ren, Collingwood et al. 2002; Bae, Kwon et al. 2003; Porteus 2006). We had made ZFNs to the GFP gene using modular assembly and found that proteins made to the same target site using the B2H system were better in terms of both efficacy

and toxicity (Porteus 2006). We found that only one of the two modular assembly proteins had activity in a bacterial reporter assay and translated into the pair having only moderate activity in the GFP gene targeting assay. In contrast, we found that all of the proteins made by B2H were active in the bacterial reporter assay and translated into them having excellent activity in the GFP gene targeting assay. Our data also suggest that if a zinc finger protein is active in the quantitative B2H β-galactosidase assay, irrespective of its initial design method, it will have excellent activity in the gene targeting reporter assay. If both modular assembly proteins are active in the β-galactosidase assay, our data suggests that this pair would be active in gene targeting. This comparison of modular assembly and B2H was obtained from a single site and a similar comparison needs to be performed at multiple sites to validate our findings as a general principle.

Another interesting finding was the observation that optimal rates of gene targeting can be achieved by altering the repair template to prevent re-binding and subsequent re-cutting by the ZFNs. We hypothesize that this beneficial effect is the result of preventing "re-cutting mutagenesis" whereby the "repaired" gene is re-cut by the ZFNs and then subsequently mutagenized by non-homologous end-joining.

Our work suggests that 3-finger B2H-ZFNs are as well-behaved as proteins with more fingers. We used these 4 and 6-finger proteins because they

were among the best previously published proteins (Porteus, Cathomen et al. 2003; Urnov, Miller et al. 2005). In fact, the 4-finger proteins produced strikingly high rates of targeting at an endogenous target in human cells (Urnov, Miller et al. 2005). One unavoidable limitation when comparing the B2H-3-finger proteins to the 4- and 6-finger proteins is that the proteins bind to different sequences. While we compared these proteins by targeting adjacent sequences in an identical locus, the nature of zinc finger binding means that a direct comparison of proteins with different numbers of fingers is not possible. Nonetheless, our data suggest that the number of fingers is not the sole determinant of whether a ZFN will have high activity and low toxicity.

In this work, we have made measuring and minimizing ZFN cytotoxicity a major focus. Specifically, we have developed a flow cytometry based cytotoxicity assay ("Toxicity Assay") as a quantitative way to efficiently screen for ZFNs for cytotoxicity either individually or as pairs. We have found that this screening assay can be used in the HEK-293 transformed cell line, in untransformed murine embryonic stem cells and in primary human foreskin fibroblasts. The ease, high-throughput nature, sensitivity and use in multiple cell lines suggest that the Toxicity Assay is a good screening method for evaluating and comparing toxicities of different ZFNs. In addition to its increased sensitivity over the gene-targeting time course assay, the Toxicity Assay has the additional advantage that it can be used to evaluate individual ZFNs (data not shown). We

established a more direct measure of off-target DSBs using an immunofluorescence foci formation assay with 53BP1. We found that there is a direct correlation between number of foci and cytotoxicity in the Toxicity Assay. One limitation of the foci formation assay is that it is just a "snap-shot" of a dynamic process and underestimates the true number of off-target DSBs that a pair of ZFNs might be creating. Nonetheless, this assay evaluates a critical aspect of cytotoxicity of ZFNs and establishes an important method of assessing new ZFNs, particularly those that may be destined for clinical use.

In summary, the work presented here demonstrates that by using a series of assays, one can examine both the activity and toxicity of pairs of ZFNs. We believe that no single assay will provide the necessary data to fully understand the potential of any ZFN pair, especially if ZFNs are to reach their full potential of being used therapeutically. We suggest using a suite of assays as a screening strategy to elucidate highly active ZFNs with optimal toxicity profiles for use in mammalian cells. For example, the first screening test for a new ZFN would be to test the zinc finger domain using a quantitative transcriptional reporter assay such as the quantitative bacterial β-galactosidase assay. Those pairs in which both proteins show good activity in this first screen should then be analyzed further in a mammalian cell-based assay. The gene targeting reporter assay is one such cell-based assay and allows one to analyze the activity of the selected ZFNs in a mammalian cell context on a chromosomally integrated substrate. At this point, the

gene targeting reporter assay can be utilized to determine a dose response curve for each pair of nucleases and to identify the optimal effective amount for each ZFN pair. Once the optimal effective amount has been determined, a series of assays evaluating toxicity can then be performed at the identified effective amount. To analyze the toxicity profile of a given pair of ZFNs, a series of assays such as a gene targeting time-course assay, a cell-based toxicity assay, and a DSB foci formation assay can be used. The gene targeting time-course assay can identify ZFNs with poor toxicity profiles. The time-course assay, however, is less sensitive than the new cell-based toxicity assay. The DSB foci formation assay allows a more direct examination of the number of DSBs that a pair of ZFNs is creating. Ideally, the activity and toxicity of a pair of ZFNs could be analyzed using a single assay, however such an assay does not currently exist. As a result, we believe that a series of assays such as those described above will be necessary to fully analyze the potential of ZFNs made by any design strategy. After a pair of ZFNs is found to have the needed efficacy and toxicity attributes using such a suite of assays, it would then be reasonable to begin to study their efficacy in stimulating gene targeting at an endogenous locus. Overall, the results presented here support continued optimism about the potential of ZFNs to mediate high rates of gene targeting in disease causing genes for therapeutic use.

Material and Methods:

DNA Manipulations and Cloning.

All plasmids were made using standard cloning and molecular biology as previously described (Ausubel 1996; Wright 2006). ZFNs made via modular assembly were constructed and fused to the *FokI* nuclease domain as previously described (Porteus 2006). ZFNs made via the B2H selection method were fused to the *FokI* nuclease domain using PCR primers in which the 5' primer hybridized with the 5' end of finger 1 and contained a *BamHI* site (5'-ACGTGGATCCACCATGGACTACAAAGACGATGACGACCCAAAAAAGAAGCGAAAGGTA-3') and the 3' primer hybridized with the 3' end of finger 3 and contained a *SpeI* site (5'-CTTTTGACTAGTTGGTCCTTCTCGTCCTGTGTGGGTTTTTAGGTG-3'). The PCR amplified ZFP was then moved into pcDNA6 (Invitrogen, Carlsbad, CA) in frame and upstream of the FokI nuclease. The GFP-DD and GFP-RR ZFNs were made using a 3 part ligation in which pcDNA6 was digested with *BamHI* and *ApaI*, the *FokI* nuclease domain digested with *SpeI* and *ApaI*, and the zinc finger protein with *BamHI* and *SpeI*. In the GFP-DD version R487 was converted to aspartic acid and in the GFP-RR version D483 is converted to arginine using PCR mutagenesis. In the previously published ZFNs, the nuclease domain inadvertently contained two point mutations: P509T and N586S (numbering with respect to FokI) (Porteus and Baltimore 2003; Porteus 2006). In the ZFNs marked by a "*" the nuclease domain

contains those mutations and the ZFNs not marked do not contain those mutations. The wild-type nuclease domain has greater activity measured both in gene targeting assays and in cytotoxicity assays.

Bacterial 2-Hybrid Selections

Bacterial 2-Hybrid (B2H) selections were performed as previously described (Hurt, Thibodeau et al. 2003). Briefly, selection strains were constructed with the target DNA site or subsite positioned immediately upstream of a weak *lac* promoter that directs transcription of the selection markers *His3* and *aadA* genes. ZFP-Gal11p protein fusions were expressed through an inducible *lac*UV5 promoter. Selections were carried out in two stages. The first stage was performed by introducing a master randomized finger library into the appropriate selection strain containing the 3 base pair target subsite for the binding site of interest (Joung, Ramm et al. 2000). Transformed cells were plated on selection medium and colonies that grew at this first stage selection were harvested. Zinc finger protein phagemid is then made from these bacteria and used to re-infect the selection strain of interest in order to generate a pool of about 100 clones to move to the second stage selection. Colonies that survived the entire first round selection were harvested for the ZFP-encoding plasmids and used to construct the shuffled three-finger library by fusion PCR. High stringency, stage two selections were performed by introducing the appropriate shuffled three-fingered library into the appropriate selection strain containing the full-length target site. Transformed cells were plated on a series of selective plates

with increasing degrees of stringency. Candidates were chosen from the highest stringency plates and sequenced. The zinc finger proteins reported here were chosen from plates containing 0 μM isopropyl β-D-thiogalactoside(IPTG), 50 mM 3-aminotrazole (3-AT), and 80 μg/ml streptomycin. In stage one of finger 2 selection for B2H-GFP2, we were unable to generate a pool of 100 different fingers for technical reasons (SMP, MHP, JKJ, unpublished data). We hypothesize that because finger 1 and finger 3 of the framework ZFP used in this study bound with weak affinity that selection of finger 2 required such a strong interaction between the target sequence and the randomized finger that only a few clones came through the stage one selection. Dr. Keith Joung's lab is currently working on these technical issues and questions regarding this problem can be addressed to him (JJOUNG@PARTNERS.ORG). To solve this problem for this study, we assigned a recognition helix for finger 2 using the dataset of Liu *et al.,* (2002) and incorporated that module into the library used in the stage 2 selection(Liu, Xia et al. 2002). Thus, at one level, the B2H-GFP2 ZFN represents a hybrid protein because it contains a single module at finger 2. At the final level, however, it is a B2H protein because it was ultimately selected using a strategy that takes inter-finger interactions into account.

Quantitative B2H-based β-galactosidase Reporter Assays

A transcriptional assay in bacteria was used as an assay to determine the activity of the selected ZFPs as previously described (Thibodeau, Fang et al. 2004; Wright

2006). Briefly, a plasmid with the target binding site inserted upstream to the lacZ reporter gene was introduced into bacterial cells at a single copy creating a reporter strain. ZFPs were then introduced into the reporter strain, induced by IPTG and the amount of β-galactosidase activity was determined and normalized to reporter strains in which the ZFP was not expressed.

Cell Culture and Transfections

All cell culture experiments were performed in *HEK*293 cells except where identified as otherwise. Cells were cultured in a humidified incubator at 37°C with 5% CO_2 in DMEM supplemented with 10% bovine growth serum (Hyclone, Logan, UT, USA), 2mM L-glutamine, 100IU/ml penicillin, and 100mg/ml streptomycin. Stable cell lines were constructed as previously described(Porteus 2006). Transient transfections were performed using the calcium phosphate technique as previously described and produced transfection efficiencies between 10-35%(Porteus and Baltimore 2003).

Measurement of Gene Targeting Using the GFP system

Gene targeting experiments were performed in triplicate as previously described using calcium phosphate transfection (Porteus and Baltimore 2003; Porteus 2006). Transfection efficiencies were determined at day 2 post-transfection, and the rates of gene targeting were determined by flow-cytometry and analyzed on a FACS Calibur (Becton-Dickerson, San Jose, CA, USA) at day 3, 5, and 7 after transfection (day of

transfection is considered day 0). Gene targeting rates are expressed as GFP positive cells per million cells transfected because the background rate of spontaneous gene targeting is approximately one event per million cells using this system.

Testing Zinc Finger Nucleases Using the GFP Gene Targeting system

The new B2H ZFNs were tested in a stably transfected cell line. The SCID ZFNs were tested in a stably transfected cell line similar to one previously described(Urnov, Miller et al. 2005). Briefly, target sites were inverted (ZFP-left 5'-AAAGCGGCTCCG-3'and ZFP-right 5'ACTCTGTGGAAG-3') and separated by a 5 base pair spacer (5'-GGGTT-3'). The Chk2 ZFNs were tested in a stably transfected cell line similar to the SCID cell line. Specifically, the Chk2 6-finger binding site (5'-ACCCGGGTTCCCCTCGGG-3') was inserted as an inverted repeat and separated by a 6 base pair spacer downstream of a stop codon and an *I-SceI* recognition site.

Flow-cytometry Assay for Cell-survival: "Toxicity Assay"

*HEK*293 cells were transfected in triplicate by calcium phosphate technique with 200ng of a GFP expression plasmid and with a nuclease expression plasmid(s). At day two post-transfection, a fraction of transfected cells was analyzed by flow-cytometry and the percentage of GFP positive cells was determined. The remaining fraction was replated in a 6-well plate and incubated until day 4 post-transfection. At day 4, the cells were split 1:4 and continued to be cultured. At day 6 post-

transfection, the percentage of GFP positive cells was determined by flow-cytometry. To calculate the percent survival relative to Sce, a ratio of ratios was calculated. The percentage of GFP positive cells remaining at day 6 that were transfected with the Sce expression vector was divided by the percentage of GFP positive cells determined at day two after transfection. This serves as a control for the spontaneous loss of GFP expression after transient transfection. The same calculations were performed for the cells that were transfected with the nuclease expression vectors of choice. The ratio after nuclease transfection was normalized to the ratio after Sce transfection and this was the percent survival compared to Sce. In control experiments, we showed that Sce expression had no effect on cell-survival compared to cells transfected with an empty expression vector. The data presented in Figure 2E, 3D, and 4D were obtained from the same experiment and were separated into different figures for clarity. For figure 5B, primary murine embryonic stem cells were transfected using lipofectamine 2000 (Invitrogen, Carlsbad, CA, USA) according to standard manufacturer's protocol. In each condition, the cells were transfected with 800 ngs total of nuclease. For figure 5C, primary human foreskin fibroblasts were transfected with nucleases and GFP using Nucleofector technology with the basic nucleofector kit for primary fibroblasts (Amaxa Inc., Gaithersburg, MD, USA) according to the protocol published by Amaxa Inc. For each sample, $1x10^6$ cells were nucleofected with 2ug of pmaxGFP (Amaxa) and 2ug of total nuclease. Nucleofection was carried out using program U-23.

Foci Formation assay

Cell Culture

Primary Human Foreskin Fibroblasts were maintained in α-modification essential medium (Hyclone, Logan, VT, USA) supplemented with 20% fetal calf serum and 2mM L-Glutamine. Cells were maintained in a humidified incubator with 5% CO_2 at 37°C.

Transfection of Primary Fibroblasts

Foci formation assays were performed in several cells lines (data not shown). Human foreskin fibroblasts were used because of low background. Primary fibroblasts (passage 18) were transfected with nucleases and GFP using Nucleofector technology with the basic nucleofector kit for primary fibroblasts (Amaxa Inc., Gaithersburg, MD, USA) according to the protocol published by Amaxa Inc. For each sample, $1x10^6$ cells were nucleofected with 2ug of pmaxGFP (Amaxa) and 0.1, 0.5 or 1ug of each nuclease. Nucleofection was carried out using program U-23. Immediately afterwards, cells were plated in 4 well Lab-Tek II Chamber Slides (Nalge Nunc Int., Rochester, NY, USA) containing supplemented α-MEM (Hyclone) and transferred to an incubator. Twenty-four hours later media was replaced with fresh supplemented α-MEM. Forty-eight hours post nucleofection cells were fixed, stained, and visualized as discussed below. 53BP1 foci were only

counted in cells that were GFP positive as these had been transfected with the GFP and the nuclease(s).

Immunofluorescence staining.

Cells were washed twice in 1X PBS then fixed in cold 4% paraformaldehyde for 20 minutes. Cells were then washed in PBS, permeabilized in 0.5% Triton X-100 for 10 minutes, washed, and blocked in 5% bovine serum albumin for 1 hour at room temperature. The cells were incubated with rabbit anti-53BP1 (Cell Signaling, Danvers, MA, USA) for 1 hour, washed four times in PBS, and then incubated with goat anti-rabbit Rhodamine Red-X (Invitrogen, Carlsbad, CA, USA) diluted in 1.5% BSA/ PBS for 1 hour at room temperature. Cells were then washed five times in PBS and mounted in Vectashield mounting medium with 4,6-diamidino-2-phenylindole (Vector Laboratories, Burlingame, CA). Fluorescence images were captured using an epifluorescence microscope equipped with a Q-Fire CCD camera (Olympus America Inc, Melville, NY) and QCapture software (QImaging). Images were merged using ImageJ software (NIH, ver. 1.36b).

Name	Abbreviation	Target Gene	# of Fingers	Target Sequence	Design Method	Reference
GFP-ZFN1-9-MA*	GFP1-9-MA*	GFP	3	5' GAAGATGGT 3'	MA	Porteus. Mol Therapy. 2006.
GFP-ZFN1-5-MA*	GFP1-5-MA*	GFP	3	5' GAAGATGGT 3'	MA	
GFP-ZFN2-MA*	GFP2-MA*	GFP	3	5' GACGACGGC 3'	MA	Porteus. Mol Therapy. 2006.
GFP-ZFN1.3-B2H*	GFP1.3-B2H*	GFP	3	5' GAAGATGGT 3'	B2H	
GFP-ZFN1.4-B2H*	GFP1.4-B2H*	GFP	3	5' GAAGATGGT 3'	B2H	
GFP-ZFN2-B2H*	GFP2-B2H*	GFP	3	5' GACGACGGC 3'	B2H	
GFP-ZFN1-B2H	GFP1-B2H	GFP	3	5' GAAGATGGT 3'	B2H	
GFP-ZFN2-B2H	GFP2-B2H	GFP	3	5' GACGACGGC 3'	B2H	
GFP-ZFN1.4-B2H(DD)	GFP1-B2H-DD	GFP	3	5' GAAGATGGT 3'	B2H	
GFP-ZFN2-B2H(RR)	GFP2-B2H-RR	GFP	3	5' GACGACGGC 3'	B2H	
SCID-ZFN1-MA	SCID1	IL2RG	4	5' ACTCTGTGGAAG 3'	PD/MA/ED	Urnov et al., Nature. 2005.
SCID-ZFN2-MA	SCID2	IL2RG	4	5' AAAGCGGCTCCG 3'	PD/MA/ED	Urnov et al., Nature. 2005.
Chk2-ZFN1-MA	Chk2	Chk2	6	5' ACCCGGGTTCCCCTCGGG 3'	PD/MA	adapted from Tan et al., PNAS. 2003.

Table 2.1. Summary of zinc finger nucleases.

The left column lists the names of the zinc finger nucleases used in this paper. The (*) indicates that the FokI nuclease domain contains two point mutations (the mutations were inadvertently present in our prior ZFNs, see materials and methods for a description of these mutations)(Porteus and Baltimore 2003; Porteus 2006). Each ZFN has been given an abbreviated name as shown in column 2. The target sequence is the nucleotide sequence to which the ZFN was designed to bind. Design methods have been abbreviated as follows: Modular assembly (MA), Bacterial-2-Hybrid (B2H), Phage Display (PD), Empirical Design (ED). The GFP1-R-MA* and GFP2-MA* ZFNs are identical to GFP1-ZFN and GFP2-ZFN previously described(Porteus 2006). GFP-ZFN1-S-MA* differs by having a serine rather than an arginine at position 1 of the recognition helix for finger 2. The GFP-DD ZFN has been mutated at position R487 using PCR mutagenesis to convert the arginine to aspartic acid and in the GFP-RR version D483 is converted to arginine. Both GFP-DD and GFP-RR were made by the B2H selection design strategy. SCID1 and SCID2 are similar to ZFN-L and ZFN-R as previously described except that the nuclease domain has been replaced by the wild-type FokI nuclease domain that has been codon optimized for expression in humans(Urnov, Miller et al. 2005). The Chk2 ZFN was made by taking the recognition helices from the zinc finger protein described in Tan *et al*.,(2003) inserting them into a Zif268 backbone, adding a TGGEKP linker between fingers 2 and 3 and 4 and 5 and then attaching the nuclease domain to the C-terminus (Porteus, Cathomen et al. 2003).

ZFN Abbreviation	Finger-1	Finger-2	Finger-3	Finger-4	Finger-5	Finger-6	Fold-activation (bacteria)
Target Site:	**GGT**	**GAT**	**GAA**				
	-1123456	-1123456	-1123456				
GFP1-S-MA	**QSSHLTR**	**TSGNLVR**	**QSGNLAR**				1.44
GFP1.1-B2H	TKEKLDV	VAH**N**LTR	**Q**HP**N**LTR				7.1
GFP1.2-B2H	LRH**H**LEA	LMH**N**LHR	**Q**RP**N**L**G**R				5.2
GFP1.3-B2H	MKHHLDR	VKH**N**LHR	**G**GD**N**L**N**R				7.32
GFP1.4-B2H	TRQKLGV	VAH**N**LTR	**Q**HP**N**LTR				6.03
Target Site:	**GGC**	**GAC**	**GAC**				
	-1123456	-1123456	-1123456				
GFP2-MA	**DRSHLTR**	**DRSNLTR**	**DRSNLTR**				6.91
GFP2.1-B2H	AP**S**KLDR	**DRSNLTR**	E**G**G**N**L**M**R				5.9
GFP2.2-B2H	AP**S**KLDR	**DRSNLTR**	E**G**G**N**L**M**R				5.9
GFP2.3-B2H	AP**S**KLDR	**DRSNLTR**	E**G**G**N**L**M**R				5.9
GFP2.4-B2H	AP**S**KLDR	**DRSNLTR**	D**Q**G**N**LTR				5.55
Target Site:	**AAG**	**TGG**	**CTG**	**ACT**			
SCID1	RSDNLSV	RNAHRIN	RSDTLSE	ARSTRTN			
Target Site:	**CCG**	**GCT**	**GCG**	**AAA**			
SCID2	RSDTLSE	ARSTRTT	RSDSLSK	QRSNLKV			
Target Site:	**GGG**	**CTC**	**CCC**	**GTT**	**CGG**	**ACC**	
Chk2	RSDHLSR	DNRDRTK	DRKTLIE	TSSGLSR	RSDHLSE	TSSDRTK	

Table 2.2. Amino acid sequence of zinc finger nucleases.

