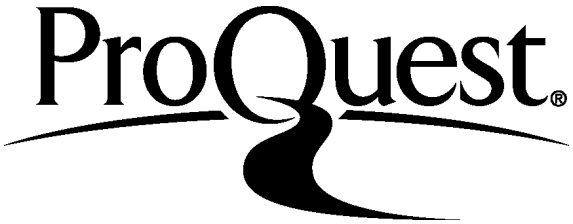
APPLICATION OF CHINESE HAMSTER OVARY CELL LINE ENGINEERING IN THE STUDY OF LIPOPROTEIN LIPASE AND ITS IMPACT ON POLYSORBATE DEGRADATION

by

Josephine Chiu

A thesis submitted to the Faculty of the University of Delaware in partial fulfillment of the requirements for the degree of Master in Chemical Engineering

Spring 2016



APPLICATION OF CHINESE HAMSTER OVARY CELL LINE ENGINEERING IN THE STUDY OF LIPOPROTEIN LIPASE AND ITS IMPACT ON POLYSORBATE DEGRADATION

by

Josephine Chiu

Approved: _____
Kelvin H. Lee, Ph.D.
Professor in charge of thesis on behalf of the Advisory Committee

Approved: _____
Abraham M. Lenhoff, Ph.D.
Chair of the Department of Chemical and Biomolecular Engineering

Approved: _____
Babatunde A. Ogunnaike, Ph.D.
Dean of the College of Engineering

Approved: _____
Ann L. Ardis, Ph.D.
Senior Vice Provost for Graduate and Professional Education

ACKNOWLEDGMENTS

I would like to acknowledge all the people who have supported me throughout my graduate studies. First, I would like to thank Kelvin Lee. His guidance and advice have been instrumental in the development of this project and in me as a scientist. I would also like to thank my committee members, Bramie Lenhoff and April Kloxin, for their support in my research and for providing invaluable advice.

I would also like to acknowledge the many members of the Lee group, who have been very encouraging and supportive the day I stepped into UD. First, I would like to thank Leila Choe for all her help in keeping my research organized and running without fail. I would also like to thank Jongyoun Baik, Lie Min, Xiaolin Zhang, Kristin Valente, Ben Kremkow, Amalie Levy, Jennifer Mantle, Madolyn Stinner, John Ruano-Salguero, Sansheng Chen, Michael Gallucci. Kristin taught me how to culture CHO cells in suspension culture and helped me in the lab when I first joined the group as a rotational student. Lie developed and ran the LPL-specific MRM assay, which had been an essential part of my project. Jongyoun, a treasure trove of CHO knowledge, taught me the importance of CHO cells and the basics of molecular biology. Ben, Amalie, and Jen often provided the much-appreciated constructive recommendations that improved the quality of my work. The dedication the Lee group shows to their research and their willingness to drop what they are doing to perfect my presentations and papers, help me troubleshoot, or just listen have been an important part of my development as a scientist.

I am also thankful to be part of the Chemistry-Biology Interface (CBI) program at Delaware, which allowed me to develop my skills as a scientist. This program allowed me to expand my knowledge beyond chemical engineering.

I would like to thank my friends from Delaware to New York for their support and encouragement. Jackie Gonzalez and Chen Yuan Kao have been great DBI comrades and shared with me the joys of commuting from DBI to Colburn. I could always count on Hyeri Han, Lawrence Ng, Gary He, Connie Chan, Susan Ng, and Yangzi Tian to be a great friend and to show me where the good food in New York is. I would like to thank Marco Dunwell, my sandwich, for his love and support. His silly antics and encouragements made my life in graduate school more enjoyable. I am blessed to have him in my life.

My greatest support has always come from my family. I would like to thank my parents, Philip and Carol, who were an endless source of love and encouragement and walked with me every step of the way. Their support has given me the strength to stride confidently towards my goals. My sisters, Eva and Emily, were also a constant source of love and support. They were with me through all my fears and frustrations of graduate school. I am blessed with the love of my family, for without them, I would not have made it this far.

TABLE OF CONTENTS

LIST OF TABLES ix
LIST OF FIGURES x
ABSTRACT xi

Chapter

1 INTRODUCTION 1

1.1 Chinese Hamster Ovary Cells in Biotechnology 1
1.2 CHO Cell Line Engineering and the Emergence of Genome-editing Tools 2
1.3 Problem in CHO Biotechnology: Formulation instability 3
1.4 Project Goals 4
1.5 Scope of work 5

REFERENCES 6

2 GENOME-EDITING TOOLS IN MAMMALIAN CELLS 10

2.1 Introduction 10
2.1 Genome-editing with ZFN, TALEN, and CRISPR/Cas9 12

2.1.1 ZFN and TALEN 13
2.1.2 CRISPR/Cas9 16

2.2 CRISPR/Cas9 in Chinese Hamster Ovary (CHO) cells 19

REFERENCES 21

3 LPL KNOCKDOWN IN CHINESE HAMSTER OVARY CELLS BY SIRNA INTERFERENCE 26

3.1 Preface 26

3.2 Introduction 27
3.3 Materials and Methods 30

3.3.1 CHO cell culture 30
3.3.2 siRNA transfection 31
3.3.3 CHO HCP preparation 31
3.3.4 LPL-specific MRM assay 32
3.3.5 PS80 degradation assay 34

3.4 Results and Discussion 34
3.5 Conclusions 40

REFERENCES 41

4 OPTIMIZATION OF POLYSORBATE DEGRADATION ASSAY BY CHINESE HAMSTER OVARY HOST CELL PROTEIN 45

4.1 Introduction 45
4.2 Materials and Methods 46

4.2.1 CHO Cell Culture 46
4.2.2 Experimental Design and Data Analysis 46
4.2.3 PS80 Degradation Assay 47

4.3 Results and Discussion 48

4.3.1 Screening and Optimization DOEs 48
4.3.2 Analysis of reproducibility 51

4.4 Concluding Remarks 55

REFERENCES 56

5 A FUNCTIONAL LPL KNOCKOUT IN CHINESE HAMSTER OVARY CELLS REDUCES POLYSORBATE DEGRADATION 57

5.1 Introduction 57
5.2 Materials and Methods 58

5.2.1 Single guide RNA (sgRNA) target design and plasmid construction 58
5.2.2 CHO cell culture and transfection 60
5.2.3 T7 Endonuclease I (T7EI) assay for targeting efficiency of CRISPR/Cas9 61
5.2.4 Gene sequencing by TOPO cloning 61
5.2.5 LPL protein expression analysis by western blot 62
5.2.6 CHO HCP Preparation and LPL-specific MRM Assay 63
5.2.7 PS80 and PS20 Degradation Assay 65

5.3 Results and Discussion 66

5.3.1 Rational design of sgRNA target sites 66
5.3.2 CRISPR/Cas9 shows targeted endonuclease activity in the *lpl* target regions in CHO cells 68
5.3.3 Generation of LPL knockout CHO-K1 cell lines 69
5.3.4 Characterization of CHO-K1 lpl knockout cell lines by western analysis and LPL-specific MRM assay 72
5.3.5 Culture performance of LPL knockout CHO-K1 cell lines 78
5.3.6 Polysorbate degradation is reduced in *lpl* knockout cell lines 81

5.4 Conclusion 84

REFERENCES 85

6 CONCLUSIONS AND FUTURE WORK 87

6.1 Summary of Conclusions 87
6.2 Recommendations for Future Work 88

6.2.1 Further investigation on the impact of LPL on formulation and product stability 88
6.2.2 Exploration of other HCPs contributing to formulation instability 90

REFERENCES 92

REFERENCES 94

Appendix

A SUPPLEMENTARY MATERIALS FOR CHAPTER 5 106

A.1 Preface 106
A.2 Supplementary table and figures for Chapter 5 106

B MATLAB SCRIPT FOR SGRNA SELECTION 108

B.1 Preface 108
B.2 MATLAB script for a complete list of possible sgRNA 109

LIST OF TABLES

Table 2.1: Comparison of ZFN, TALEN, and CRISPR/Cas9 for genome-editing. 13

Table 3.1: MRM assay parameters. 33

Table 4.1: Experimental set-up of PS80 degradation assay 47

Table 4.2: Full factorial design with corresponding responses for reproducibility. 49

Table 5.1: sgRNA target regions and target sequences for *lpl* 58

Table 5.2: sgRNA gBlock construct for *lpl*. 59

Table 5.3: MRM assay parameters. 65

Table 5.4: Corresponding functional site of *lpl* exons. 68

Table 6.1: HCPs likely to co-purify with therapeutic products 91

Table A.1: Sequencing primers. 106

LIST OF FIGURES

Figure 2.1: Gene knockout by CRISPR/Cas9 12

Figure 3.1: Impact of siRNA treatment on relative LPL expression, measured by MRM 36

Figure 3.2: Cell culture performance following siRNA interference. 37

Figure 3.3: Effect of low LPL expression on the formation of free fatty acids from enzymatic hydrolysis of PS80 during incubation. 39

Figure 4.1: One-way analysis of relative error by material. 51

Figure 4.2: One-way analysis of relative error by incubation time. 52

Figure 4.3: The Prediction Profiler maps the changes in the response (error) 53

Figure 4.4: One-way analysis of relative error by both material and incubation time . 54

Figure 5.1: Analysis of indels in *lpl*. 69

Figure 5.2: Genome editing of *lpl* in CHO-K1 cells by CRISPR/Cas9. 71

Figure 5.3: Characterization of LPL protein expression from CHO *lpl* knockout cell lines by western. 74

Figure 5.4: Peptides used for LPL MRM detection. 76

Figure 5.5: Characterization of LPL protein expression from CHO *lpl* knockout cell lines by MRM. 77

Figure 5.6: Cell culture performance of CHO-K1 *lpl* knockout cell lines. 80

Figure 5.7: Effect of LPL on the formation of free fatty acids from enzymatic hydrolysis of (A) PS80 and (B) PS20 83

Figure A.1: Transfection efficiency of Lonza Nucleofector. 107

ABSTRACT

Biopharmaceutical products, such as monoclonal antibodies (mAbs) and other recombinant therapeutic proteins, had global sales of $140 billion in 2013. Chinese hamster ovary (CHO) cells are one of the most commonly used cell platforms to express mAbs and recombinant therapeutic proteins. Tools for cell line engineering, such as short interfering RNA (siRNA) interference and clustered regularly interspaced short palindromic repeat (CRISPR)/CRISPR-associated protein-9 nuclease (CRISPR/Cas9), which emerged over the last decade, were extensively used to generate recombinant CHO cell lines with desired properties. With the availability of the CHO-K1 and Chinese hamster genome sequences, these tools allow for the rational design of CHO cell engineering through sequence-specific approaches.

This thesis describes an application of CHO cell engineering to address the impact of host cell protein (HCP) on formulation stability, a recurring problem observed in the biopharmaceutical industry. One particular HCP, lipoprotein lipase (LPL), is hypothesized to contribute to the degradation of polysorbate 20 (PS20) and polysorbate 80 (PS80), two surfactants used to stabilize recombinant therapeutic proteins in formulation, through enzymatic hydrolysis. The goals of this research are to (i) generate engineered CHO-K1 cell lines with a functional *lpl* knockout and (ii) study the impact of LPL on polysorbate degradation. Two cell line engineering tools, siRNA interference and CRISPR/Cas9, are applied in this work. Reduction in LPL expression by siRNA interference did not affect cell culture performance and is associated with lower PS80 degradation. The results of this study suggest that LPL

may play a significant role in PS80 hydrolysis, which motivates further investigation with a fully functional *lpl* knockout. Design of experiment methods are applied for optimizing the conditions for the polysorbate degradation assay and conditions are optimized to improve the reproducibility of the assay. These optimized conditions are used in subsequent polysorbate degradation studies. Rational design of targeted gene-editing by CRISPR/Cas9 leads to the generation of five CHO-K1 cell lines with a functional *lpl* knockout. In this work, incubation with HCPs from the CHO-K1 *lpl* knockout cell lines was found to reduce polysorbate degradation. The findings from this work support our hypothesis that LPL contributes to polysorbate degradation through enzymatic hydrolysis. The tools and approach used in this thesis demonstrate the application of cell line engineering tools for extensive genetic manipulation and genomic analysis of engineered CHO cell lines and their potential use in addressing biopharmaceutical problems.

Chapter 1

INTRODUCTION

1.1 Chinese Hamster Ovary Cells in Biotechnology

Biopharmaceutical had global sales of $140 billion in 2013, with $63 billion contributed by sales of monoclonal antibodies (mAbs), a type of recombinant therapeutic protein (Walsh, 2014). Six of the top ten product sales in 2013 are mAbs and they are projected to remain the most prominent class of biopharmaceuticals in the future (Walsh, 2014). Chinese hamster ovary (CHO) cell is one of the commonly used cell platforms to express recombinant therapeutic proteins, with 46% of therapeutic protein products on the market in 2012 produced in CHO cells (Kremkow & Lee, 2013).

The first recombinant therapeutic protein to gain approval from the Food and Drug Administration was humulin, which was recombinant human insulin produced in *Escherichia coli (E. coli)* (Walsh, 2012). The production of the human growth hormone in *E. coli* soon followed. However, the production of proteins that requires complex folding and post-translational modifications cannot be fulfilled with *E. coli* because it does not have the machinery necessary for such processes. Certain recombinant therapeutic proteins undergo post-translational modifications to be biologically active. Successful productions of therapeutic proteins with post-

translational modifications, such as erythropoietin and tissue-type plasminogen activator (tPA), have been demonstrated in CHO cells (Jayapal et al., 2007; Walsh, 2014). The adaptability and ease of genetic manipulation made CHO cells an important expression platform for the production of complex recombinant therapeutic proteins.

1.2 CHO Cell Line Engineering and the Emergence of Genome-editing Tools

The most traditional way of CHO cell line engineering is based on the stable integration of the gene of interest (GOI). Stable integration of the GOI in the CHO genome facilitates the constant expression of the GOI. The first step to inserting the GOI into the CHO genome is to clone the complementary DNA of the gene into an expression vector. The vector is then transfected into CHO cells for expression of the GOI. Incorporation of the transfected genetic material into the CHO genome is encouraged by applying selective pressure, like treatment with antibiotics. However, this method does not target specific sites and the gene is integrated at random locations in the genome, which can disrupt existing genetic material of the cell. Cells with desired traits, such as high specific productivity and cell growth, are selected through clonal isolation and expansion. Improvements in cloning techniques, such as expression vector design and clonal selection methods, and optimization of bioprocess increased specific productivity from 0.05 to 2–10 g/L (Datta et al., 2013; Huang et al., 2010; Wurm, 2004). Although this method is effective, there is little understanding of how the growth and the productivity of the therapeutic product are affected by the transgene expression and changes to the engineered CHO cell.

Recent advances that led to the availability of the CHO-K1 and the Chinese hamster genome (Brinkrolf et al., 2013; Xu et al., 2011) aids the rational design of engineered CHO cell lines. The knowledge of the CHO genome aids the development of cell lines with improved productivity and cell viability, while reducing both cell line selection and process development times (Datta et al., 2013). The emergence of small interfering RNA (siRNA) interference and genome-editing tools allow for studies of targeted genes. siRNA interference has been applied in CHO cells to regulate cellular mechanisms, such as apoptosis (Lim et al., 2006) and lactate production (Kim & Lee, 2007). Clustered regularly interspaced short palindromic repeat (CRISPR)/CRISPR-associated protein-9 nuclease (CRISPR/Cas9) and transcription activator-like effector nuclease (TALEN) are two genome-editing tools have been applied in CHO cells to knockout and knockin gene function (Lee et al., 2015; Ronda et al., 2014; Sun et al., 2015; Valton et al., 2012). These tools, which emerged over the last decade, are extensively used to generate recombinant CHO cell lines with desired properties.

1.3 Problem in CHO Biotechnology: Formulation instability

Biopharmaceutical formulations typically include surfactants, such as polysorbate 20 (PS20) and polysorbate 80 (PS80), to prevent surface adsorption and protein aggregation of recombinant therapeutic proteins (Kishore, Kiese, et al., 2011; Maggio, 2012). In recent years, reports of polysorbate degradation in long-term storage of recombinant therapeutic proteins have surfaced due its potential impact on product quality (Doshi et al., 2015; Kerwin, 2008; Kishore, Pappenberger, et al., 2011;

Labrenz, 2014; Tomlinson et al., 2015). Polysorbates degrade by two types of mechanism: autoxidation or hydrolysis. Ester hydrolysis of polysorbate produces free fatty acids while autoxidation of polysorbate produces free fatty acids short chain alkanes, ketones, aldehydes, and acids (Doshi et al., 2015). Recent studies on the mechanism of polysorbate degradation suggest that the degradation observed may be catalyzed by an impurity in the downstream process (Labrenz, 2014; Tomlinson et al., 2015).

1.4 Project Goals

The purpose of this work is to apply CHO cell line engineering tools on a known biopharmaceutical problem. In particular, the impact of lipoprotein lipase (LPL) on polysorbate stability is investigated. The objectives of this work are as follows:

(1) Apply siRNA interference to CHO-K1 cells to reduce LPL expression. The impact of reduced LPL expression on cell culture performance and polysorbate degradation were explored to determine the practicality of a permanent *lpl* knockout study.

(2) Apply CRISPR/Cas9 to (i) generate engineered CHO-K1 cell lines with a functional *lpl* knockout and (ii) study the impact of LPL on polysorbate degradation. The observations made in this study will generate discussion on the importance of addressing protein impurity on formulation and product stability.

1.5 Scope of work

This thesis is composed of several sections, which describes the application of siRNA interference and CRISPR/Cas9 on CHO-K1 cells to investigate the impact of LPL on polysorbate stability. In Chapter 3, CHO-K1 cells are treated with LPL-specific siRNA to investigate the impact of reduced LPL expression on cell culture performance and PS80 degradation. The observations made in Chapter 3 warrant a functional *lpl* knockout study with CRISPR/Cas9. Chapter 4 presents an optimization study used to identify the conditions that offer improved reproducibility of the PS80 degradation assay. This work established the experimental conditions for all subsequent polysorbate studies and demonstrated improvement in the assay's variability. Chapter 5 concludes this research by describing the application of CRISPR/Cas9 to knock out the function of *lpl*. Cell culture performance, LPL protein expression, and polysorbate degradation are investigated with CHO-K1 *lpl* knockout cell lines generated using this tool.

REFERENCES

Brinkrolf, K., Rupp, O., Laux, H., Kollin, F., Ernst, W., Linke, B., Kofler, R., Romand, S., Hesse, F., Budach, W. E., Galosy, S., Muller, D., Noll, T., Wienberg, J., Jostock, T., Leonard, M., Grillari, J., Tauch, A., Goesmann, A., Helk, B., Mott, J. E., Puhler, A., & Borth, N. (2013). Chinese hamster genome sequenced from sorted chromosomes. *Nat Biotechnol, 31*(8), 694-695.

Datta, P., Linhardt, R. J., & Sharfstein, S. T. (2013). An 'omics approach towards CHO cell engineering. *Biotechnol Bioeng, 110*(5), 1255-1271.

Doshi, N., Demeule, B., & Yadav, S. (2015). Understanding Particle Formation: Solubility of Free Fatty Acids as Polysorbate 20 Degradation Byproducts in Therapeutic Monoclonal Antibody Formulations. *Mol Pharm, 12*(11), 3792-3804.

Huang, Y. M., Hu, W., Rustandi, E., Chang, K., Yusuf-Makagiansar, H., & Ryll, T. (2010). Maximizing productivity of CHO cell-based fed-batch culture using chemically defined media conditions and typical manufacturing equipment. *Biotechnol Prog, 26*(5), 1400-1410.

Jayapal, K. P., Wlaschin, K. F., & Hu, W. S. (2007). Recombinant Protein Therapeutics from CHO Cells — 20 Years and Counting. *Chemical Engineering Progress, 103*, 40-47.

Kerwin, B. A. (2008). Polysorbates 20 and 80 used in the formulation of protein biotherapeutics: structure and degradation pathways. *J Pharm Sci, 97*(8), 2924-2935.

Kim, S. H., & Lee, G. M. (2007). Down-regulation of lactate dehydrogenase-A by siRNAs for reduced lactic acid formation of Chinese hamster ovary cells producing thrombopoietin. *Appl Microbiol Biotechnol, 74*(1), 152-159.

Kishore, R. S., Kiese, S., Fischer, S., Pappenberger, A., Grauschopf, U., & Mahler, H. C. (2011). The degradation of polysorbates 20 and 80 and its potential impact on the stability of biotherapeutics. *Pharm Res, 28*(5), 1194-1210.

Kishore, R. S., Pappenberger, A., Dauphin, I. B., Ross, A., Buergi, B., Staempfli, A., & Mahler, H. C. (2011). Degradation of polysorbates 20 and 80: studies on thermal autoxidation and hydrolysis. *J Pharm Sci, 100*(2), 721-731.

