

**INVESTIGATION OF PREFERENTIAL CELLULAR REPAIR  
MECHANISMS IN K562 CELLS AS A RESULT OF TARGETED  
GENOME ENGINEERING**

A Thesis  
Presented to  
The Academic Faculty

by

Caleb Appleton

In Partial Fulfillment  
of the Requirements for the Degree  
Bachelors of Science in the  
School of Biomedical Engineering

Georgia Institute of Technology  
May 2015

**COPYRIGHT 2015 BY CALEB APPLETON**

**INVESTIGATION OF PREFERENTIAL CELLULAR REPAIR  
MECHANISMS IN K562 CELLS AS A RESULT OF TARGETED  
GENOME ENGINEERING**

Approved by:

Dr. Gang Bao, Advisor  
School of Biomedical Engineering  
*Georgia Institute of Technology*

Dr. Julia Kubanek  
School of Biology  
*Georgia Institute of Technology*

Date Approved: [Date Approved by Committee]



[To the students of the Georgia Institute of Technology]

## **ACKNOWLEDGEMENTS**

I would like to thank all of the members of the Bao Laboratory of Biomolecular Engineering and Nanomedicine. First, I would like to thank Dr. Bao for allowing me the opportunity to pursue my research goals and interests. I would like to thank Brian Wile for his help and guidance on this project. Finally, I would like to thank Eli Fine for his constant mentorship and patience: helping me work through setbacks and tolerating my silly mistakes.

# TABLE OF CONTENTS

	Page
ACKNOWLEDGEMENTS	iv
LIST OF FIGURES	vi
LIST OF SYMBOLS AND ABBREVIATIONS	vii
SUMMARY	vii
<u>CHAPTER</u>	
1 Introduction	
Background	1
Literature Review	2
2 Materials and Methods	
Nuclease Construction and Design	7
Donor Design	7
Nucleofection	7
Cell Sorting	8
DNA Extraction	8
Restriction Fragment Length Polymorphism	8
Illumina Sequencing	8
3 Findings	10
Results	10
Discussion	12
REFERENCES	14

## **LIST OF FIGURES**

	Page
Figure 1.1: Method of NHEJ vs HDR analysis using SMRT sequencing	6
Figure 2.1: Comparison of HDR and NHEJ rates among TALENS and Zinc Fingers	11
Figure 2.2: HDR and NHEJ Analysis of CRISPR nickases	11

## **LIST OF SYMBOLS AND ABBREVIATIONS**

NHEJ	Non-Homologous End Joining
HDR	Homology Directed Repair
ZFN	Zinc Finger Nuclease
TALEN	Transcription Activator Like Effector Nuclease
CRISPR	Clustered Regularly Interspaced Palindromic Repeats
Bp	Base Pairs
DSB	Double Strand Break
ssDSB	Site Specific Double Strand Break



## **SUMMARY**

**The performed research examines the response of K562 cells to targeted introduction of double strand DNA breaks. The research provides a comparison of repair via non-homologous end joining and homology directed repair at an endogenous loci. For the experiments, a number of DNA modification techniques were utilized: transcription activator like effector nucleases, zinc finger nucleases and clustered regularly interspaced palindromic repeats. Findings showed that when given a donor template for repair, all gene modification methods showed repair by both HDR and NHEJ, however the rates were not always robust. Highest rates of gene modification were observed in paired CRISPR nickases. The study highlights the need for better understanding of gene repair strategies at endogenous loci.**

# CHAPTER 1

## INTRODUCTION

The idea of genetic engineering has been something long contemplated by both researchers and science fiction writers as a way to optimize the human body. While the idea of creating a “superhuman” is unrealistic, drastic advances in genome editing have been made in the field of genetic engineering and gene therapy over the past decade. These advances have brought genome engineering to the forefront of cellular therapy as the ability to introduce slight modifications to the genome provides for immense therapeutic and scientific possibilities. These advances include the discovery and continual optimization of Zinc Finger Nucleases (ZFNs), Transcription Activator Like Effector Nucleases (TALENs), and Clustered Regularly Interspaced Palindromic Repeats (CRISPRs). These nucleases (ZFNs and TALENs) and paired nickases (CRISPRs) can be delivered *in vitro* as a way of creating targeted double strand breaks in the human genome as well as supplying a donor piece of DNA with a corrected sequence for the cell to use in its natural repair pathways.