The left column lists the ZFN abbreviations as in Table 1. The amino acid sequence for the recognition helices are listed and the numbering above is with respect to the beginning of the α-helix and reflects standard nomenclature. Positions -1, 3, and 6 of the recognition helix make major contacts with DNA(Pavletich and Pabo 1991). The corresponding amino acid sequences use the single-letter code. The target sequence to which each zinc finger protein (ZFP) was designed to bind is listed above each listed ZFP in bold and is in the 3' to 5' direction. To compare the amino acid sequence of the MA-ZFPs to that of the B2H-ZFPs, common amino acids are highlighted in bold. We assigned and fixed finger 2 of all B2H-GFP2 because we were unable to generate a complex finger 2 pooled library. Thus, the B2H-GFP2 ZFN is a hybrid of B2H selection and modular assembly, but we consider it a B2H protein because in the final selection on the full 9 base pair target site, fingers 1 and 3 were selected in the context of each other and the fixed finger 2. The leucine at position 4 is not highlighted because this residue is conserved in all of our 3-finger proteins. The B2H-ZFPs highlighted in grey were randomly chosen and made into ZFNs and used in subsequent experiments. GFP-2.1-B2H, GFP-2.2-B2H, and GFP-2.3-B2H are identical and will be called GFP2-B2H hence forth. More comprehensive experiments are in progress to compare a panel of ZFNs made by the B2H strategy. These were picked randomly and complete comparison of all of the proteins is ongoing. The right most column lists the fold-activation of each 3-finger GFP-ZFN in the quantitative bacterial β-galactosidase reporter assay (Thibodeau, Fang et al. 2004; Wright 2006).

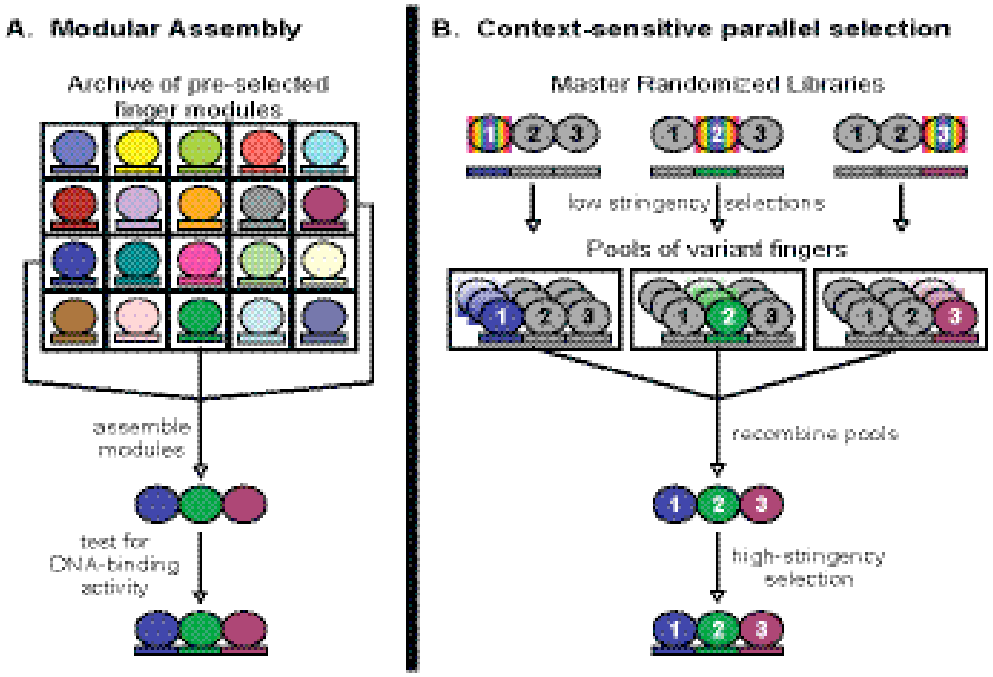
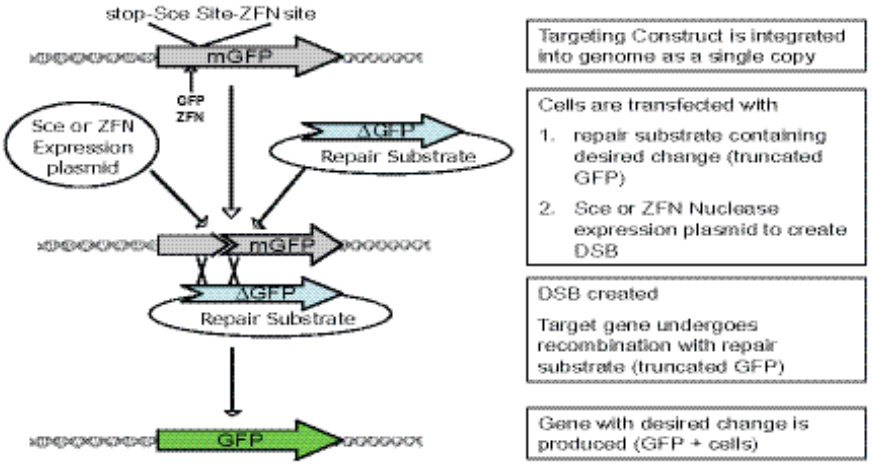


Figure 2.1. Comparison of zinc finger protein design strategies.

(A) The modular assembly approach to zinc finger protein design involves linking finger modules from an archive of pre-selected individual finger modules (pictured as the matrix of colored circles depicting individual fingers binding their respective target sites depicted as color-matched rectangles) together in tandem arrays to create 3-finger proteins. This approach does not take inter-finger interactions into account. A parallel selection strategy (bacterial-2-hybird (B2H) strategy) takes inter-finger interactions into account. Master randomized libraries were created from a synthetic framework zinc finger protein with known DNA binding specificity (grey circles). Three master libraries were created allowing randomization of each of the three finger positions (rainbow colored circles). Two rounds of selections are performed. This first round is performed using a low stringency selection and generates pools of variant fingers that bind to their respective subsite with acceptable affinity in the context of the framework zinc finger protein. These pools are randomly recombined to form a collection of 3-finger zinc finger proteins for which higher stringency selections are carried out in the context of the other individually selected fingers for a given full target DNA site. Overall, the strategy first creates a pool of fingers for a 3-basepair subsite and then selects for the best finger(s) in the context of the desired overall protein for a 9-basepair full target site.

A **GFP Gene Targeting Assay**



B **Flow Cytometry Cell Survival Assay "Pirate Assay"**

1. **Co-transfect cells with GFP expression plasmid and nuclease expression plasmid.**
2. **Day 2: Flow cytometry to measure baseline number of GFP positive cells.**
3. **Day 6: Flow cytometry to measure number of GFP positive cells.**
4. **Calculate ratio of ratios:**

$$\frac{\text{ZFN Day6/ZFN Day2}}{\text{Sce Day6/Sce Day2}} \times 100$$

Ratio=100 no toxicity

Ratio<100 toxicity

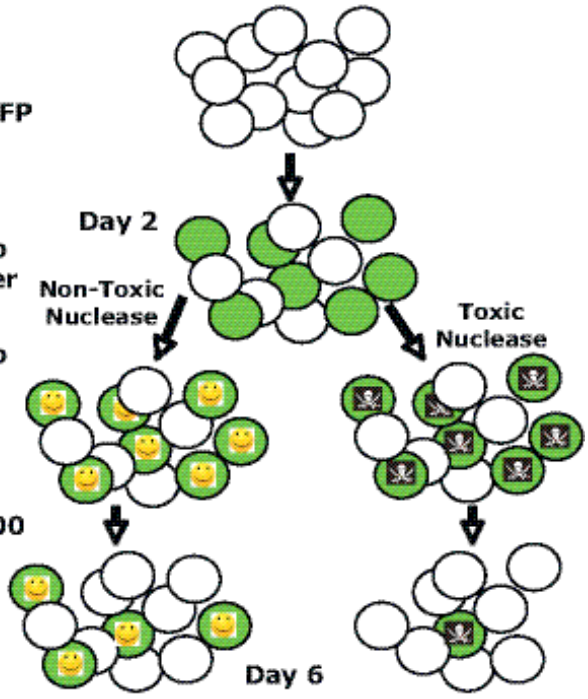


Figure 2.2. Assays used to analyze quality of ZFNs.

(A) GFP gene targeting assay. The GFP gene targeting system is as previously described and is designed to mimic the correction of an endogenous gene(Porteus 2006). A stable cell line is generated in which the target construct is integrated into the genome as a single copy. The target construct consists of a mutated GFP gene (mGFP). The GFP gene has been mutated by the insertion of a stop codon, *I-SceI* endonuclease recognition site (18 base pairs), and the ZFN site of interest (SCID or Chk2). This stable cell line is GFP negative. The GFP-ZFNs are designed to recognize a target site (arrow) adjacent to the insertion described above. Because ZFNs with different numbers of fingers can never be directly compared at the same target sequence, this system of comparing ZFNs at adjacent sites in the same chromosomally integrated locus provides the best alternative. After stable cell lines are generated with the appropriate target construct, cells are transfected with (1) a repair substrate (ΔGFP) and (2) the Sce or ZFN nuclease expression plasmid(s). For figure 2H, RS+ is identical to GFP Donor except that the wild-type sequence 5'-ACCATCTTCTTCAAGGACGACGGC-3' is replaced by 5'-AC*A*AT*T*TT*T*TTCAAGGA*T*GA*T*GGC-3', which introduces five wobble mutations (italics) that abrogate ZFN cutting but do not change the amino acid sequence of GFP. Gene targeting occurs when the nuclease creates a double-strand break in the integrated target and the break is repaired by the repair substrate (ΔGFP) by homologous recombination. When these occur, the mutated GFP gene is converted to wild-type, functional GFP is produced, and the cell can be identified as GFP positive by flow-cytometry. The rate of gene targeting is determined by normalizing the fraction of GFP positive cells to the transfection efficiency. (B) Flow cytometry cytotoxicity assay "Toxicity Assay". Cells are co-transfected with a full-length GFP expression plasmid and either an *I-SceI* expression plasmid or ZFN expression plasmid(s) at day 0. At day 2, flow-cytometry is performed to establish the baseline number of cells that are GFP positive (transfection efficiency). The cells are then cultured until day 6, at which time they are again analyzed by flow-cytometry for GFP expression. Loss of GFP positive cells occurs through two mechanisms. The first is the spontaneous loss of the GFP expression plasmid in dividing cells. The ratio of GFP positive cells on day 6 to day 2 in the cells transfected with *I-SceI* determines this rate. The second is through cytotoxicity from the nuclease. If the nuclease has cytotoxicity (depicted as sword and skull pirate flag), it will cause these cells to either die or to stop dividing and thus the percentage of GFP positive cells (which is a marker of cells that have also been transfected with the nuclease) will decrease faster than if the nuclease has no cytotoxicity (as for *I-SceI* and depicted by "smiley faces" in the figure). To determine the degree of cytotoxicity, a ratio of ratios is calculated

as shown in the figure. If the ratio is <100, it demonstrates that the nuclease has cytotoxic effects on the cells. The lower the ratio, the greater the cytotoxicity. As a positive control for toxicity, we use Caspase Activated Dnase (CAD) that causes non-specific double-strand breaks in the cell.

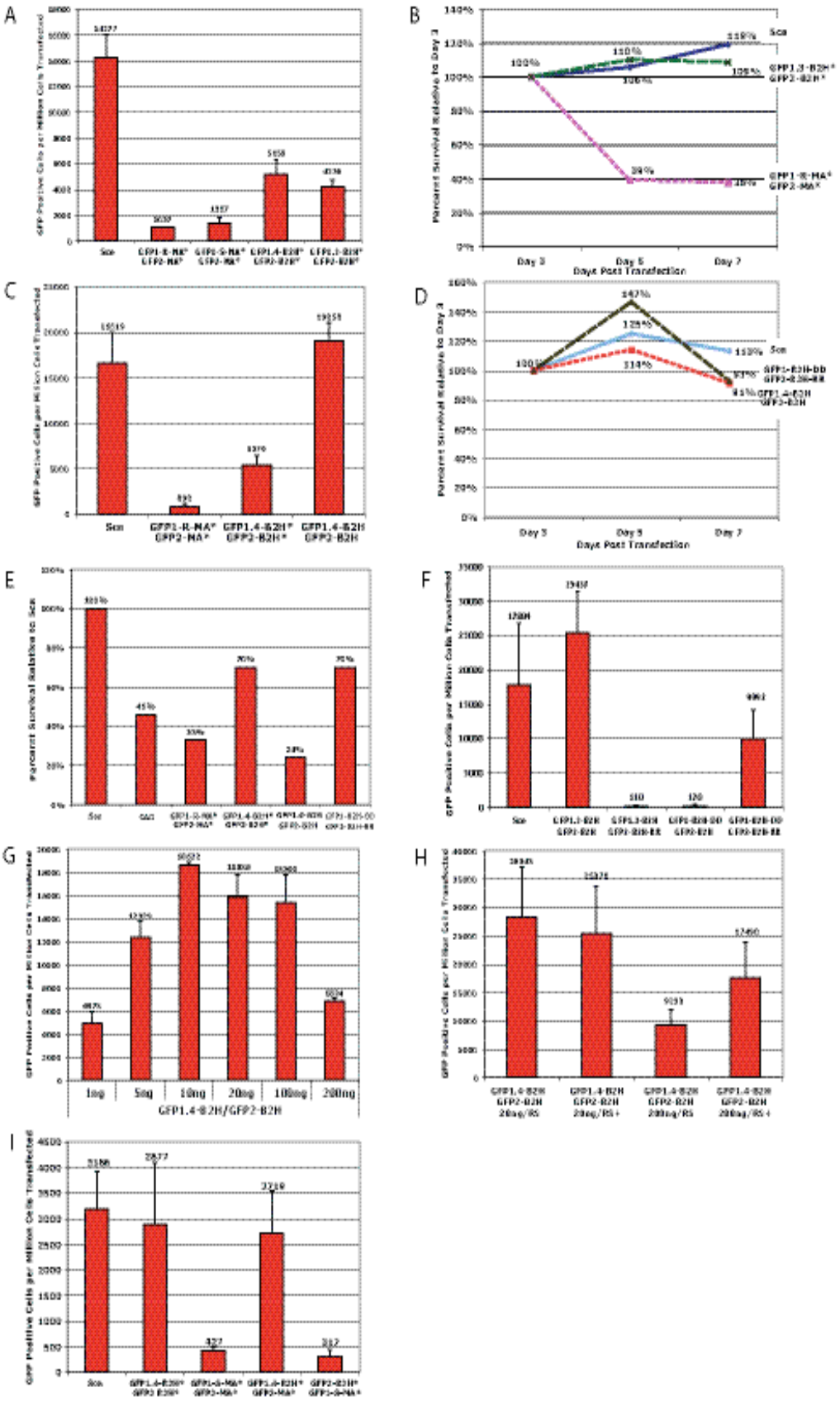

A
B
C
D
E
F
G
H
I
Scn
Days Post Transfection
Day 3
Day 5
Day 7
1mg
5mg
10mg
20mg
100mg
200mg
GFP1.4-B2H/GFP2-B2H

Figure 2.3. Comparison of 3-fingered ZFNs made by modular assembly (MA) to ZFNs made by bacterial 2-hybrid (B2H) selection.

(A) Stimulation of gene targeting using GFP-ZFNs with a mutated nuclease domain (*) made by MA and B2H compared to Sce with 100 ng of transfected DNA for each nuclease. The background rate of gene targeting in *HEK* 293 cells is one in a million. Therefore, the fold stimulation in gene targeting is represented by the number of GFP positive cells per million cells transfected. (B) Gene targeting time course assay with 100 ng of transfected DNA for each nuclease. This assay measures survival of cells that have undergone gene targeting relative to those that have not undergone targeting. The stimulation of gene targeting using GFP-ZFNs was determined at day three post-transfection. At day 5 and 7 post-transfection, the percentage of targeting events relative to day 3 post-transfection was determined and plotted. Because gene targeting events are genomic changes, a decrease in the number of targeting events indicates toxicity associated with the ZFN (C) Stimulation of gene targeting using GFP-ZFNs with mutated (*) and non-mutated nuclease domain. (D) Gene targeting time course assay using 100 ng transfected DNA for each nuclease comparing non-mutated nuclease domain to Sce and "B2H-DD/RR" ZFNs. (E) Toxicity Assay for all iterations of 3-finger GFP-ZFNs tested in the gene targeting assay relative to Sce. Caspase Activated DNAse (CAD), a nuclease that is involved in DNA fragmentation during the apoptotic response, serves as a positive control for cytotoxicity. A number less than 100% shows decreased cell-survival compared to Sce and demonstrates a toxic effect. (F) Stimulation of gene targeting by "B2H-DD/RR" ZFNs at day 3 post-transfection using 100 ng of transfected DNA for each nuclease. The "DD" ZFN was paired with its corresponding wild-type ZFN partner and likewise the "RR" ZFN was paired with its corresponding wild-type partner. (G) Titration of transfected DNA of B2H-GFP-ZFNs in gene targeting assay at day 3 post-transfection using increasing amounts of transfected DNA (H) Gene targeting using Sce and GFP-ZFNs with normal repair substrate and an "protected repair template" (RS+) using 100 ng of transfected DNA of each nuclease. RS+ identical to GFP Donor except that the wild-type sequence 5'-ACCATCTTCTTCAAGGACGACGGC-3' is replaced by 5'-AC*A*AT*T*TT*T*TTCAAGGA*T*GA*T*GGC-3', which introduces five wobble mutations (italics) that abrogate ZFN cutting but do not change the amino acid sequence of GFP. (I) Gene targeting assay performed by mixing and matching MA-GFP-ZFNs and B2H-GFP-ZFNs using 100 ng of transfected DNA of each nuclease. Error bars are standard deviation of three samples.

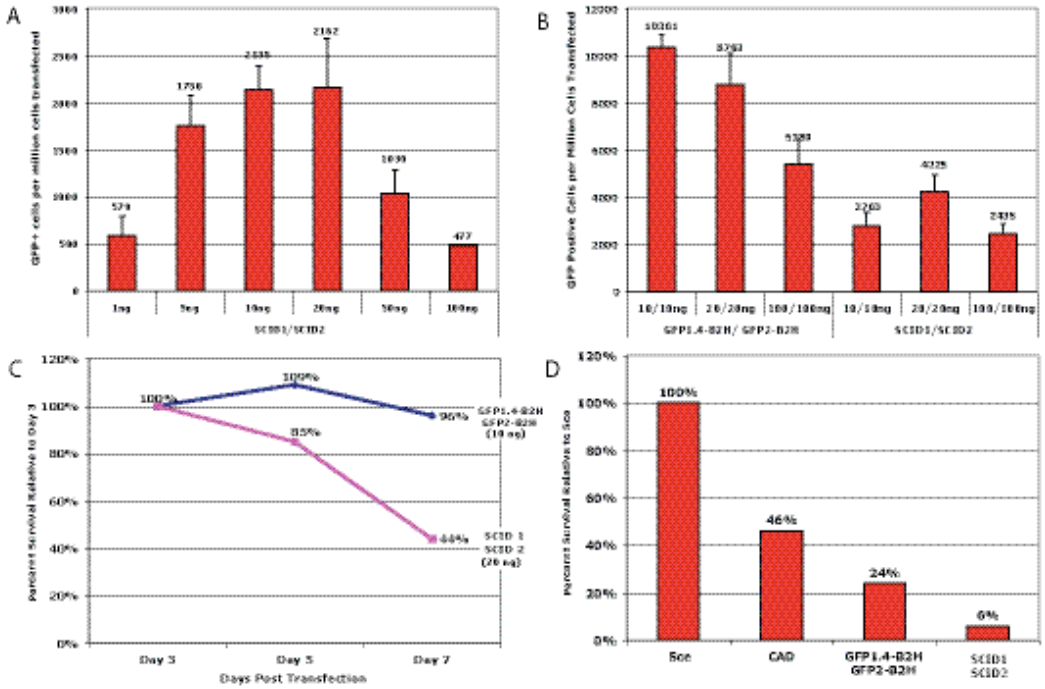


Figure 2.4. Comparison of 3-finger to 4-finger ZFNs.