Kremkow, B., & Lee, K. H. (2013). Next-generation sequencing technologies and their impact on CHO cell-based biomanufacturing. *Pharm Bioprocess, 1*(5), 455-465.

Labrenz, S. R. (2014). Ester hydrolysis of polysorbate 80 in mAb drug product: evidence in support of the hypothesized risk after the observation of visible particulate in mAb formulations. *J Pharm Sci, 103*(8), 2268-2277.

Lee, J. S., Kallehauge, T. B., Pedersen, L. E., & Kildegaard, H. F. (2015). Site-specific integration in CHO cells mediated by CRISPR/Cas9 and homology-directed DNA repair pathway. *Sci Rep, 5*, 8572.

Lim, S. F., Chuan, K. H., Liu, S., Loh, S. O., Chung, B. Y., Ong, C. C., & Song, Z. (2006). RNAi suppression of Bax and Bak enhances viability in fed-batch cultures of CHO cells. *Metab Eng, 8*(6), 509-522.

Maggio, E. T. (2012). Polysorbates, Peroxides, Protein Aggregation, and Immunogenicity: A Growing Concern. *J Excip Food Chem, 3*(2), 45-53.

Ronda, C., Pedersen, L. E., Hansen, H. G., Kallehauge, T. B., Betenbaugh, M. J., Nielsen, A. T., & Kildegaard, H. F. (2014). Accelerating genome editing in CHO cells using CRISPR Cas9 and CRISPy, a web-based target finding tool. *Biotechnol Bioeng, 111*(8), 1604-1616.

Sun, T., Li, C., Han, L., Jiang, H., Xie, Y., Zhang, B., Qian, X., Lu, H., & Zhu, J. (2015). Functional knockout of FUT8 in Chinese hamster ovary cells using CRISPR/Cas9 to produce a defucosylated antibody. *Engineering in Life Sciences, 15*(6), 660-666.

Tomlinson, A., Demeule, B., Lin, B., & Yadav, S. (2015). Polysorbate 20 Degradation in Biopharmaceutical Formulations: Quantification of Free Fatty Acids, Characterization of Particulates, and Insights into the Degradation Mechanism. *Mol Pharm, 12*(11), 3805-3815.

Valton, J., Dupuy, A., Daboussi, F., Thomas, S., Maréchal, A., Macmaster, R., Melliand, K., Juillerat, A., & Duchateau, P. (2012). Overcoming Transcription Activator-like Effector (TALE) DNA Binding Domain Sensitivity to Cytosine Methylation. *J Biol Chem, 287*(46), 38427-38432.

Walsh, G. (2012). New biopharmaceuticals: a review of new biologic drug approvals over the years, featuring highlights from 2010 and 2011. *BioPharm International, Process Development Forum*.

Walsh, G. (2014). Biopharmaceutical benchmarks 2014. *Nat Biotechnol, 32*(10), 992-1000.

Wurm, F. M. (2004). Production of recombinant protein therapeutics in cultivated mammalian cells. *Nat Biotechnol, 22*(11), 1393-1398.

Xu, X., Nagarajan, H., Lewis, N. E., Pan, S., Cai, Z., Liu, X., Chen, W., Xie, M., Wang, W., Hammond, S., Andersen, M. R., Neff, N., Passarelli, B., Koh, W., Fan, H. C., Wang, J., Gui, Y., Lee, K. H., Betenbaugh, M. J., Quake, S. R., Famili, I., Palsson, B. O., & Wang, J. (2011). The genomic sequence of the Chinese hamster ovary (CHO)-K1 cell line. *Nat Biotechnol, 29*(8), 735-741.

Chapter 2

GENOME-EDITING TOOLS IN MAMMALIAN CELLS

2.1 Introduction

Genome-editing tools can be used to knock out or knock in gene expression to improve product quality and increase cellular production of therapeutic proteins. Targeted genome-editing is a controlled method used to induce genomic modifications, such as deletions and insertions, at specific points in the gene of interest, and has already shown its applicability in both bacterial and eukaryotic cells (Gaj et al., 2013). Studies that were previously difficult to perform, such as the verification of specific gene functions or the elucidation of how the genotype influences the phenotype, can be facilitated by genome-editing (Gaj et al., 2013; H. Kim & Kim, 2014). Due to its versatility and applicability in a broad range of organisms, the tool is widely adopted in biomedical research, biotechnology, and medicine (Hsu et al., 2014).

Three particular genome-editing tools are zinc-finger nuclease (ZFN), transcription activator-like effector nuclease (TALEN), and clustered regularly interspaced short palindromic repeats (CRISPR)-Cas9. Although the structural make-up among the three tools is different, the general mechanisms they use to edit genes are very similar. Two distinct domains are present in all three tools: a DNA-specific

recognition domain and a nuclease domain. As the name implies, the DNA-specific recognition domain recognizes a particular nucleotide sequence, called the target sequence, and complementarily binds to it. Once the target sequence is recognized by the DNA-binding domain, the nuclease domain cleaves the target sequence to create a double-strand break (DSB).

DSBs in the DNA induce cellular responses for DNA repair, such as the nonhomologous end joining (NHEJ) pathway or the homology-directed repair (HDR) pathway. Both pathways repair DSBs by ligating the two broken ends together, but the repair pathway taken is influenced by the stage of the cell cycle (Chapman et al., 2012; Thompson, 2012). NHEJ is error-prone (1%) and often introduces small insertions and deletions (indels) of various lengths at the ligation site (Gaj et al., 2013; Sander & Joung, 2014). Indels can lead to a frameshift in the coding regions of the gene, disrupting the subsequent translations of the gene, which may result in a functional knockout of the target gene (Figure 2.1). The HDR pathway is more accurate and repairs DSBs in the presence of a homologous template of the DNA, such as a donor DNA, a single-strand oligodeoxynucleotide, or a sister chromatid (Chapman et al., 2012; H. Kim & Kim, 2014; Thompson, 2012). For example, the introduction of a DNA template that is both homologous to the target site and containing a gene of interest, post-DSB enables insertion of the gene of interest into the target site via the HDR pathway (Hockemeyer et al., 2009; Hockemeyer et al., 2011; Mali, Yang, et al., 2013). Gene modifications, such as gene deletion, gene insertions, or point mutations, can be achieved following a DSB.

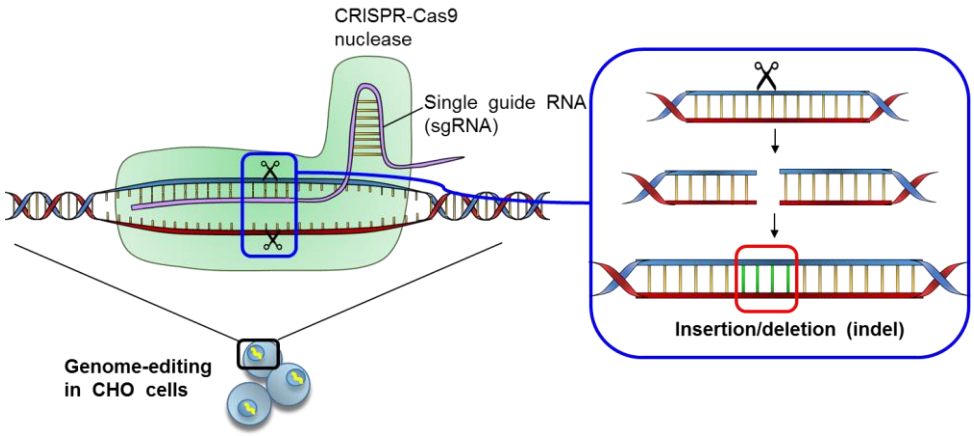


Figure 2.1: Gene knockout by CRISPR/Cas9. The sgRNA contains an RNA sequence complementary to the target gene. It guides the Cas9 nuclease to the target site. The Cas9 nuclease acts as a pair of scissors to produce DSB on the target site. Repair by NHEJ can either introduce random nucleotides or delete nucleotides around the break site.

2.1 Genome-editing with ZFN, TALEN, and CRISPR/Cas9

ZFNs, TALENs, and CRISPR/Cas9 have the potential to be used in regulating endogenous gene expression (Sander & Joung, 2014). Although these tools share the same genome-editing mechanism, they differ in aspects such as nuclease composition, DNA-binding mechanism, target sequence specification, and mutation signatures (Table 2.1). ZFN and TALEN rely on protein-DNA interactions for gene-targeting, while CRISPR/Cas9 uses an RNA/DNA-pairing principle. Due to the protein-DNA interaction necessary for gene-targeting, the success of ZFN and TALEN relies on a rational design and construction of the nuclease fusion protein, which is often time-consuming and costly (Pennisi, 2013). The CRISPR/Cas9 system has comparable efficiency at lower cost and requires only design one component, the single guide RNA (sgRNA), which is simple to design (Wei et al., 2013).

Table 2.1: Comparison of ZFN, TALEN, and CRISPR/Cas9 for genome-editing.

	ZFN	**TALEN**	**CRISPR/Cas9**
Target-binding principle	Protein-DNA recognition	Protein-DNA recognition	RNA-DNA complementary pairing
Nuclease	*Fok*I	*Fok*I	Cas9
Designed component	ZFP - *Fok*I fusion protein	TALE - *Fok*I fusion protein	Single guide RNA (sgRNA)
Mechanism	ZFP recognizes target DNA and *Fok*I generates DSB	TALE recognizes target DNA and *Fok*I generates DSB	sgRNA recognizes target DNA and Cas9 generates DSB
Mutation efficiency	Low (~24%)	High (>99%)	High (~90%)
Length of Target DNA	18-36 bp	2 x 17 bp + spacer (14-18)	23 bp
Specification of Target Site	5'-GNN-3'	Start with T and end with A	PAM (NGG) adjacent
Off-target effects	High	Low	Variable
Cytotoxicity	Variable to high	Low	Low

2.1.1 ZFN and TALEN

ZFN, discovered in 1994, was the first engineered nuclease used for genome engineering and had been used to modify genes in many different organisms, including bacteria, plants, insects, zebrafish, mouse, and mammalian cells (H. Kim & Kim, 2014; Y. G. Kim & Chandrasegaran, 1994). ZFNs are composed of two domains, a DNA-binding zinc finger protein (ZFP) domain and the nuclease domain,

*Fok*I. The ZFP domain is composed of zinc ions that stabilize the fold of the protein and an array of ZFP modules, with each ZFP module designed to recognize three nucleotides. The length of ZFN's target sequence varies with the number of ZFPs in the array, where 3-ZFP and 6-ZFP recognize nine nucleotides and 18 nucleotides, respectively. The sequence specificity is determined by the amino acid composition of the ZFP, therefore, each 3-nucleotide combination requires a different ZFP (Cermak et al., 2011). Methods to assemble ZFNs include the bacterial one-hybrid system and oligomerized pool engineering (OPEN), where both methods use bacterial selection to identify zinc finger combinations (Urnov et al., 2010). To complete the construction of ZFN, the ZFP array is flanked with the *Fok*I cleavage domain at the end of the array. The *Fok*I nuclease is only active as a dimer, so two ZFP arrays are necessary to induce a DSB. Upon dimerization of the ZFN pair, the *Fok*I nuclease creates a DSB on the targeted gene.

Despite several publications of successful genomic modification using ZFNs (Duda et al., 2014; Hockemeyer et al., 2009), the construction of engineered ZFP arrays has been challenging(Ramirez et al., 2008). ZFN construction requires the application of protein engineering techniques, such as the systematic evolution of ligands by exponential enrichment (SELEX) and ligation reactions, both of which can be time consuming and challenging. The construction of ZFP arrays is subject to target sequence constraints including the presence of 5'-GNN-3' repeats, where N is any nucleotide (Ramirez et al., 2008). Sequence recognition can be impacted by neighboring interactions between an individual ZFP and surrounding ZFPs, which can

affect the overall specificity of the ZFN (Gaj et al., 2013). Additionally, a limited number of sequences can be targeted because not all possible ZFP combinations have not been discovered. Overall, production of highly selective ZFN proteins is costly, laborious, and time-consuming (Wei et al., 2013).

TALENs emerged in 2009 as another genome-editing tool (Boch et al., 2009). The efficiency of TALENs to modify genes have been demonstrated in a broad range of organisms, including yeast, plants, insect, fish, mice, and mammalian cells(H. Kim & Kim, 2014). TALENs are similar to ZFNs and are composed of two domains, a DNA-binding transcription activator-like effectors (TALE) domain and a *Fok*I nuclease domain. TALEN, like ZFN, relies on protein-DNA interactions for sequence recognition. TALEs are proteins naturally secreted by *Xanthomonas* spp. bacteria to alter gene transcription in host plant cells (Joung & Sander, 2013), and are composed of 33-35 amino acid repeats, each recognizing a single nucleotide. The nucleotide specificity of TALE is determined by repeat variable diresidues (RVDs), which comprise the twelfth and thirteenth amino acids. TALENs are constructed as an array of 20 TALEs and flanked with the *Fok*I cleavage domain at the end of the array (Cermak et al., 2011). TALEN recognizes DNA using the central repeat domain of TALEs and cleaves DNA with the *Fok*I nuclease. Like ZFNs, the *Fok*I nuclease cleaves only when assembled as a dimer, so two TALEN arrays are necessary to induce a DSB.

The emergence of TALENs provided an alternative to ZFNs for genome editing. TALE domains are comparatively easier to design than ZFP because the

sequence recognition relies only on the RVD components, unlike ZFP designs which are based on the whole amino acid composition. Following construction, the delivery of the TALEN's DNA sequence to the host cell using viral vectors, such as lentivirus, can also be difficult due to the size of the DNA sequence (Holkers et al., 2013). Additionally, the construction TALE repeats requires the use of nonstandard molecular biology cloning methods, such as "Golden Gate" cloning and solid-phase assembly(Joung & Sander, 2013), which can be laborious and time-consuming.

2.1.2 CRISPR/Cas9

In 2012, Jinek and coworkers reported a new genome-editing technique using the CRISPR/Cas9 system adapted from *Streptococcus pyogenes (S. pyogenes)* (Jinek et al., 2012). In *S. pyogenes*, the CRISPR/Cas9 system uses an RNA-guided system for targeting and cleaving DNA sequences, where the system uses RNA-DNA base pairing for sequence recognition. The gene silencing CRISPR system is a naturally occurring adaptable immune response used by bacteria and archaea to protect themselves from foreign nucleic acids such as viruses or plasmids (Wiedenheft et al., 2012). The immune response acts in two phases, the immunization phase and the immunity phase. In the immunization phase, the host cell incorporates fragments of foreign DNA sequences (spacers) into the CRISPR locus. In the immunity phase, the CRISPR/Cas9 system uses a multistep mechanism to defend against foreign pathogens(Mali, Esvelt, et al., 2013). This defensive mechanism involves three main

components — the bacterial endonuclease Cas9, CRISPR RNA (crRNA), and transactivating crRNA (tracrRNA). The spacers are transcribed from the CRISPR locus to produce the crRNA, which is an RNA sequence transcribed from the foreign DNA (protospacer), with CRISPR repeat regions integrated within the sequence. The tracrRNA, another RNA sequence transcribed from the CRISPR locus, hybridizes to the repeat regions of the crRNA, and Cas9 recognizes the structure formed by the hybridization to construct a Cas9:crRNA-tracrRNA complex (Mali, Esvelt, et al., 2013; Sander & Joung, 2014). The system requires the DNA sequence to be adjacent to a protospacer adjacent motif (PAM) for the sequence to be cleaved. The PAM sequence varies for different bacteria or archaea. The Cas9:crRNA-tracrRNA complex recognizes the complementary target DNA sequence through the protospacer of the crRNA, and Cas9 creates a DSB before the PAM sequence (Makarova et al., 2011).

This system has been adapted for targeted genome editing using the Cas9 nuclease and a single guide RNA (sgRNA), a union of a 20-nucleotide RNA designed to recognize the target sequence and a shortened tracrRNA (Jinek et al., 2012). The sgRNA guides Cas9 to the target DNA site using complementary base-pairing rules. The engineered CRISPR/Cas9 system requires the target DNA to be adjacent to a PAM sequence. *S. pyogenes'* CRISPR/Cas9 system recognizes the PAM sequence 5'-NGG, where N refers to any nucleotide. In the CRISPR/Cas9 system adapted by Jinek et al., Cas9 nuclease activity can be directed to any target sequence with the form $(N)_{20}$NGG by altering the 5' end of the 20-nucleotide sgRNA to match the $(N)_{20}$ of the target sequence(Jinek et al., 2012). The ability of the CRISPR/Cas9 system to induce

targeted edits in genomes has been demonstrated in bacteria, humans, zebrafish, yeast, mice, and many other species(Sander & Joung, 2014).

The length of the target sequence in CRISPR/Cas9 is shorter than those designed for ZFNs or TALENs, which has the disadvantage of an increased probability of off-target effects (Fu et al., 2013). However, off-target effects can be reduced through techniques such as the use of paired nickases (Ran et al., 2013). A nickase is a mutation of the Cas9 nuclease that contains only one active nuclease domain. The unmodified Cas9 nuclease contains two conserved nuclease domains, HNH and RuvC, which cleave sense and anti-sense strands, respectively. An amino acid mutation at position D10A or H840A of Cas9 inactivates RuvC or HNH to generate a modified Cas9 with only one active nuclease domain (Ran et al., 2013). With one active domain, the nickase cleaves only one strand of the DNA, which can be repaired quickly without indel formation. Paired nickases generate nicks on both sides of the DNA by using two sgRNAs and Cas9 nickases. Paired single-strand breaks induced on opposite DNA strands and separated by 4- to 100-nucleotides, create indels as efficiently as the wildtype Cas9 (Ran et al., 2013). The use of two sgRNAs for sequence recognition was found to reduce off-target effects by 50- to 1,500-fold (Ran et al., 2013).

Because the promoter can limit possible target sequences, the choice of promoter also needs to be taken into consideration when choosing the target sequence and designing the sgRNA. For example, the RNA polymerase III-dependent U6 promoter requires a G at the 5' end while the T7 promoter requires GG at the 5'

site(Sander & Joung, 2014). In addition to the requirement for adjacent PAM sequences, the target sequence can only be of the form $G(N)_{19}NGG$ or $GG(N)_{18}NGG$ when using U6 or T7 promoters, respectively. Such restrictions limit the number of targetable sites per gene.

2.2 CRISPR/Cas9 in Chinese Hamster Ovary (CHO) cells

CHO cells are commonly used as one of the cell platforms for biotherapeutic production. Historically, due to the lack of genomic information, CHO cell lines with desirable traits were often identified only after intensive screening, mutagenesis, drug treatment, or media optimization. With the sequencing of the CHO-K1 genome recently completed (Xu et al., 2011), a systems biology approach towards rational design of CHO cell lines with desired attributes, such as improved productivity and product quality, can be implemented using targeted genome-editing. The CRISPR/Cas9 system provides a systematic way to identify mutations that contribute towards the desired attributes.

Several groups have demonstrated the application of CRISPR/Cas9 in CHO cells (Grav et al., 2015; Lee, Kallehauge, et al., 2015; Ronda et al., 2014; Sun et al., 2015), which results in gene deletion or gene insertion. Functional knockouts of FUT8 yield CHO cell lines that produce defucosylated antibodies (Grav et al., 2015; Ronda et al., 2014; Sun et al., 2015), while BAX and BAK knockouts yield CHO cell lines

with high viability under long culture times (Grav et al., 2015). Simultaneous gene disruptions of up to three genes, with 58% of all clones containing two or more knockouts (Grav et al., 2015), have also been shown, demonstrating the feasibility of multiplex genome-editing in CHO cells. HDR-mediated gene insertions have also been explored in CHO cells, where mCherry was integrated at targeted genes (*COSMC, Mgat1, LdhA*). Although HDR-based gene repair occurs infrequently in mammalian cells (Lee, Grav, et al., 2015), targeting efficiencies between 7.4% and 27.8% were observed (Lee, Kallehauge, et al., 2015).

The completion of the CHO-K1 genome and the availability of genome-editing tools open new opportunities for rational design in CHO cell engineering. For example, questions regarding the mechanism between genotype and phenotype in CHO cells can be addressed by the CRISPR/Cas9 system. Gene modifications on the target genes facilitated by CRISPR/Cas9 can also address present problems in the biopharmaceutical industries, such as host cell protein impurity. The application of CRISPR/Cas9 in CHO cells will yield a vast amount of information that could lead to engineering of an improved CHO cell line.

REFERENCES

Boch, J., Scholze, H., Schornack, S., Landgraf, A., Hahn, S., Kay, S., Lahaye, T., Nickstadt, A., & Bonas, U. (2009). Breaking the Code of DNA Binding Specificity of TAL-Type III Effectors. *Science, 326*(5959), 1509-1512.