Much of the current genetic engineering research via the designer nucleases and nickases pathway has focused on the knockdown or correction of mutated genes in which targeted breaks in the genome are made and a template is provided for their repair<sup>1</sup>. In turn, large amounts of research funds and hours have been devoted to the study of correcting point mutations in the genome that cause prevalent systemic diseases like cystic fibrosis and sickle cell anemia. However, the technique of correcting mutated genes also offers the unique opportunity to explore various cellular pathways. By

knocking down or introducing mutations that inactivate various genes or promoters *in vitro*, the role of those specific targets can be studied.

In order for further research into gene editing and genome engineering to be successful, it is first necessary to better understand the mechanisms by which the cell repairs itself in response to directed DNA cleavage. The described research seeks to elucidate the rates by which cells undergo different repair mechanisms at an endogenous loci. A better understanding of these rates will allow for the optimization of various genome engineering approaches in order to preferentially upregulate specific repair mechanisms.

### **Literature Review:**

One of the most common single gene disorders in the world, Sickle-cell anemia, is a blood disorder that results from a point mutation in the human  $\beta$ -globin gene that results in an amino acid shift from glutamic acid to valine <sup>2</sup>. As such, those with the disease have red blood cells that form in a sickle shape instead of the customary disk in unaffected individuals. This deformation leads to an increased rate of apoptosis and decreased cell life in sickle-cell individuals <sup>2</sup>. Additionally, the disease drastically reduces life expectancy with an average life span of just forty-eight years for females and forty-two for males and also leads to a reduction in overall quality of life <sup>3</sup>. Currently there exists no cure for the disease; rather, current therapies seek to manage symptoms and improve quality of life rather than treating or curing the disease itself.

As such, many research hours have been invested in not just mitigating symptoms but rather providing a functional cure enabling patients to experience a standard quality of life. Since sickle-cell results from an individual nucleotide mutation in a single gene, it

provides a favorable template for genetic therapy approaches. Specifically, targeted engineered genome editing can be used to correct the mutated genetic sequence and in turn provide a pathway for a functional cure. Efficient and effective genome engineering and editing has long been a goal of scientists as it presents the opportunity for large advances in both therapeutic and research fields. Genome engineering provides the opportunity for customized therapy to repair genomic diseases as well as improved methods for inactivating or upregulating certain genes for more streamlined study. Genome engineering via engineered endonucleases became a real possibility with the discovery of zinc finger nucleases (ZFNs) in 1985; however it wasn't until 1984 that the first engineered ZFN was found to selectively bind to a native gene locus *in vivo*.<sup>4, 5</sup>

Since that important discovery, many advances have been made in the field of engineered nucleases. Currently there is a stage II clinical trial underway that utilizes ZFNs to inactivate the CCR5 gene: the CCR5 gene codes for a protein on the surface of white blood cells (t-cells) that is crucial in the mechanism by which HIV infects the cells<sup>6</sup>. While significant advances have been made in the construction and employment of ZFNs, it remains extremely difficult and expensive to design effective ZFNs<sup>1</sup>. Thus other methods for modifying the genome via nuclease activity have been sought, ultimately resulting in the discovery of transcription like effector nucleases (TALENs) and clustered regularly interspaced palindromic repeats (CRISPRs ) in 2009 and 2012 respectively<sup>7, 8</sup>.

All three platforms work in a similar fashion by combining a DNA cleavage domain with a complex that recognizes a specific sequence in the genome. For this purpose, the FokI binding domain is used in ZFN and TALEN construction whereas Cas9 is used for CRISPR design. Primarily, these nucleases are used to create site-specific