(A) Titration of transfected DNA of 4-finger SCID-ZFNs in gene targeting assay at day 3 post-transfection. Titration demonstrates that maximal targeting occurs using 10-20 ng of each SCID-ZFN (B) Comparison of 3-finger and 4-finger ZFNs with increasing amounts of DNA at day 3 post-transfection. (C) Gene targeting time course assay for 3-fingered GFP-ZFNs and 4-fingered SCID-ZFNs as described in Figure 2B using the amount of transfected DNA determined in (B) that resulted in the highest rate of gene targeting (10 ng of transfected DNA for the GFP-B2H ZFNs and 20 ng of transfected DNA for the SCID ZFNs). (D) Toxicity Assay for 3-fingered GFP-ZFNs and 4-fingered SCID-ZFNs using 100 ng of transfected DNA for each nuclease. CAD (100ng) serves as a positive control for toxicity. Error bars are standard deviation of three samples.

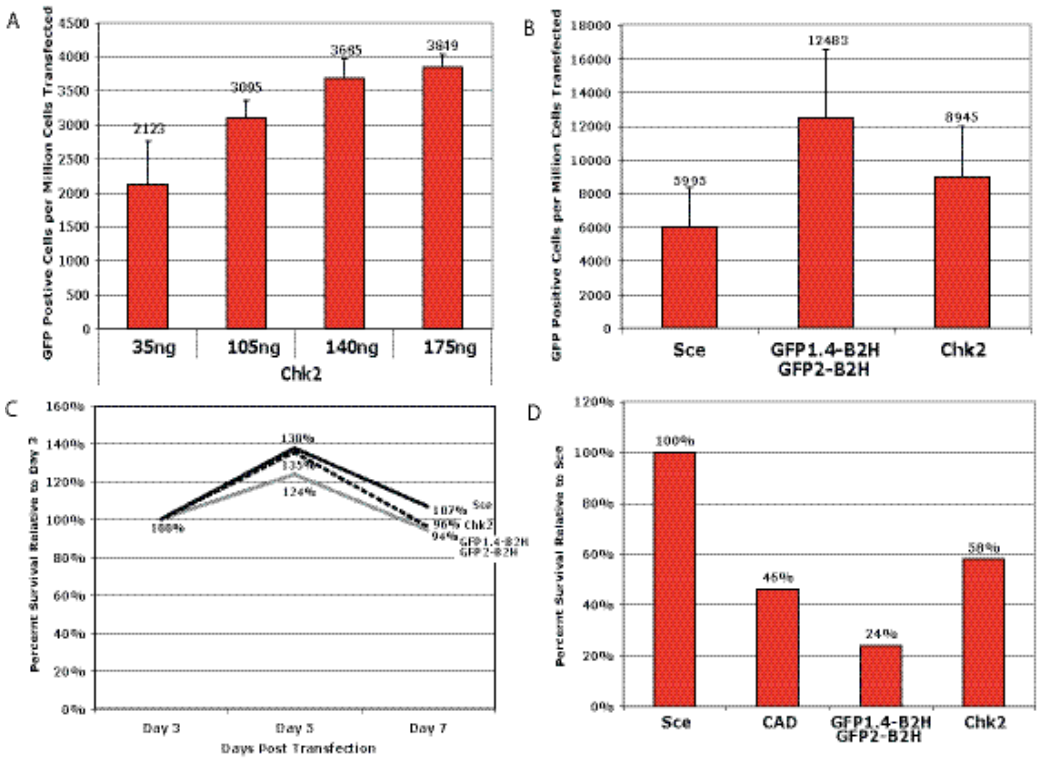


Figure 2.5. Comparison of 3-finger to 6-finger ZFNs.

(A) Titration of transfected DNA of 6-finger Chk2-ZFN in gene targeting assay at day 3 post-transfection. (B) Stimulation of gene targeting of 3-finger B2H-GFP-ZFNs and 6-finger Chk2-ZFN with 100 ng transfected DNA of each nuclease(s) (C) Gene targeting time course assay for 3-finger GFP-ZFNs and 6-finger Chk2-ZFN with 100 ng of transfected DNA for each nuclease(s) at day 3 post-transfection (D) Toxicity Assay for 3-finger GFP-ZFNs and 6-finger Chk2-ZFN with 100 ng of transfected DNA for each nuclease(s). Error bars are standard deviation of three samples.

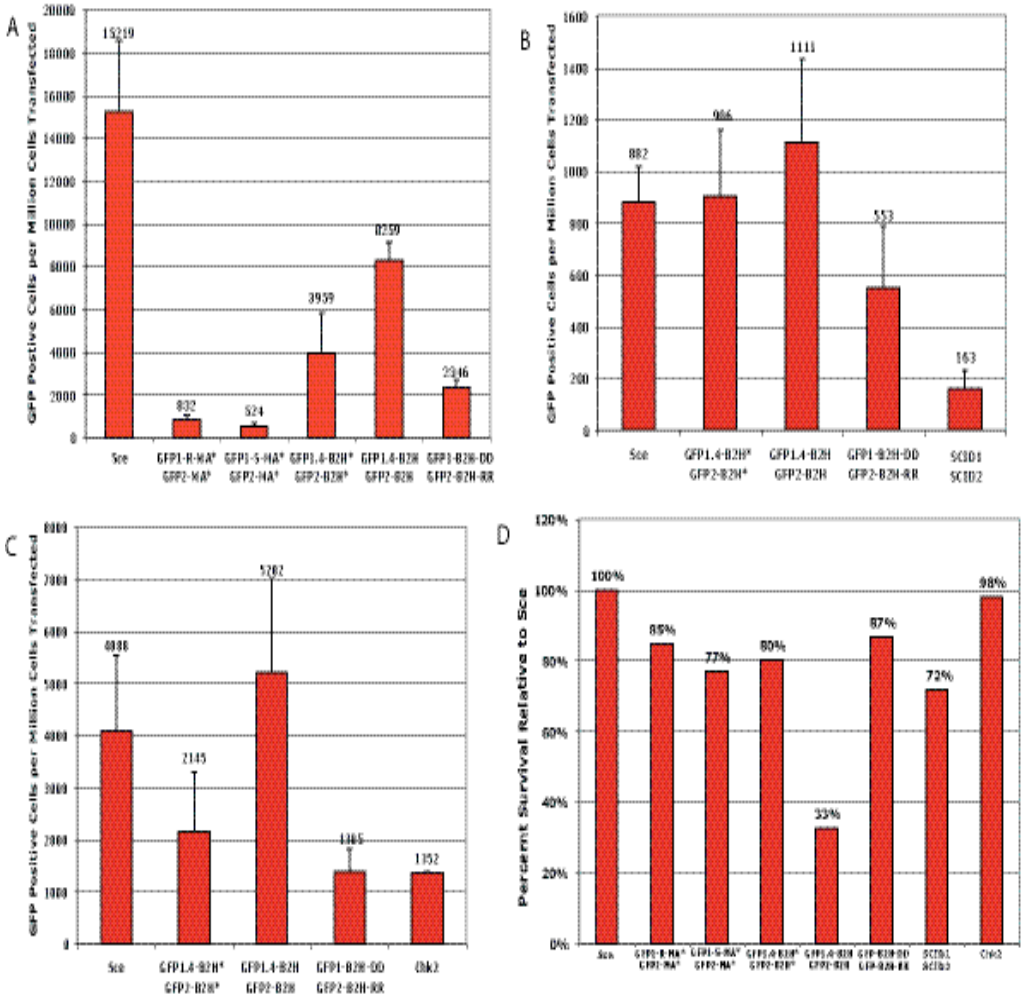


Figure 2.6. Functionality and toxicity of ZFNs at normalized protein levels.

(A) Western blot analysis (data not shown) was performed to determine the amount of DNA needed to transfect into HEK293 cells in order to get equivalent amounts of ZFN protein expression. The gene targeting assay was performed with the determined amounts of DNA for all tested 3-fingered ZFNs. The amounts of DNA transfected for each nuclease were as follows: 50 ng GFP2-MA* and GFP2-B2H*, 250 ng GFP1-R-MA*, GFP1-S-MA*, GFP1.4-B2H*, and GFP1.4-B2H, 100 ng GFP2-B2H, 40 ng B2H-GFP1-DD, 150 ng B2H-GFP2-RR (B) Stimulation of gene targeting with normalized amounts of protein as determined by Western blot analysis for indicated 3 and 4-fngered ZFNs. The amounts of DNA transfected for each nuclease were the same as listed above in (B) and 5 ng SCID1 and SCID2. (C) Stimulation of gene targeting with normalized amounts of protein as determined by Western blot analysis for

indicated 3 and 6-fngered ZFNs. The amounts of DNA transfected for each nuclease were the same as listed above in (B) and 35 ng Chk2 (D) Toxicity Assay for ZFNs at normalized ZFN protein levels. The amounts of DNA transfected for each nuclease were the same as listed above in (B), C), and (D). In summary, for the same protein level the B2H 3-finger ZFNs have higher targeting activity than either the 4 or 6-finger proteins. Error bars are standard deviation of three samples.

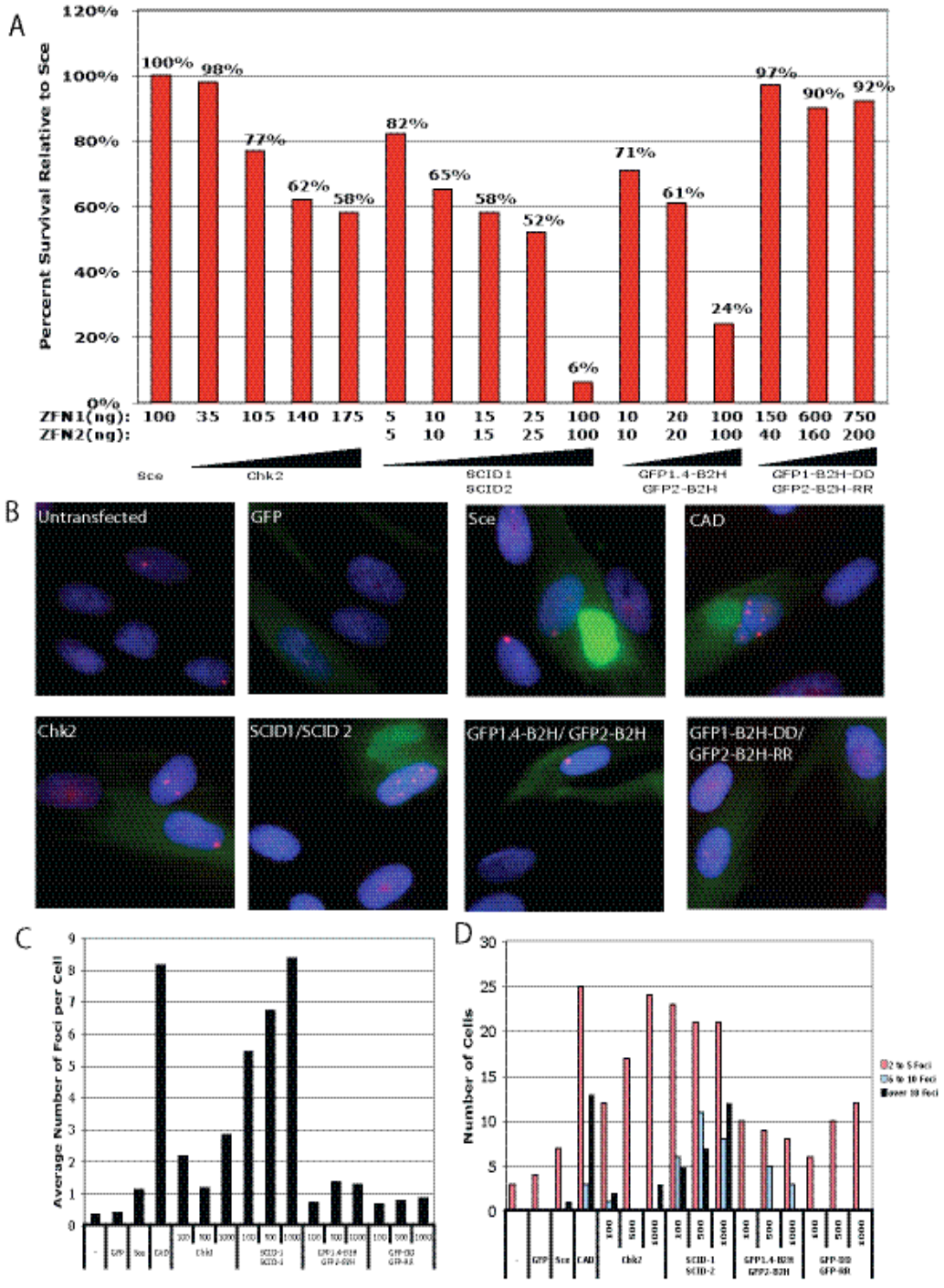

A
Percernt Survival Relative to Sce
120%
100%
80%
60%
40%
20%
0%
100% 98% 77% 62% 58% 82% 65% 58% 52% 6% 71% 61% 24% 97% 90% 92%
ZFN1(ng): 100 35 105 140 175 5 10 15 25 100 10 20 100 150 600 750
ZFN2(ng): 5 10 15 25 100 10 20 100 40 160 200
Sce
Chk2
SCID1
SCID2
GFP1.4-B2H
GFP2-B2H
GFP1-B2H-DD
GFP2-B2H-RR
B
Untransfected
GFP
Sce
CAD
Chk2
SCID1/SCID 2
GFP1,4-B2H/ GFP2-B2H
GFP1-B2H-DD/
GFP2-B2H-RR
C
Average Number of Foci per Cell
D
Number of Cells

Figure 2.7. Measuring cytotoxicity of ZFNs

(A) Toxicity Assay with increasing amounts of ZFNs. Except for GFP-DD/GFP-RR pair, all pairs showed increasing toxicity with increasing amounts of DNA transfected. (B) Toxicity Assay performed in murine embryonic stem cells with 800ng of one nuclease (Sce, CAD, or Chk2) or 400ng of each nuclease (B2H-GFP-ZFNs and SCID-ZFNs). (C) Toxicity Assay performed in human foreskin fibroblasts with 2 ug of one nuclease (Sce, CAD, or Chk2) or 1ug of each nuclease (B2H-GFP-ZFNs and SCID-ZFNs). (D) Representative cells for each experimental condition after 53BP1 staining using 1 μg of transfected DNA for each nuclease. 53BP1 foci are seen in red, DAPI staining in blue, and GFP positive cells in green. Foci were counted in transfected cells that were GFP positive. (E) Average number of 53BP1 foci per transfected cell with increasing amount of nucleases. (-) indicates untransfected, GFP indicates transfection of GFP alone, Sce serves as a negative control for foci formation, and CAD serves as a positive control for 53BP1 foci formation. (F) The number of cells with a given number of 53BP1 foci after transfection with increasing amounts of nuclease(s) of interest. As the amount of ZFN transfected is increased, the number of cells with higher numbers of foci increases for all ZFNs. A significant number of cells with the highest number of foci (>10) are only seen with the SCID-ZFNs. Error bars are standard deviation of three samples.

CHAPTER III: ATTENUATION OF ZINC FINGER NUCLEASE TOXICITY BY SMALL MOLECULE REGULATION OF PROTEIN LEVELS

Abstract

Zinc finger nucleases (ZFNs) have been used successfully to create genome specific double strand breaks and thereby stimulate gene targeting by several thousand fold. ZFNs are chimeric proteins composed of a specific DNA binding domain linked to a non-specific DNA cleavage domain. By changing key residues in the recognition helix of the specific DNA binding domain, one can alter the ZFN binding specificity and thereby change the sequence to which a ZFN pair is being targeted. For these and other reasons, ZFNs are being pursued as potential reagents for genome modification including use in gene therapy. In order to reach their full potential, it is important to attenuate the cytotoxic effects currently associated with ZFNs. Here we evaluate two potential strategies for reducing toxicity by regulating protein levels. Both strategies involve creating ZFNs with shortened half-lives and then regulating protein level with small molecules. First, we destabilize ZFNs by linking a ubiquitin moiety to the N-terminus and regulate ZFN levels using a proteasome inhibitor. Second, we

destabilize ZFNs by linking a modified destabilizing FKBP12 domain to the terminus and regulate ZFN levels by using a small molecule that blocks the destabilization effect of the N-terminal tag. We show that by regulating protein levels we can maintain high rates of ZFN mediated gene targeting while reducing ZFN toxicity.

Introduction

Homologous recombination is a natural mechanism that cells use for a variety of processes including double strand break (DSB) repair (Wyman, Ristic et al. 2004). To repair a DSB by homologous recombination, the cell usually uses the sister chromatid as a donor-template but can use other pieces of DNA such as extrachromosomal DNA. Gene targeting uses homologous recombination to make a precise genomic change and is commonly used experimentally in a variety of cells including yeast and murine embryonic stem cells. However, the spontaneous rate of homologous recombination is too low in mammalian somatic cells (10^{-6}) to be commonly used experimentally or therapeutically (Orr-Weaver, Szostak et al. 1981; Porteus and Baltimore 2003; Yamazoe, Sonoda et al. 2004; Capecchi 2005). The rate of gene targeting, however, can be increased (to about 10^{-2}) by creating a gene specific DSB (Choulika, Perrin et al. 1995; Smih, Rouet et al. 1995; Brenneman, Gimble et al. 1996; Sargent, Brenneman et al. 1997; Porteus and Baltimore 2003).

Zinc Finger Nucleases (ZFNs) can create site-specific DSBs and have been shown to increase the rate of gene targeting by over 5 orders of magnitude (Durai, Mani et al. 2005; Porteus and Carroll 2005; Urnov, Miller et al. 2005). ZFNs are chimeric proteins that consist of a specific DNA binding domain made up of tandem zinc finger binding motifs fused to a non-specifc cleavage domain from the FokI restriction endonuclease (Chandrasegaran and Smith 1999). By changing key residues in the DNA binding domain, ZFN binding specificity can be altered providing a generalized strategy for delivering a site-specific DSB. However, ZFNs have been shown to have cytotoxic effects (Porteus and Baltimore 2003; Porteus 2006; Pruett-Miller, Connelly et al. 2008). Several studies suggest that this toxicity is caused by "off-target" DSBs. For example, a zinc finger protein containing no nuclease domain was not toxic when transfected into HEK293 cells (unpublished data). Similarly, Beumer et al. (2006) have shown that ZFNs containing point mutations to inactivate the nuclease domain do not exhibit cytotoxicity in flies (Beumer, Bhattacharyya et al. 2006). There have been two published strategies for reducing the number of "off-target" breaks: (1) increase the specificity of the ZFN by protein engineering or (2) force heterodimerization of the ZFN pairs (Miller, Holmes et al. 2007; Szczepek, Brondani et al. 2007; Cornu, Thibodeau-Beganny et al. 2008; Pruett-Miller, Connelly et al. 2008). Here, we explore a third strategy to reduce cytotoxicity by small molecule regulation of ZFN protein levels.

By creating ZFNs from zinc finger DNA binding domains that are more specific, toxicity is reduced. While on-target cutting is generated by heterodimerization of a ZFN pair at its target site (at least 18 base pairs), off-target cutting can be mediated by either homodimer pairs or heterodimer pairs. Modifications in the nuclease to prevent homodimerization results in ZFNs with reduced toxicity (Miller, Holmes et al. 2007; Szczepek, Brondani et al. 2007; Pruett-Miller, Connelly et al. 2008). We found, however, that this reduction can come at a cost of reduced activity in stimulating gene targeting (Pruett-Miller, Connelly et al. 2008; Wilson et al., manuscript submitted).

The hypothesis that reduced toxicity could also be obtained by being able to regulate ZFN expression comes from experiments showing that high levels of ZFN expression cause toxicity without increasing targeting rates. Conversely, very low levels of ZFN expression cause reduced targeting rates (Pruett-Miller, Connelly et al. 2008). This "Goldilocks" phenomenon means that being able to titrate the amount of ZFN protein is critical to optimizing ZFN mediated gene targeting.