Cermak, T., Doyle, E. L., Christian, M., Wang, L., Zhang, Y., Schmidt, C., Baller, J. A., Somia, N. V., Bogdanove, A. J., & Voytas, D. F. (2011). Efficient design and assembly of custom TALEN and other TAL effector-based constructs for DNA targeting. *Nucleic Acids Res, 39*(12), e82.

Chapman, J. R., Taylor, M. R., & Boulton, S. J. (2012). Playing the end game: DNA double-strand break repair pathway choice. *Mol Cell, 47*(4), 497-510.

Duda, K., Lonowski, L. A., Kofoed-Nielsen, M., Ibarra, A., Delay, C. M., Kang, Q., Yang, Z., Pruett-Miller, S. M., Bennett, E. P., Wandall, H. H., Davis, G. D., Hansen, S. H., & Frodin, M. (2014). High-efficiency genome editing via 2A-coupled co-expression of fluorescent proteins and zinc finger nucleases or CRISPR/Cas9 nickase pairs. *Nucleic Acids Res, 42*(10), e84.

Fu, Y., Foden, J. A., Khayter, C., Maeder, M. L., Reyon, D., Joung, J. K., & Sander, J. D. (2013). High-frequency off-target mutagenesis induced by CRISPR-Cas nucleases in human cells. *Nat Biotechnol, 31*(9), 822-826.

Gaj, T., Gersbach, C. A., & Barbas, C. F., 3rd. (2013). ZFN, TALEN, and CRISPR/Cas-based methods for genome engineering. *Trends Biotechnol, 31*(7), 397-405.

Grav, L. M., Lee, J. S., Gerling, S., Kallehauge, T. B., Hansen, A. H., Kol, S., Lee, G. M., Pedersen, L. E., & Kildegaard, H. F. (2015). One-step generation of triple knockout CHO cell lines using CRISPR/Cas9 and fluorescent enrichment. *Biotechnol J, 10*(9), 1446-1456.

Hockemeyer, D., Soldner, F., Beard, C., Gao, Q., Mitalipova, M., DeKelver, R. C., Katibah, G. E., Amora, R., Boydston, E. A., Zeitler, B., Meng, X., Miller, J. C., Zhang, L., Rebar, E. J., Gregory, P. D., Urnov, F. D., & Jaenisch, R. (2009). Efficient targeting of expressed and silent genes in human ESCs and iPSCs using zinc-finger nucleases. *Nat Biotechnol, 27*(9), 851-857.

Hockemeyer, D., Wang, H., Kiani, S., Lai, C. S., Gao, Q., Cassady, J. P., Cost, G. J., Zhang, L., Santiago, Y., Miller, J. C., Zeitler, B., Cherone, J. M., Meng, X., Hinkley, S. J., Rebar, E. J., Gregory, P. D., Urnov, F. D., & Jaenisch, R. (2011). Genetic engineering of human pluripotent cells using TALE nucleases. *Nat Biotechnol, 29*(8), 731-734.

Holkers, M., Maggio, I., Liu, J., Janssen, J. M., Miselli, F., Mussolino, C., Recchia, A., Cathomen, T., & Goncalves, M. A. (2013). Differential integrity of TALE nuclease genes following adenoviral and lentiviral vector gene transfer into human cells. *Nucleic Acids Res, 41*(5), e63.

Hsu, P. D., Lander, E. S., & Zhang, F. (2014). Development and applications of CRISPR-Cas9 for genome engineering. *Cell, 157*(6), 1262-1278.

Jinek, M., Chylinski, K., Fonfara, I., Hauer, M., Doudna, J. A., & Charpentier, E. (2012). A programmable dual-RNA-guided DNA endonuclease in adaptive bacterial immunity. *Science, 337*(6096), 816-821.

Joung, J. K., & Sander, J. D. (2013). TALENs: a widely applicable technology for targeted genome editing. *Nat Rev Mol Cell Biol, 14*(1), 49-55.

Kim, H., & Kim, J. S. (2014). A guide to genome engineering with programmable nucleases. *Nature Reviews Genetics, 15*(5), 321-334.

Kim, Y. G., & Chandrasegaran, S. (1994). Chimeric restriction endonuclease. *Proc Natl Acad Sci USA, 91*, 883-887.

Lee, J. S., Grav, L. M., Lewis, N. E., & Faustrup Kildegaard, H. (2015). CRISPR/Cas9-mediated genome engineering of CHO cell factories: Application and perspectives. *Biotechnol J, 10*(7), 979-994.

Lee, J. S., Kallehauge, T. B., Pedersen, L. E., & Kildegaard, H. F. (2015). Site-specific integration in CHO cells mediated by CRISPR/Cas9 and homology-directed DNA repair pathway. *Sci Rep, 5*, 8572.

Makarova, K. S., Haft, D. H., Barrangou, R., Brouns, S. J., Charpentier, E., Horvath, P., Moineau, S., Mojica, F. J., Wolf, Y. I., Yakunin, A. F., van der Oost, J., & Koonin, E. V. (2011). Evolution and classification of the CRISPR-Cas systems. *Nat Rev Microbiol, 9*(6), 467-477.

Mali, P., Esvelt, K. M., & Church, G. M. (2013). Cas9 as a versatile tool for engineering biology. *Nat Methods, 10*(10), 957-963.

Mali, P., Yang, L., Esvelt, K. M., Aach, J., Guell, M., DiCarlo, J. E., Norville, J. E., & Church, G. M. (2013). RNA-guided human genome engineering via Cas9. *Science, 339*(6121), 823-826.

Pennisi, E. (2013). The CRISPR Craze. *Science, 341*, 833-836.

Ramirez, C. L., Foley, J. E., Wright, D. A., Muler-Lerch, F., Rahman, S. H., Cornu, T. I., Winfrey, R. J., Sander, J. D., Fu, F., Townsend, J. A., Cathomen, T., Voytas, D. F., & Joung, J. K. (2008). Unexpected failure rates for modular assembly of engineered zinc fingers. *Nat Methods, 5*(5), 374-375.

Ran, F. A., Hsu, P. D., Lin, C. Y., Gootenberg, J. S., Konermann, S., Trevino, A. E., Scott, D. A., Inoue, A., Matoba, S., Zhang, Y., & Zhang, F. (2013). Double nicking by RNA-guided CRISPR Cas9 for enhanced genome editing specificity. *Cell, 154*(6), 1380-1389.

Ronda, C., Pedersen, L. E., Hansen, H. G., Kallehauge, T. B., Betenbaugh, M. J., Nielsen, A. T., & Kildegaard, H. F. (2014). Accelerating genome editing in CHO cells using CRISPR Cas9 and CRISPy, a web-based target finding tool. *Biotechnol Bioeng, 111*(8), 1604-1616.

Sander, J. D., & Joung, J. K. (2014). CRISPR-Cas systems for editing, regulating and targeting genomes. *Nat Biotechnol, 32*(4), 347-355.

Sun, T., Li, C., Han, L., Jiang, H., Xie, Y., Zhang, B., Qian, X., Lu, H., & Zhu, J. (2015). Functional knockout of FUT8 in Chinese hamster ovary cells using CRISPR/Cas9 to produce a defucosylated antibody. *Engineering in Life Sciences, 15*(6), 660-666.

Thompson, L. H. (2012). Recognition, signaling, and repair of DNA double-strand breaks produced by ionizing radiation in mammalian cells: the molecular choreography. *Mutat Res, 751*(2), 158-246.

Urnov, F. D., Rebar, E. J., Holmes, M. C., Zhang, H. S., & Gregory, P. D. (2010). Genome editing with engineered zinc finger nucleases. *Nat Rev Genet, 11*(9), 636-646.

Wei, C., Liu, J., Yu, Z., Zhang, B., Gao, G., & Jiao, R. (2013). TALEN or Cas9 - rapid, efficient and specific choices for genome modifications. *J Genet Genomics, 40*(6), 281-289.

Wiedenheft, B., Sternberg, S. H., & Doudna, J. A. (2012). RNA-guided genetic silencing systems in bacteria and archaea. *Nature, 482*(7385), 331-338.

Xu, X., Nagarajan, H., Lewis, N. E., Pan, S., Cai, Z., Liu, X., Chen, W., Xie, M., Wang, W., Hammond, S., Andersen, M. R., Neff, N., Passarelli, B., Koh, W., Fan, H. C., Wang, J., Gui, Y., Lee, K. H., Betenbaugh, M. J., Quake, S. R., Famili, I., Palsson, B. O., & Wang, J. (2011). The genomic sequence of the Chinese hamster ovary (CHO)-K1 cell line. *Nat Biotechnol, 29*(8), 735-741.

Chapter 3

LPL KNOCKDOWN IN CHINESE HAMSTER OVARY CELLS BY SIRNA INTERFERENCE

3.1 Preface

To examine the impact of reduced lipoprotein lipase (LPL) expression on CHO culture performance, such as cell growth and cell viability, short interfering ribonucleic acid (siRNA) was used to knock down LPL expression. We also analyzed the impact of reduced LPL expression on polysorbate 80 (PS80) degradation. LPL expression by cells treated with siRNA was reduced by 56-72% compared to the wildtype control. The reduction in LPL expression did not significantly affect cell culture performance, where the cell densities of siRNA-transfected cells were reduced by 10-19% and minimal changes to total extracellular protein expression were observed. These results warrant subsequent experiments on a functional knockout of lpl using genome-editing tools, such as CRISPR/Cas9. Lie Min developed and performed the multiple-reaction monitoring (MRM) assay for the relative quantification of LPL expression levels in all samples. Kristin Valente performed the PS80 degradation study for siRNA-treated and non-treated CHO cells.

3.2 Introduction

Polysorbate is a class of surfactant that is commonly included in biopharmaceutical formulations to stabilize biotherapeutic proteins by preventing surface adsorption and protein aggregation (Kishore, Pappenberger, et al., 2011; Tomlinson et al., 2015). Polysorbate 20 (PS20) and polysorbate 80 (PS80) are two surfactants commonly used to stabilize and prolong the shelf-life of biotherapeutic proteins (Marichal-Gallardo & Alvarez, 2012; Tomlinson et al., 2015), where PS20 and PS80 are chemically diverse mixtures of polyoxyethylene sorbitan fatty acid esters. Monoclonal antibody (mAb) formulations typically contain 0.01-0.05% of surfactants; higher percentages of surfactants can cause protein unfolding (Marichal-Gallardo & Alvarez, 2012). Polysorbates are known to degrade by autoxidation and hydrolysis into degradant molecules, such as free fatty acid, peroxides, acetaldehyde, acetone, and ethanol (Kishore, Kiese, et al., 2011) , which causes an increase in the concentration of degradant molecules in biopharmaceutical formulation. The degradation of polysorbate could influence the quality of biotherapeutic proteins due to (a) the absence of polysorbate necessary for protein stability or (b) a buildup of degradants, which could potentially have an additional impact on the protein's stability. Many studies discussed the mechanism of chemical degradation of polysorbates, such as autoxidation and acid and based-catalyzed hydrolysis (Doshi et al., 2015; Ilko et al., 2015; Kerwin, 2008; Kishore, Kiese, et al., 2011; Kishore, Pappenberger, et al., 2011; Maggio, 2012; Tomlinson et al., 2015). However, the impact of biologic-based impurities on polysorbate degradation is not as frequently discussed. We report here our study of polysorbate degradation through an enzyme-catalyzed ester hydrolysis.

Biologic-based impurities that catalyze polysorbate degradation are from host cell proteins (HCPs) impurities. HCP impurity biopharmaceutical formulations can play a negative role in formulation stability (Labrenz, 2014). Despite multiple of purification steps, such as Protein A capture or ion-exchange chromatography, low levels (1–100 ppm) of HCPs may remain in the final purified product (Doneanu et al., 2012; Valente, 2014). Residual HCPs may contribute to polysorbate degradation. For example, an enzyme-like mechanism of ester hydrolysis of PS80 in formulation development samples has been reported and hypothesized to be caused by the enzymatic activity of triacylglycerol lipase, an enzyme that hydrolyzes triacylglycerol (Labrenz, 2014). The biological function of lipase enzymes is ester hydrolysis of triglycerides into fatty acids and glycerol. A lipase of interest, lipoprotein lipase (LPL) has been characterized as a difficult-to-remove HCP (Valente et al., 2015). It has been observed to co-purify with multiple mAb products across Protein A capture and non-affinity polishing resins, exhibit variable expression with cell age, and associate with different mAbs (Levy et al., 2014; Valente et al., 2015; Valente et al., 2014). The enzyme LPL hydrolyzes the triacylglycerol component of lipoproteins, an assembly of protein and lipid that transport fat in water, into three free fatty acids and glycerol. The chemical structure of polysorbate consists of an ester bond with a long hydrocarbon chain, similar to the triglyceride structure that LPL recognizes. Previous research has shown that LPL produced in *E. coli* degrades PS80 at 37 °C (Levy, 2014). Therefore, we hypothesize that LPL contained in the CHO-derived product may contribute to the degradation of polysorbates through enzymatic hydrolysis during storage.

Short interfering ribonucleic acid (siRNA) interference offers a way to study the impact of LPL on the stability of polysorbates through transient reduction of CHO

LPL expression. siRNAs are 21-22 nucleotide RNA strands designed to interfere with the mRNA of the target gene. The siRNA forms a complex with the RNA-induced silencing complex (RISC) and guide it to the target mRNA, which is then cleaved by the RISC. Cleavage of the target mRNA inhibits translation of the target protein. Several studies have demonstrated the use of siRNA interference to study gene functions in CHO cells (Wu, 2009). For example, the knockdown of BAK and BAX, two proteins known to regulate apoptosis, in a recombinant CHO cell line by siRNA interference improved cell culture duration by 5 days and recombinant human interferon-ɣ productivity by 35% (Lim et al., 2006). The knockdown of lactate dehydrogenase-A, an enzyme that catalyzes the conversion of pyruvate to lactate, improved specific productivity of human thrombopoietin in recombinant CHO cells (Kim & Lee, 2007). Both studies demonstrated the application of siRNA interference technology to study gene functions in CHO cells. We will apply this particular technology for the knockdown of LPL to study the impact of LPL in mAb formulation stability.

Mass spectrometry-based methods, such as multiple-reaction monitoring (MRM), have been developed for targeted protein quantification (Liebler & Zimmerman, 2013). For quantification of a specific protein, MRM provides higher sensitivity than other methods, such as western analysis. The assay uses a selection of proteotypic peptides, whose sequences are unique to the target protein, and are further fragmented to optimally stable precursor-fragment ion transitions by collision-induced dissociation (Gstaiger & Aebersold, 2009; Liebler & Zimmerman, 2013). MRM assays for protein quantification were applied to a number of studies (Janas et al., 2012; Menetret et al., 2007). For example, MRM measurements of Argonaute proteins

were applied to calculate the stoichiometry and dynamics of miRNA-mRNA complexes (Janas et al., 2012), while MRM measurements of SecY were applied to determine the binding relationship between *E. coli* ribosomes and SecY (Menetret et al., 2007).

This study applied siRNA technology to knock down the expression of LPL, an HCP that has been shown to be difficult to remove during purification, and is hypothesized to affect the stability of biotherapeutic formulation through polysorbate degradation. Here, we investigate the application of siRNA interference to reduce LPL expression and explore the impact of decreased LPL expression on cell culture performance, such as cell growth and total protein production, and PS80 degradation.

3.3 Materials and Methods

3.3.1 CHO cell culture

A null CHO-K1 cell line (ATCC CCL-61), previously adapted to serum-free suspension culture (Valente, 2014), was cultured in 125-mL shake flasks containing 25-mL SFM4CHO medium (GE Healthcare Life Sciences, Little Chalfont, Buckinghamshire, UK). Cultures were seeded at 5×10^4 cells/mL and incubated for 3 - 5 days in a 37°C cell culture incubator with 5% CO_2 and 80% relative humidity.

3.3.2 siRNA transfection

Serum-free suspension CHO cells were treated with Opti-MEM medium (Life Technologies, Carlsbad, CA, USA) and transfected with one of the three *lpl*-specific siRNAs (5'-GCAACAATGTGGGCTATGA-3', 5'-CCTTTCTCCTGATGATGCA-3', and 5'-GAAATGATGTGGCCAGGTT-3') or a non-specific siRNA control (all obtained from Sigma-Aldrich Chemical Co., St. Louis, MO, USA). Cells were transfected using Lipofectamine 2000 (Life Technologies) according to the manufacturer's instruction and incubated in 50 mL CultiFlask bioreactors (Sartorius Stedim Biotech, Göttingen, Germany) for 4 – 6 hours. Following transfection, cells were diluted in SFM4CHO medium and cultured for 48 hours for siRNA-mediated silencing. The cells were counted using a Fuchs Rosenthal hemocytometer and cell viability was determined by the Trypan blue exclusion method. CHO harvested cell culture fluid (HCCF) from wildtype and siRNA-treated cells were separated from the cells by centrifugation (180 x g, 10 min) and stored at -20 °C.

3.3.3 CHO HCP preparation

HCP from CHO HCCF were precipitated with methanol by the addition of 2 volumes of methanol to 1 volume of CHO HCCF, followed by overnight incubation at -20 °C. The precipitated protein was recovered by centrifugation (5000 rpm, 65 min, 4 °C). Residual detergent was removed by DetergentOUT GBS10-800 detergent removal kit (G-Biosciences, St. Louis, MO, USA) according to the manufacturer's protocol. Trypsin digestion was performed as described (Valente, 2014). Peptide

pellets were resolubilized in 0.1% trifluoroacetic acid (TFA, Fisher Scientific, Fair lawn, NJ, USA), loaded onto C18 ZipTips (EMD Millipore, Billerica, MA, USA) and eluted in 50% acetonitrile with 0.1% TFA (Fisher Scientific). Concentrations of resolubilized HCPs were measured with a Bradford assay.

3.3.4 LPL-specific MRM assay

High pH reversed phase high performance liquid chromatography (RP-HPLC) was performed using an UltiMate 3000 nLC system (Dionex, Sunnyvale, CA, USA). Digested CHO HCPs were loaded onto a C18 trap column (Dionex) and washed with 150 µL of 2% acetonitrile (Mallinckrodt Chemicals, Phillipsburg, NJ, USA) in 0.1% formic acid (Pierce Chemical). Peptides were eluted onto a 0.66 µL C18 column (Dionex) by a 26 column linear gradient from 2 – 49% acetonitrile, followed by an additional 6 column volumes of 49% acetonitrile. All operations were performed at 2.6 min residence time and both mobile phases included 0.1% formic acid. Column eluate was directly injected into a QTrap 4000 (AB Sciex, Foster City, CA, USA) through a nanoSpray II source (AB Sciex) with an uncoated fused-silica Pico tip (New Objective, Woburn, MA, USA). The instrument was operated in positive ESI ion mode, with spray voltage at 2400 V and a source temperature of 150 °C, with MRM triggered enhanced resolution scan and enhanced product ion scans. Database searches were performed using ProteinPilot software v4.0 (AB Sciex) against the proteome database of the CHO genome (Xu et al. 2011). Possible MRM transitions were generated with Skyline v2.5.0.6157 (MacLean et al. 2010) and monitored through Analyst 1.6.2 (AB Sciex) with parameters specified in Table 3.1. Raw MRM data

were integrated for peak area and normalized to yeast alcohol dehydrogenase (YAD) transitions peak areas. All analyses were performed with three technical replicates and two biological replicates. The LPL-specific MRM assay was developed and performed by Lie Min.

Table 3.1: MRM assay parameters.

	Target peptide sequence	Precursor (m/z)	Product (m/z)	Product ion	Scan time (ms)	CE (V)
LPL	ITGLDPAGPNFEYAEAPSR	1002.987	793.424	$y72^+$	20	72.9
	ITGLDPAGPNFEYAEAPSR	1002.987	430.271	$y42^+$	20	67.9
	ITGLDPAGPNFEYAEAPSR	1002.987	359.204	$y32^+$	20	52.9
	EPDSNVIVVDWLYR	852.933	1162.662	$y92^+$	20	44.4
	EPDSNVIVVDWLYR	852.933	1063.594	$y82^+$	20	44.4
	EPDSNVIVVDWLYR	852.933	950.509	$y72^+$	20	44.4
	ITGLDPAGPNFEYAEAPSR	668.994	793.424	$y73^+$	20	37.8
	ITGLDPAGPNFEYAEAPSR	668.994	630.331	$y63^+$	20	37.8
	ITGLDPAGPNFEYAEAPSR	668.994	430.241	$y43^+$	20	37.8
	LSPDDADFVDVLHTFTR	974.476	661.352	$y52^+$	20	56.3
	GLGDVDQLVK	522.29	873.478	$y82^+$	20	30.5
	GLGDVDQLVK	522.29	701.429	$y62^+$	20	30.5
	GLGDVDQLVK	522.29	602.361	$y52^+$	20	30.5
YAD	ANELLINVK	507.3031	699.4763	$y62^+$	20	32.9
	ANELLINVK	507.3031	586.3923	$y52^+$	20	32.9
	ANGTTVLVGMPAGAK	693.8741	730.3916	$y82^+$	20	42.8
	ANGTTVLVGMPAGAK	693.8741	631.3232	$y72^+$	20	42.8
	EALDFFAR	484.7454	655.3198	$y52^+$	20	31.7
	EALDFFAR	484.7454	540.2929	$y42^+$	20	31.7
	VVGLSTLPEIYEK	483.2729	778.3981	$y63^+$	20	32
	VVGLSTLPEIYEK	483.2729	552.3028	$y43^+$	20	32

3.3.5 PS80 degradation assay

CHO HCCF harvested from siRNA-transfected cells and wildtype control were buffer-exchanged into pH 6.8 10 mM $CaCl_2$ (Sigma-Aldrich Chemical Co) buffer. PS80 (Fisher Scientific) was added to the buffer-exchanged CHO HCCF to a final concentration of 23 mM (0.03% w/w) and was incubated at 37 °C for 24 hours with mixing. The concentration of free fatty acid produced by degradation of PS80 was measured using the EnzyChrom Free Fatty Acid Kit (BioAssay Systems, Hayward, CA, USA) according to the manufacturer's protocol.