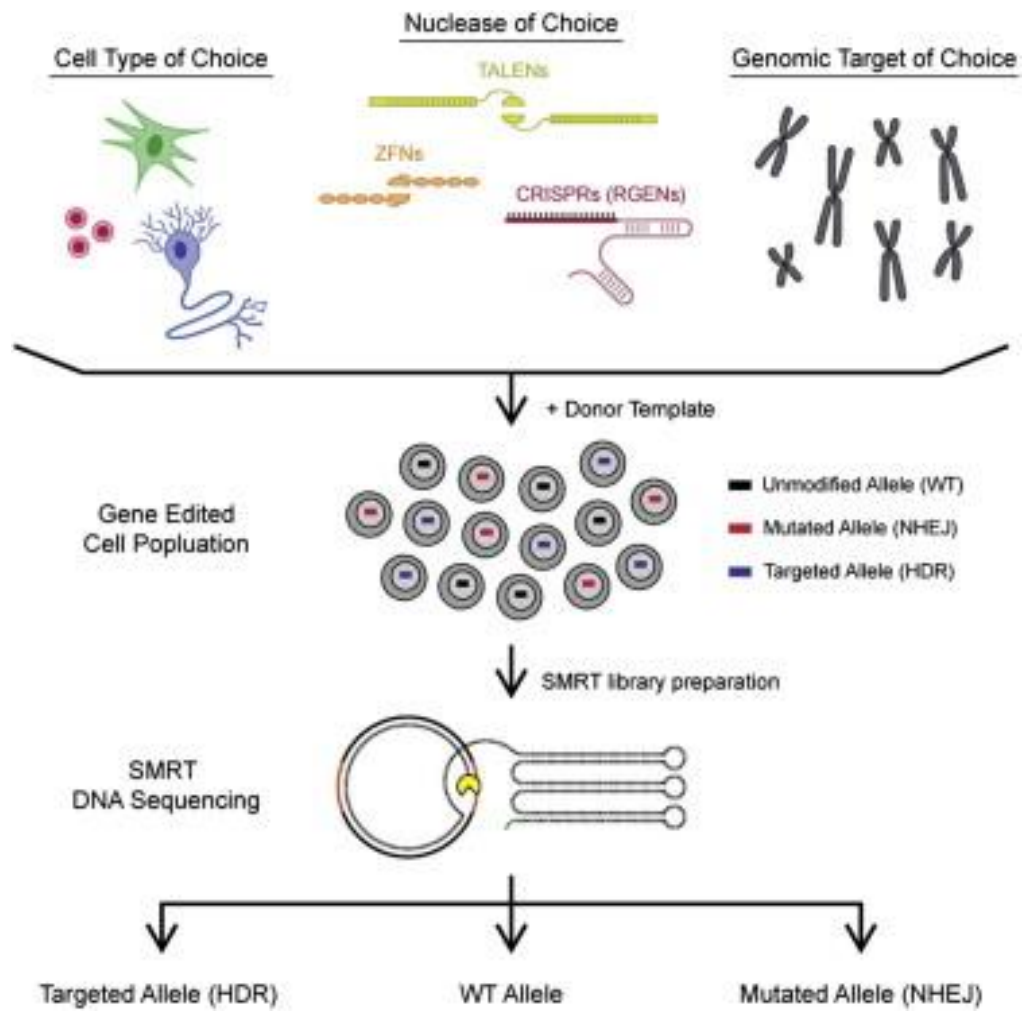
double strand breaks (DSBs) at targets of interest in the genome. Once a break is created, the cell repairs it via one of two repair pathways: non-homologous end joining (NHEJ) or homology directed repair (HDR). Both methods have their merits and drawbacks. NHEJ is an error prone repair mechanism that occurs at a much higher rate than HDR and introduces small insertions and deletions that can be useful for gene disruption and knockout. In NHEJ the overhangs resulting from the DSBs are ligated together creating insertions and deletions in the genome that often create a frameshift or premature stop codon<sup>9</sup>. While it is the predominant cellular repair mechanism, NHEJ is not ideal for therapeutic approaches that require a corrected DNA sequence to be obtained rather than simply a gene knockout or knockdown. Whereas with HDR, once the DSB occurs, the cell utilizes a homologous region of DNA to guide repair. In genome engineering, a donor template with arms of homology flanking the mutation site and corrections introduced at the point of mutation is supplied alongside the engineered endonuclease. Thus, for the correction of sickle-cell anemia it is paramount that cells undergo HDR not NHEJ to drive gene repair. The ongoing research seeks to elucidate the relative rates of NHEJ vs. HDR and employ techniques to increase the percentage of cells undergoing repair by HDR.