DSB-mediated gene targeting occurs within 60 hours of transfection of DNA (Porteus and Baltimore 2003). Expression of ZFNs outside this window can only increase toxicity without increasing targeting. We hypothesized that if we could narrow the time of ZFN protein levels, we could reduce toxicity while maintaining high rates of targeting. In this study, we use two previously

described strategies to regulate protein levels and apply them to ZFNs. We show that by regulating protein levels, we reduce the number of "off-target" DSBs and reduce toxicity, while maintaining high ZFN-stimulated gene targeting activity.

Results

The N-end rule and ubiquitin fusion degradation method as a strategy to regulate ZFN protein levels

The ability to regulate ZFN protein levels could theoretically give optimal rates of gene targeting with minimal toxicity. Degradation signals or "degrons" are specific domains that confer instability on a protein (Varshavsky 1997). The N-end rule correlates the *in vivo* half-life of a protein to the N-terminal amino acid; some residues are destabilizing while other residues are stabilizing (Bachmair, Finley et al. 1986; Varshavsky 1997). Normal N-terminal processing precludes simply adding a desired residue to the N-terminus of a protein. By adding a ubiquitin moiety (Ub) to the N-terminus of a protein, the N-terminal amino acid of a protein can be somewhat controlled. In eukaryotes, the Ub-X-POI (where POI is protein of interest) fusion is cleaved by Ub-specific processing proteases immediately before X (where X is an amino acid residue). This cleavage leaves the X residue as the N-terminal amino acid and thus affects

protein stability. It has been established by several groups that an N-terminal arginine is a degradation signal (Varshavsky 1996; Dantuma, Lindsten et al. 2000).

It is also possible to create poorly cleavable or uncleavable Ub-X-POI fusions. If the ubiquitin protein is not cleaved from the POI, the protein can undergo ubiquitin fusion degradation (Johnson, Ma et al. 1995). That is, the Ub-X-POI fusion can be further ubiquitinated and thereby "marked" for degradation by the proteasome. This allows for another strategy to create short-lived POIs. By substituting the last residue of the ubiquitin moiety from glycine to valine and using a valine linker (Ub-VV-POI), the ubiquitin moiety can no longer be cleaved from the POI (Dantuma, Lindsten et al. 2000).

We created a pair of Ub-VV-ZFNs and Ub-R-ZFNs fusion proteins (Figure 3.1A) to destabilize a pair of previously validated ZFNs targeting the GFP gene (Pruett-Miller, Connelly et al. 2008). The Ub-VV-ZFNs were made to take advantage of the potential destabilizing effect of a covalently linked N-terminal ubiquitin, and the Ub-R-ZFNs were made to take advantage of the potential destabilizing effect of an N-terminal arginine. Expression of the ZFN chimeras in transiently transfected HEK293 cells was examined by Western blot analysis (Figure 3.1B). The size of the Ub-VV-ZFNs corresponded with the expected size of an uncleaved fusion protein. The size of the Ub-R-ZFNs corresponded with

the size of the unmodified ZFNs, confirming that the ubiquitin moiety was cleaved.

The addition of the proteasome inhibitor MG132 can increase the levels of Ub-X-POI fusion proteins (Dantuma, Lindsten et al. 2000). We therefore examined the expression of ubiquitin linked ZFNs and unmodified ZFNs in the presence and absence of MG132 (Figure 3.1B). The addition of the proteasome inhibitor had little effect on the unmodified ZFNs. In contrast, addition of MG132 to cells transfected with ubiquitin modified ZFNs produced a striking increase in the expression. It is interesting to note that even the untreated Ub-R-ZFNs had an increase in protein levels relative to the unmodified ZFNs (Figure 3.1B, compare Ub-R-ZFNs, +MG132 at 24 hours to unmodified ZFNs at 24 hours). We hypothesize that the addition of the ubiquitin moiety to the N-terminus aided in protein folding of the ZFNs and thus produced higher expression. Interestingly, although in the presence of MG132, expression of the Ub-R-ZFNs is higher than the unmodified ZFNs, it appears as though the Ub-R-ZFNs levels decrease more rapidly suggesting that the N-terminus arginine is in fact destabilizing.

We compared the activity of the ubiquitin linked ZFNs to the unmodified ZFNs using a GFP gene targeting assay. In this assay, gene targeting is measured by the correction of a chromosomally integrated mutated GFP target gene (Porteus and Baltimore 2003). We normalized the gene targeting rate for each

condition to the rate obtained for the optimal amount of the unmodified ZFNs as previously determined (Pruett-Miller, Connelly et al. 2008). We first compared activities of the Ub-VV-ZFNs and Ub-R-ZFNs with increasing amounts of DNA in the absence of proteasome inhibitor to that of the unmodified proteins (Figure 3.2A and 3.2B). In the absence of drug, both the Ub-VV-ZFNs and the Ub-R-ZFNs, at all DNA concentrations tested, produced lower amounts of gene targeting compared to rates produced using the unmodified pair. The Ub-VV-ZFNs produced lower rates of gene targeting in the absence of drug compared to the Ub-R-ZFNs.

We next evaluated the rate of gene targeting produced by the Ub-modified proteins in the presence of MG132 compared to when the drug was absent (Figure 3.2C and 3.2D). Both the Ub-VV-ZFNs and Ub-R-ZFNs produced increased rates of gene targeting in the presence of MG132 compared to when the drug was absent. The rate of gene targeting produced by the Ub-VV-ZFNs in the presence of drug was not as high, however, as rates produced by the unmodified protein. In contrast, in the presence of drug, the Ub-R-ZFNs produced equivalent rates of gene targeting as compared to the unmodified proteins.

In order to determine if the ubiquitin modification of these ZFNs reduced the cytotoxicity associated with unmodified ZFNs, we used a flow cytometry based cell survival assay (the "toxicity assay") (Pruett-Miller, Connelly et al. 2008). In this assay, we use a non-toxic endonuclease, *I-SceI* (hereafter called

Sce), as the gold-standard for a non-toxic nuclease to which we normalize relative amounts of toxicity. The percent of surviving cells transfected with a potentially toxic nuclease is compared to the percent of surviving cells transfected with Sce. A lower percent of surviving cells is a sign of greater toxicity. As shown in figure 3.2E, the percent survival relative to Sce of the unmodified ZFNs is about 50%. Both the Ub-VV-ZFNs and the Ub-R-ZFNs examined in this experiment produced lower toxicity and therefore a higher percentage of survival compared to the unmodified proteins. At 20 nanograms of Ub-R-ZFNs in the presence of drug, there was no observable toxicity in this assay. This is also the amount at which equivalent rates of gene targeting were obtained relative to the unmodified proteins (Figure 3.2D). In summary, we found that we could minimize toxicity but with a reduced targeting efficiency using the VV-linked versions. In contrast, using the R-linked versions, we could decrease toxicity without losing targeting efficiency.

The destabilization domain method as a strategy to regulate ZFN protein levels

An alternative strategy to using ubiquitin involves linking a destabilization domain to the POI. This destabilization domain was engineered by making mutants of the FKBP12 protein, which is constitutively and rapidly degraded in

mammalian cells (Banaszynski, Liu et al. 2005). Fusion of this destabilization domain to another protein confers instability to the fusion protein. In order to stabilize the protein, Banaszynski et al. (2006), developed a synthetic ligand (called Shield1) that binds the destabilization domain and protects the fusion protein from degradation (Banaszynski, Liu et al. 2005).

We made a pair of chimeric proteins that linked the destabilization domain (dd) to the N-terminus of the ZFNs targeting the GFP gene ("dd-ZFNs", Figure 3A). We examined the expression of the dd-ZFNs and unmodified ZFNs in transfected HEK293 cells by Western blot analysis (Figure 3.3B). In the absence of Shield1, the dd-ZFNs were destabilized as shown by reduced expression relative to the unmodified ZFNs. Upon addition of Shield1 for the first 24 hours post transfection, however, the dd-ZFNs were stabilized to relatively equivalent levels of expression as the unmodified ZFNs at 24 hours. The amount of protein expressed at 32 hours post transfection after drug treatment, however, was substantially reduced when compared to the unmodified ZFNs at the same time point (Figure 3.3B).

To examine the activity of the dd-ZFNs, we used the GFP gene targeting assay. At high concentrations of DNA, the rate of gene targeting stimulated by the dd-ZFNs in the absence of drug almost reached the rate stimulated by the unmodified ZFNs (Figure 3.4A). Because of this high rate of targeting in the absence of drug, we chose to continue the experiments with 5 or 20 nanograms of

transfected DNA. We conducted a series of experiments to characterize the timing and dosing of the drug in order to determine the drug conditions needed to obtain optimal rates of gene targeting. After 24 hours of exposure to Shield1, the rate of gene targeting induced by the dd-ZFNs is equivalent to the rate stimulated by the unmodified ZFNs (Figure 3.4B). We found that additional exposure, beyond 24 hours, to the drug did not further increase these rates (data not shown).

We next evaluated the dosing of the Shield1 drug with respect to gene targeting. At 1000nM of Shield1, we observed statistically equivalent rates of gene targeting, but that there was a dose-dependent decrease in targeting as the dose was lowered. With Shield1 present at 1000nM for the first 24 hours, we observed that using either 5 or 20 nanograms of the dd-ZFNs could produce rates of gene targeting statistically equivalent to the optimal rates obtained with the unmodified ZFNs.

To determine if linking the destabilization domain to the ZFNs reduced the cytotoxicity associated with the unmodified ZFNs, we used the toxicity assay. As expected, decreasing the amount of ZFN results in a decrease in cytotoxicity as demonstrated by an increase in percent survival relative to Sce. Strikingly, in the presence of drug at 5, 20, or even 50 nanograms of dd-ZFNs, toxicity relative to Sce appears to be negligible.

Previous studies have suggested that the cytotoxicity associated with unmodified ZFNs is due to the creation of off-target DSBs (Pruett-Miller,

Connelly et al. 2008). When a DSB occurs, a signaling cascade is activated including the phosphorylation of H2AX and the recruitment of an array of proteins, including 53BP1, to the site of DSB that can be detected as foci by immunofluorescence (Jeggo and Lobrich 2005; Pruett-Miller, Connelly et al. 2008). We have previously shown that ZFNs that produce larger numbers of foci are more toxic than ZFNs that produce fewer foci. Although the unmodified ZFN pair used in this study shows cytoxicity in the toxicity assay, this pair did not show an increased number of foci per cell relative to Sce when this assay was performed in human foreskin fibroblasts. To sensitize the assay, we used cells mutated in Ku80, a gene important in the nonhomologous end-joining pathway of DSB repair, which are known to have delayed repair of DSBs (Khanna and Jackson 2001). In this cell line, GFP transfected cells and cells transfected with Sce alone had an average of about 4 foci per cell (Figure 3.5). As a further control, we transfected cells with a plasmid encoding a Caspase Activated DNAse (CAD). CAD-transfected cells had an average of about 12 foci per cell (Figure 5). To aid in our comparison of the unmodified ZFNs and dd-ZFNs, we used higher amounts of DNA than determined in figure 3.3 in order to amplify the number of DSBs visualized. We did however maintain the 1:4 ratio (5ng:20ng vs. 75ng:300ng) of dd-ZFN DNA concentration with respect to unmodified ZFN DNA concentration for this comparison. Cells transfected with the unmodified ZFNs had an average of 10 foci per cell (comparable to the CAD transfected cells,

Figure 3.5B). In contrast, the dd-ZFN transfected cells had only about 4 foci per cell (comparable to the GFP-alone and Sce transfected cells). In summary, linking the destabilization domain of a modified FKBP12 protein to the N-terminus of ZFNs resulted in a way to regulate the expression level of the ZFN that maintained high rates of gene targeting while minimizing toxicity.

Discussion

Homologous recombination is the most precise way to manipulate the genome and is a powerful experimental tool in several different systems. ZFNs have been shown to increase the rate of gene targeting in a wide variety of experimental systems previously not amenable to genome manipulation by homologous recombination (Orr-Weaver, Szostak et al. 1981; Porteus and Baltimore 2003; Yamazoe, Sonoda et al. 2004; Capecchi 2005). In addition to the problem of designing ZFNs to recognize target sites (Ramirez, Foley et al. 2008), another limitation has been concern about off-target effects (Porteus and Baltimore 2003; Pruett-Miller, Connelly et al. 2008). Improvements in toxicity have been attained by increasing the specificity of ZFNs and by modifications of the nuclease domain (Miller, Holmes et al. 2007; Szczepek, Brondani et al. 2007; Cornu, Thibodeau-Beganny et al. 2008; Pruett-Miller, Connelly et al. 2008). However, further reduction in toxicity is needed. In whole organisms such as flies and zebrafish, high levels of ZFN expression led to abnormal developmental

mutations or “monsters” (Beumer, Bhattacharyya et al. 2006; Doyon, McCammon et al. 2008; Meng, Noyes et al. 2008). Reducing ZFN toxicity by regulating ZFN expression could hypothetically help attenuate these abnormalities. In this work, we show that small molecule regulation of ZFN expression can result in an improved toxicity profile without sacrificing gene targeting activity.

The standard strategy to control protein levels is to use transcriptional based methods (TetOn or TetOff: Clonetech, Ponasterone System; Stratagene, Dimerizer System; Ariad). Gene targeting induced by ZFNs is already a three-component system (ZFN-1, ZFN-2, and a repair donor molecule). Adding an inducible transcriptional regulator as a fourth component to make the system more complex was not desirable, particularly as the technology moves into cell types that are more difficult to transfect or infect. The ERT2 domain, a modified ligand binding domain from the estrogen receptor, has been successfully used to control protein activity by modulating its location. Unfortunately, we found that attaching the ERT2 domain to ZFNs did not work (data not shown). An alternative strategy is to use a post-translational method of regulating ZFN level. In this strategy, a destabilized ZFN is created by adding a destabilizing domain and then levels of ZFNs are controlled by adding a small molecule to block the destabilization effects. By fusing a ubiquitin domain to the N-terminus through a non-cleavable linker (Ub-VV-ZFN), we made ZFNs that could be regulated by proteasome inhibition, which resulted in decreased toxicity. When we fused the

ubiquitin domain to the N-terminus of the ZFN through a cleavable linker leaving a destabilizing arginine at the N-terminus, we created ZFNs that were regulated by proteasome inhibition resulting in decreased toxicity and maintained high rates of gene targeting. Because proteasome inhibitors such as bortezimib are FDA approved for use in humans, this strategy has long-term promise. We did find the window of exposure to MG132, the proteasome inhibitor used in this study, in which we got good induction without cytotoxicity was narrow. Finally, when we fused a modified FKBP12 domain to the N-terminus of the ZFN, we created ZFNs that were regulated by the small molecule Shield1, which resulted in reduced ZFN toxicity and maintained high rates of targeting. Despite using amounts of Shield1 for prolonged periods (up to 48 hours), we did not observe any discernable toxicity. Moreover, by expression microarray analysis, Shield-1 has almost no effect on gene expression (Maynard-Smith, Chen et al. 2007). Thus, the Shield1/FKBP12 system may ultimately be the better system despite Shield1 not being currently FDA approved for use in humans.

Regulating ZFN expression also gave insight into the kinetics of gene targeting. Previously, we had found that maximal gene targeting was measured at 60 hours after transfection. In this work, we demonstrate that ZFNs need only be expressed for less than 32 hours after transfection to attain maximal gene targeting (measured at 72 hours post-transfection). These experiments define a window for ZFN expression, here defined as 0-32 hours but perhaps slightly

shorter, in which expression of ZFNs beyond the window does not increase targeting but does increase toxicity.

Previously, we used human diploid fibroblasts to measure 53BP1 foci created by off-target DSBs. The unmodified ZFNs used in this study did not show significantly increased numbers of foci in that cell line, presumably because the cells were efficient at repairing DSBs. To sensitize the assay, we used murine Ku80$^{-/-}$ cells that are deficient in DSB repair. By using these sensitized cells, the background of 53BP1 foci did increase, but we were able to detect extra off-target DSBs than the unmodified versions. As ZFNs continued to improve, the use of sensitized assays to quantitate these improvements will be an important strategy.

In summary, we have found that small molecule regulation of ZFN expression is an effective way to reduce cytotoxicity without compromising targeting efficiency. This strategy may be particularly beneficial to using ZFN mediated genome modification in a wide variety of cell types, including human stem cells.

Materials and Methods

DNA Manipulations and Cloning

All plasmids were made using standard cloning techniques and molecular biology as previously described (Ausubel 1996). The unmodified ZFNs were selected by the B2H design strategy and fused to the *FokI* nuclease domain as described

earlier and called “GFP1.4-B2H” and “GFP2-B2H” (Pruett-Miller, Connelly et al. 2008). For the Ub-X-ZFN versions, the ubiquitin open reading frame was amplified by PCR from pUb-R-GFP (Dantuma, Lindsten et al. 2000) with sense primer 5’-ACTGGGATCCTCTAGATCCACCATGCAGATCTTCGTGAAG-3’ and the antisense primers 5’-ACTGGGATCCAAGCTTCCC**CACCAC**ACCTCTGAGACGGAGTAC-3’ for the Ub-**VV**-ZFNs, or 5’-ACTGGGATCCAAGCTTCCC**TCT**GCCACCTCTGAGACGGAGTAC-3’ for the Ub-**R**-ZFNs (restriction site underlined, variable codons in bold) and cloned into the ZFN expression plasmid using the BamHI site. Directionality was determined by XbaI digest. To create the dd-ZFNs, the L106P destabilization domain was PCR amplified using primers 5’-ACGTGCGGCCGCACCATGGGAGTGCAGGTGGAAACCATCTCC – 3’ AND 5’-ACTGGGATCCTTCCGGTTTTAGAAGCTCCAC-3’. The resulting fragment was digested with NotI and BamHI and cloned in-frame to the N-terminus of the GFP-ZFNs in a CMV expression vector. For all constructs the N-terminal domains and junctions were confirmed by sequencing.

Cell Culture and Transfection

All cell culture experiments were performed in *HEK*293 cells except where identified. Cells were cultured in a humidified incubator at 37°C with 5% CO_2 in DMEM supplemented with 10% bovine growth serum (Hyclone, Logan, UT,

USA), 2mM L-glutamine, 100IU/ml penicillin, and 100mg/ml streptomycin. Stable cell lines were constructed as previously described (Porteus 2006). Transient transfections were performed using the calcium phosphate technique as previously described and produced transfection efficiencies between 10-35% (Porteus, Cathomen et al. 2003).

Proteasome Inhibitor

For experiments using MG132 (carboxybenzyl-leucyl-leucyl-leucinal; Sigma-Aldrich, St. Louis, MO),10uM drug was added to cells from 18-22 hours post-transfection unless otherwise noted.

Shield1

For experiments using Shield1 (Clontech, Mountain View, CA), 1000nM drug was added to cells from time of transfection and left on for 24 hours unless otherwise noted.

Measurement of Gene Targeting Using the GFP System

Gene targeting experiments were performed in triplicate as previously described using calcium phosphate transfection (Porteus 2006). Transfection efficiencies were determined at day 2 post-transfection, and the rates of gene targeting were determined by flow-cytometry and analyzed on a FACS Calibur (Becton-Dickerson, San Jose, CA, USA) at day 3 (day of transfection is considered day 0). Gene targeting rates are calculated as GFP positive cells per million cells transfected

because the background rate of spontaneous gene targeting is approximately one event per million cells using this system. Gene targeting rates are then normalized to the percent gene targeting obtained using 20ng of ZFN-1 and ZFN-2 as these conditions have given the highest rates of gene targeting for the unmodified proteins.

Flow Cytometry Based Assay for Cell Survival: "Toxicity Assay"

Toxicity assays were performed as previously described (Pruett-Miller, Connelly et al. 2008). Briefly, *HEK*293 cells were transfected in triplicate by calcium phosphate technique with 200ng of a GFP expression plasmid and with varying amounts of each nuclease expression plasmid (two plasmids total). At day two post-transfection, a fraction of transfected cells was analyzed by flow-cytometry and the percentage of GFP positive cells was determined. At day 6 post-transfection, the percentage of GFP positive cells was determined by flow-cytometry. To calculate the percent survival relative to Sce, a ratio of ratios was calculated as previously described (Pruett-Miller, Connelly et al. 2008). The ratio after nuclease transfection was normalized to the ratio after Sce transfection and this determined the percent survival compared to Sce. In control experiments, we showed that Sce expression had no effect on cell survival compared to cells transfected with an empty expression vector.