3.4 Results and Discussion

siRNA interference is preferable to genome editing when the impact of gene deletion is unknown. A partial knockdown of protein expression serves as a preliminary study to help determine whether the elimination of the gene function will impair cell function. Additionally, genome-editing mediated knockout often require clone isolation for accurate evaluation of the gene function, where clonal expansion requires long culture times. The use of siRNA interference for the knockdown of LPL expression serves as a preliminary study to examine the impact of reduced LPL expression on CHO cell culture performance and on polysorbate degradation before proceeding with a gene function knockout using the genome-editing tool, CRISPR/Cas9.

Three siRNA sequences were selected based on their specificity to the CHO-K1 *lpl* mRNA, as determined by a proprietary algorithm developed by Sigma-Aldrich

Chemical Company. Each siRNA targets a different region of the *lpl* mRNA. The *lpl*-specific siRNA and a non-specific siRNA control were separately transfected in CHO cells. Cell growth, total extracellular protein expression, LPL expression, and PS80 degradation results from siRNA-transfected CHO cells were compared to the wildtype CHO control. LPL expression from siRNA-transfected cells and wildtype cells were measured by an MRM assay, which is specific for the detection of LPL. The relative LPL expressions of the five cases were compared to the wildtype control (Figure 3.1). Cells treated with *lpl*-specific siRNA exhibited 56-72% reduction in relative LPL expression. The reduction in LPL expression did not adversely affect cell culture performance significantly, with viable cell density of siRNA-transfected cells reduced by 10-19% (Figure 3.2A) and minimal change to total extracellular protein expression was observed (Figure 3.2B). Although LPL expressions in siRNA-transfected cells were reduced, the cell cultures were able to maintain properties that are important for biopharmaceutical processing, such as viable cell density and protein production. These results formed the basis for subsequent experiments on a functional knockout of *lpl* using the genome-editing tool, CRISPR/Cas9.

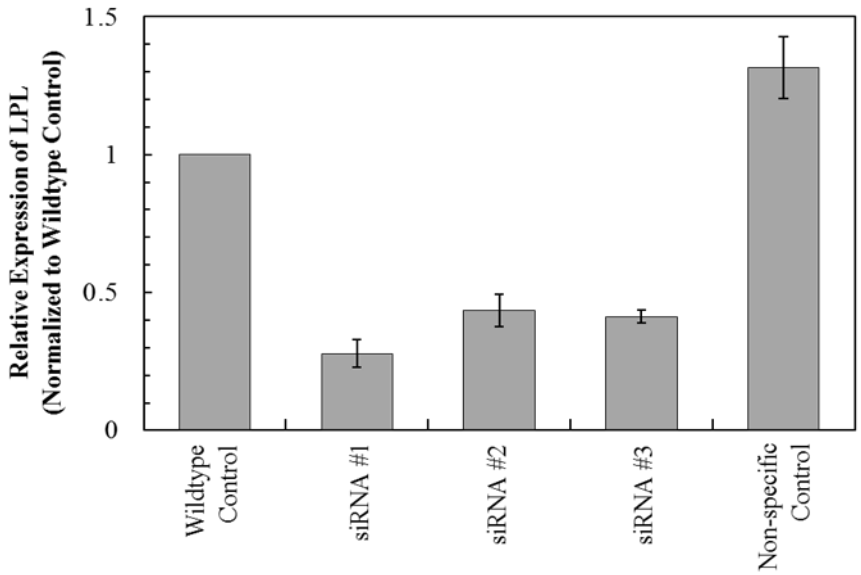


Figure 3.1: Impact of siRNA treatment on relative LPL expression, measured by MRM. Error bars represent the standard error of the mean from two biological replicates.

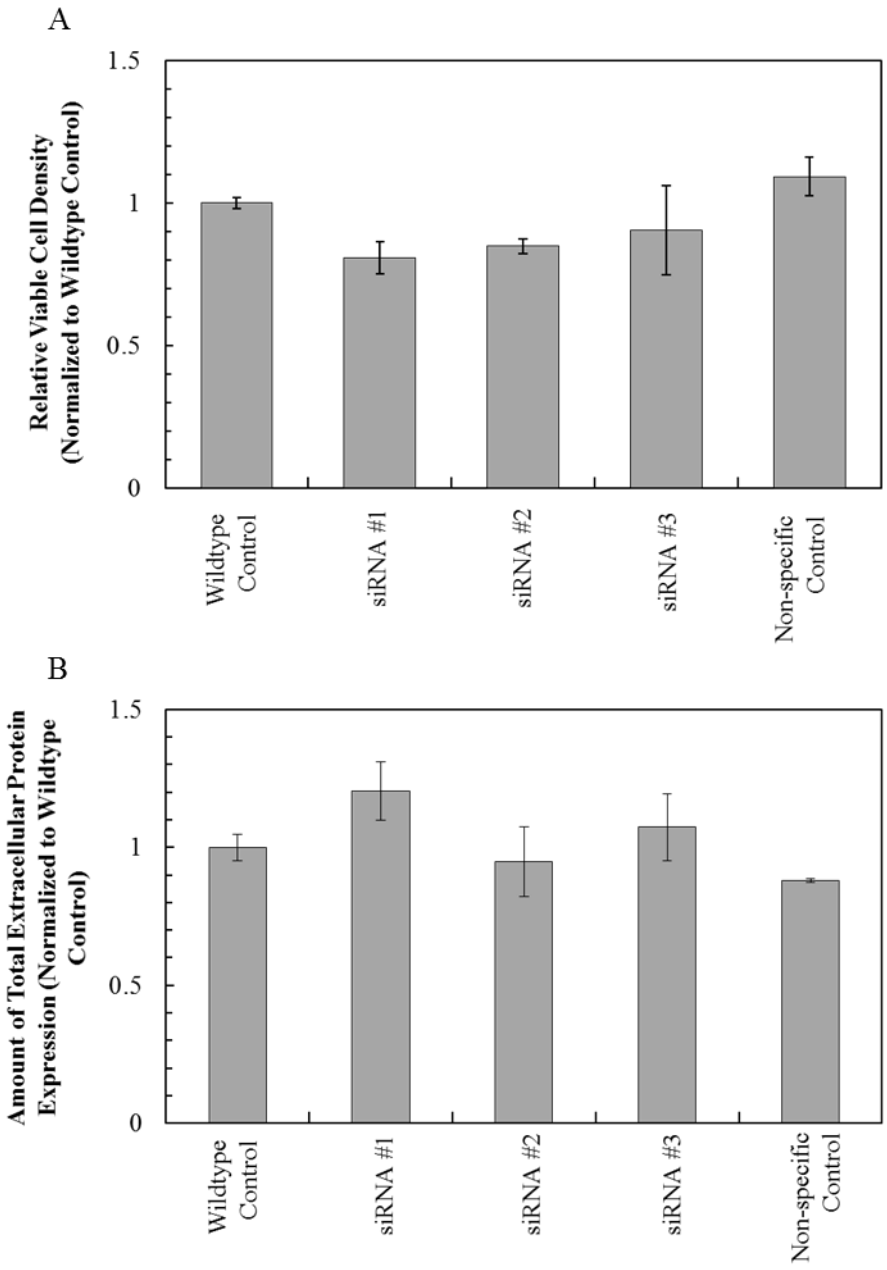


Figure 3.2: Cell culture performance following siRNA interference. (A) Viable cell density and (B) total extracellular protein expression at day 3 after siRNA interference. Error bars represent the standard error of the mean from two biological replicates.

The impact of reduced LPL on formulation stability, more specifically PS80 degradation, was examined by measuring the free fatty acid concentration. PS80 degradation is measured indirectly by measuring free fatty acid concentration because enzymatic hydrolysis of PS80 yields oleic acid and an alcohol component. PS80 with a final concentration of 0.03% w/w was incubated with concentrated CHO HCCF obtained from wildtype CHO cells or HCCF pooled from cells treated with *lpl*-specific siRNA. The samples were incubated at an elevated temperature of 37 °C to accelerate the formation of free fatty acids. The concentrations of free fatty acids from each sample were examined following 24 hours of incubation (Figure 3.3). CHO HCCF from CHO wildtype control degraded 19% (44.3 μM) of total PS80, while the amount of PS80 degraded by CHO HCCF from the pooled siRNA samples was minimal and under the limit of detection (13%). Although the exact protein composition of HCCF used in the PS80 degradation assay is unknown, the negligible amount of PS80 degradation following reduction of LPL expression suggests that LPL may play a significant role in PS80 hydrolysis.

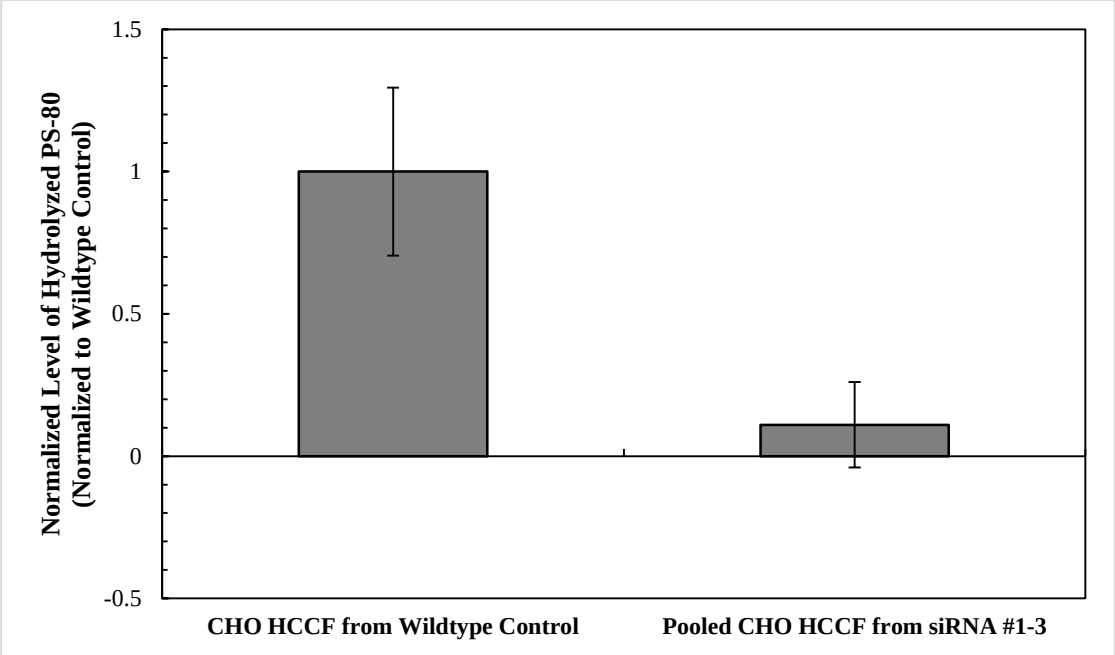


Figure 3.3: Effect of low LPL expression on the formation of free fatty acids from enzymatic hydrolysis of PS80 during incubation. Error bars represent the standard error of the mean from two biological replicates.

3.5 Conclusions

LPL is an HCP that has been shown to associate with multiple mAbs and co-purify with mAbs on chromatographic resins. LPL that had been retained in the final drug product can affect the stability of polysorbates, which are present in biopharmaceutical formulations. LPL knockdown by siRNA interference demonstrated that reduced LPL expression was not detrimental to cell growth and protein expression. This work also suggests that LPL may contribute to PS80 degradation through enzymatic hydrolysis. Based on these results, we continued the study through a functional LPL knockout using CRISPR/Cas9, where we investigated the impact of LPL on polysorbate degradation with five different CHO-LPL knockout clones.

REFERENCES

Doneanu, C. E., Xenopoulos, A., Fadgen, K., Murphy, J., Skilton, S. J., Prentice, H., Stapels, M., & Chen, W. (2012). Analysis of host-cell proteins in biotherapeutic proteins by comprehensive online two-dimensional liquid chromatography/mass spectrometry. *mAbs, 4*(1), 24-44.

Doshi, N., Demeule, B., & Yadav, S. (2015). Understanding Particle Formation: Solubility of Free Fatty Acids as Polysorbate 20 Degradation Byproducts in Therapeutic Monoclonal Antibody Formulations. *Mol Pharm, 12*(11), 3792-3804.

Gstaiger, M., & Aebersold, R. (2009). Applying mass spectrometry-based proteomics to genetics, genomics and network biology. *Nat Rev Genet, 10*(9), 617-627.

Ilko, D., Braun, A., Germershaus, O., Meinel, L., & Holzgrabe, U. (2015). Fatty acid composition analysis in polysorbate 80 with high performance liquid chromatography coupled to charged aerosol detection. *Eur J Pharm Biopharm, 94*, 569-574.

Janas, M. M., Wang, B., Harris, A. S., Aguiar, M., Shaffer, J. M., Subrahmanyam, Y. V., Behlke, M. A., Wucherpfennig, K. W., Gygi, S. P., Gagnon, E., & Novina, C. D. (2012). Alternative RISC assembly: binding and repression of microRNA-mRNA duplexes by human Ago proteins. *RNA, 18*(11), 2041-2055.

Kerwin, B. A. (2008). Polysorbates 20 and 80 used in the formulation of protein biotherapeutics: structure and degradation pathways. *J Pharm Sci, 97*(8), 2924-2935.

Kim, S. H., & Lee, G. M. (2007). Down-regulation of lactate dehydrogenase-A by siRNAs for reduced lactic acid formation of Chinese hamster ovary cells producing thrombopoietin. *Appl Microbiol Biotechnol, 74*(1), 152-159.

Kishore, R. S., Kiese, S., Fischer, S., Pappenberger, A., Grauschopf, U., & Mahler, H. C. (2011). The degradation of polysorbates 20 and 80 and its potential impact on the stability of biotherapeutics. *Pharm Res, 28*(5), 1194-1210.

Kishore, R. S., Pappenberger, A., Dauphin, I. B., Ross, A., Buergi, B., Staempfli, A., & Mahler, H. C. (2011). Degradation of polysorbates 20 and 80: studies on thermal autoxidation and hydrolysis. *J Pharm Sci, 100*(2), 721-731.

Labrenz, S. R. (2014). Ester hydrolysis of polysorbate 80 in mAb drug product: evidence in support of the hypothesized risk after the observation of visible particulate in mAb formulations. *J Pharm Sci, 103*(8), 2268-2277.

Levy, N. E. (2014). Host cell protein impurities and protein-protein interactions in downstream purification of monoclonal antibodies. *Ph.D. dissertation, University of Delaware*.

Levy, N. E., Valente, K. N., Choe, L. H., Lee, K. H., & Lenhoff, A. M. (2014). Identification and characterization of host cell protein product-associated impurities in monoclonal antibody bioprocessing. *Biotechnol Bioeng, 111*(5), 904-912.

Liebler, D. C., & Zimmerman, L. J. (2013). Targeted quantitation of proteins by mass spectrometry. *Biochemistry, 52*(22), 3797-3806.

Lim, S. F., Chuan, K. H., Liu, S., Loh, S. O., Chung, B. Y., Ong, C. C., & Song, Z. (2006). RNAi suppression of Bax and Bak enhances viability in fed-batch cultures of CHO cells. *Metab Eng, 8*(6), 509-522.

Maggio, E. T. (2012). Polysorbates, Peroxides, Protein Aggregation, and Immunogenicity: A Growing Concern. *J Excip Food Chem, 3*(2), 45-53.

Marichal-Gallardo, P. A., & Alvarez, M. M. (2012). State-of-the-art in downstream processing of monoclonal antibodies: process trends in design and validation. *Biotechnol Prog, 28*(4), 899-916.

Menetret, J. F., Schaletzky, J., Clemons, W. M., Jr., Osborne, A. R., Skanland, S. S., Denison, C., Gygi, S. P., Kirkpatrick, D. S., Park, E., Ludtke, S. J., Rapoport, T. A., & Akey, C. W. (2007). Ribosome binding of a single copy of the SecY complex: implications for protein translocation. *Mol Cell, 28*(6), 1083-1092.

Tomlinson, A., Demeule, B., Lin, B., & Yadav, S. (2015). Polysorbate 20 Degradation in Biopharmaceutical Formulations: Quantification of Free Fatty Acids, Characterization of Particulates, and Insights into the Degradation Mechanism. *Mol Pharm, 12*(11), 3805-3815.

Valente, K. N. (2014). Optimization and application of proteomic methods for characterization of host cell protein impurities from Chinese hamster ovary cells. *Ph.D. dissertation, University of Delaware*.

Valente, K. N., Lenhoff, A. M., & Lee, K. H. (2015). Expression of difficult-to-remove host cell protein impurities during extended Chinese hamster ovary cell culture and their impact on continuous bioprocessing. *Biotechnol Bioeng, 112*(6), 1232-1242.

Valente, K. N., Schaefer, A. K., Kempton, H. R., Lenhoff, A. M., & Lee, K. H. (2014). Recovery of Chinese hamster ovary host cell proteins for proteomic analysis. *Biotechnol J, 9*(1), 87-99.

Wu, S. C. (2009). RNA interference technology to improve recombinant protein production in Chinese hamster ovary cells. *Biotechnol Adv, 27*(4), 417-422.

Chapter 4

OPTIMIZATION OF POLYSORBATE DEGRADATION ASSAY BY CHINESE HAMSTER OVARY HOST CELL PROTEIN

4.1 Introduction

This chapter presents a design of experiment approach to optimize sample preparation setup to improve the reproducibility of the polysorbate 80 (PS80) assay. The factors considered to affect the reproducibility of the assay were the material of the container (glass, polypropylene, polypropylene in thermomixer), incubation time (6 hr, 12 hr, 24 hr), total sample volume (40 µL, 270 µL, 500 µL), and percentage of PS80 in the sample (0.03%, 0.075%, 0.12%). A full factorial design is incorporated to span the different combination of conditions that is explored. This work uses a design of experiment (DOE) approach to systematically determine the optimal condition that minimizes the assay's limit of quantitation.

4.2 Materials and Methods

4.2.1 CHO Cell Culture

A null CHO-K1 cell line (ATCC CCL-61), previously adapted to serum-free suspension culture (Valente, 2014), was cultured in 125-mL shake flasks containing 25-mL SFM4CHO medium (GE Healthcare Life Sciences, Little Chalfont, UK). Cultures were seeded at 5×10^4 cells/mL and incubated for 4 days in a 37°C cell culture incubator with 5% CO_2 and 80% relative humidity. CHO harvested cell culture fluid (HCCF) were separated from the cells by centrifugation (180 x g, 10 min) and stored at -20°C.

4.2.2 Experimental Design and Data Analysis

The factors varied considered to affect the reproducibility of the assay were the material of the container, incubation time, total sample volume, and initial PS80 concentration in the sample. The material of the container, total sample volume, and PS80 were three conditions screened by a three-factor full factorial design with two levels for each factor. One center point was included in the center of the design space. The design space consists of 15 different experimental set-ups (Table 4.1). Experimental design and statistical analysis were performed using JMP Pro 9 (SAS Institute Inc., Cary, NC, USA). Box plots were generated in JMP Pro 9 for visual interpretation.

Table 4.1: Experimental set-up of PS80 degradation assay. HCP samples in either glass or polypropylene container with varying sample volume and PS80 concentration.