As the two native pathways have such drastically different outcomes, it is necessary to quantify the rates of each repair pathway when evaluating a gene therapy agent. There exist a number of ways to measure either NHEJ or HDR however it is more difficult to measure the two repair pathways in tandem at an endogenous loci. A rather simplistic gel-based assay utilizes the Surveyor nuclease or T7-EndonucleaseI to measure NHEJ repair frequency<sup>10</sup>. Known as the Surveyor nuclease assay, this approach utilizes

the enzymes ability to selectively cleave mismatched DNA duplexes; while the assay confirms nuclease activity at any endogenous locus it does not provide any information on the rates of HDR repair and thus is not optimal for repair techniques that necessitate a donor sequence<sup>10</sup>. Porteus et al. developed a method to measure HDR rates *in vitro* through the introduction of a fluorescent reporter in the donor strand<sup>11</sup>. This method requires the longitudinal measuring of fluorescence expression in the cells via fluorescence activated cell sorting (FACS) and only provides relative not exact rates of HDR<sup>11</sup>. One study utilized a reporter called the Traffic Light Reporter that can measure NHEJ and HDR rates using flow cytometry<sup>12</sup>. This method has been used to confirm that ZFNicks (ZFNs designed to create single strand breaks instead of DSBs) elevate HDR rates in comparison to NHEJ<sup>13</sup>. Additional work by Mali et al. showed comparative rates of NHEJ and HDR in cells transfected with CRISPR nickases using the TLR<sup>14</sup>. While this is beneficial in comparing global rates of repair, the TLR is unable to provide information about the rates at a specific gene locus and thus has not been used in human primary cells<sup>12</sup>.

To date, only one paper has successfully measured both rates of repair at an endogenous gene loci; however it relies on a cost-prohibitive and analytically intensive method known as single molecule real time (SMRT). sequencing (**Figure 1**)<sup>15</sup>. The described study presents the of the rates of HDR and NHEJ at the human  $\beta$ -globin gene by combining two wet-lab techniques while utilizing a more cost-effective method of sequencing known as Illumina for validation.

**Figure 1:** Method of NHEJ vs HDR analysis using SMRT sequencing (Hendel 2014)



## **CHAPTER 2**

### **MATERIALS AND METHODS**

#### **2.1 Nuclease Construction and design**

Engineered endonucleases and nickases were designed to flank the mutated gene with varying spacer lengths. TALEN construction was performed using the Golden Gate reaction<sup>16</sup>. This allows for quick and efficient construction of a large number of TALENs. CRISPRs and CRISPRnickases were constructed using the Gibson Assembly method.

#### **2.2 Donor Design**

Donors were designed to have 400 bp arms of homology to the genome. In addition to correcting the sickle-cell mutation, also included in the donors was a restriction enzyme cut site so that an RFLP assay could be performed post nucleofection. An initial donor named, 'EcoRI Donor' was utilized initially that allowed for post treatment HDR quantification via RFLP utilizing the EcoRI as the donor incorporated the EcoRI restriction enzyme site. However, later tests utilized a "M4" donor provided by collaborators that allowed for post treatment assay using a StuI restriction site.

#### **2.3 Nucleofection**

Nucleofections were performed in K562 cells with the 4D Nucleofector-X unit (Lonza). Cells were nucleofected with 1 ug of total DNA in 2 uL: 800 ng of donor template (here with 400 bp arms of homology), 200 ng of nuclease/nickase and 50ng of GFP.

Nucleofections were performed using pulse sequence FF-120 with 200,000 cells per 22 uL SF-Cell Line solution (Lonza) and reaction mix. Cells were cultured in a 24 well plate with 500 uL of RPMI-1640 (Life Technologies) +10% fetal bovine serum (Gibco). Media was changed at 48 hours to prevent cells from becoming too confluent.



## **2.4 Cell Sorting**

Fluorescence activated cell sorting was performed on all samples 72 hours post nucleofection using the C6 Flow Cytometer (BD Accuri). GFP fluorescence was measured using a 488-nm laser. Gates were created to determine cell viability as well as transfection efficiency.

## **2.5 DNA Extraction**

DNA was extracted from cell samples using QuickExtract DNA Extraction Solution (Epicentre). 80 uL of QuickExtract was used per sample.