Immunodection of ZFNs

For time course blots, cells were harvested at indicated times post-transfection. Each sample was counted and lysate volumes were adjusted to give equal amounts of cells per volume. Equal amounts of total lysates were subjected to SDS-PAGE, wet transferred to PVDF membranes and incubated with specific antibodies. ZFNs were detected using an anti-Flag M2 monoclonal antibody (1:10,000, Sigma-Aldrich), and β-actin was detected using a rabbit anti-actin antibody (1:5,000, Sigma-Aldrich). The blots were further incubated with HRP-conjugated secondary antibodies and visualized using Western blotting luminal reagent (Santa Cruz Biotechnology, Santa Cruz, CA).

Sensitized 53BP1 Foci Formation Assay

Cell Culture

Ku80$^{-/-}$ mouse 3T3 cells were maintained in Dulbecco's Modified Eagle's Medium (Hyclone, Logan, UT) supplemented with 20% fetal calf serum and 2 mmol/l L-glutamine. The cells were maintained in a humidified incubator with 5% CO_2 at 37°C.

Transfection of 3T3 Ku80$^{-/-}$ Cell Line

Mouse 3T3 cells that are Ku80$^{-/-}$ were used in the presented studies because they repair DNA breaks more slowly, providing a more sensitive assay for monitoring DNA damage. Ku80$^{-/-}$ cells were seeded in 4-well Lab-Tek II

Chamber Slides (Nalge Nunc, Rochester, NY) at 40,000 cells per well. 24 hours later, cells in each well were lipofected with 200 ng GFP DNA and 75ng or 300ng of each nuclease. Stuffer DNA was added when necessary to raise the total DNA per well to 800ng. Lipofectamine 2000 Reagent (Invitrogen) was used to transfect cells using Invitrogen's suggested protocol. 1000nM Shield1 was added to drug-treated wells at the time of transfection. 24 hours later, Shield1 was removed and the medium was replaced with fresh, supplemented Dulbecco's Modified Eagle's Medium. 48 hours after lipofection the cells were fixed, stained and visualized. 53BP1 foci were counted only in cells that were brightly GFP positive because these were the ones transfected with the GFP and the nuclease(s).

Immunofluorescence Staining

Immunofluorescence staining was carried out as performed in (Pruett-Miller, Connelly et al. 2008). Briefly, cells were washed in phosphate buffered saline, fixed in cold 4% paraformaldehyde, washed again, and then permeabilized with .5% Triton X-100. Cells were re-washed, blocked in 5% bovine serum albumin, and then incubated with rabbit anti-53BP1 (Cell Signaling, Danvers, MA). After another set of washes, cells were incubated with goat anti-rabbit Rhodamine Red-X (Invitrogen, Carlsbad, CA). Cells were washed again and then mounted in Vectashield mounting medium containing 4,6-diamidino-2-phenylinodole (Vector Laboratories, Burlingame, CA). Images were captures

using an epifluorescence microscope equipped with a Q-Fire charge-coupled device camera (Olympus America, Melville, NY) and QCapture Software (QImaging, British Columbia, Canada). Images were merged using ImageJ Software (NIH, ver. 1.40g).

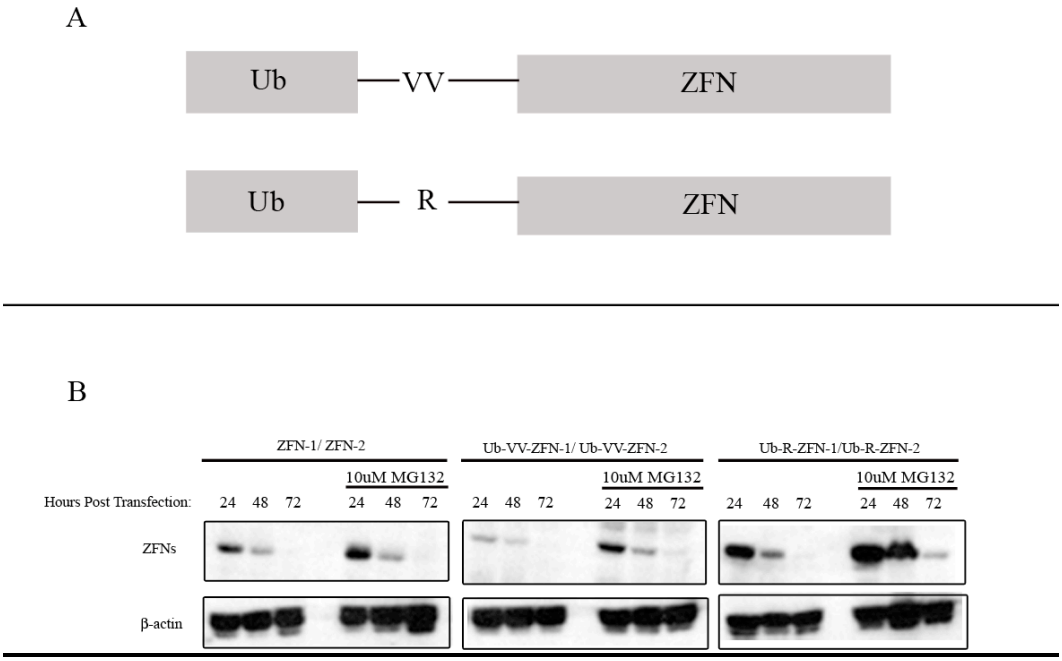


Figure 3.1. Characterization of Ub-X-ZFNs that display drug-dependent stability.

(A) Genetic fusion of a ubiquitin moiety to a ZFN with either a "VV" linker or an "R" linker. (B) Expression profile of unmodified and Ub-X-ZFN proteins in the presence and absence of the proteasome inhibitor 10uM MG132 from 18-22 hours post-transfection. HEK293FT cells were transiently transfected with vectors encoding either ZFN-1/-2, Ub-VV-ZFN-1/-2, or UB-R-ZFN-1/-2. ZFNs were detected with an anti-Flag antibody. β-actin serves as a loading control.

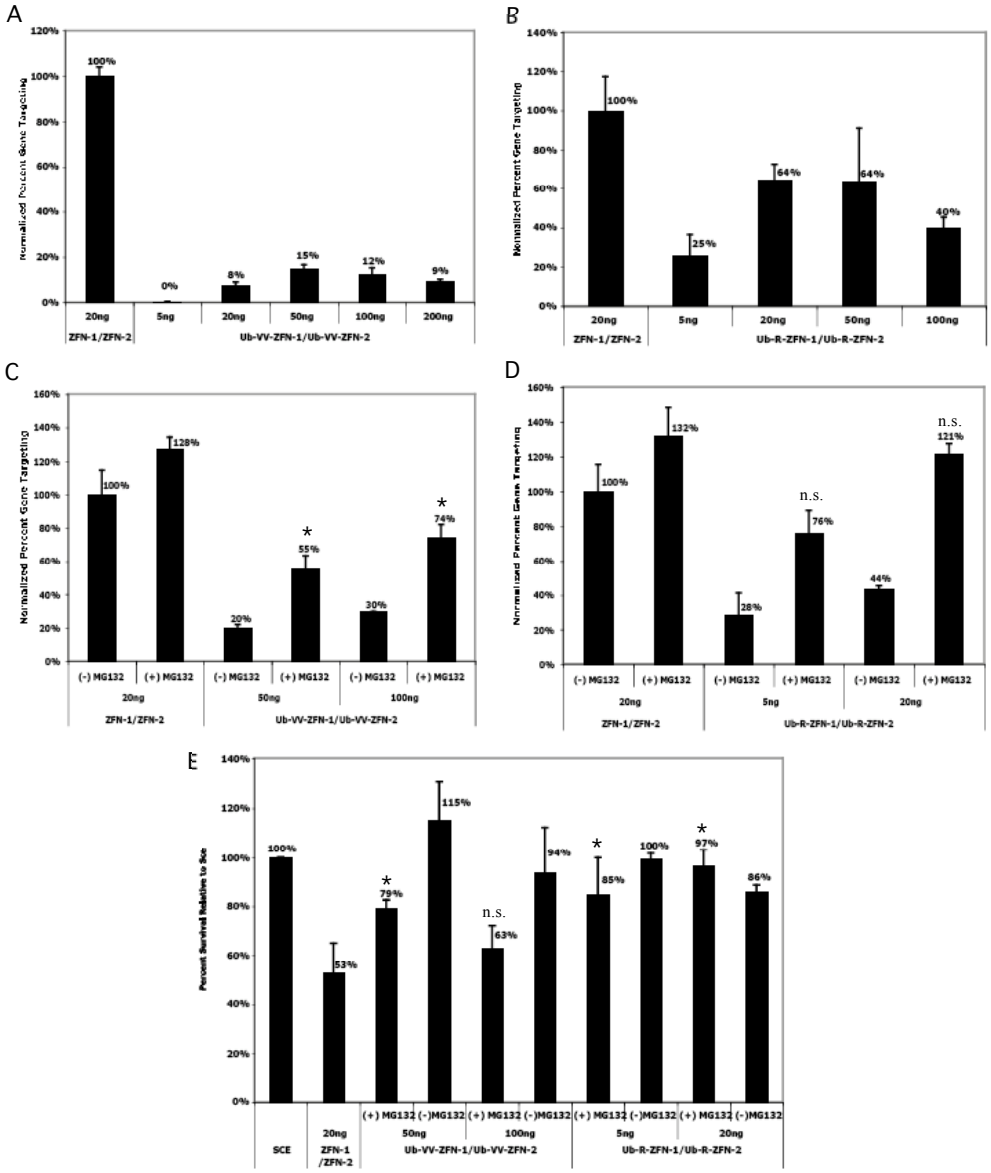


Figure 3.2. Analysis of Ub-X-ZFNs.

Unless otherwise indicated, rates of gene targeting at day 3 were normalized to the rate of gene targeting achieved using 20ng of the unmodified ZFNs without drug treatment as this was previously determined to be the conditions used to obtain optimal gene targeting with the unmodified ZFNs. (A) Titration of

transfected DNA of Ub-VV-ZFNs in the gene targeting assay at day 3 after transfection with increasing amounts of transfected DNA. (B) Titration of transfected DNA of Ub-R-ZFNs in the gene targeting assay at day 3 after transfection with increasing amounts of transfected DNA. (C and D) Gene targeting in the absence and presence of 10uM MG132 for given ZFN pairs at stated DNA concentrations. (E) Toxicity assay for all iterations of Ub-modified and unmodified ZFNs tested in the gene targeting assay relative to Sce. A value of <100% indicates decreased cell survival as compared with Sce, and demonstrates a toxic effect. Statistical analysis was performed using the Student's T-test comparing ZFN-1/ZFN-2 at 20ng with no drug treatment to Ub-modified ZFNs treated with MG132. (*) indicates a P-value of <.05 and n.s. indicates no statistical significance or a P-value of >.05. Error bars are the standard deviation in measurement of three samples.

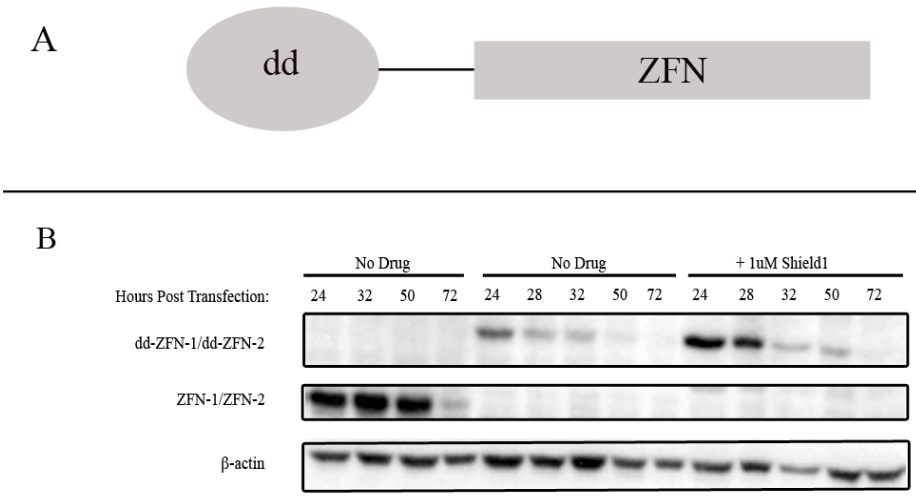


Figure 3.3. Characterization of dd-ZFNs that display Shield1-dependent stability.

(A) Genetic fusion of a destabilization domain derived from an FKBP12 mutant to a ZFN. (B) Expression profile of unmodified and dd-ZFN proteins in presence and absence of 1000nM drug for 24 hours starting at time of transfection. HEK293FT cells were transiently transfected with vectors encoding either ZFN-1/-2, dd-ZFNs. ZFNs were detected with an anti-Flag antibody. β-actin serves as a loading control.

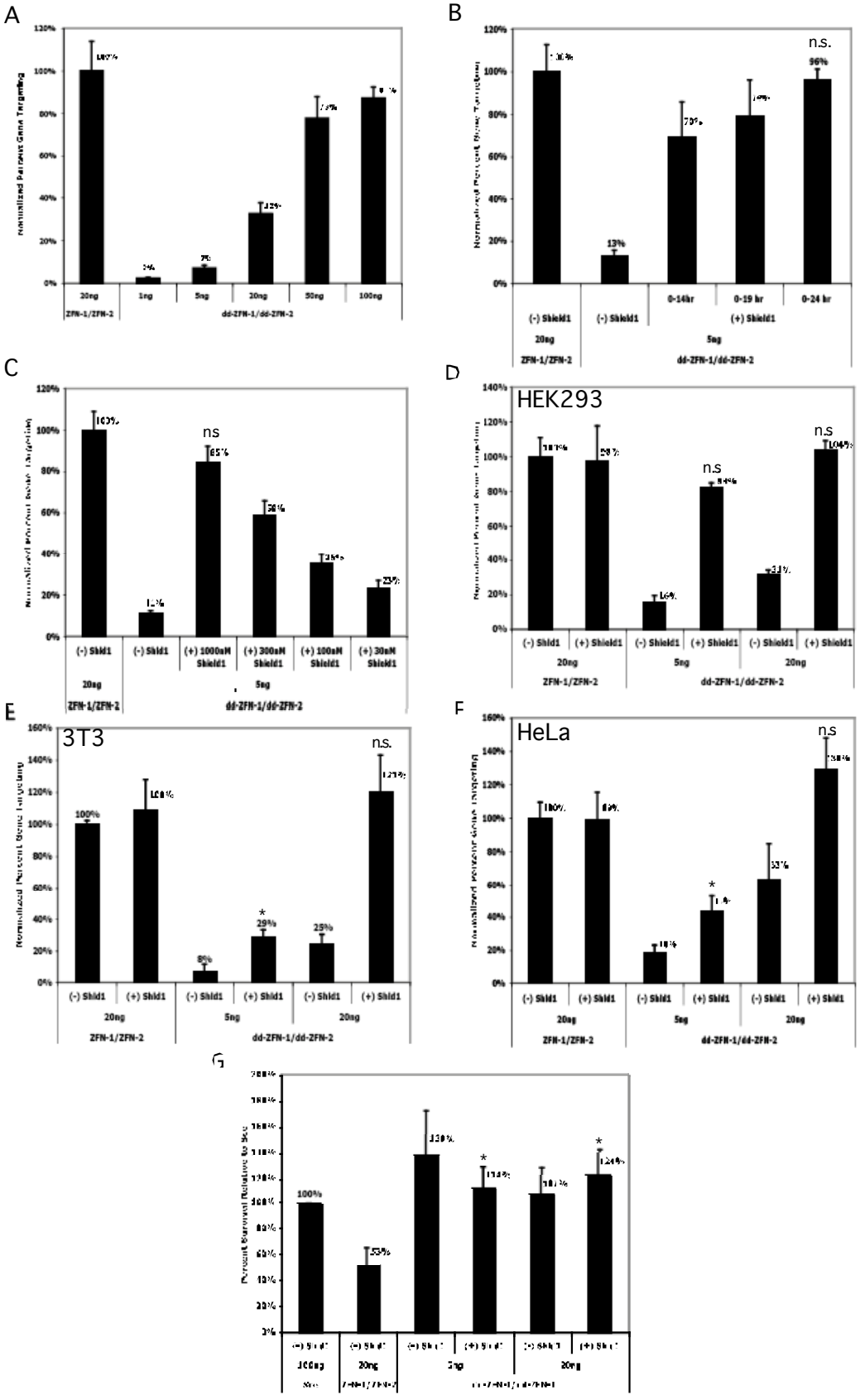
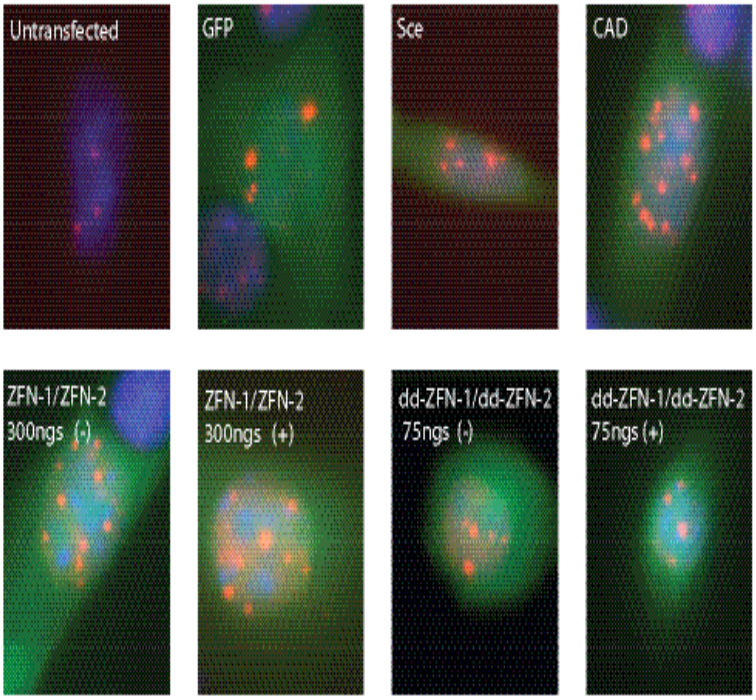

A
B
n.s.
C
ns
D
HEK293
n.s
E
3T3
n.s.
F
HeLa
n.s
G
(-) Shld1
(+) Shld1
20ng
5ng
ZFN-1/ZFN-2
dd-ZFN-1/dd-ZFN-2

Figure 3.4. Analysis of dd-ZFNs.

Unless otherwise indicated, rates of gene targeting at day 3 were normalized to the rate of gene targeting achieved using 20ng of the unmodified ZFNs without drug treatment as this was previously determined to be the conditions used to obtain optimal gene targeting with the unmodified ZFNs. (A) Titration of transfected DNA of dd-ZFNs in the gene targeting assay at day 3 after transfection with increasing amounts of transfected DNA. (B) Time-course experiment for length of exposure of 1000nM Shield1 using 5 ng of dd-ZFNs. Hours are given relative to the time of transfection, where "0" is the time of transfection. (C) Drug dose response curve for Shield1 with 5ng of dd-ZFNs. D) Gene targeting in the absence and presence of 1000nM Shield1 for given ZFN pairs at stated DNA concentrations in HEK293 cells (E) Gene targeting in the absence and presence of 1000nM Shield1 for given ZFN pairs at stated DNA concentrations in 3T3 cells. (F) Gene targeting in the absence and presence of 1000nM Shield1 for given ZFN pairs at stated DNA concentrations in HeLa cells. (G) Toxicity assay for all iterations of dd-modified and unmodified ZFNs tested in the gene targeting assay relative to Sce. A value of <100% indicates decreased cell survival as compared with Sce, and demonstrates a toxic effect. Statistical analysis was performed using the Student's T-test comparing ZFN-1/ZFN-2 at 20ng with no drug treatment to dd-modified ZFNs treated with Shield1. (*) indicates a P-value of <.05 and n.s. indicates no statistical significance or a P-value of >.05.Error bars are the standard deviation in measurement of three samples.