Material	Sample Volume, µL	PS80 Concentration
Incubator, Glass	40	0.03%
		0.12%
	270	0.075%
	500	0.03%
		0.12%
Incubator, Polypropylene	40	0.03%
		0.12%
	270	0.075%
	500	0.03%
		0.12%
Thermomixer, Polypropylene	40	0.03%
		0.12%
	270	0.075%
	500	0.03%
		0.12%

4.2.3 PS80 Degradation Assay

CHO HCCF harvested CHO-K1 cells were subjected to concentration and buffer exchange using a 4-mL capacity centrifugal filtration device with 10 kDa MWCO membranes (EMD Millipore). The centrifugal filtration device was wetted twice with 4 mL ddH_2O and centrifuged at 5000 RPM for 15 minutes. It was then filled with 4 mL of CHO HCCF and centrifuged at 5000 RPM for 20 minutes. The sample was then subjected to three rounds of buffer exchange, where 3.9 mL of 50

mM Bis-tris, 10 mM $CaCl_2$, pH 6.8 buffer was added to the centrifugal filtration device and centrifuged at 5000 RPM, 20 minutes for two rounds and then 15 minutes for the third round. The protein concentration of buffer-exchanged HCCF was measured by Bradford assay.

Buffer exchanged CHO HCCF (10 μg/mL) was incubated with 50 mM Bis-tris, 10 mM $CaCl_2$, pH 6.8 buffer and PS80 at 37 °C with mixing in the thermomixer set at 700 rpm and mixing in the incubator set at 400 rpm. The degradation of PS80 was measured using the EnzyChrom Free Fatty Acid Kit (BioAssay Systems, Hayward, CA, USA), which measures the concentration of fatty acid released following polysorbate hydrolysis.

4.3 Results and Discussion

4.3.1 Screening and Optimization DOEs

The impact of four factors (material, sample volume, PS80 concentration, incubation time) on reproducibility of the PS80 degradation assay was screened by a full factorial design (Table 4.2) to determine the conditions that are most reproducible. The reproducibility is represented by the standard error of the mean of each condition.

Table 4.2: Full factorial design with corresponding responses for reproducibility.

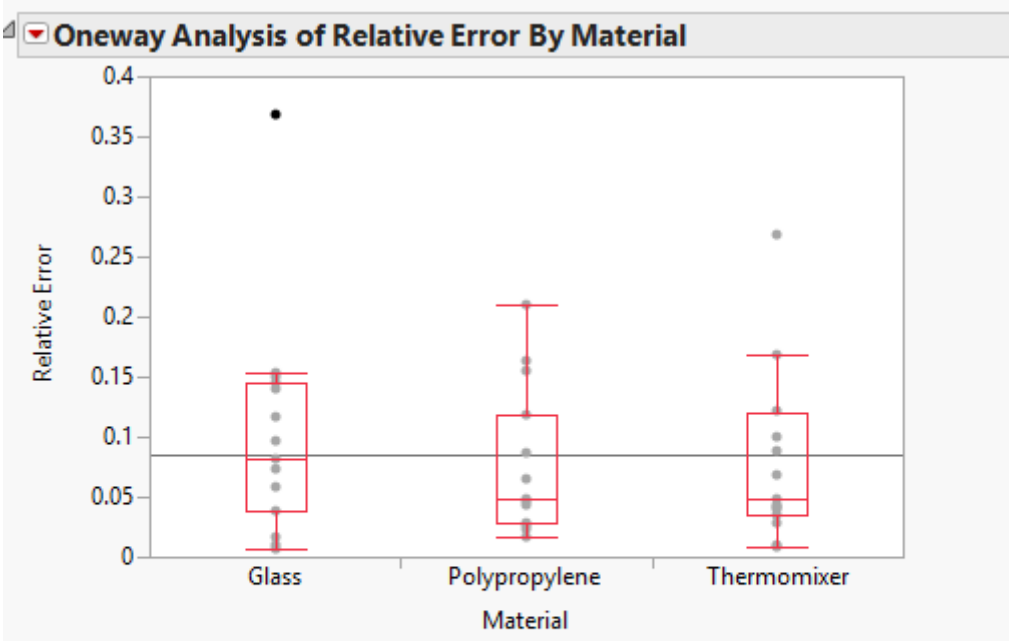
Material	Sample Volume, uL	PS80, %	Time Incubation, hr	Error
Glass	40	0.03	6	0.082
Glass	40	0.03	12	0.038
Glass	40	0.03	24	0.368
Glass	40	0.12	6	0.007
Glass	40	0.12	12	0.017
Glass	40	0.12	24	0.149
Glass	270	0.075	6	0.038
Glass	270	0.075	12	0.073
Glass	270	0.075	24	0.096
Glass	500	0.03	6	0.058
Glass	500	0.03	12	0.117
Glass	500	0.03	24	0.153
Glass	500	0.12	6	0.140
Glass	500	0.12	12	0.009
Glass	500	0.12	24	0.146
Polypropylene	40	0.03	6	0.087
Polypropylene	40	0.03	12	0.048
Polypropylene	40	0.03	24	0.043
Polypropylene	40	0.12	6	0.210
Polypropylene	40	0.12	12	0.066
Polypropylene	40	0.12	24	0.154
Polypropylene	270	0.075	6	0.048
Polypropylene	270	0.075	12	0.065
Polypropylene	270	0.075	24	0.119
Polypropylene	500	0.03	6	0.027
Polypropylene	500	0.03	12	0.028
Polypropylene	500	0.03	24	0.163
Polypropylene	500	0.12	6	0.046
Polypropylene	500	0.12	12	0.024
Polypropylene	500	0.12	24	0.017

Table 4.2 continued: Full factorial design with corresponding responses for reproducibility.

Material	Sample Volume, uL	PS80, %	Time Incubation, hr	Error
Thermomixer	40	0.03	6	0.008
Thermomixer	40	0.03	12	0.068
Thermomixer	40	0.03	24	0.040
Thermomixer	40	0.12	6	0.048
Thermomixer	40	0.12	12	0.029
Thermomixer	40	0.12	24	0.169
Thermomixer	270	0.075	6	0.044
Thermomixer	270	0.075	12	0.042
Thermomixer	270	0.075	24	0.121
Thermomixer	500	0.03	6	0.121
Thermomixer	500	0.03	12	0.034
Thermomixer	500	0.03	24	0.268
Thermomixer	500	0.12	6	0.088
Thermomixer	500	0.12	12	0.009
Thermomixer	500	0.12	24	0.100

4.3.2 Analysis of reproducibility

A one-way analysis was performed using JMP to determine which material contribute the least variability (Figure 4.1). The lowest relative error is observed with polypropylene. Incubation with polypropylene led to a low median error (4.78%), and low relative error at 25, 75, and 90 percentile. Higher relative error is seen for incubation with glass and polypropylene in thermomixer (thermomixer). The material that contributed to the highest variability was glass. Incubation with glass led to a high median error (8.16%), and high relative error at 25, 75, and 90 percentile.

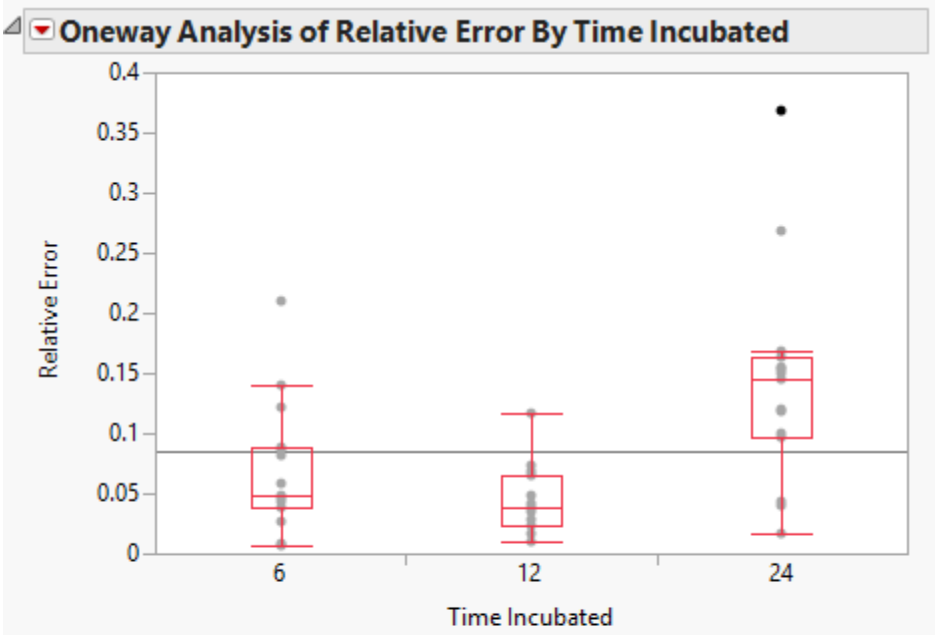


Quantiles

Level	Minimum	10%	25%	Median	75%	90%	Maximum
Glass	0.006993	0.008313	0.03805	0.08161	0.145598	0.238869	0.367852
Polypropylene	0.017283	0.021273	0.028169	0.047852	0.118796	0.181686	0.20994
Thermomixer	0.00834	0.008877	0.034294	0.048413	0.120679	0.208314	0.267918

Figure 4.1: One-way analysis of relative error by material.

A one-way analysis was performed using JMP to determine the optimal incubation time (Figure 4.2). The lowest relative error is observed following 12 hours of incubation. Incubation for 12 hours led to low median error (3.85%), and low relative error at 25, 75, and 90 percentile. The highest relative error is observed following 24 hours of incubation, where a high median (14.55%) and high relative error at 25, 75, and 90 percentile are observed.



Quantiles

Level	Minimum	10%	25%	Median	75%	90%	Maximum
6	0.006993	0.007801	0.03805	0.048413	0.088401	0.167949	0.20994
12	0.009193	0.009218	0.023933	0.038462	0.065589	0.090952	0.117211
24	0.017283	0.030879	0.096215	0.145598	0.16285	0.307892	0.367852

Figure 4.2: One-way analysis of relative error by incubation time.

A one-way analysis was not performed for PS80% because low PS80 concentration (0.03% w/w) is applied in biopharmaceutical formulation to prevent surface adsorption and protein aggregation. Higher PS80 concentration can cause protein unfolding (Marichal-Gallardo & Alvarez, 2012). The relative errors of all conditions were fitted to a polynomial equation using JMP's Prediction profiler, which maps the changes to the response (relative error) against each factor (Figure 4.3). The Prediction Profiler shows that the total sample volume did not affect the relative error.

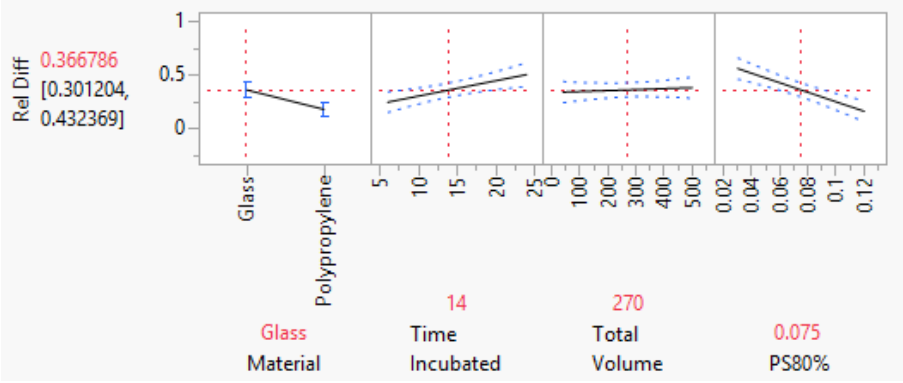
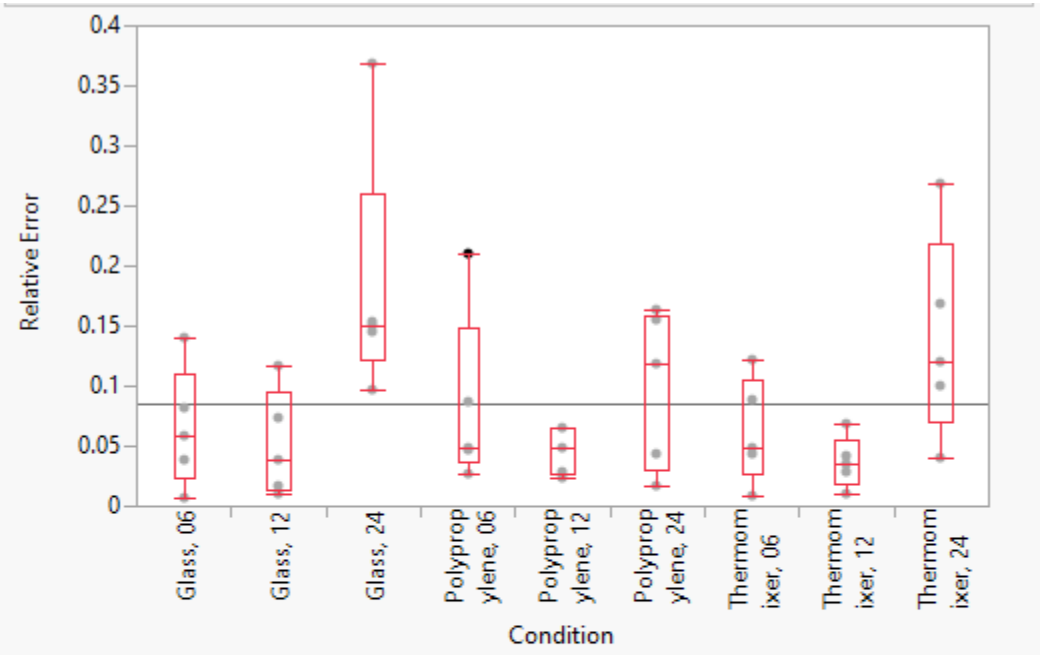


Figure 4.3: The Prediction Profiler maps the changes in the response (error) against material, time incubated, total volume, and PS80 concentration.

An additional one-way analysis was performed for both material and incubation time (Figure 4.4). Incubation with a polypropylene container for 12 hours produced the lowest limit of quantitation (4.77%). Optimized conditions of the PS80 degradation assay were shown to depend on the time of incubation. A high relative error is observed following 24 hours of incubation across all conditions. The variability in the samples could originate from chemical degradation of PS80, like autoxidation. The initiation of autoxidation in PS80 could occur by various means, such as residual peroxides and radical initiation in the presence of oxygen by light.



Quantiles

Level	Minimum	10%	25%	Median	75%	90%	Maximum
Glass, 06	0.006993	0.006993	0.022522	0.057984	0.110783	0.139955	0.139955
Glass, 12	0.009193	0.009193	0.012916	0.038462	0.095329	0.117211	0.117211
Glass, 24	0.096215	0.096215	0.120906	0.14934	0.260366	0.367852	0.367852
Polypropylene, 06	0.026984	0.026984	0.036652	0.047852	0.148594	0.20994	0.20994
Polypropylene, 12	0.023933	0.023933	0.026051	0.047722	0.065372	0.065589	0.065589
Polypropylene, 24	0.017283	0.017283	0.030055	0.118796	0.158576	0.16285	0.16285
Thermomixer, 06	0.00834	0.00834	0.026011	0.048413	0.104853	0.121305	0.121305
Thermomixer, 12	0.009235	0.009235	0.018996	0.034294	0.055004	0.068041	0.068041
Thermomixer, 24	0.039943	0.039943	0.070088	0.120679	0.218248	0.267918	0.267918

Figure 4.4: One-way analysis of relative error by both material and incubation time.

4.4 Concluding Remarks

This research used a systematic approach for the optimization of the incubation conditions in the PS80 degradation assay. Although these optimized conditions are specific for PS80, the DOE approach presented here can be used to minimize the limit of quantitation for any assay. These optimized conditions presented here will be used for subsequent polysorbate studies.

REFERENCES

Marichal-Gallardo, P. A., & Alvarez, M. M. (2012). State-of-the-art in downstream processing of monoclonal antibodies: process trends in design and validation. *Biotechnol Prog, 28*(4), 899-916.

Valente, K. N. (2014). Optimization and application of proteomic methods for characterization of host cell protein impurities from Chinese hamster ovary cells. *Ph.D. dissertation, University of Delaware*.

Chapter 5

A FUNCTIONAL LPL KNOCKOUT IN CHINESE HAMSTER OVARY CELLS REDUCES POLYSORBATE DEGRADATION

5.1 Introduction

In Chapter 3, we demonstrated that knockdown of LPL expression using siRNA interference led to reduced polysorbate degradation and did not impair cell culture. This motivates further investigation with a fully functional *lpl* knockout. In this study, we demonstrate the application of CRISPR/Cas9 on altering the *lpl* gene, preventing the production of LPL protein. The aims of this study are to (i) generate engineered CHO-K1 cell lines with functional *lpl* knockout and (ii) study the impact of LPL on formulation stability. Here, we explored the impact of a functional LPL knockout on CHO cell culture performance, such as cell growth and total protein production, and on polysorbate (PS80 and PS20) degradation.

5.2 Materials and Methods

5.2.1 Single guide RNA (sgRNA) target design and plasmid construction

The Cas9 expression vector was obtained from Addgene (Addgene #41815). The sgRNA target selections for *lpl* were filtered using a bioinformatics tool, CRISPy (Ronda et al., 2014). The *lpl* target sites are presented in Table 5.1. The target sgRNA expression vectors were constructed by cloning 455 bp gBlocks from IDT into pCR-Blunt-II TOPO vector (Invitrogen, Carlsbad, CA, USA) according to manufacturer's recommendations. The gBlock sequences (Integrated DNA Technologies, Coralville, IA, USA) for sgRNA construction are listed in Table 5.2. This 455 bp fragment contains all components necessary for sgRNA expression, specifically the U6 promoter (gray), target sequence (green), single guide RNA scaffold (blue), and termination signal (red). The sgRNA expression vectors were transformed into *Escherichia coli* cells One Shot TOP10 competent cells (Invitrogen) according to manufacturer's recommendations. Transformed cells containing Cas9 and sgRNA expression vectors were selected on 100 µg/mL ampicillin LB plates and 50 µg/mL kanamycin LB plates, respectively. The plasmids were purified using QIAprep Spin Miniprep Kit (Qiagen, Hilden, Germany) and the construct sequences were verified by Sanger sequencing (DBI sequencing center, Newark, DE, USA).

Table 5.1: sgRNA target regions and target sequences for *lpl* with the corresponding number of possible 1 base pair (bp) off-targets.

	Region	Target sequence	1 bp off-targets
sg1	Exon 5	5′-GGGTCACCTGGCCGAAGTAT-3′	23
sg2	Exon 6	5′-GAGATGAATGGATCGCTCAT-3′	23
sg3	Exib 8	5′-GCTTTGTCATCGAGAAGATT-3′	51

Table 5.2: sgRNA gBlock construct for *lpl*. The components for sgRNA expression are the U6 promoter (gray), target sequence (green), single guide RNA scaffold (blue), and termination signal (red).

gB_lpl_1	TGTACAAAAAAGCAGGCTTTAAAGGAACCAATTCAGTCGACTGGATCCGG TACCAAGGTCGGGCAGGAAGAGGGCCTATTTCCCATGATTCCTTCATATT TGCATATACGATACAAGGCTGTTAGAGAGATAATTAGAATTAATTTGACT GTAAACACAAAGATATTAGTACAAAATACGTGACGTAGAAAGTAATAATT TCTTGGGTAGTTTGCAGTTTTAAAATTATGTTTTAAAATGGACTATCATA TGCTTACCGTAACTTGAAAGTATTTCGATTTCTTGGCTTTATATATCTTG TGGAAAGGACGAAACACC**GGGTCACCTGGCCGAAGTAT**GTTTTAGAGCTA GAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGG CACCGAGTCGGTGCTTTTTTTCTAGACCCAGCTTTCTTGTACAAAGTTGG CATTA
gB_lpl_2	TGTACAAAAAAGCAGGCTTTAAAGGAACCAATTCAGTCGACTGGATCCGG TACCAAGGTCGGGCAGGAAGAGGGCCTATTTCCCATGATTCCTTCATATT TGCATATACGATACAAGGCTGTTAGAGAGATAATTAGAATTAATTTGACT GTAAACACAAAGATATTAGTACAAAATACGTGACGTAGAAAGTAATAATT TCTTGGGTAGTTTGCAGTTTTAAAATTATGTTTTAAAATGGACTATCATA TGCTTACCGTAACTTGAAAGTATTTCGATTTCTTGGCTTTATATATCTTG TGGAAAGGACGAAACACC**GAGATGAATGGATCGCTCAT**GTTTTAGAGCTA GAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGG CACCGAGTCGGTGCTTTTTTTCTAGACCCAGCTTTCTTGTACAAAGTTGG CATTA
gB_lpl_3	TGTACAAAAAAGCAGGCTTTAAAGGAACCAATTCAGTCGACTGGATCCGG TACCAAGGTCGGGCAGGAAGAGGGCCTATTTCCCATGATTCCTTCATATT TGCATATACGATACAAGGCTGTTAGAGAGATAATTAGAATTAATTTGACT GTAAACACAAAGATATTAGTACAAAATACGTGACGTAGAAAGTAATAATT TCTTGGGTAGTTTGCAGTTTTAAAATTATGTTTTAAAATGGACTATCATA TGCTTACCGTAACTTGAAAGTATTTCGATTTCTTGGCTTTATATATCTTG TGGAAAGGACGAAACACC**GCTTTGTCATCGAGAAGATT**GTTTTAGAGCTA GAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGG CACCGAGTCGGTGCTTTTTTTCTAGACCCAGCTTTCTTGTACAAAGTTGG CATTA

5.2.2 CHO cell culture and transfection

A null CHO-K1 cell line (ATCC CCL-61), previously adapted to serum-free suspension culture (Valente, 2014), was cultured in 125-mL shake flasks containing 25-mL SFM4CHO medium (GE Healthcare Life Sciences, Little Chalfont, UK). Cultures were seeded at 5×10^4 cells/mL and incubated for 3 - 5 days in a 37°C cell culture incubator with 5% CO_2 and 80% relative humidity.