## **2.6 Restriction Fragment Length Polymorphism Assay**

To determine initial HDR rates RFLP assay is performed. First samples are subjected to PCR with primes designed just outside the arms of homology. PCRs were completed using Accuprime polymerase (Invitrogen). Primers that encompassed the region of interest were used at a 0.2 uM concentration. PCR amplification utilized a 5 minute denaturing step at 95°C followed by 40 cycles of 95°C for 30 seconds, 67°C for 45 s and 68°C for 120s and a final extension of 68°C for five minutes. PCR products were confirmed by running on a 2% agarose gel and then cleaned-up using the Promega Wizard PCR Clean-up kit (Promega Wizard). Cleaned-up products were then digested using 20 units of StuI at 37°C for one hour and then run on a gel next to uncut controls to measure comparative rates. A T7EI assay was performed as detailed above to determine rates of NHEJ. RFLP and T7EI assays were used as cursory measures of HDR and NHEJ only and not for quantitative purposes in this study.

## **Illumina Sequencing:**

Illumina sequencing was also performed to measure the rates of NHEJ vs HDR and confirm accuracy of proposed method. Analysis of Illumina samples was performed by Eli Fine in accordance with a modified method from that proposed in the Haydal paper<sup>15</sup>.

## CHAPTER 3

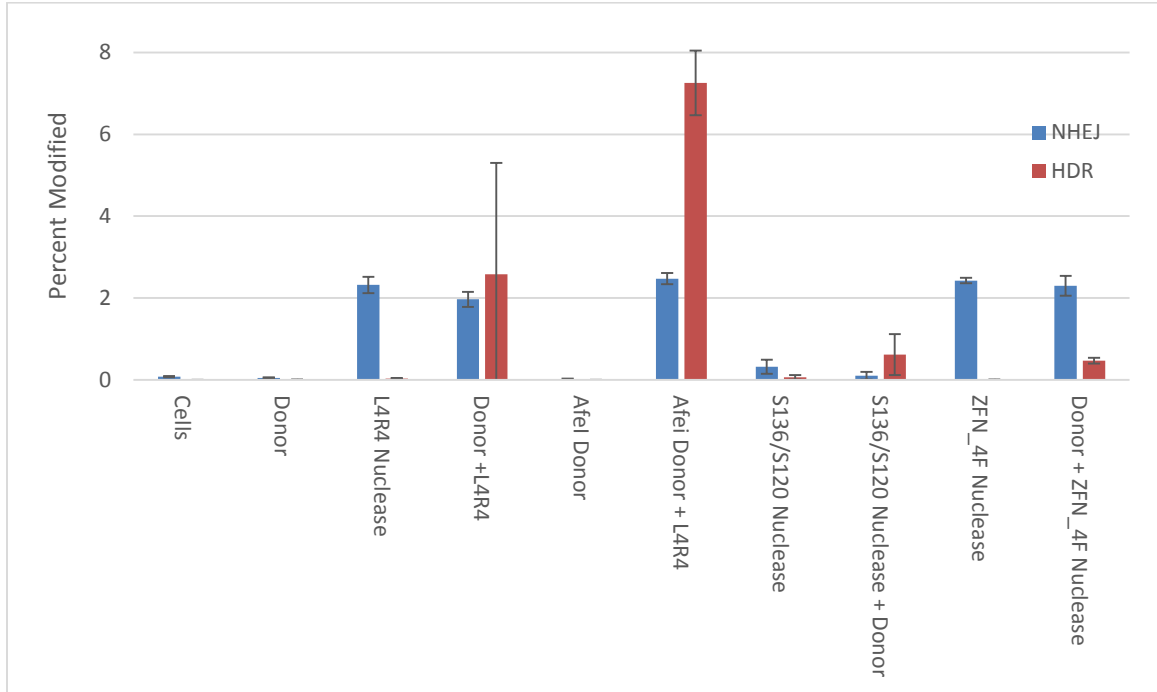
### RESULTS AND DISCUSSION

#### **Results:**

In an attempt to determine baseline HDR and NHEJ rates *in vitro*, DNA harvested from K562 72 hours post nucleofection was analyzed via Illumina sequencing. Several nucleases as well as donor combinations were studied (**Figure 3.1**). As expected, minimal HDR was observed in any of the samples containing no nuclease and NHEJ rates fell between 0-1% in all cases not including donor. Further, approximately equivalent rates of NHEJ were seen in treatment groups that received a nuclease pair but not a donor strand of DNA. Repair via HDR was witnessed in all samples with a nuclease and a donor. The most robust rates were observed using the AfeI donor provided by our collaborator and the L4R4 TALEN pair at 7.25%. Substantial variation was witnessed in a number of the samples as can be seen by the large standard deviations.