A


Untransfected
GFP
Sce
CAD
ZFN-1/ZFN-2
300ngs (-)
ZFN-1/ZFN-2
300ngs (+)
dd-ZFN-1/dd-ZFN-2
75ngs (-)
dd-ZFN-1/dd-ZFN-2
75ngs (+)

B

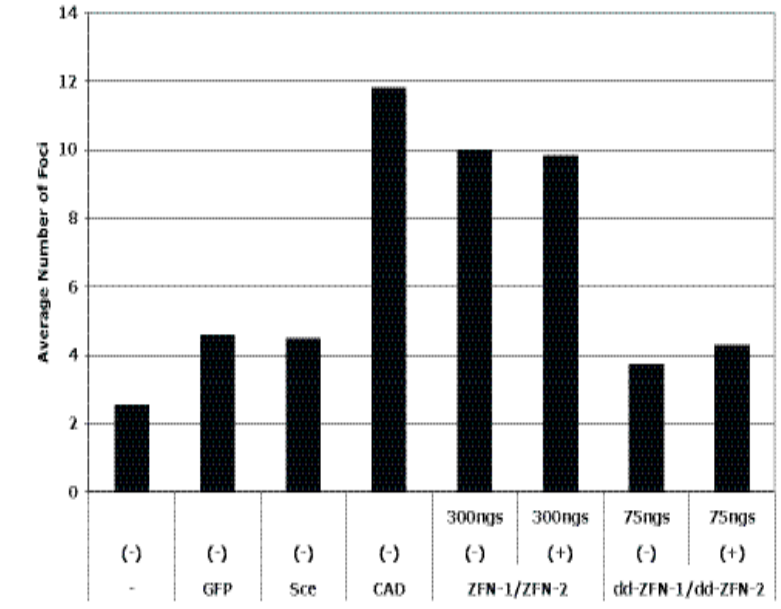

Average Number of Foci
14
12
10
8
6
4
2
0
(-)
-
(-)
GFP
(-)
Sce
(-)
CAD
300ngs
(-)
300ngs
(+)
ZFN-1/ZFN-2
75ngs
(-)
75ngs
(+)
dd-ZFN-1/dd-ZFN-2

Figure 3.5. Visualization of ZFN-induced DSBs by sensitized 53BP1 foci formation assay.

(A) Representative cells for each experimental condition after 53BP1 staining using indicated amounts of transfected DNA of each nuclease in the presence or absence of drug. 53BP1 foci are seen in red, 4,6-diamidino-2-phenylindole staining in blue, and GFP-positive cells in green. The foci were counted in transfected cells that were GFP-positive. (-) indicates untransfected, GFP indicates transfection of GFP alone, Sce serves as a negative control for ZFN-induced foci formation, and Caspase Activated DNAse (CAD) serves as a positive control for 53BP1 foci formation. ((-)) indicates no Shield1 treatment and ((+)) indicates 1000nM Shield1 treatment for 24 hours after transfection. (B) The average number of 53BP1 foci per transfected cell in $Ku80^{-/-}$ murine 3T3 cells for each experimental condition in (A).

CHAPTER IV: CONCLUSIONS AND FUTURE DIRECTIONS

Advances in ZFP Design Strategy

Exciting progress has been made in the study of ZFNs over the last decade, but many issues remain. One important issue that is currently being debated is the method by which ZFPs are designed. This issue is vital to the advancement of ZFNs for use in genome modification because a ZFN is only as good as its ZFP. The specificity of the ZFN is conferred entirely by the ZFP, and it is therefore essential that the ZFP be designed to be very specific for its target site. Several design strategies have been proposed, most of which fall into one of two main categories: (1) a modular assembly approach or (2) a combinatorial selection based approach. The modular assembly approach involves linking together single zinc finger motifs with precharacterized specificities (Beerli and Barbas 2002; Liu, Xia et al. 2002; Bae, Kwon et al. 2003; Segal, Beerli et al. 2003; Mandell and Barbas 2006). Although the modular assembly strategy is relatively easy to perform, the use of this method to create functional ZFN pairs has been met with limited success (Ramirez, Foley et al. 2008). When successful,

this design strategy has been shown to produce ZFNs with low activities and high toxicities (Cornu, Thibodeau-Beganny et al. 2008; Pruett-Miller, Connelly et al. 2008). Sangamo Biosciences, a small biotech company, holds the IP rights to a library of two finger cassettes that were designed by a combination of modular assembly, selection based approaches, and empirical design. The ZFNs produced from this library have been used in many of the proof-of-principle experiments for using ZFNs to mediate genome modification (Urnov, Miller et al. 2005; Lombardo, Genovese et al. 2007; Moehle, Rock et al. 2007; Perez, Wang et al. 2008; Santiago, Chan et al. 2008). Unfortunately, this library is not currently available to the academic community. Several combinatorial selection based approaches are, however, available to the academic community and have yielded multifinger ZFPs with high DNA-binding affinities and specificities (Greisman and Pabo 1997; Isalan, Klug et al. 2001; Hurt, Thibodeau et al. 2003). In this work, we used a bacterial 2-hybrid combinatorial selection based strategy to select for ZFPs and fused each protein to the FokI nuclease domain to create ZFNs (Hurt, Thibodeau et al. 2003; Pruett-Miller, Connelly et al. 2008).

The work presented here demonstrates that the use of a bacterial 2-hybrid based design strategy can produce ZFNs that are both more active and less toxic than ZFNs that target the same site but that were designed using the modular assembly design strategy. Cornu and colleagues observed similar results when comparing multiple pairs of ZFNs designed by the B2H strategy to different target

sites, and showed that ZFNs designed in this way have both high affinities and specificities for their target sites (Cornu, Thibodeau-Beganny et al. 2008). Moreover, we compared the rates of gene targeting produced by the best published pair of 4-fingered ZFNs to the 3-finger B2H designed ZFNs and showed that higher finger number alone does not produce higher rates of gene targeting (Pruett-Miller, Connelly et al. 2008). The tested 3-finger proteins also exhibited lower rates of cytotoxicity than the tested 4-finger ZFNs. Although a more comprehensive study needs to be performed in order to assure that this phenomenon is generally applicable, we propose that the B2H design strategy for creating 3-finger ZFPs is the optimal platform currently available for producing ZFNs with high rates of gene targeting and low rates of toxicity.

Unfortunately, all of the published combinatorial selection based strategies, including the B2H strategy, are quite difficult to perform, and require more expertise and time relative to the modular assembly approach. For example, the B2H design strategy involves two rounds of selection, which could be performed from start to finish in about 6 months in labs with some degree of expertise. We have been involved in making this strategy more available to the academic community and have reduced the time to production of ZFPs considerably by performing the first round of selection and making the selected pools available (Maeder, Thibodeau-Beganny et al. 2008). In Maeder et al (2008), we further validated the B2H strategy as an approach that produces ZFPs

and subsequently ZFNs that are more effective than ZFNs produced by the current modular assembly approaches.

Reduction of ZFN Toxicity

The issue of ZFN toxicity and off-target DSBs is one of the major hurdles in using ZFNs to stimulate genome modification in human cells because DSBs are known to be mutagenic and can lead to translocations and tumorigenesis. Therefore, ZFN systems are needed that minimize off-target DSBs and cytotoxicity. As discussed earlier, the best way to reduce ZFN toxicity is to design highly specific ZFPs. However, we have demonstrated that even with the best published ZFNs, each pair has an optimal range or "sweet spot" in which they function. That is, the rate of gene targeting increases as the concentration of ZFN increases, but only up to a point at which gene targeting decreases due to the increasing toxicity of the ZFNs at higher concentrations. The "sweet spot" is the optimal range in which the ZFNs function with minimal toxicity and can be determined by performing titration experiments for both toxicity and functionality. In order for ZFNs to stimulate high rates of targeting with minimal toxicity, it will be important to find the "sweet spot" in which each pair of ZFNs functions. This will be especially important for those ZFNs that are destined for clinical use.

With this concept in mind, we developed two additional strategies for further reducing toxicity caused by off-target breaks through the regulation of ZFN protein levels with small molecules. The first strategy utilizes a system in which a ubiquitin moiety is N-terminally linked to the ZFN. Using a proteasome inhibitor (MG132), the protein levels were regulated in order to expose cells to the ZFNs for a shorter period of time, which resulted in a reduction in toxicity while maintaining activity. We also used a system whereby a previously described destabilization domain based on an FKBP12 mutant was N-terminally linked to the ZFN (Banaszynski, Liu et al. 2005). Using the small molecule Shield1, we were able to regulate protein levels and reduce ZFN toxicity, while maintaining ZFN stimulated gene targeting activity. ZFN systems that reduce toxicity like the ones described here may be essential to using ZFNs in primary human cells, especially human embryonic stem cells.

Another strategy for reducing the number of off-target breaks is to reduce the number of sites to which a pair of ZFNs can bind. For example, ZFN-1 can homodimerize with itself, and create a new target site. The same is true for ZFN-2. Several labs have shown that mutations that force heterodimerization of ZFN pairs can reduce toxicity (Miller, Holmes et al. 2007; Szczepek, Brondani et al. 2007; Pruett-Miller, Connelly et al. 2008). That is, if ZFNs are not allowed to homodimerize, toxicity is reduced, most likely due to the creation of fewer off-target DSBs. However, in our studies, the obligate heterodimers were not only

less toxic but also less active in the GFP gene targeting assay (Pruett-Miller, Connelly et al. 2008). From these results, we cannot determine if the reduction in toxicity was due to a decrease in the nuclease activity or because homodimerization of ZFNs was reduced resulting in fewer off-target breaks. However, this reduction in activity was not observed in the other studies (Miller, Holmes et al. 2007; Szczepek, Brondani et al. 2007).

Increasing ZFN Efficacy

In recent years, the field of gene therapy has received a great deal of negative attention because of the adverse events that have resulted from some clinical trials. In particular, although several patients benefited from the SCID and chronic granulomatous disease (CGD) trials, unexpected insertional activation of protooncogenes resulted in leukemia or a myeloproliferative condition in a few patients (Cavazzana-Calvo, Lagresle et al. 2005; Ott, Schmidt et al. 2006). The clinical trials in which these adverse effects occurred were using the gene addition strategy for treatment of disease in which they used an integrating virus to insert a functional copy of the mutated, disease-causing gene somewhere in the genome.

ZFNs are being pursued as reagents to stimulate gene corrective gene therapy. The major benefits to this type of therapy are safety and efficacy. By correcting the gene at its endogenous locus, there is no theoretical risk of

insertional oncogenesis, which makes this strategy safer. Also, because the defective gene is being corrected at its endogenous locus, it can remain under control of its natural elements giving it a good efficacy profile by retaining its endogenous expression pattern. Unlike the traditional gene addition gene therapy strategies, which can potentially only treat homozygous recessive diseases, ZFN stimulated gene correction could theoretically treat homozygous recessive, heterozygous dominant or even homozygous dominant diseases by correcting one or both of the mutant alleles, repsectively. ZFNs are showing great promise as biallelic genome modifiers. Urnov and colleagues demonstrated that ZFNs could stimulate 13% monoallelic modification and 6% biallelic modification in human cells (Urnov, Miller et al. 2005). Other studies have shown that ZFNs can stimulate targeted mutatgenesis at both alleles via abberent NHEJ and thereby create knockouts in both mammalian cells and zebrafish (Doyon, McCammon et al. 2008; Meng, Noyes et al. 2008; Santiago, Chan et al. 2008).

In order for ZFN stimulated gene correction to be highly efficacious, the gene must be corrected precisely and permanently. ZFN stimulated gene correction involves the use of a repair substrate that has homology to the gene to be repaired. ZFNs are designed so that they will cut their target site within the genome at sites near the desired change, and the repair substrate serves as a template from which the ZFN-induced DSB can be repaired. Because the repair template must contain some degree of homology to the gene to be repaired, the

ZFN target site is often also found on the repair substrate. This ability to cut the repair substrate could lead to at least two ways in which gene targeting could be less efficacious than if the repair substrate did not contain the ZFN full-length binding site. First, ZFNs can bind and cut the repair substrate itself and thereby reduce the amount of available repair substrate. Second, once the repair substrate has been used to repair a mutant gene, the DNA content of the repair substrate is essentially transferred to the defective gene in order to repair it. If the repair substrate "transfers" the ZFN full-site, ZFNs can rebind and cut already targeted genes, and possibly lead to a reduction in gene targeting efficacy. That is, cells in which gene targeting had already occurred would lose their targeted events because of recutting. In this work, we observed that optimal rates of gene targeting could be achieved by altering the repair template to prevent re-binding and subsequent re-cutting by the ZFNs. We hypothesize that this beneficial effect is the result of preventing "re-cutting mutagenesis" whereby the targeted gene is re-cut by the ZFNs and then subsequently mutagenized by NHEJ.

Future Directions

General ZFN design

One of the greatest appeals to using ZFNs to create target specific modifications of the genome is that they can theoretically be designed to recognize any target sequence. However, the design strategies that are currently available to academic researchers are most successful at designing ZFPs that bind 5'-GNNGNNGNN-3' targets. The Zif268 crystal structure gives some insight into this phenomenon. In the Zif268 crystal structure, all three of the fingers have the same amino acid residues at positions -1 and 2; arginine and aspartic acid respectively. The arginine makes two hydrogen bonds to the guanine at the 3' position in the primary DNA strand of each finger, while the aspartic acid stabilizes this interaction by also making two hydrogen bonds to the guanidinium group of arginine. The aspartic acid at position 2 of the recognition helix also makes minor contacts with the adenine or cytosine on the complementary strand just outside the primary triplet subsite (Pavletich and Pabo 1991). It is these cross finger interactions that may be the limiting factor in ZFN design, especially since a large number of the phage display and combinatorial selection based strategies rely on the "framework" of Zif268 to design novel ZFNs. In order to design ZFNs that bind with great affinity and specificity to non-pure "GNN" sites, it will

be important to better understand these cross-finger interactions in order to improve current design strategies to create novel ZFNs to a greater number of target sites. Ideally, ZFN pairs will be designed that are sufficiently specific to create DSBs at any single DNA target site within a genome.

Using ZFNs for gene knockout in cell culture and in model organisms

Several recent studies have demonstrated the successful use of ZFNs for creating targeted knockouts via aberrant NHEJ in both mammalian cultured cells and zebrafish (Doyon, McCammon et al. 2008; Meng, Noyes et al. 2008; Santiago, Chan et al. 2008). Although these proof-of-principle experiments may lay the groundwork for a new widely used gene knockout strategy, several issues need to be further examined. First, just as with other gene disruption strategies, it will be important to establish the potential for confounding off-target effects. Second, ZFN accessibility to DNA due to chromatin status and/or developmental expression pattern may be an issue when targeting a wide variety of target sites in a wide variety of cell types. Third, when dealing with cultured cells that contain chromosomal aberrations and are often aneuploid, it will be important to determine that all target sites on all chromosomes are targeted within a cell to

ensure complete knockout. One exciting advantage that this technique might provide is that multiple genes could be knocked out with relative ease.

Using ZFNs for targeted engineering of the human genome in stem cells for clinical application

Traditional gene therapy has involved using *in vivo* administration of viral vectors that allow random integration of a functional copy of a defective gene into the human genome of precursor stem cells. Several problems arise from such a strategy. First, the *in vivo* administration of viral vectors can elicit a host immune response. Second, multiple cell types, besides the disease associated cell type, within the patient can be infected, further increasing the risk. This also means using high MOIs of the viral vector to ensure infections of the correct cell type. Also, the risk of germline transduction and insertional activation of proto-oncogenes are also worrisome.

In contrast to the *in vivo* administration of viral vectors, ZFN-based strategies as a form of gene corrective gene therapy would involve the modification of patient-specific stem cells by ZFN-induced gene targeting in culture, followed by selection and expansion of desired modified cells *ex vivo* and re-introduction into the patient. Ideally, the modified cells would then repopulate

the patient's afflicted tissues and alleviate or cure the disease. In order for this type of therapy to reach clinical application, several obstacles must be overcome.

As discussed above, the most important limiting factor towards using ZFN-based strategies is the creation of highly specific ZFPs. In order to evaluate if a pair of ZFNs can move forward toward clinical use, we propose that a panel of assays be performed in order to show minimal off-target DSBs and low or no toxicity. We also believe that it will be important to use both cell-based strategies and mouse-model systems to show that these systems do not cause oncogenic mutations or translocations.

Another major issue impeding the success of ZFNs is that of delivery. In contrast to traditional gene therapy in which only a functional copy of the gene needs to be delivered into a cell, the ZFN-based strategy involves simultaneous delivery of two different ZFNs and a repair substrate. Lombardo and colleagues successfully used an integrase-defective lentiviral vector to deliver both ZFNs and the template DNA for gene correction in several different cell types, including human ES cells (Lombardo, Genovese et al. 2007). However, even when using non-integrating viral vectors, the delivery of DNA expression cassettes with strong promoters, including the ZFNs or the repair template, has a potential risk of insertional activation of nearby genes. Even if the DNA was engineered to be promoterless, random integration of the delivered DNA could still theoretically inactivate important genes. In order to reduce the risk of insertional mutagenesis,

ZFNs could potentially be delivered as proteins. This type of delivery would also theoretically reduce the potential for off-target breaks by reducing the length of time the ZFNs were present. Finally, as with other types of therapy, ZFN-based gene therapy has a risk of evoking an immune response, particularly because the FokI nuclease domain was derived from a bacterial nuclease.

Overall Conclusions

ZFN technology has come a long way in the last decade. Although, I would not say that it is in its infancy, I would say it is going through its toddler phase. While great steps have been made to advance the field, many important questions and technical difficulties remain. The work presented in this thesis has advanced the field by validating the B2H selection strategy as a highly effective strategy for designing ZFPs and subsequently ZFNs. We have also shown that increasing the number of fingers alone does not produce more active or less toxic ZFNs, and we therefore conclude that the 3-finger platform is the optimal platform currently available with which to continue designing ZFNs in the academic community. We have also demonstrated several strategies for reducing ZFN toxicity and increasing efficacy, and developed a panel of assays by which to analyze both the activity and toxicity of novel ZFNs.

APPENDIX A

BACTERIAL 2-HYBRID SELECTION PROTOCOL

I. Reporter strain construction

Annealing Oligos

1ul oligo 1 (10pmol/ul)

1ul oligo 2 (10pmol/ul)

20ul ligase buffer

178 ul dH20

200ul total

place at 95C for 5 mins, then let cool to room temp

Backbone: KJ1712

2ug KJ1712

3ul 10X Buffer 4

1ul SapI

fill to 30 ul with dH20

37C for 1-2 hours, run out on a .6% agarose gel, gel purify into 30ul EB

Pfu the BKB

2ul 10mM dCTP

2ul 10X Pfu Buffer

10ul gel isolated BKB

1.2ul Cloned poly Pfu enzyme

4.8ul dH20

20ul total

ligation

2ul BKB

8ul annealed oligos

1ul 10X ligase buffer

1ul ligase

12 ul total

ligate o/n at RT, plate on LB-Kan plates, pick colonies based on stimulation

this is a very low copy plasmid, so grow up in 4mls of rich media

miniprep, digest with NotI/HindIII (look for absense of 400bp band)

sequence with OK181 (this is the reverse primer, so you'll have to look at the reverse compliment to find your site: target site will be right after EcoRI site)

II. Crosses

2 days prior to cross, streak our KJ1C on LB-Tet12.5

day prior, transform with 5 ul of reporter plasmids into (50ul) chemically competent CSH100 (see SP103 for protocol to make competent), plate entire reaction on LB-Kan30

Day prior, also start KJ1C o/n culture in LB-Tet12.5 (2ml)

Day of crosses:

1. Scrape plates with sterile wooden sticks and transfer to 10ml LB.
2. Vortex on low speed
3. If culture is confluent of close to it (as it should be when scaping a plate that has a complete lawn) subculture into 5 mls LB
    a. Subculture to be about OD 600 of .1, so that they can get back into log phase and regrow their pilli
4. Also subculture KJ1Cs (~500ul) into 10mls of LB
5. Also start a blank of 10mls
6. Let 3,4, and 5 from above grow for 2 hours at 37C

a. Not on wheel, to let pilli reform

7. Mix 1 ml of reporter CSH100 culture with 1ml of KJ1Cs
8. Also, for controls mix 1ml reporter CSH100 culture with 1 ml blank LB
9. Further controls, LB+LB and LB+KJ1C
10. Let grow for 1 hours at 37C, not on wheel
11. Move to wheel for 90 mins (or shaker at 105 rpm)
12. Plate 300ul of each experimental culture (CSH100+KJ1C) on both LB-TK and LB-TKS plates
13. Plate all controls on one plate in 20ul spots on LB-TK an LB-TKS plates to check for contamination
14. Let grow o/n at 37C
    a. Should see a 10 fold reduction in TK to TKS plates
        i. TK- single recombination
        ii. TKS double recombination event (we want)
15. Pick 2 individual colonies and streak side by side on one LB-TKS plate for individual colonies (primary restreak)
16. Repeat step 15 (secondary restreak)

III. Confirm Strains: genetic screen

a. Use one colony to start o/n culture in 4ml LB-Kan30 of both A and B candidates and to patch onto an M9-glu + His plate to perform genetic test (should be able to grow in the absence of pro and arg
    1. Use a 96 well plate and multichannel pipet
    2. Add 100ul of NM media to each well
    3. Then pick and swirl individual colony in appropriate well
    4. Let grow 30 mins to an hour (not necessary)
    5. Plate on LB+Kan and M9 plates in 5 ul spots (see SP107 for details)
    6. Use the remaining media in the 96 well plate to start o/n LB-kan30 cultures
    7. Let grow o/n in shaker (250 rpm, 37 C)
        a. LB-Kan30 is positive control to make sure you inoculated the well
        b. If it doesn't grow on M9 plate, use B clone or redo cross