1×10^6 cells were simultaneously transfected with expression vectors for Cas9 and sgRNA targeting *lpl* using Nucleofector Kit V (Lonza, Basel, Switzerland) according to manufacturer's protocol. A transfection with pmaxGFP vector (Lonza) was applied to evaluate the transfection efficiency. Transfected cells were cultured for 48 hours to allow CRISPR/Cas9-mediated gene knockout. Single cell clones from cells treated with CRISPR and sgRNA plasmids were generated by limiting dilution into a 96-well plate to a target density of 0.5 cells/well. Single clones were grown at 37°C cell culture incubator with 5% CO_2 and 80% relative humidity for three weeks before expanded in 6-well plates. Genomic DNA was extracted from the remaining cell population using the QIAamp DNA Mini Kit (Qiagen) for a T7 endonuclease I (T7EI) assay.

The cells were counted using either a Fuchs-Rosenthal hemocytometer or Countess II (Invitrogen). Cell viability was determined by the Trypan blue exclusion method. CHO harvested cell culture fluid (HCCF) from wildtype and *lpl* knockout clones were separated from the cells by centrifugation (180 x g, 10 min) and stored at -20°C.

5.2.3 T7 Endonuclease I (T7EI) assay for targeting efficiency of CRISPR/Cas9

The overall targeting efficiency of CRISPR/Cas9 can be assessed using a T7EI assay. T7EI is an enzyme that recognizes and cleaves heteroduplex DNA, a double-stranded DNA with mismatched bases. Insertions or deletions (indels) derived from CRISPR-Cas9 activity were assessed by PCR amplification. Genomic regions around the CRISPR target site were amplified from the genomic DNA extracts using Phusion high-fidelity polymerase (New England Biolabs, Ipswich, MA, USA), by touchdown PCR (98 °C for 1 min; 11x: 98 °C for 10 s, 72–66 °C (-0.5 °C/cycle) for 15 s, 72 °C for 8 s; 30x: 98 °C for 10 s, 66 °C for 15 s, 72 °C for 8 s; 72 °C for 10 min) using PCR primers listed in Table A.2. The PCR products were subjected to a reannealing process to enable heteroduplex formation (95°C for 5 min; 95-85°C ramped at -2°C/s; 85-25°C ramped at -0.1°C/s; and held at 4 °C). Annealed PCR products were subsequently digested with T7EI (New England Biolabs) at 37°C for 20 minutes. The digests were analyzed on a 2% TAE gel and the percentage of indels was estimated from analysis of the cut and uncut gel bands with ImageJ.

5.2.4 Gene sequencing by TOPO cloning

The genomic regions covering the three *lpl* sgRNA target sites were PCR-amplified from the genomic extracts as described in the T7EI assay. PCR products were purified using the QIAquick PCR Purification Kit (Qiagen). Purified PCR products were TOPO-cloned into the pCR-Blunt II-TOPO vector using the Zero Blunt

TOPO PCR Cloning Kit (Invitrogen) and transformed into *Escherichia coli* cells One Shot TOP10 competent cells (Invitrogen). Transformed cells were then plated on lysogeny broth (LB)-kanamycin agar plates and grown at 37 °C overnight. Single colonies were picked and grown in LB medium with 50 ug/mL kanamycin at 37 °C overnight. Plasmids from single colonies were extracted using QIAprep Spin Miniprep Kit (Qiagen) and the sequences of the plasmids were verified by Sanger sequencing (DBI sequencing center, Newark, DE, USA) using the M13 (-27) reverse primer.

5.2.5 LPL protein expression analysis by western blot

HCP from CHO HCCF were precipitated with methanol by the addition of 2 volumes of methanol to 1 volume of CHO HCCF, followed by overnight incubation at -20 °C. The precipitated protein was recovered by centrifugation (5000 rpm, 65 min, 4 °C) and resolubilized in phosphate buffered saline (PBS). Concentrations of resolubilized HCPs were measured with a Bradford assay. 30 µg of CHO HCPs were first treated with SDS and DTT, followed by an incubation at 95 °C for 5 minutes, before loading to Mini-PROTEAN TGX precast gel (Biorad, Hercules, CA, USA). Proteins were separated by electrophoresis by applying a constant voltage, 150 V, for 1 hour, and were then transferred to 0.45 µm polyvinylidene difluoride membrane (Invitrogen) by applying a constant voltage, 100 V, for 1 hour. Membranes were blocked in 3% non-fat dry milk in Tris-buffered saline with TWEEN 20, pH 8.0 (TBST, Sigma-Aldrich, St. Louis, MO, USA) for 1 hour. A membrane was incubated overnight at 4°C with LPL N-terminus primary antibody (1:500 dilution), which

recognizes LPL residues 27-79 (Santa Cruz Biotechnology, Dallas, TX, USA). A second membrane was incubated overnight at 4°C with LPL C-terminus primary antibody (1:250 dilution), which recognizes LPL residues 297-326 (Abcam, Cambridge, UK). Following 3 x TBST washes, the membranes were incubated with alkaline phosphatase conjugated mouse anti-rabbit IgG (1:5,000 dilution, Sigma-Aldrich). Bound antibodies were detected using enhanced chemi-fluorescence (ECF, GE Healthcare Life Sciences) substrate following the manufacturer's instructions and imaged using a Typhoon FLA-7000 scanner.

5.2.6 CHO HCP Preparation and LPL-specific MRM Assay

HCP from CHO HCCF were precipitated with methanol by the addition of 2 volumes of methanol to 1 volume of CHO HCCF, followed by overnight incubation at -20 °C. The precipitated protein was recovered by centrifugation (5000 rpm, 65 min, 4 °C). Residual detergent was removed by DetergentOUT GBS10-800 detergent removal kit (G-Biosciences, St. Louis, MO, USA) according to the manufacturer's protocol. Trypsin digestion was performed as described (Valente, 2014). Peptide pellets were resolubilized in 0.1% trifluoroacetic acid (TFA, Fisher Scientific, Fair lawn, NJ, USA), loaded onto C18 ZipTips (EMD Millipore, Billerica, MA, USA) and eluted in 50% acetonitrile with 0.1% TFA (Fisher Scientific, Hampton, NH, USA). Concentrations of resolubilized HCPs were measured with a Bradford assay.

LC-MRM assay was performed on a QTrap 4000 (AB Sciex, Foster City, CA, USA) equipped with an UltiMate 3000 nLC system (Dionex, Sunnyvale, CA, USA).

Digested CHO HCPs were injected onto a C18 trap column (Dionex), and washed with 2% acetonitrile with 0.1% formic acid (Mallinckrodt Chemicals, Phillipsburg, NJ, USA) for 5 min with a flow rate of 30 µl/min, then eluted onto a C18 column (Acclaim PepMap100, 75 µm X 150 mm, 3µm, 100 Å, Dionex) by a program of 2 – 49% acetonitrile with 0.1% formic acid in 50 min, followed by 49% acetonitrile with 0.1% formic acid for 20 min. Column eluate was directly injected into a QTrap 4000 through a nanoSpray II source (AB Sciex) with an uncoated fused-silica Pico tip (New Objective, Woburn, MA, USA). The instrument was operated in positive ESI ion mode, with spray voltage at 2400 V and a source temperature of 150 °C, with MRM triggered enhanced resolution scan and enhanced product ion scans. MRM transitions were generated with Skyline v2.5 (MacLean et al. 2010) and monitored through Analyst 1.6.2 (AB Sciex) with parameters specified in Table 5.3. Raw MRM data were integrated for peak area with Skyline and normalized to `ITGLDPAGPNFEYAEAPSR` with c-terminal $^{13}C^{15}N$-labeled R. All analysis was performed with three technical replicates.

Table 5.3: MRM assay parameters. The ITGLDPAGPNFEYAEAPSR peptide targets LPL residues 151-173, while GLGDVDQLVK targets LPL residues 232-242.

	Target peptide sequence	Precursor (m/z)	Product (m/z)	ID	Scan time (ms)	CE (V)
LPL	ITGLDPAGPNFEYAEAPSR	1002.987	793.424	+2y7	20	72.9
	ITGLDPAGPNFEYAEAPSR	1002.987	753.355	+2y14+2	20	67.9
	ITGLDPAGPNFEYAEAPSR	1002.987	359.204	+2y3	20	52.9
	ITGLDPAGPNFEYAEAPSR	1007.991	803.392	+2y7	20	72.9
	ITGLDPAGPNFEYAEAPSR	1007.991	758.359	+2y14+2	20	67.9
	ITGLDPAGPNFEYAEAPSR	1007.991	369.212	+2y3	20	52.9
	GLGDVDQLVK	522.29	873.478	+2y8	20	30.5
	GLGDVDQLVK	522.29	701.429	+2y6	20	30.5
	GLGDVDQLVK	522.29	602.361	+2y5	20	30.5
SPARC	NVLVTLYER	553.814	780.425	+2y6	20	35.4
	NVLVTLYER	553.814	681.357	+2y5	20	35.4
	NVLVTLYER	553.814	467.225	+2y3	20	35.4

5.2.7 PS80 and PS20 Degradation Assay

CHO HCCF harvested from *lpl* knockout cells and the wildtype control were subjected to concentration and buffer exchange using a 4-mL capacity centrifugal filtration device with 10 kDa MWCO membranes (EMD Millipore). The centrifugal filtration device was wetted twice with 4 mL ddH_2O and centrifuged at 5000 RPM for 15 minutes. They were then filled with 4 mL of HCCF from each clone and centrifuged at 5000 RPM for 20 minutes. The samples were then subjected to three rounds of buffer exchange, where 3.9 mL of 50 mM Bis-tris, 10 mM $CaCl_2$, pH 6.8 buffer was added to each centrifugal filtration device and centrifuged at 5000 RPM, 20

minutes for two rounds and then 15 minutes for the third round. Protein concentrations of buffer-exchanged HCPs were measured by Bradford assay.

Buffer exchanged HCCF were incubated with 50 mM Bis-tris, 10 mM $CaCl_2$, pH 6.8 buffer and 246 μM (0.03% w/w) PS80 or 269 μM (0.03% w/w) PS20 at 37 °C for 12 hours with mixing at 400 rpm. The degradation of PS80 and PS20 were measured using the EnzyChrom Free Fatty Acid Kit (BioAssay Systems, Hayward, CA, USA), which measures the concentration of fatty acid released following polysorbate hydrolysis. The LPL inhibitor study includes 2 μM apolipoprotein C-III (apoC-III, Sigma-Aldrich) in the reaction mixture.

5.3 Results and Discussion

5.3.1 Rational design of sgRNA target sites

A rational design of the sgRNA target site is necessary to knockout LPL function while minimizing off-target genome-editing on other CHO genes (Fu et al., 2014; Kuscu et al., 2014; Ran et al., 2013). The design for CRISPR/Cas9-mediated genome editing requires a selection of target sites that are located in the exons of the *lpl* gene. These targets are filtered based on their location in the *lpl* gene and sequence similarity to the rest of the CHO-K1 genome (Ronda et al., 2014). A MATLAB script was written to facilitate the compilation of all possible target sequences in *lpl* (Appendix B). The compiled target sequences were compared to the results obtained by “CRISPy” (Ronda et al., 2014), an interactive bioinformatics tool that displays the

sequences' exon locations and calculates the number of off-target mismatches the sequence has, relative to the CHO-K1 genome. Each target sequence follows the format, $GN_{19}NGG$, where N is any nucleotide. The U6 promoter (G-), commonly used to express small RNAs, and Cas9 recognition site (-NGG) restrict other nucleotides of the target sequence.

The exon and its corresponding function were taken into consideration when selecting target sequences. The *lpl* gene is composed of 10 exons and 9 introns (Braun & Severson, 1992). Six exons with corresponding functional sites have been identified in the LPL protein, as shown in Table 5.4. The active site of LPL, encoded by exon 5, is essential for the hydrolysis of lipids to occur. The heparin-binding site, encoded by exon 6, serves as a bridge between the protein and lipoprotein. A study has shown that an inactive heparin-binding site diminishes LPL activity (Lutz et al., 2001). Newly synthesized LPL undergoes N-linked glycosylation in the endoplasmic reticulum to be catalytically active. An LPL protein that is not N-linked glycosylated is catalytically inactive and accumulates in the endoplasmic reticulum (Braun & Severson, 1992). Additionally, glycosylation is often required for the protein to be correctly folded and packaged. Improperly-folded LPL proteins tend to aggregate and remain in the endoplasmic reticulum, where rapid degradation of inactive LPL occurs. The selected target sequences for functional *lpl* knockout, as listed in Table 5.1, are present only in the *lpl* gene in the CHO-K1 genome and contain the lowest number of off-targets mismatch and are predicted to disrupt at least one important functional site. Exon 5, which encodes for the active site of LPL, exon 6, which encodes for the heparin binding domain, and exon 8, which encodes for the N-linked glycosylation site, were the three target domains for CRISPR/Cas9 genome editing.

Table 5.4: Corresponding functional site of *lpl* exons.

Exon	Functional site
1	5’ untranslated region and signal peptide
2	N-linked glycosylation site (Asn-Xaa-Ser)
4	Interfacial lipid binding region
5	Active site (Gly-Xaa-Ser-Xaa-Gly)
6	Heparin-binding domain
8	N-linked glycosylation site (Asn-Xaa-Thr)

5.3.2 CRISPR/Cas9 shows targeted endonuclease activity in the *lpl* target regions in CHO cells

Three sgRNA expression vectors, each specific for their respective *lpl* exon target, were constructed and individually co-transfected with a Cas9 expression vector in CHO-K1 cells. The efficacy of CRISPR-mediated genome-editing was determined three days post-transfection with CHO-K1 cells transfected with pmaxGFP plasmids. The transfection efficiency of Lonza’s nucleofector kit was determined to be 91.3% with cells at 62.6% viability (Figure A.1).

The percentages of cells with indels were measured by a T7EI digestion assay (Figure 5.1). PCR products from the cell populations treated with CRISPR/Cas9 were stained with SYBR Green and visualized with a DNA gel imaging system. Indel frequencies (%) were analyzed using ImageJ. The indel percentages observed in CRISPR-treated cell populations were 7.0%, 1.1%, and 0.0% for cells transfected with sg1, sg2, and sg3, respectively. Indels not observed in cells treated with sg3 may be contributed by the instability of sgRNA formation, which depends on folding energy

and G-quadruplex formation (Moreno-Mateos et al., 2015). CRISPR/Cas9 is a relatively new technology, so while there are many tools that offer computational predictions of the best sgRNA, the parameters affecting the specificity and efficiency of the system are still unclear and have not been optimized.

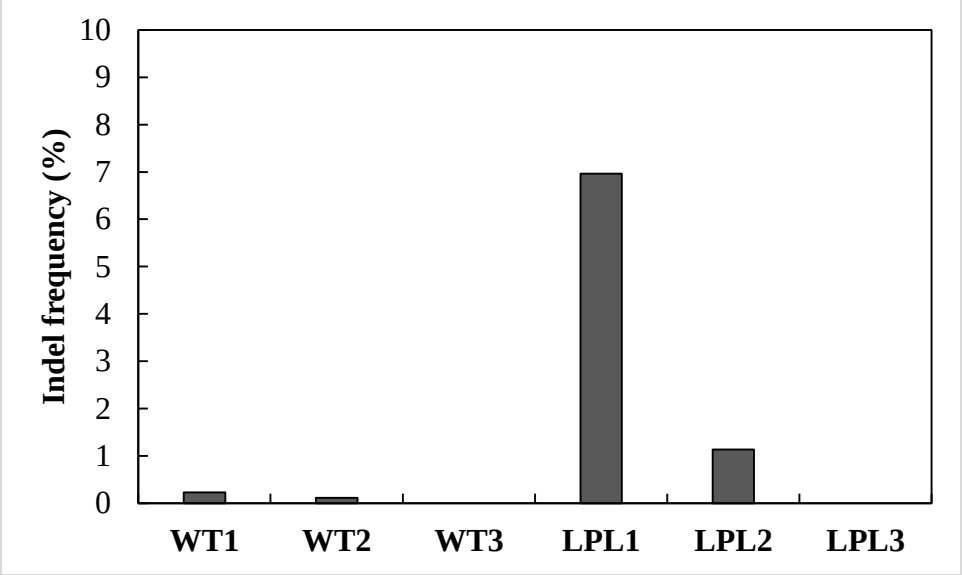


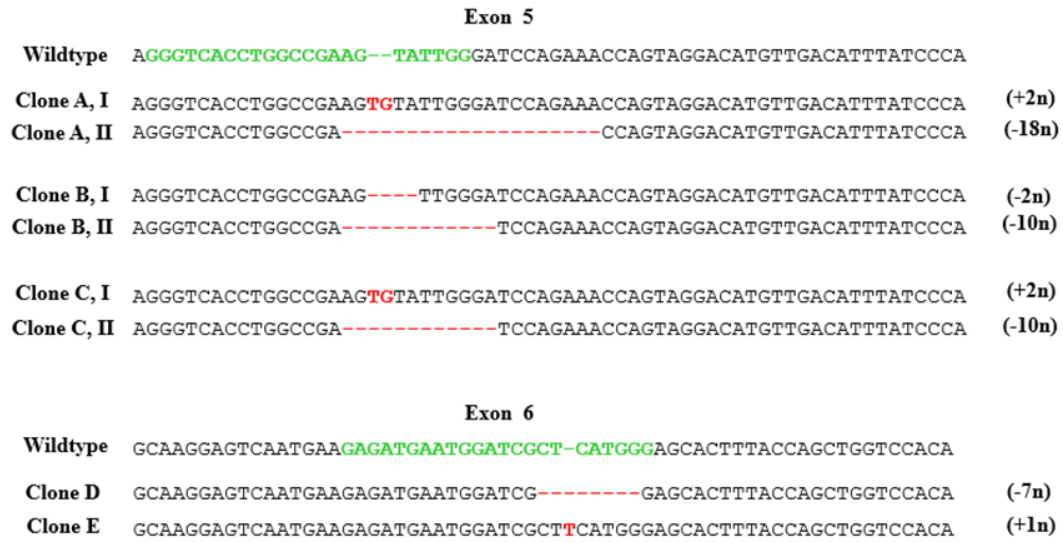
Figure 5.1: Analysis of indels in *lpl*. The frequency of indels generated at *lpl* target sites upon transient transfection of Cas9 and their respective sgRNAs. The percentage of indels in the cell population is measured by analyzing the intensity of both fragmented and unfragmented bands. WT1, WT2, and WT3 represent the genomic regions of target sites 1 (exon 5), 2 (exon 6), and 3 (exon 8), respectively, in wildtype cells. LPL1, LPL2, LPL3 represent the genomic regions of target sites in cells transfected with sg1, sg2, and sg3, respectively. The results are from one biological replicate.

5.3.3 Generation of LPL knockout CHO-K1 cell lines

Single cells from cell populations transfected with sg1 and sg2 were expanded to generate clonal cell lines by limiting dilution. Genomic DNA from expanded clonal

cell lines was extracted and the *lpl* gene region of interest was PCR amplified. The sequences were then sent to the DBI sequencing center for Sanger sequence analysis. Three clones (A, B, C) targeting exon 5 and two clones (D, E) targeting exon 6 showed indels in LPL (Figure 5.2A). Clones A, B, and C show two different mutations, while clones D and E show one mutation in *lpl*. This suggests that clones A, B, and C may contain two subpopulations with changes in the *lpl* gene, while clones D and E contain only one subpopulation. Based on the indels, changes to the amino acid composition and the molecular weight of the LPL protein were predicted (Figure 5.2B). Four of the five clones (clones B, C, D, and E) contain indels in the *lpl* gene that could lead to a frameshift mutation and, if expressed, a truncated LPL protein. The truncated LPL protein is predicted to be 197 to 251 amino acids shorter than the native LPL protein, with regions near the C-terminus deleted. Clone A contain one allele that leads to a 6-amino acid deletion. However, the predicted amino acid composition of LPL for clone A revealed that the deletion occurs at the active site. More specifically, the serine and second glycine of the Gly-Xaa-Ser-Xaa-Gly active site are deleted. The five CHO *lpl* knockout cell lines were selected for further characterization.