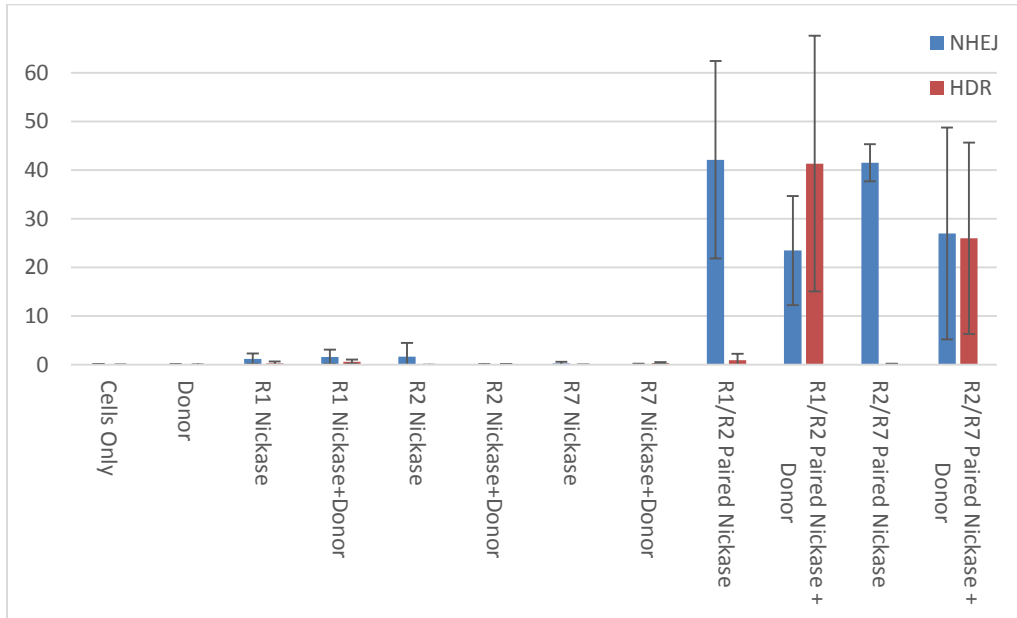
A comparison of CRISPR nickases and paired CRISPR nickases was also performed through Illumina analysis of DNA extracted from K562 cells 72 hours post treatment (**Figure 3.2**). Four different nickases were compared with two groups of paired nickases. Once again minimal to no HDR was observed in samples that were not nucleofected with a DNA donor. Additionally, NHEJ rates were consistent among all control samples. Robust activity was observed in R1/R2 paired nickases with the M4 donor with 41.34% of cells experiencing repair via the HDR pathway. Additionally, the R2/R7 saw substantial repair via HDR at 26%. Of additional interest is the fact that NHEJ rates dropped drastically after the introduction of a donor strand. However, once again large variation between samples inside treatment groups was seen.

**Figure 3.1: Comparison of HDR and NHEJ rates among TALENS and Zinc Fingers**



Two TALEN pairs and a Zinc Finger pair were analyzed for NHEJ and HDR rates via Illumina sequencing. All samples use the M4 donor unless otherwise noted. All error bars represent standard deviation (n=3 for all treatment groups).

**Figure 3.2: HDR and NHEJ Analysis of CRISPR nickases**



Four CRISPR nickases and two paired nickases were analyzed for HDR and NHEJ rates via Illumina sequencing. All treatment groups with a donor used the M4 donor. Error bars represent standard deviation (n=3 for all treatment groups).

**Discussion:**

A comparison study was completed to analyze the two main mechanisms by which a cell can repair post DNA cleavage at an endogenous loci *in vitro*. A number of gene editing strategies were investigated: TALENs, CRISPR nickases, paired CRISPR nickases and ZNFs. In K562 cells, it was shown that relatively low rates of total gene repair were exhibited when treated with either TALENs or CRISPR nickases. This is in contrast to results previously published by the Bao lab which saw high modification rates by the same nucleases in 293T cells. This finding suggests that more work may be necessary in designing the nucleases to better target the site of interest. Further, more work should be done across cell lines as traditionally rates in immortalized lines such as 293T and K562 cells have exhibited much more robust repair rates than human primary cell lines.