IV. Make reporter strains competent to get a-gal4 into cells

a. Add Kan30 to 10mls LB-15mM MgCl2 o/n culture (grow o/n to make sure not contaminated)

b. Inoculate with 200ul of o/n reporter culture

c. Let grow for 1-1.5 hours at 37C in shaker (250rpm)
    i. For some reason, some don't grow, so don't worry if this happens
d. Transfer cells to a 15 ml conical tube and spin @ 2500rpm, 4C, 25min
e. Resuspend in 300ul of MES solution A/15% glycerol
    i. 10mM MnCl2-4H20 (that is manganese chloride)
    ii. 50mM CaCl2
    iii. 10mM MES, pH 6.3
    iv. 15% glycerol
    v. filter sterilize
        1. no need to pH final solution
        2. wrap in foil and store at 4C
f. aliquot into 150ul aliquots
g. quick freeze on ethanol/dry ice bath and store at -80C
h. label tubes

V. PCR templates to make sure strains are correct
a. Sterilely transfer 100ul of culture to an ependorf tube
b. Spin max 1 minute
c. Discard supe and resuspend in 100ul of dH2O

d. Boil cells @ 95C, 10 minutes

e. Spin max 1 minute

f. Transfer 50 ul of supe to a fresh tube and store @ -20

VI. Make glycerol stocks:

a. To remainder of the o/n culture, add 430 ul of 50% sterile glycerol solution for every 1 ml of culture

VII. PCR

a. Use PCR boiled DNA

5ul DNA

5ul 10X PCR buffer

2ul MgCl2 (50mM)

5ul DNTPs(2.5mM)

1ul OK 5 (10pmol/ul)

1ul OK163 (10pmol/ul)

.5 taq

30.5 dH2O

50ul total

cycle

95C for 5 min

95C for 30 sec

60C for 1 min

72C for 2 min

cycle 25 X

72C for 10 min

Sequence with OK 181 to confirm cell lines

VIII. transform cells with KJ1267 (CAM resistant; agal-4)

a. 1ul into 50ul appropriate chemically competent cells

b. plate on on LB-Kan30 Cam30 plates

c. grow o/n at 37C

d. pick individual colony and grow up in LB-CK

e. next day, make glycerol stocks

i. 1ml culture

ii. 500ul 50% glycerol

f. store -80C

Selection 1A Protocol

1. Day prior to selection, start 20 ml o/n culture of appropriate selection strain in NM-CK-I50
2. Approx. 18 hours after inoculation, colonies should be close to confluent (OD600 >1.0… colonies have a yellow tint, but not all)
3. take out all but 5 mls of o/n culture
4. infect with 25 microliters of master randomized library to appropriate finger
5. briefly swirl to disperse phage
6. let infection go at RT for 30 min. (no swirling)
7. add 20mls pre-warmed NM-CK-I50 media to infection
8. shake at 105 rpm for 1.5 hours
9. transfer to 50 ml conical vial and spin at RT at 2500rpm
10. discard supe and resuspend pellet in 2.5 mls of pre-warmed NM-CK-I50
11. record volume (it's always a little more than 2.5 mls usually about 2.7)
12. make 10 fold dilutions of infection from 10E-1 to 10E-8
    a. in a 96-well plate, add 90 microliter of NM-CK-I50 to all wells and add 10 microliters of infection to 10E-1 dilution (X3 for each infection)

b. mix up and down and add 10 microliters of 10E-1 dilution to 10E-2 dilution

c. remember to change tips between dilutions for accuracy

13. make 5 microliter spots of 10E-3-10E-8 dilutions in triplicate (so each dilution of each infection will have 9 spots total) on

    a. LB-CK (bacteria titer)

    b. LB-CCK (infected bacteria titer)

    c. NM-CCK-I50 (infected bacteria titer on minimal media)

14. plate remainder of infection (2.5mls) on Big 245mmX245mm plates of NM-CCK-I50-3AT10 using sterile glass beads to spread.

15. When plates are completely dry, turn plates over and tap beads into lid (do not discard beads, you will use these to harvest later)

16. Let LB plates incubate at 37 degrees overnight

17. Let NM-CCK-I50 incubate at 37 degrees for 24 hours

18. Let Big NM plates incubate at 37 degrees for 24, then move to RT for 18 hours before harvesting

19. Next day, titer LB plates as follow

    a. (Number of colonies averaged from three spottings)/(15 microliters per individual spotting) = Y

    b. (Y) X (total volume of NM that the pellot was resuspended in) = Z

c. (Z) X (dilution factor at which you are counting the colonies) = titer

Harvest and Rescue Protocol

1. After plates have been incubated for 24 hours at 37 degrees and 18 hours at RT, turn plates over and tap beads back onto plate.
2. Add 15 mls pre-warmed NM-CK-I50 media to plate (put media around beads…it helps speed the harvest)
3. When all colonies are scraped off of agar, tilt plate and suck up media with a 10ml pipet, and transfer to a 15 ml conical vial.
4. Add about 1 ml of harvested cells to 90 mls of 2XYT-Kan30 (we want an OD600 of about .1, so that bacteria can double several times to allow pilli to reform after harvest)
5. Save 3 mls of harvested cells for glycerol stocks (add 432 mls of 50% glycerol to 1 ml harvested cells) and store at –80C. Label tubes KJSP__ Sel 1A harvest and date
6. Allow 90 ml culture to swirl at 105 rpms for 1 hour
7. Add 30 microliters of M13K07 helperphage (titer 1E11) to each culture
8. Swirl to disperse helperphage

9. Let infection go at RT for 30 minutes with no swirling
10. Add Kan to a final concentration of Kan100 (add another 213 microliters).
11. Let shake at 105 rpms for six hours at 37C
12. Filter sterilize through a PES .22 micrometer Millipore filter system (cat #SCGPU01RE--- these work the best)
13. Label KJSP__ or ZFC__ and store at 4C. (eventually transfer 40 mls to a 50cc conical tube and store at –80C.

Selection 1B

1. Grow o/n culture of selection strain in 20 mls of NM-CK-I50 (15-22 hours= OD600 approx. 1.0)
2. Make 10-fold dilutions in triplicate of enriched phage library from 10E0-10E-5 (done in 96-well plate)
3. infect 50 microliters of o/n culture with 10 microliters of phage dilutions (done in 96-well plate)
4. let infections go at RT for 30 min
5. add 190 microliters of NM-CK-I50 to each well
6. incubate at 37C for 2 hours

7. spot 5 microliter spots on NM-CCK-I50 and NM-CCK-I50-3AT10-Strep20 plates in triplicate (9 spots total for each dilution)
8. incubate at 37C for 48 hours
9. determine titer as follows
    a. average # of colonies/15 microliters (3X5microliter spots) = X
    b. (X)*250 microliter total volume = Y
    c. Y/10 microliters of phage in original infection = Z
    d. Z*dilution factor = TITER
10. From this titer determine how many microliters of phage are needed to get 300 colonies on the NM-CCK-I50-3AT10-Strep20 plate as follows
    a. 300/TITER = A
    b. A*10 = B (multiply by ten, and do infections like normal 5 ml infections, but only plate 1/10 (250 microliters)
11. For infections, day prior, start 20 ml o/n culture of appropriate selection strain
12. remove all but 5 mls of o/n culture
13. infect the remaining 5 mls with appropriate amount of phage as determined by 10b and briefly swirl to disperse
14. let infection go at RT for 30 minutes (no swirling)
15. add 20 mls NM-CK-I50 to infection and swirl at 105 rpms for 2 hours
16. spin down at RT at 2500 rpm for 25 min

17. resuspend in 2.5 mls of NM-CK-I50 (pre-warmed)
18. plate 1/10 of volume (250 microliters) on NM-CCK-I50-3AT10-Strep20 small round plates
19. incubate at 37C for 48 hours
20. pick 95 individual colonies and inoculate a 96-well block in 1 ml TB+Carb50 (96^{th} well is control)
21. grow o/n at 900rpm at 37C for about 3 hours
22. then, use stamp to inoculate another 96-well block of TB-Carb50 with same 95 colonies (for next day minipreps)
    a. To sterilize stamp, dip in 95% EtOH and flame (be sure to flame backstop to prevent contamination)
    b. Let cool briefly and dip in Block 1 and carefully remove being careful not to touch the sides of the wells.
    c. Slow move to Block B and let stamp rest on backstop for about 5 seconds while recovering block A.
    d. Remove stamp from Block B being careful not to touch sides of wells.
    e. Cover with Airpore breathable strip.
23. Let both blocks grow for at least 12 hours
24. Next day, add 432 microliters of 50% glycerol (15% total) and snap freeze block A

25. store at –80C
26. to block B, suck out all media (1ml per well) and mix in flask to create pooled bacteria into a clean flask (use 2 ml pipet… a 1 ml multichannel pipet might also work).
27. aliquot 10mls from pooled bacteria into a 15cc tube and spin at 3000rpm for 15 min and do a miniprep
28. to remaining 90mls of pooled bacteria, spin down at 3000rpms and resuspend in 6 mls of P1 buffer and store at –80 (snap freeze) to be used for Midi-prep (label tube ZFC__ pooled bacteria in P1 and date.

Single Finger Amplification and Fusion PCR

1. amplify individual fingers as follows

	1(ul)	2(ul)	3(ul)
F1_pooled DNA (from step 27	.5		
F2_pooled DNA		.5	
F3_pooled DNA			.5

10X PCR buffer	5	5	5
MgCl2 (50mM)	2	2	2
dNTPs (2mM)	5	5	5
SP29A (30pmol/ul)	1		
MP154. OK.122(30 pmol/ul)	1		
MP154. OK.120 (30 pmol/ul		1	
MP154. OK.123 (30 pmol/ul)		1	
*MP154. OK.121 (30 pmol/ul)			1
SP29D (30 pmol/ul)			1
*Taq polymerase	.5	.5	.5
dH2O	35	35	35

*hot start (add when Thermocycler reaches 95)

Cycle 65X25

95 for 5’

95 for 1’

65 for 45s

72 for 1’

X 25 cycles

72 for 10’

4 forever

Run out on a 1.5% agarose gel

Gel purify and elute with 30 ul of EB (Qiagen kit)

Fusion PCR

F1 amplified (from above)	1
F2 amplified	1
F3 amplified	1
10XPCR buffer	5
MgCl2 (50mM)	2
dNTPs (2mM)	5
SP29A (30 pmol/ul)	1

SP29D (30pmol/ul)	1
*Taq polymerase	.5
dH2O	32.5

*hot start

Cycle 55X25

95 for 5’

95 for 1’

55 for 45s

72 for 1’

X 25 cycles

72 for 10’

4 forever

Run out on a 1.3% agarose gel

Gel purify and elute with 30 ul of EB (Fragment ~330bp)

Making Cell Library

Digest M266 (KJ1514 with AscI/NotI sites flanking ~300bp insert)

M266 (100ng/ul)	20
10 X buffer 4	3
AscI	1
NotI	1
dH20	5
	30 ul

Digest o/n at 37, CIP last hour, gel purify on a .6% agarose gel

Digest Three finger cassette (TFC) from fusion PCR

TFC	30	
10 X buffer 4	4	
AscI	1	Digest o/n at 37, gel purify
1.3%		
NotI	1	gel
dH2O	4	
	40ul	

Ligation

	V	V+I
M266/AscI/NotI	1	1
TFC/Asc/NotI	-	3
10 X ligase buffer	1	1
T4 ligase	1	1
dH2O	7	4
	10ul	10ul

ligate o/n at RT

Transform 100ul of Xl-1Blue Chem. Comp. Cells with 4 ul of ligation

Plate all on LB-AMP plate

37 o/n

If good colony stimulation, pick at least 5 colonies digest with XbaI/HindIII

If insert looks correct (330bp), send to sequencing w/OK.61

We need a complexity of at least (95 colonies)(95 colonies)(95 colonies) = 8.57E5

Use Super Electroporation competent cells (Stratagene cat#200158)

EtOH precipitate ligation, resuspend in 10ul of dH2O and use 5 ul of cleaned up ligation and 50 ul of super comp cells

Follow manufacturer's protocol

1. After electroporation, recover in 10ml of prewarmed SOC to each tube and inoculate @37 degrees for 1 hour (no antibiotics)
2. Pre-amplification titering: do 3 independent dilns 10-1 through 10-6 in 2XYT and spot 5ul spots on LB/tet12.5/carb100 (to determine complexity) and LB/carb100/kan70 plates (to check for phage contamination)
3. Amplify: transfer each library culture to 90 mls 2XYT/carb50/tet12.5 and shake at 250 rpm, 37, 2 hours.
4. POST-amplification tittering: repeat step 2 (pre-amplification tittering) after amplification.
5. Harvest: spin down cultures in sterile centrifuge bottles at 4 degrees, 4000rpm, 30'
6. Drain of supe and resuspend pellet (on ice) in 2 ml 2XYT/15% glycerol
7. label: SP149 cell library and date and freeze on dry ice for 30 min then to –80 freezer
8. Next day: determine pre-amplification titer as follows

a. (# of colonies/ul)(DF at which colonies grew)(10 ml volume) = complexity

Growing Library as Phage

1. thaw tube of frozen cell library and inoculate into 10mls 2XYT/tet12.5/carb50 in a 250 ml flask
2. place at 37 for 1.5 hours at 105 rpm
3. add 5ul of M13K07 phage (titer 1E11/ul)
    a. assume OD600=.1 = (7E7 bacteria/ml)(10ml) = 7E8
    b. (150 fold over sampling) (7E8)=1.05E11
    c. (1.05E11)(3) = 3.15 E11
    d. 3.15E11/1E11 titer = 3.15 round up to 5 ul
4. let infection go for 15 min at RT
5. Then place at 37 for 1.5 hours, 105 rpm
6. remove aliquot and dilute 10-1 through 10-8 and spot 5 ul spots on LB/tet/carb100 plates and LB/carb100/kan70 plates (to make sure every cell was infected)
7. add 90ml 2XYT/carb 50 to each flask
8. place at 37, 250 rpm for 1 hour
9. add Kan70

10. let incubate at 37 for 18 hours at 250 rpm
11. Next day, transfer cultures to 50 ml conical vial
12. spin at 4 degrees, 4000 rpm, 30 min
13. pour supe through a PES filter unit and store at 4 degrees

PEG Precipitation

1. put 40mls of filtered library and 10mls 5X PEG/NaCl (17.5%/12.5%) solution in each sterile 50 mls spin tube
2. Mix well and place on ice a minimum of 2.5 hours (I usually go overnight)
3. Spin at 4 degrees for 45 mins at 10,000 rpm
4. keep phage on ice and pour off each supe and resuspend each library in .5 ml/40 mls culture in 2XYT/ 15% glycerol

Titering PEG-precipitated Phage Library

1. Grow o/n culture of selection strain in 20 mls of NM-CK-I50 (15-22 hours= OD600 approx. 1.0)
2. Make 10-fold dilutions in triplicate of enriched phage library from 10E0- 10E-5 (done in 96-well plate)

3. infect 50 microliters of o/n culture with 10 microliters of phage dilutions (done in 96-well plate)
4. let infections go at RT for 30 min
5. add 190 microliters of NM-CK-I50 to each well
6. incubate at 37C for 2 hours
7. spot 5 microliter spots on NM-CCK-I50 and LB-CCK plates in triplicate (incubate at o/n 37C for 48 hours)
8. determine titer as follows
    a. average # of colonies/15 microliters (3X5microliter spots) = X
    b. (X)*250 microliter total volume = Y
    c. Y/10 microliters of phage in original infection = Z
    d. Z*dilution factor = TITER

Real Infections on Selective Media

1. day prior to infection start 20 ml NM-CK-I50 culture 18 hours before infection
2. Day of infection

    a. infect appropriate 5 mls o/n culture with appropriate amount of phage

        *also do a control with the wrong phage on selection strain

*determine the amount of PEG-precipitated phage as follows (cell-based complexity)(6-fold oversampling)(#of NM plates)=X

* X / PEG precipitated titer = number of ul to add to infection normally less than 1 ul and requires that you do a dilution to be able to pipet accurately

b. allow infection to go for 30' at RT

c. add 20mls prewarmed NM-CK-I50 media and shake at 105 rpm, 37, for 2 hours

d. spin down at 2500, 25' at RT and discard supe

e. resuspend in 200ul per plate (ie normally 2 plates, so 400ul total) of NM-CK_I50 and record volume after resuspended

f. perform 10-1 through 10-7 dilutions on LB-CK, LB-CCK, and NM-CCK-I50 plates to determine infection conditions

g. plate 200ul on each of the following plates

i. NM-CCK-I50-3AT25-Strep40

ii. NM-CCK-I50-3AT40-Strep60

iii. NM-CCK-3AT40-Strep60

iv. NM-CCK-3AT50-Strep80

h. grow 2-5 days at 37

3. determine the number of colonies on each plate

4. pick at least 5 colonies from the highest stringency plate that had growth

5. Grow o/n in 2 ml TB/amp and miniprep with Promega Wizard Kit

6. Digest with XbaI/HindIII and if correct, send to sequencing w/ OK.61.

BIBLIOGRAPHY

Aiuti, A., S. Slavin, et al. (2002). "Correction of ADA-SCID by stem cell gene therapy combined with nonmyeloablative conditioning." Science **296**(5577): 2410-3.

Alkhatib, G., C. Combadiere, et al. (1996). "CC CKR5: a RANTES, MIP-1alpha, MIP-1beta receptor as a fusion cofactor for macrophage-tropic HIV-1." Science **272**(5270): 1955-8.

Alwin, S., M. B. Gere, et al. (2005). "Custom zinc-finger nucleases for use in human cells." Mol Ther **12**(4): 610-7.

Ausubel, F., Brent, R, Kingston, RE, Moore, DD, Seidman, JG, Smith, JA *et al.* (1996). "Current protocols in Molecular Biology."

Ausubel, F. M., *et al* (1996). (1996). *Current Protocols in Molecular Biology*. Boston, Wiley.

Bachmair, A., D. Finley, et al. (1986). "In vivo half-life of a protein is a function of its amino-terminal residue." Science **234**(4773): 179-86.

Bae, K. H., Y. D. Kwon, et al. (2003). "Human zinc fingers as building blocks in the construction of artificial transcription factors." Nat Biotechnol **21**(3): 275-80.

Banaszynski, L. A., C. W. Liu, et al. (2005). "Characterization of the FKBP.rapamycin.FRB ternary complex." J Am Chem Soc **127**(13): 4715-21.

Beerli, R. R. and C. F. Barbas, 3rd (2002). "Engineering polydactyl zinc-finger transcription factors." Nat Biotechnol **20**(2): 135-41.

Beerli, R. R., B. Dreier, et al. (2000). "Positive and negative regulation of endogenous genes by designed transcription factors." Proc Natl Acad Sci U S A **97**(4): 1495-500.

Beerli, R. R., D. J. Segal, et al. (1998). "Toward controlling gene expression at will: specific regulation of the erbB-2/HER-2 promoter by using polydactyl zinc finger proteins constructed from modular building blocks." Proc Natl Acad Sci U S A **95**(25): 14628-33.

Belfort, M. and R. J. Roberts (1997). "Homing endonucleases: keeping the house in order." Nucleic Acids Res **25**(17): 3379-88.

Beumer, K., G. Bhattacharyya, et al. (2006). "Efficient gene targeting in Drosophila with zinc-finger nucleases." Genetics **172**(4): 2391-403.

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

Bibikova, M., D. Carroll, et al. (2001). "Stimulation of homologous recombination through targeted cleavage by chimeric nucleases." Mol Cell Biol **21**(1): 289-97.

Bibikova, M., M. Golic, et al. (2002). "Targeted chromosomal cleavage and mutagenesis in Drosophila using zinc-finger nucleases." Genetics **161**(3): 1169-75.

Bitinaite, J., D. A. Wah, et al. (1998). "FokI dimerization is required for DNA cleavage." Proc Natl Acad Sci U S A **95**(18): 10570-5.

Brenneman, M., F. S. Gimble, et al. (1996). "Stimulation of intrachromosomal homologous recombination in human cells by electroporation with site-specific endonucleases." Proc Natl Acad Sci U S A **93**(8): 3608-12.