A



B

Clone	Nucleotide Change	AA length	MW (kDa)
Wildtype	0	473	53.0
Clone A, I	+2	226	24.7
Clone A, II	-18	467	52.2
Clone B, I	-2	228	25.0
Clone B, II	-10	222	24.4
Clone C, I	+2	226	24.7
Clone C, II	-10	222	24.4
Clone D	-7	276	29.9
Clone E	+1	275	29.9

Figure 5.2: Genome editing of *lpl* in CHO-K1 cells by CRISPR/Cas9. (A) Sanger sequences of *lpl* exons 5 and 6 obtained from five CHO *lpl* knockout clones. Subpopulations were designated as I and II. The change to *lpl* is denoted by + or -, indicating addition or deletion of nucleotides, respectively. The target sequences are highlighted in green. Specific changes to the gene are highlighted in red. (B) Predicted length and expected molecular weight of LPL protein. AA = amino acid, MW = molecular weight, measured in kDa.

5.3.4 Characterization of CHO-K1 lpl knockout cell lines by western analysis and LPL-specific MRM assay

To investigate whether the genome modifications introduced by CRISPR/Cas9 resulted in the expression of a native LPL protein, the five CHO-K1 *lpl* knockout cell lines with validated genotypes described above were analyzed by western to assess their relative expression of the N-terminal and C-terminal regions of the LPL protein. CHO HCCF from each cell culture was collected on Day 4 and LPL protein expressions were analyzed by western using two LPL antibodies, N-terminus LPL antibody and C-terminus LPL antibody (Figure 5.3A). The relative expression of N-terminus and C-terminus LPL were analyzed using ImageJ (Figure 5.3B).

Western analysis shows reduced expression of LPL near the expected molecular weight of 53 kDa from clone A, with 40% relative expression of the N-terminus LPL and 22% relative expression of the C-terminus LPL. Clone A is predicted to express two modified LPL proteins with molecular weights 24.7 and 52.2 kDa. Previous sequencing analysis suggests that this protein could be a modified LPL with 6-amino acid deletion of molecular weight 52.2 kDa. The presence of the modified 24.7 kDa LPL protein was not observed in clone A. Neither native LPL expression, at 53 kDa, nor modified LPL expressions were detected in clone B. Clone B is predicted to express two modified LPL proteins molecular weights 25.0 and 24.4 kDa. Similarly, neither native LPL expression nor modified LPL expression were detected in clones C and E. Native LPL expression was not detected in clone D. However, expression of the modified LPL protein was detected using the N-terminus LPL antibody, as indicated by the white arrow in Figure 5.3A. Clone D is predicted to

express a modified LPL protein of molecular weight 29.9 kDa. Expression of the modified LPL protein in clone D was not detected using the C-terminus LPL antibody. This is expected because the modified LPL protein in clone D consists of the first 276 amino acids of the native LPL protein, while the C-terminus antibody can only detect LPL-297 to LPL-326 of the native LPL protein.

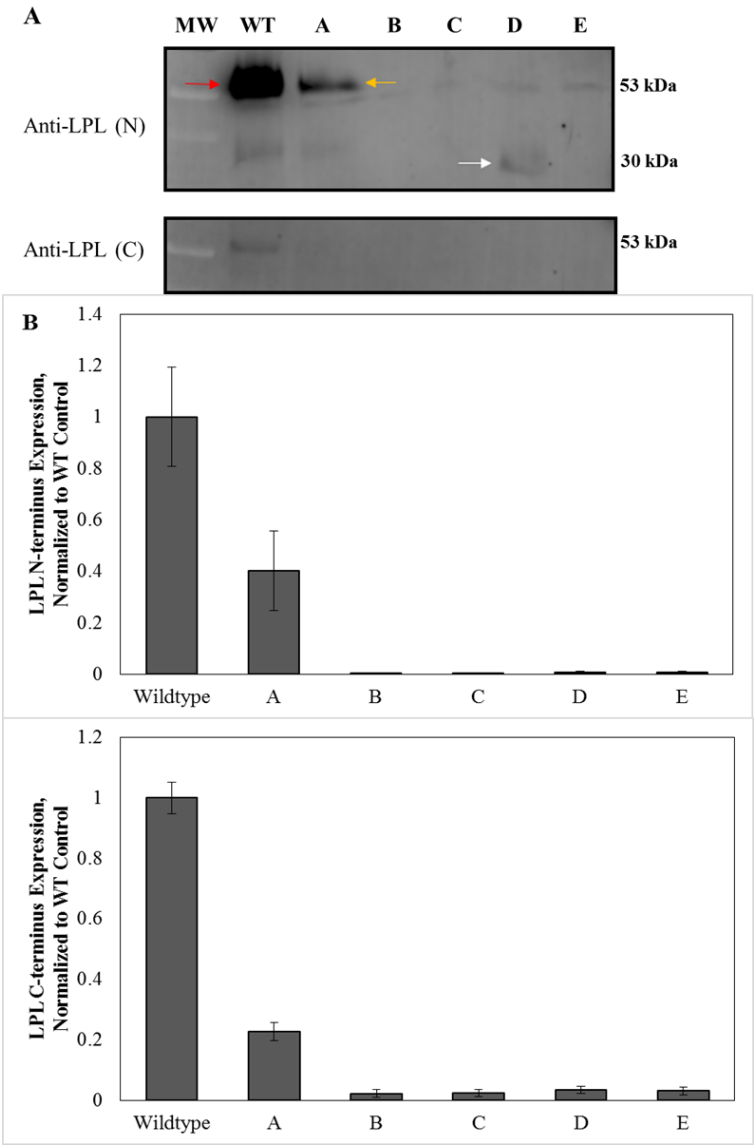


Figure 5.3: Characterization of LPL protein expression from CHO *lpl* knockout cell lines by western. (A) Western of LPL protein with N-terminus LPL antibody (Anti-LPL(N)) and C-terminus LPL antibody (Anti-LPL (C)). The predicted molecular weight of LPL is 53 kDa. The red arrow indicates the native LPL band. The yellow arrow indicates the altered LPL protein expressed by clone A, with predicted molecular weight of 52.2 kDa. The white arrow indicates the altered LPL protein expressed by clone D, with predicted molecular weight of 29.9 kDa. (B) LPL expression measured by N-terminus and C-terminus LPL antibody, relative to wildtype control. Error bars represent the standard error of the mean from three biological replicates.

LPL expression from *lpl* knockout cells and wildtype cells were also measured by an MRM assay, which is specific for the detection of LPL. The peptide fragments used for the detection of LPL expression are ITGLDPAGPNFEYAEAPSR (ITG) and GLGDVDQLVKC (GLG). The positions of the peptide fragment and areas affected by CRISPR/Cas9 are shown in Figure 5.4. The relative ITG and GLG expressions of the five *lpl* knockout were compared to the wildtype control (Figure 5.5A, B). The expression of both ITG and GLG peptides in clone A was 5% of the wildtype control. The presence of both peptides in clone A is consistent with the sequencing and western blot data. The absence of or low expression of ITG and GLG peptides in clones B, C and E were also consistent with the sequencing and western data. The expressions of ITG and GLG peptides in clone D were 132% and 92% of the wildtype control, respectively. The presence of both peptides in clone D is consistent with both the sequencing and western blot data.

Western and MRM results both support the absence of a native LPL expression in all five CHO-K1 *lpl* knockout cell lines. The successful knockout at the protein level of LPL occurred through frameshift mutations in the exons (clones B, C, D, E) which yielded either no expression or truncated expression of LPL proteins.

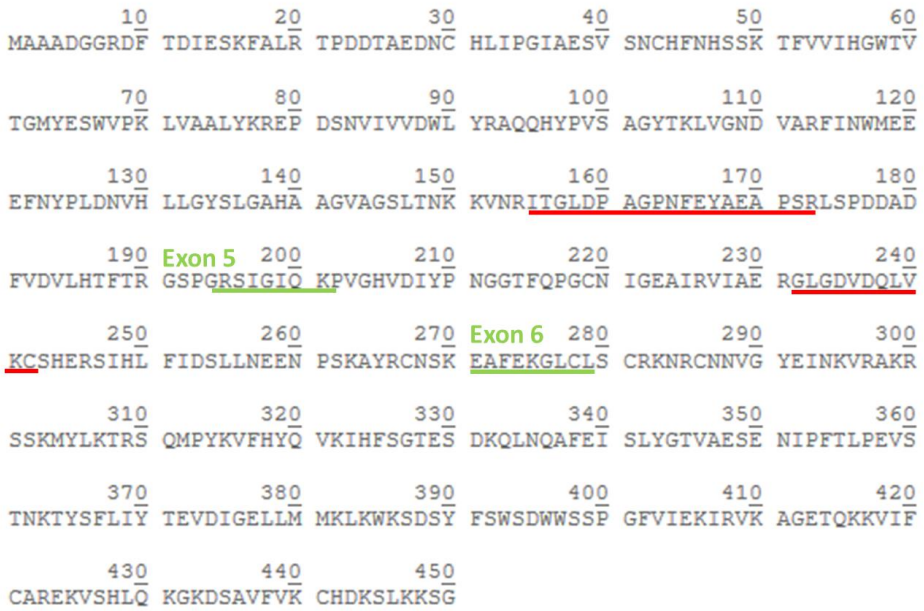


Figure 5.4: Peptides used for LPL MRM detection. Peptides are indicated by red lines. Green lines illustrate the location where frameshift or amino acid deletion occurs. Exon 5 is targeted by clones A, B, and C. Exon 6 is targeted by clones D and E.

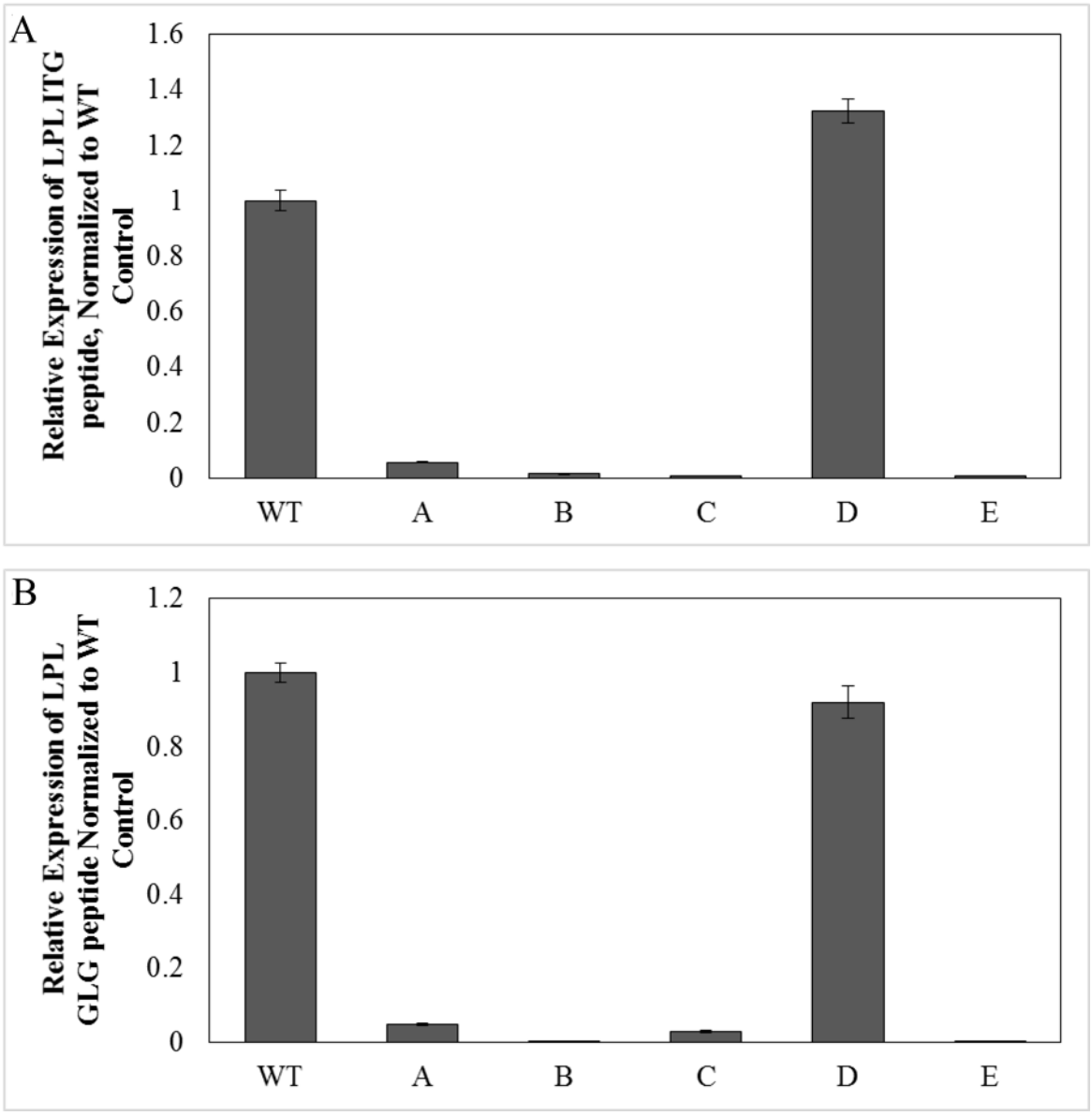


Figure 5.5: Characterization of LPL protein expression from CHO *lpl* knockout cell lines by MRM. LPL expressions are normalized to the wildtype control. (A) Expression of LPL ITG peptide. (B) Expression of LPL GLG peptide. Error bars represent the standard error of the mean from three technical replicates.

5.3.5 Culture performance of LPL knockout CHO-K1 cell lines

Previous study of LPL knockdown by siRNA interference (Chapter 3) showed that reduced LPL expression did not significantly affect cell culture performance. The five CHO-K1 *lpl* knockout cell lines (A, B, C, D, E) were cultivated parallel to the wildtype CHO-K1 control to evaluate the effect of a functional *lpl* knockout on culture performance. The cell culture performance study includes comparisons on integrated viable cell density, viability, and total extracellular protein between *lpl* knockout cell lines and the wildtype control. Viable cell density and percentage of cell viability were measured daily for 10 days, while total extracellular protein was measured on day 4. The five CHO-K1 *lpl* knockout cell lines showed similar integrated viable cell density (IVCD) to the wildtype control during the first six days of cultivation and by Day 10 the IVCD of *lpl* knockouts ranged from -16% to +5% of the wildtype control (Figure 5.6A). High levels of cell viability were maintained for 8 days across all cultures, with three *lpl* knockout cell lines (A, C, D) able to maintain longer cell culture durations with over 50% viability (Figure 5.6B). Changes to total extracellular protein in the HCCF on day 4 of cultivation were not statistically significant ($p > 0.05$) with the exception of clones B ($p = 0.013$) and E ($p = 0.046$), which produced 29% and 17% more protein than the wildtype control, respectively (Figure 5.6C).

The reduction of IVCD in *lpl* knockout cell lines suggests that LPL activity affects cell growth. This is supported by previous studies that showed cell growth is promoted by the addition of free fatty acids and phospholipids in mammalian cell culture (Prasad, 1980; Schmid et al., 1991). Minimal changes to IVCD observed across the cell lines indicate that the absence of LPL does not impair cell culture.

Variable total protein production between the cell lines does not give a clear indication of LPL's role in HCP productivity. With the absence of native LPL expression, the *lpl* knockout cell lines were able to maintain properties that are important for biopharmaceutical processing, including IVCD, viability, and total extracellular protein expression.

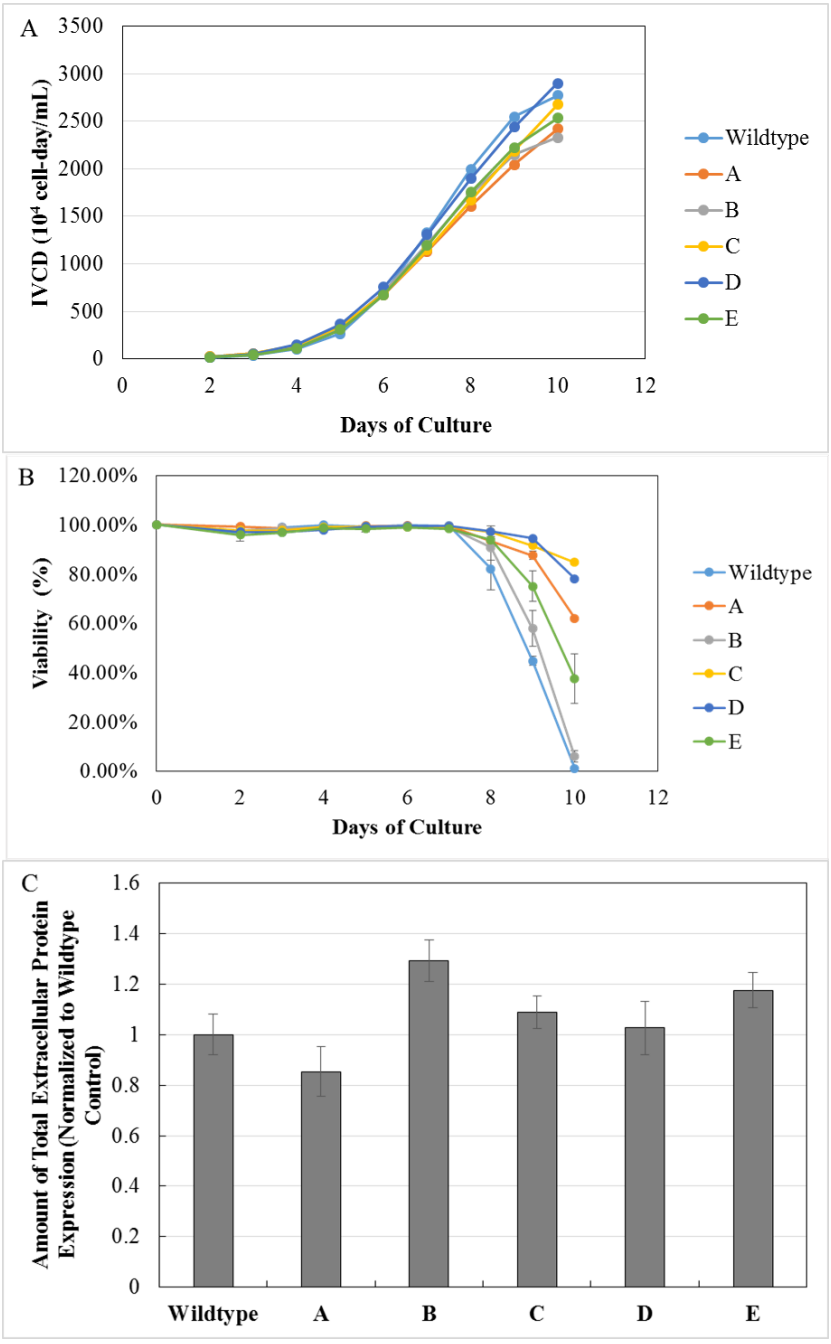


Figure 5.6: Cell culture performance of CHO-K1 *lpl* knockout cell lines. (A) Integrated viable cell density (IVCD) and (B) viability cell culture profile of CHO-K1 wildtype, and clones A, B, C, D, and E. Error bars represent the standard error of the mean from two biological replicates. (C) Total extracellular protein expression at Day 4 of cell culture. Error bars represent the standard error of the mean from three biological replicates.

5.3.6 Polysorbate degradation is reduced in *lpl* knockout cell lines

The impact of LPL on formulation stability, more specifically polysorbate degradation, was examined indirectly by measuring the free fatty acid concentration in polysorbate samples incubated with CHO HCCF. Polysorbate degradation is measured indirectly by measuring free fatty acid concentration since enzymatic hydrolysis of PS80 yields oleic acid and an alcohol component, while enzymatic hydrolysis of PS20 yields lauric acid and an alcohol component. mAb formulations typically contain 0.01-0.05% w/w surfactants; higher percentages of surfactants can cause protein unfolding (Marichal-Gallardo & Alvarez, 2012). PS80 or PS20 with a final concentration of 0.03% w/w was incubated with concentrated CHO HCCF obtained from wildtype CHO-K1 cells or CHO-K1 *lpl* knockout cells. Samples used in an LPL inhibitor study were incubated with either PS80 or PS20 and 2 μM apolipoprotein C-III, which has been shown to inhibit LPL activity (Bobik, 2008; Larsson et al., 2013; McConathy et al., 1992). The samples were incubated at an elevated temperature of 37 °C to accelerate the formation of free fatty acids. The free fatty acid concentrations of each sample were examined following 12 hours of incubation (Figure 5.7).