An analysis of CRISPR nickases and paired CRISPR nickases was also performed in K562. Similar to earlier published results, a small but inconsistent rate of cellular repair was observed when DNA damage was introduced by single CRISPR nickases<sup>17</sup>. Further, CRISPR nickase pairs showed robust triggering of DNA repair mechanisms both with and without a donor template. However, once a donor template was introduced, a substantial increase in repair via the HDR pathway was observed suggesting that CRISPR nickases are an effective vehicle for gene modification.

In both studies, the error within treatment groups was substantial, particularly at high rates of genome modification. Future work should look at the optimization of nucleofection and cell culture techniques to reduce variation between samples treated in

parallel. Additionally, further work should be done to explore the rates of NHEJ and HDR in a broader range of nucleases and nickases. This would, in-turn, allow for better optimization of nuclease and nickase design for preferential repair via one of the two mechanisms. The approach in this study only looked at repair at one endogenous loci: the human beta globin gene. Of interest is whether the nucleases and nickases perform in a similar fashion at other genes of interest.

## REFERENCES

1. Segal, D.J. & Meckler, J.F. Genome Engineering at the Dawn of the Golden Age. *Annual review of genomics and human genetics* (2013).
2. Rees, D.C., Williams, T.N. & Gladwin, M.T. Sickle-cell disease. *The Lancet* **376**, 2018-2031 (2010).
3. Platt, O.S. et al. Mortality In Sickle Cell Disease -- Life Expectancy and Risk Factors for Early Death. *New England Journal of Medicine* **330**, 1639-1644 (1994).
4. Choo, Y., Sánchez-García, I. & Klug, A. In vivo repression by a site-specific DNA-binding protein designed against an oncogenic sequence. *Nature* **372**, 642-645 (1994).
5. Gaj, T., Gersbach, C.A. & Barbas III, C.F. ZFN, TALEN, and CRISPR/Cas-based methods for genome engineering. *Trends in biotechnology* (2013).
6. Perez, E.E. et al. Establishment of HIV-1 resistance in CD4+ T cells by genome editing using zinc-finger nucleases. *Nature biotechnology* **26**, 808-816 (2008).
7. Boch, J. et al. Breaking the code of DNA binding specificity of TAL-type III effectors. *Science* **326**, 1509-1512 (2009).
8. Jinek, M. et al. A programmable dual-RNA-guided DNA endonuclease in adaptive bacterial immunity. *Science* **337**, 816-821 (2012).
9. Lieber, M.R., Ma, Y., Pannicke, U. & Schwarz, K. Mechanism and regulation of human non-homologous DNA end-joining. *Nature reviews Molecular cell biology* **4**, 712-720 (2003).
10. Guschin, D.Y. et al. in Engineered zinc finger proteins 247-256 (Springer, 2010).
11. Porteus, M.H. & Baltimore, D. Chimeric nucleases stimulate gene targeting in human cells. *Science* **300**, 763-763 (2003).
12. Certo, M.T. et al. Tracking genome engineering outcome at individual DNA breakpoints. *Nature methods* **8**, 671-676 (2011).
13. Ramirez, C.L. et al. Engineered zinc finger nickases induce homology-directed repair with reduced mutagenic effects. *Nucleic acids research* **40**, 5560-5568 (2012).
14. Mali, P. et al. CAS9 transcriptional activators for target specificity screening and paired nickases for cooperative genome engineering. *Nature biotechnology* **31**, 833-838 (2013).
15. Hendel, A. et al. Quantifying Genome-Editing Outcomes at Endogenous Loci with SMRT Sequencing. *Cell reports* **7**, 293-305 (2014).
16. Cermak, T. et al. Efficient design and assembly of custom TALEN and other TAL effector-based constructs for DNA targeting. *Nucleic acids research* **39**, e82-e82 (2011).
17. Tsai, S.Q. et al. Dimeric CRISPR RNA-guided FokI nucleases for highly specific genome editing. *Nature biotechnology* **32**, 569-576 (2014).