Burma, S., B. P. Chen, et al. (2006). "Role of non-homologous end joining (NHEJ) in maintaining genomic integrity." DNA Repair (Amst) **5**(9-10): 1042-8.

Capecchi, M. R. (2005). "Gene targeting in mice: functional analysis of the mammalian genome for the twenty-first century." Nat Rev Genet **6**(6): 507-12.

Cavazzana-Calvo, M., S. Hacein-Bey, et al. (2000). "Gene therapy of human severe combined immunodeficiency (SCID)-X1 disease." Science **288**(5466): 669-72.

Cavazzana-Calvo, M., C. Lagresle, et al. (2005). "Gene therapy for severe combined immunodeficiency." Annu Rev Med **56**: 585-602.

Chandrasegaran, S. and J. Smith (1999). "Chimeric restriction enzymes: what is next?" Biol Chem **380**(7-8): 841-8.

Chevalier, B. S., T. Kortemme, et al. (2002). "Design, activity, and structure of a highly specific artificial endonuclease." Mol Cell **10**(4): 895-905.

Chevalier, B. S. and B. L. Stoddard (2001). "Homing endonucleases: structural and functional insight into the catalysts of intron/intein mobility." Nucleic Acids Res **29**(18): 3757-74.

Choo, Y. and A. Klug (1994). "Selection of DNA binding sites for zinc fingers using rationally randomized DNA reveals coded interactions." Proc Natl Acad Sci U S A **91**(23): 11168-72.

Choo, Y. and A. Klug (1994). "Toward a code for the interactions of zinc fingers with DNA: selection of randomized fingers displayed on phage." Proc Natl Acad Sci U S A **91**(23): 11163-7.

Choulika, A., A. Perrin, et al. (1995). "Induction of homologous recombination in mammalian chromosomes by using the I-SceI system of Saccharomyces cerevisiae." Mol Cell Biol **15**(4): 1968-73.

Cornu, T. I., S. Thibodeau-Beganny, et al. (2008). "DNA-binding specificity is a major determinant of the activity and toxicity of zinc-finger nucleases." Mol Ther **16**(2): 352-8.

Dantuma, N. P., K. Lindsten, et al. (2000). "Short-lived green fluorescent proteins for quantifying ubiquitin/proteasome-dependent proteolysis in living cells." Nat Biotechnol **18**(5): 538-43.

Deng, H., R. Liu, et al. (1996). "Identification of a major co-receptor for primary isolates of HIV-1." Nature **381**(6584): 661-6.

Dervan, P. B. and B. S. Edelson (2003). "Recognition of the DNA minor groove by pyrrole-imidazole polyamides." Curr Opin Struct Biol **13**(3): 284-99.

Desjarlais and J. M. Berg (1992). "Redesigning the DNA-binding specificity of a zinc finger protein: a data base-guided approach." Proteins **13**(3): 272.

Desjarlais, J. R. and J. M. Berg (1992). "Redesigning the DNA-binding specificity of a zinc finger protein: a data base-guided approach." Proteins **12**(2): 101-4.

Desjarlais, J. R. and J. M. Berg (1992). "Toward rules relating zinc finger protein sequences and DNA binding site preferences." Proc Natl Acad Sci U S A **89**(16): 7345-9.

Desjarlais, J. R. and J. M. Berg (1993). "Use of a zinc-finger consensus sequence framework and specificity rules to design specific DNA binding proteins." Proc Natl Acad Sci U S A **90**(6): 2256-60.

Diakun, G. P., L. Fairall, et al. (1986). "EXAFS study of the zinc-binding sites in the protein transcription factor IIIA." Nature **324**(6098): 698-9.

Doyon, Y., J. M. McCammon, et al. (2008). "Heritable targeted gene disruption in zebrafish using designed zinc-finger nucleases." Nat Biotechnol **26**(6): 702-8.

Dreier, B., R. R. Beerli, et al. (2001). "Development of zinc finger domains for recognition of the 5'-ANN-3' family of DNA sequences and their use in the construction of artificial transcription factors." J Biol Chem **276**(31): 29466-78.

Dreier, B., R. P. Fuller, et al. (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**(42): 35588-97.

Dreier, B., D. J. Segal, et al. (2000). "Insights into the molecular recognition of the 5'-GNN-3' family of DNA sequences by zinc finger domains." J Mol Biol **303**(4): 489-502.

Durai, S., M. Mani, et al. (2005). "Zinc finger nucleases: custom-designed molecular scissors for genome engineering of plant and mammalian cells." Nucleic Acids Res **33**(18): 5978-90.

Elrod-Erickson, M. and C. O. Pabo (1999). "Binding studies with mutants of Zif268. Contribution of individual side chains to binding affinity and specificity in the Zif268 zinc finger-DNA complex." J Biol Chem **274**(27): 19281-5.

Gaspar, H. B., K. L. Parsley, et al. (2004). "Gene therapy of X-linked severe combined immunodeficiency by use of a pseudotyped gammaretroviral vector." Lancet **364**(9452): 2181-7.

Gimble, F. S. (2000). "Invasion of a multitude of genetic niches by mobile endonuclease genes." FEMS Microbiol Lett **185**(2): 99-107.

Greisman, H. A. and C. O. Pabo (1997). "A general strategy for selecting high-affinity zinc finger proteins for diverse DNA target sites." Science **275**(5300): 657-61.

Hacein-Bey-Abina, S., C. Von Kalle, et al. (2003). "LMO2-associated clonal T cell proliferation in two patients after gene therapy for SCID-X1." Science **302**(5644): 415-9.

Hirata, R. K. and D. W. Russell (2000). "Design and packaging of adeno-associated virus gene targeting vectors." J Virol **74**(10): 4612-20.

Hurt, J. A., S. A. Thibodeau, et al. (2003). "Highly specific zinc finger proteins obtained by directed domain shuffling and cell-based selection." Proc Natl Acad Sci U S A **100**(21): 12271-6.

Isalan, M. and Y. Choo (2001). "Engineering nucleic acid-binding proteins by phage display." Methods Mol Biol **148**: 417-29.

Isalan, M., A. Klug, et al. (2001). "A rapid, generally applicable method to engineer zinc fingers illustrated by targeting the HIV-1 promoter." Nat Biotechnol **19**(7): 656-60.

Jamieson, A. C., S. H. Kim, et al. (1994). "In vitro selection of zinc fingers with altered DNA-binding specificity." Biochemistry **33**(19): 5689-95.

Jamieson, A. C., H. Wang, et al. (1996). "A zinc finger directory for high-affinity DNA recognition." Proc Natl Acad Sci U S A **93**(23): 12834-9.

Jantz, D., B. T. Amann, et al. (2004). "The design of functional DNA-binding proteins based on zinc finger domains." Chem Rev **104**(2): 789-99.

Jeggo, P. A. and M. Lobrich (2005). "Artemis links ATM to double strand break rejoining." Cell Cycle **4**(3): 359-62.

Johnson, E. S., P. C. Ma, et al. (1995). "A proteolytic pathway that recognizes ubiquitin as a degradation signal." J Biol Chem **270**(29): 17442-56.

Joung, J. K., E. I. Ramm, et al. (2000). "A bacterial two-hybrid selection system for studying protein-DNA and protein-protein interactions." Proc Natl Acad Sci U S A **97**(13): 7382-7.

Kandavelou, K., M. Mani, et al. (2005). ""Magic" scissors for genome surgery." Nat Biotechnol **23**(6): 686-7.

Khanna, K. K. and S. P. Jackson (2001). "DNA double-strand breaks: signaling, repair and the cancer connection." Nat Genet **27**(3): 247-54.

Kim, C. A. and J. M. Berg (1995). "Serine at position 2 in the DNA recognition helix of a Cys2-His2 zinc finger peptide is not, in general, responsible for base recognition." J Mol Biol **252**(1): 1-5.

Kim, Y. G., J. Cha, et al. (1996). "Hybrid restriction enzymes: zinc finger fusions to Fok I cleavage domain." Proc Natl Acad Sci U S A **93**(3): 1156-60.
Kim, Y. G. and S. Chandrasegaran (1994). "Chimeric restriction endonuclease." Proc Natl Acad Sci U S A **91**(3): 883-7.
Kim, Y. G., J. Smith, et al. (1998). "Chimeric restriction enzyme: Gal4 fusion to FokI cleavage domain." Biol Chem **379**(4-5): 489-95.
Kuan, J. Y. and P. M. Glazer (2004). "Targeted gene modification using triplex-forming oligonucleotides." Methods Mol Biol **262**: 173-94.
Lander, E. S., L. M. Linton, et al. (2001). "Initial sequencing and analysis of the human genome." Nature **409**(6822): 860-921.
Li, L., L. P. Wu, et al. (1992). "Functional domains in Fok I restriction endonuclease." Proc Natl Acad Sci U S A **89**(10): 4275-9.
Liu, P. Q., E. J. Rebar, et al. (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**(14): 11323-34.
Liu, Q., Z. Xia, et al. (2002). "Validated zinc finger protein designs for all 16 GNN DNA triplet targets." J Biol Chem **277**(6): 3850-6.
Lloyd, A., C. L. Plaisier, et al. (2005). "Targeted mutagenesis using zinc-finger nucleases in Arabidopsis." Proc Natl Acad Sci U S A **102**(6): 2232-7.
Lombardo, A., P. Genovese, et al. (2007). "Gene editing in human stem cells using zinc finger nucleases and integrase-defective lentiviral vector delivery." Nat Biotechnol **25**(11): 1298-306.
Maeder, M. L., S. Thibodeau-Beganny, et al. (2008). "Rapid "open-source" engineering of customized zinc-finger nucleases for highly efficient gene modification." Mol Cell **31**(2): 294-301.
Mandell, J. G. and C. F. Barbas, 3rd (2006). "Zinc Finger Tools: custom DNA-binding domains for transcription factors and nucleases." Nucleic Acids Res **34**(Web Server issue): W516-23.
Mani, M., J. Smith, et al. (2005). "Binding of two zinc finger nuclease monomers to two specific sites is required for effective double-strand DNA cleavage." Biochem Biophys Res Commun **334**(4): 1191-7.
Maynard-Smith, L. A., L. C. Chen, et al. (2007). "A directed approach for engineering conditional protein stability using biologically silent small molecules." J Biol Chem **282**(34): 24866-72.
Meng, X., M. B. Noyes, et al. (2008). "Targeted gene inactivation in zebrafish using engineered zinc-finger nucleases." Nat Biotechnol **26**(6): 695-701.
Miller, J. C., M. C. Holmes, et al. (2007). "An improved zinc-finger nuclease architecture for highly specific genome editing." Nat Biotechnol **25**(7): 778-85.

Moehle, E. A., J. M. Rock, et al. (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**(9): 3055-60.

Moore, J. K. and J. E. Haber (1996). "Cell cycle and genetic requirements of two pathways of nonhomologous end-joining repair of double-strand breaks in Saccharomyces cerevisiae." Mol Cell Biol **16**(5): 2164-73.

Morton, J., M. W. Davis, et al. (2006). "Induction and repair of zinc-finger nuclease-targeted double-strand breaks in Caenorhabditis elegans somatic cells." Proc Natl Acad Sci U S A **103**(44): 16370-5.

Naldini, L., U. Blomer, et al. (1996). "In vivo gene delivery and stable transduction of nondividing cells by a lentiviral vector." Science **272**(5259): 263-7.

Nardelli, J., T. J. Gibson, et al. (1991). "Base sequence discrimination by zinc-finger DNA-binding domains." Nature **349**(6305): 175-8.

Nightingale, S. J., R. P. Hollis, et al. (2006). "Transient gene expression by nonintegrating lentiviral vectors." Mol Ther **13**(6): 1121-32.

Notarangelo, L. D., S. Giliani, et al. (2000). "Of genes and phenotypes: the immunological and molecular spectrum of combined immune deficiency. Defects of the gamma(c)-JAK3 signaling pathway as a model." Immunol Rev **178**: 39-48.

Orr-Weaver, T. L., J. W. Szostak, et al. (1981). "Yeast transformation: a model system for the study of recombination." Proc Natl Acad Sci U S A **78**(10): 6354-8.

Ott, M. G., M. Schmidt, et al. (2006). "Correction of X-linked chronic granulomatous disease by gene therapy, augmented by insertional activation of MDS1-EVI1, PRDM16 or SETBP1." Nat Med **12**(4): 401-9.

Pavletich, N. P. and C. O. Pabo (1991). "Zinc finger-DNA recognition: crystal structure of a Zif268-DNA complex at 2.1 A." Science **252**(5007): 809-17.

Perez, E. E., J. Wang, et al. (2008). "Establishment of HIV-1 resistance in CD4+ T cells by genome editing using zinc-finger nucleases." Nat Biotechnol **26**(7): 808-16.

Philippe, S., C. Sarkis, et al. (2006). "Lentiviral vectors with a defective integrase allow efficient and sustained transgene expression in vitro and in vivo." Proc Natl Acad Sci U S A **103**(47): 17684-9.

Porteus, M. H. (2006). "Mammalian gene targeting with designed zinc finger nucleases." Mol Ther **13**(2): 438-46.

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

Porteus, M. H. and D. Carroll (2005). "Gene targeting using zinc finger nucleases." Nat Biotechnol **23**(8): 967-73.

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

Pruett-Miller, S. M., J. P. Connelly, et al. (2008). "Comparison of zinc finger nucleases for use in gene targeting in mammalian cells." Mol Ther **16**(4): 707-17.

Ramirez, C. L., J. E. Foley, et al. (2008). "Unexpected failure rates for modular assembly of engineered zinc fingers." Nat Methods **5**(5): 374-5.

Rebar, E. J. and C. O. Pabo (1994). "Zinc finger phage: affinity selection of fingers with new DNA-binding specificities." Science **263**(5147): 671-3.

Ren, D., T. N. Collingwood, et al. (2002). "PPARgamma knockdown by engineered transcription factors: exogenous PPARgamma2 but not PPARgamma1 reactivates adipogenesis." Genes Dev **16**(1): 27-32.

Rouet, P., F. Smih, et al. (1994). "Expression of a site-specific endonuclease stimulates homologous recombination in mammalian cells." Proc Natl Acad Sci U S A **91**(13): 6064-8.

Santiago, Y., E. Chan, et al. (2008). "Targeted gene knockout in mammalian cells by using engineered zinc-finger nucleases." Proc Natl Acad Sci U S A **105**(15): 5809-14.

Sargent, R. G., M. A. Brenneman, et al. (1997). "Repair of site-specific double-strand breaks in a mammalian chromosome by homologous and illegitimate recombination." Mol Cell Biol **17**(1): 267-77.

Schultz, L. B., N. H. Chehab, et al. (2000). "p53 binding protein 1 (53BP1) is an early participant in the cellular response to DNA double-strand breaks." J Cell Biol **151**(7): 1381-90.

Scott, C. T. (2005). "The zinc finger nuclease monopoly." Nat Biotechnol **23**(8): 915-8.

Sedivy, J. M. and P. A. Sharp (1989). "Positive genetic selection for gene disruption in mammalian cells by homologous recombination." Proc Natl Acad Sci U S A **86**(1): 227-31.

Segal, D. J., R. R. Beerli, et al. (2003). "Evaluation of a modular strategy for the construction of novel polydactyl zinc finger DNA-binding proteins." Biochemistry **42**(7): 2137-48.

Segal, D. J., B. Dreier, et al. (1999). "Toward controlling gene expression at will: selection and design of zinc finger domains recognizing each of the 5'-GNN-3' DNA target sequences." Proc Natl Acad Sci U S A **96**(6): 2758-63.

Shi, Y. and J. M. Berg (1995). "A direct comparison of the properties of natural and designed zinc-finger proteins." Chem Biol **2**(2): 83-9.

Smih, F., P. Rouet, et al. (1995). "Double-strand breaks at the target locus stimulate gene targeting in embryonic stem cells." Nucleic Acids Res **23**(24): 5012-9.

Smith, J., M. Bibikova, et al. (2000). "Requirements for double-strand cleavage by chimeric restriction enzymes with zinc finger DNA-recognition domains." Nucleic Acids Res **28**(17): 3361-9.

Sonoda, E., H. Hochegger, et al. (2006). "Differential usage of non-homologous end-joining and homologous recombination in double strand break repair." DNA Repair (Amst) **5**(9-10): 1021-9.

Stoddard, B. L. (2005). "Homing endonuclease structure and function." Q Rev Biophys **38**(1): 49-95.

Szczepek, M., V. Brondani, et al. (2007). "Structure-based redesign of the dimerization interface reduces the toxicity of zinc-finger nucleases." Nat Biotechnol **25**(7): 786-93.

Taghian, D. G. and J. A. Nickoloff (1997). "Chromosomal double-strand breaks induce gene conversion at high frequency in mammalian cells." Mol Cell Biol **17**(11): 6386-93.

Thibodeau, S. A., R. Fang, et al. (2004). "High-throughput beta-galactosidase assay for bacterial cell-based reporter systems." Biotechniques **36**(3): 410-5.

Thukral, S. K., M. L. Morrison, et al. (1992). "Mutations in the zinc fingers of ADR1 that change the specificity of DNA binding and transactivation." Mol Cell Biol **12**(6): 2784-92.

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

Vargas, J., Jr., G. L. Gusella, et al. (2004). "Novel integrase-defective lentiviral episomal vectors for gene transfer." Hum Gene Ther **15**(4): 361-72.

Varshavsky, A. (1996). "The N-end rule: functions, mysteries, uses." Proc Natl Acad Sci U S A **93**(22): 12142-9.

Varshavsky, A. (1997). "The N-end rule pathway of protein degradation." Genes Cells **2**(1): 13-28.

Venter, J. C., M. D. Adams, et al. (2001). "The sequence of the human genome." Science **291**(5507): 1304-51.

Wah, D. A., J. Bitinaite, et al. (1998). "Structure of FokI has implications for DNA cleavage." Proc Natl Acad Sci U S A **95**(18): 10564-9.

Watanabe, S., K. Terashima, et al. (2007). "Hematopoietic stem cell-engrafted NOD/SCID/IL2Rgamma null mice develop human lymphoid systems and induce long-lasting HIV-1 infection with specific humoral immune responses." Blood **109**(1): 212-8.

West, S. C., C. Chappell, et al. (2000). "Double-strand break repair in human cells." Cold Spring Harb Symp Quant Biol **65**: 315-21.

Wolfe, S. A., H. A. Greisman, et al. (1999). "Analysis of zinc fingers optimized via phage display: evaluating the utility of a recognition code." J Mol Biol **285**(5): 1917-34.

Wolfe, S. A., L. Nekludova, et al. (2000). "DNA recognition by Cys2His2 zinc finger proteins." Annu Rev Biophys Biomol Struct **29**: 183-212.

Wright, D., Thibodeau-Beganny, S., Sander, J.D., Winfrey, R.J., Hirsh, A.J., Eichtinger, M., Fu, F., Porteus, M.H., Dobbs, D., Voytas, D.F., and Joung, J.K. (2006). "Standardized reagents and protocols for engineering zinc finger nucleases by modular assembly." Nature Protocols **1**: 1637-1652.

Wright, D. A., J. A. Townsend, et al. (2005). "High-frequency homologous recombination in plants mediated by zinc-finger nucleases." Plant J **44**(4): 693-705.

Wu, H., W. P. Yang, et al. (1995). "Building zinc fingers by selection: toward a therapeutic application." Proc Natl Acad Sci U S A **92**(2): 344-8.

Wurtz, N. R. and P. B. Dervan (2000). "Sequence specific alkylation of DNA by hairpin pyrrole-imidazole polyamide conjugates." Chem Biol **7**(3): 153-61.

Wyman, C. and R. Kanaar (2004). "Homologous recombination: down to the wire." Curr Biol **14**(15): R629-31.

Wyman, C. and R. Kanaar (2006). "DNA double-strand break repair: all's well that ends well." Annu Rev Genet **40**: 363-83.

Wyman, C., D. Ristic, et al. (2004). "Homologous recombination-mediated double-strand break repair." DNA Repair (Amst) **3**(8-9): 827-33.

Yamane-Ohnuki, N., S. Kinoshita, et al. (2004). "Establishment of FUT8 knockout Chinese hamster ovary cells: an ideal host cell line for producing completely defucosylated antibodies with enhanced antibody-dependent cellular cytotoxicity." Biotechnol Bioeng **87**(5): 614-22.

Yamazoe, M., E. Sonoda, et al. (2004). "Reverse genetic studies of the DNA damage response in the chicken B lymphocyte line DT40." DNA Repair (Amst) **3**(8-9): 1175-85.

Yanez-Munoz, R. J., K. S. Balaggan, et al. (2006). "Effective gene therapy with nonintegrating lentiviral vectors." Nat Med **12**(3): 348-53.