CHO HCCF from CHO wildtype control degraded 34.3% (84.6 μM) of the total PS80 during the 12-hour incubation. Incubation of CHO wildtype HCCF with apolipoprotein C-III, an inhibitor of LPL, reduced PS80 degradation by 28.2% (Figure 5.7A), which supports our previous observations regarding the impact of LPL on PS80 degradation (Chapter 3). A PS80 degradation assay with HCCF from the various *lpl* knockout lines showed that only 18.2% (44.7 μM) to 20.5% (50.4 μM) of total PS80 was degraded. This represents a reduction of 40.5% to 47.1% in PS80 degradation,

compared to the CHO-K1 wildtype control (Figure 5.7A). Minimal changes to PS80 degradation by CHO HCCF from CHO-K1 *lpl* knockout were seen with the addition of apolipoprotein C-III (Figure 5.7A) and were within the assay's limit of quantitation (5%).

A similar analysis was applied to our study in PS20 degradation. HCCF from CHO wildtype control degraded 38.0% (102.1 μM) of total PS20. Incubation of CHO wildtype HCCF with apolipoprotein C-III, an inhibitor of LPL, reduced PS20 degradation by 20.2% (Figure 5.7B). PS20 incubated with CHO HCCF from CHO-K1 *lpl* knockout reveals that only 16.3% (44.0 μM) to 21.4% (57.5 μM) of total PS20 was degraded, which represents a reduction of 43.7% to 57.0% in PS20 degradation, compared to the CHO-K1 wildtype control (Figure 5.7B). Changes to PS20 degradation by CHO HCCF from CHO-K1 *lpl* knockout observed with the addition of apolipoprotein C-III (Figure 5.7B) were minimal and within the assay's limit of quantitation (5%). Inhibitory effects of apolipoprotein C-III in either polysorbate degradation assays were not observed with the *lpl* knockout cell lines, which suggest that LPL is not catalytically active, if present.

A previous study by Valente, et al. had consistently identified LPL to make up 0.1% of the total CHO HCCF concentration. Although the exact protein composition of HCCF used in each polysorbate degradation assay is unknown, the reduction in polysorbate degradation (40.5% to 57%) by CHO-K1 *lpl* knockout HCCF suggests that LPL, though not abundantly present in CHO HCCF, plays a significant role in polysorbate hydrolysis.

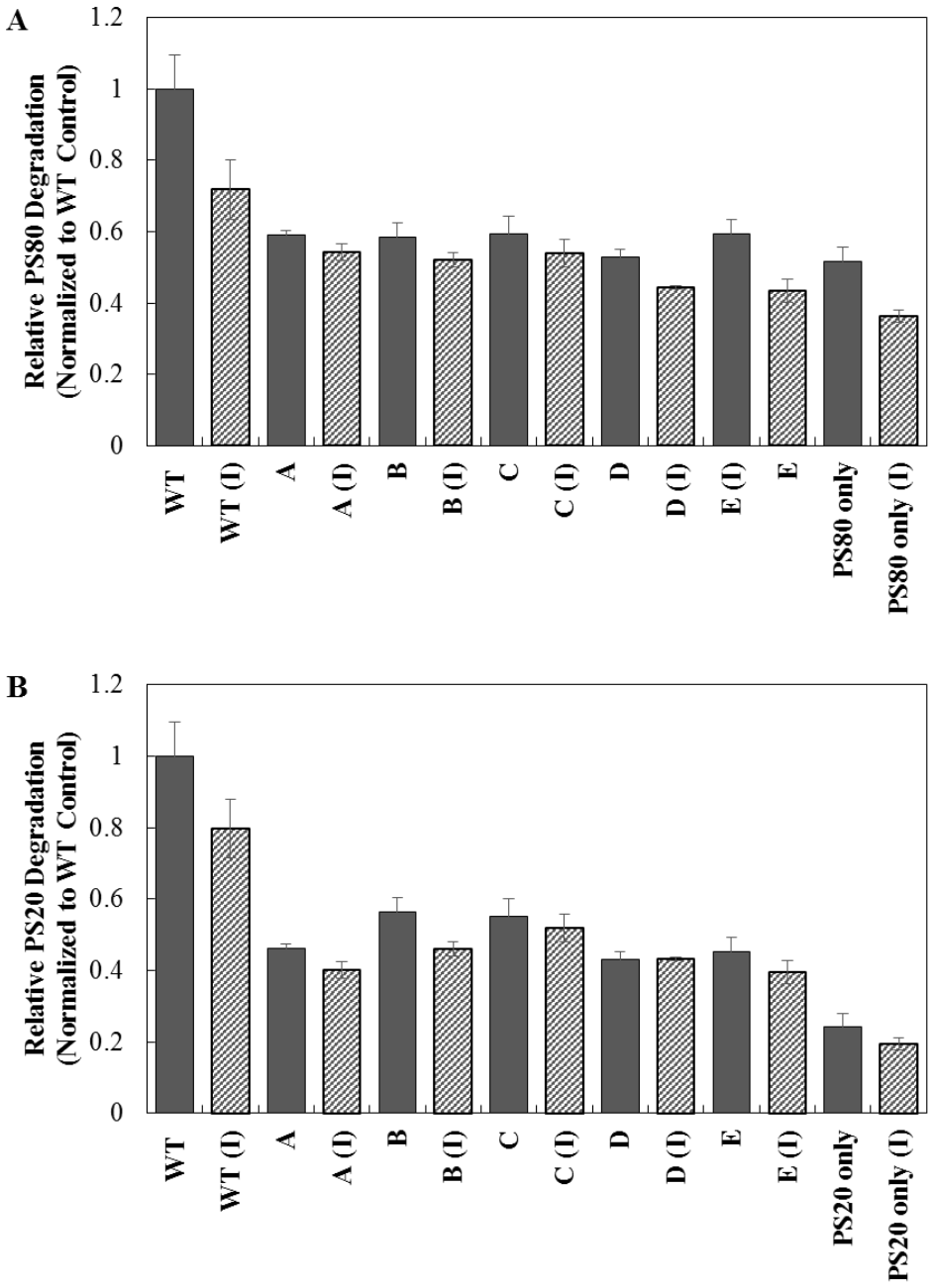


Figure 5.7: Effect of LPL on the formation of free fatty acids from enzymatic hydrolysis of (A) PS80 and (B) PS20 by CHO HCCF from CHO-K1 wildtype control (WT), CHO-K1 *lpl* knockout cell lines (A, B, C, D, E), and apolipoprotein C-III (I). Error bars represent the standard error of the mean from three biological replicates.

5.4 Conclusion

The availability of both the highly adaptable CRISPR/Cas9 genome-editing tool and CHO-K1 genome sequence allowed us to engineer CHO-K1 cell lines with functional *lpl* knockout. In this study, we report the generation of five CHO-K1 *lpl* knockout cell lines and the impact of LPL on polysorbate degradation, and ultimately, formulation stability. The absence of a native LPL in our five CHO-K1 *lpl* knockout cell lines have been confirmed through sequencing and protein quantification assays. Polysorbate degradation studies with these cell lines have confirmed that LPL natively produced in CHO-K1 cells can degrade either PS20 or PS80 by enzymatic ester hydrolysis. In this study, we found that the absence of native LPL expression is correlated with reduced polysorbate degradation, thus possibly improve polysorbate, and biopharmaceutical, stability. The results reported here demonstrate the application of CRISPR/Cas9 for extensive genetic manipulation and genomic analysis of engineered CHO cell lines to address biopharmaceutical problems.

REFERENCES

Bobik, A. (2008). Apolipoprotein CIII and atherosclerosis: beyond effects on lipid metabolism. *Circulation, 118*(7), 702-704.

Braun, J. E. A., & Severson, D. L. (1992). Regulation of the synthesis, processing and translocation of lipoprotein lipase. *Biochem J, 287*, 337-347.

Fu, Y., Sander, J. D., Reyon, D., Cascio, V. M., & Joung, J. K. (2014). Improving CRISPR-Cas nuclease specificity using truncated guide RNAs. *Nat Biotechnol, 32*(3), 279-284.

Kuscu, C., Arslan, S., Singh, R., Thorpe, J., & Adli, M. (2014). Genome-wide analysis reveals characteristics of off-target sites bound by the Cas9 endonuclease. *Nat Biotechnol, 32*(7), 677-683.

Larsson, M., Vorrsjo, E., Talmud, P., Lookene, A., & Olivecrona, G. (2013). Apolipoproteins C-I and C-III inhibit lipoprotein lipase activity by displacement of the enzyme from lipid droplets. *J Biol Chem, 288*(47), 33997-34008.

Lutz, E. P., Merkel, M., Kako, Y., Melford, K., Radner, H., Breslow, J. L., Bensadoun, A., & Goldberg, I. J. (2001). Heparin-binding defective lipoprotein lipase is unstable and causes abnormalities in lipid delivery to tissues. *J. Clin. Invest., 107*, 1183-1192.

Marichal-Gallardo, P. A., & Alvarez, M. M. (2012). State-of-the-art in downstream processing of monoclonal antibodies: process trends in design and validation. *Biotechnol Prog, 28*(4), 899-916.

McConathy, W. J., Gesquiere, J. C., Bass, S. H., Tartar, A., Fruchart, J. C., & Wang, C. S. (1992). Inhibition of LPL activity by synthetic peptides of apolipoprotein C-III. *J. Lipid. Res, 33*, 995-1003.

Moreno-Mateos, M. A., Vejnar, C. E., Beaudoin, J.-D., Fernandez, J. P., Mis, E. K., Khokha, M. K., & Giraldez, A. J. (2015). CRISPRscan: designing highly efficient sgRNAs for CRISPR-Cas9 targeting in vivo. *Nat Methods, 12*(10), 982-988.

Prasad, K. N. (1980). Butyric acid: A small fatty acid with diverse biological functions. *Life Sciences, 27*, 1351-1358.

Ran, F. A., Hsu, P. D., Lin, C. Y., Gootenberg, J. S., Konermann, S., Trevino, A. E., Scott, D. A., Inoue, A., Matoba, S., Zhang, Y., & Zhang, F. (2013). Double nicking by RNA-guided CRISPR Cas9 for enhanced genome editing specificity. *Cell, 154*(6), 1380-1389.

Ronda, C., Pedersen, L. E., Hansen, H. G., Kallehauge, T. B., Betenbaugh, M. J., Nielsen, A. T., & Kildegaard, H. F. (2014). Accelerating genome editing in CHO cells using CRISPR Cas9 and CRISPy, a web-based target finding tool. *Biotechnol Bioeng, 111*(8), 1604-1616.

Schmid, G., Zilg, H., Eberhard, U., & Johannsen, R. (1991). Effect of free fatty acids and phospholipids on growth of and product formation by recombinant baby hamster kidney (rBHK) and Chinese hamster ovary (rCHO) cells in culture. *J. Biotechnology, 17*, 155-167.

Valente, K. N. (2014). Optimization and application of proteomic methods for characterization of host cell protein impurities from Chinese hamster ovary cells. *Ph.D. dissertation, University of Delaware*.

Chapter 6

CONCLUSIONS AND FUTURE WORK

6.1 Summary of Conclusions

This thesis examined the application of two CHO cell line engineering tools, siRNA interference and CRISPR/Cas9, to rationally design a cell line that can be used to address problems present in the biopharmaceutical industry. The first part of this research demonstrated the use of siRNA interference to reduce lipoprotein (LPL) expression in a null CHO-K1 cell line to address its impact on formulation stability. Knockdown of LPL expression did not adversely affect cell growth or total extracellular protein expression. Incubation of polysorbate 80 (PS80) with host cell protein (HCP) containing reduced LPL content revealed that reduced LPL expression is associated with reduced PS80 degradation. siRNA interference proved to be a versatile tool to examine the function of independent genes by temporarily silencing the phenotype.

The findings from our siRNA interference study motivated subsequent studies to examine the application of CRISPR/Cas9 to generate a CHO-K1 cell line with a functional *lpl* knockout and to determine the impact of LPL on polysorbate degradation. The availability of the CHO-K1 genome aided our quest to rationally design a CHO-K1 *lpl* knockout cell line. This led to the generation of five different CHO-K1 cell lines with *lpl* knockouts. Characterization of these cell lines revealed that the altered LPL is truncated and the relative LPL expression in the cell lines was

reduced. Subsequent polysorbate degradation studies demonstrated that LPL plays a significant role in the enzymatic hydrolysis of polysorbate 20 (PS20) and PS80. This research thesis demonstrates that rational design with CHO cell line engineering tools can efficiently advance the process for extensive genetic manipulation and genomic analysis of engineered CHO cell lines to address biopharmaceutical problems.

6.2 Recommendations for Future Work

The polysorbate degradation studies performed in this thesis can be extended to gain insight on the impact of biologic-based impurities on formulation stability. Extensions of this work could improve formulation stability and the safety of biopharmaceuticals.

6.2.1 Further investigation on the impact of LPL on formulation and product stability

Chapters 3 and 5 of this thesis demonstrated the effect of LPL on polysorbate degradation under accelerated conditions and explored the implication of these results on formulation stability. A polysorbate study under relevant conditions used in industry can provide the most accurate representation of LPL's contribution to polysorbate stability. Relevant conditions include polysorbate incubation with mAbs and purification of mAbs to reduce HCP impurity. Protein A affinity chromatography is often the first step for mAb purification, with two-thirds of commercial mAb purification begin with Protein A column chromatography (Marichal-Gallardo &

Alvarez, 2012). The product can then be further polished with an ion exchange chromatography (IEX), such as cation exchange chromatography (CEX) or anion exchange chromatography (AEX). The type of IEX used is based on the characteristic of the impurities (Stein & Kiesewetter, 2007). A polysorbate degradation study between the wildtype and *lpl* knockout clones with mAbs following the proper purification and polishing steps may provide some information on the rate of polysorbate degradation.

HCP profile of retained impurities produced in *lpl* knockout cell lines can also provide insight on the changes of protein expression following gene disruption of *lpl*. Variable expression of HCP can increase the difficulty in removing these impurities because the composition of HCPs expressed upstream was shown to affect HCP clearance during downstream purification (Wu, 2013). Proteomics methods such as shotgun proteomics and 2-dimensional gel electrophoresis have been applied in various studies to detect the HCP composition of CHO cell lines (Valente et al., 2015; Valente et al., 2014). These methods can be applied to determine the HCP profile of *lpl* knockout cell lines. One function of LPL is to hydrolyze triglyceride into free fatty acids and glycerol. The enzyme activity and binding of LPL to triglycerides are maintained by heparan sulfate proteoglycan core protein (Berryman & Bensadoun, 1995). From this information, we expect to see increased expression of similar lipases, such as hepatic lipase and pancreatic lipase, and decreased expression of the heparan sulfate proteoglycan core protein.

From Chapters 3 and 5 of this thesis, we found that the disruption of *lpl* did not impair the cell growth of a null CHO-K1 cell line. However, changes to other characteristics that are important for biopharmaceutical production, including specific

productivity and product quality of recombinant therapeutics, are unknown. A mAb-producing CHO cell line will provide results that are most relevant to the biopharmaceutical industry. However, such cell line will be difficult to obtain. Other available CHO cell lines, such as secreted alkaline phosphatase or tissue plasminogen activator, could serve as a proof-of-concept demonstration.

6.2.2 Exploration of other HCPs contributing to formulation instability

LPL was one HCP we found to contribute to polysorbate degradation, accounting for 40% of both polysorbate 20 and polysorbate 80 degradation (Chapter 5). Many other HCPs were found to co-purify with mAbs across Protein A capture and non-affinity polishing resins (Levy, 2014; Valente, 2014). HCPs that are likely to co-purify with therapeutic products during polishing chromatography operations and contribute to formulation instability are listed in Table 6.1. The retention of HCPs on the four polishing chromatographic resins, cation exchange (CEX), anion exchange (AEX), hydrophobic interaction chromatography (HIC), and multimodal cation (MMC) are also indicated in the table (Valente, 2014). Basement membrane-specific heparan sulfate proteoglycan core protein (HSPG) is responsible for the maintenance of enzyme activity and binding of LPL (Berryman & Bensadoun, 1995). It was also shown to retain in elution fractions following AEX, HIC, and MMC (Table 6.1). HSPG's association with LPL and its retention with multiple resins are two pieces of evidence that support the evaluation of HSPG and its impact on polysorbate degradation. Cathepsin D is an enzyme that was shown to be associated with hydrolysis of lipids and is retained in elution fractions following HIC. Triacylglycerol

lipase was identified to retain in high-concentration drug products and hypothesized to contribute to polysorbate degradation by Labrenz (Labrenz, 2014). Polysorbate studies with these HCPs can provide information that can be used to improve formulation and product stability.

Table 6.1: HCPs likely to co-purify with therapeutic products (Valente, 2014) and contribute to formulation instability. Four polishing chromatographic resins are cation exchange (CEX), anion exchange (AEX), hydrophobic interaction chromatography (HIC), multimodal cation (MMC). Cross-interaction chromatography (CIC) is used to evaluate product association under HIC conditions. Liquid chromatography (LC) and polyacrylamide gel electrophoresis (PAGE) were two techniques used to separate and identify the protein composition. (a) triacylglycerol lipase was not identified in Valente's study but was identified by Labrenz (2014). Chromatography data for triacylglycerol lipase is not available.

Protein Identification	CEX		AEX			HIC		MMC	
	LC	PAGE	LC Early	LC Late	PAGE	LC	CIC	LC	PAGE
Basement membrane-specific heparan sulfate proteoglycan core protein				x	x	x		x	
Cathepsin D							x		
Triacylglycerol lipase[a]	[redacted]								

REFERENCES

Berryman, D. E., & Bensadoun, A. (1995). Heparan Sulfate Proteoglycans Are Primarily Responsible for the Maintenance of Enzyme Activity, Binding, and Degradation of Lipoprotein Lipase in Chinese Hamster Ovary Cells. *J of Biol Chem, 270*(41), 24525–24531.

Labrenz, S. R. (2014). Ester hydrolysis of polysorbate 80 in mAb drug product: evidence in support of the hypothesized risk after the observation of visible particulate in mAb formulations. *J Pharm Sci, 103*(8), 2268-2277.

Levy, N. E. (2014). Host cell protein impurities and protein-protein interactions in downstream purification of monoclonal antibodies. *Ph.D. dissertation, University of Delaware*.

Marichal-Gallardo, P. A., & Alvarez, M. M. (2012). State-of-the-art in downstream processing of monoclonal antibodies: process trends in design and validation. *Biotechnol Prog, 28*(4), 899-916.

Stein, A., & Kiesewetter, A. (2007). Cation exchange chromatography in antibody purification: pH screening for optimised binding and HCP removal. *J Chromatogr B Analyt Technol Biomed Life Sci, 848*(1), 151-158.

Valente, K. N. (2014). Optimization and application of proteomic methods for characterization of host cell protein impurities from Chinese hamster ovary cells. *Ph.D. dissertation, University of Delaware*.

Valente, K. N., Lenhoff, A. M., & Lee, K. H. (2015). Expression of difficult-to-remove host cell protein impurities during extended Chinese hamster ovary cell culture and their impact on continuous bioprocessing. *Biotechnol Bioeng, 112*(6), 1232-1242.

Valente, K. N., Schaefer, A. K., Kempton, H. R., Lenhoff, A. M., & Lee, K. H. (2014). Recovery of Chinese hamster ovary host cell proteins for proteomic analysis. *Biotechnol J, 9*(1), 87-99.

Wu, P. (2013). Practical experiences integrating upstream and downstream processing. *Am Pharml Rev, 16*, 84-86.

REFERENCES

Berryman, D. E., & Bensadoun, A. (1995). Heparan Sulfate Proteoglycans Are Primarily Responsible for the Maintenance of Enzyme Activity, Binding, and Degradation of Lipoprotein Lipase in Chinese Hamster Ovary Cells. *J of Biol Chem, 270*(41), 24525–24531.

Bobik, A. (2008). Apolipoprotein CIII and atherosclerosis: beyond effects on lipid metabolism. *Circulation, 118*(7), 702-704. doi: 10.1161/CIRCULATIONAHA.108.794081

Boch, J., Scholze, H., Schornack, S., Landgraf, A., Hahn, S., Kay, S., . . . Bonas, U. (2009). Breaking the Code of DNA Binding Specificity of TAL-Type III Effectors. *Science, 326*(5959), 1509-1512.

Braun, J. E. A., & Severson, D. L. (1992). Regulation of the synthesis, processing and translocation of lipoprotein lipase. *Biochem J, 287*, 337-347.

Brinkrolf, K., Rupp, O., Laux, H., Kollin, F., Ernst, W., Linke, B., . . . Borth, N. (2013). Chinese hamster genome sequenced from sorted chromosomes. *Nat Biotechnol, 31*(8), 694-695. doi: 10.1038/nbt.2645

Cermak, T., Doyle, E. L., Christian, M., Wang, L., Zhang, Y., Schmidt, C., . . . Voytas, D. F. (2011). Efficient design and assembly of custom TALEN and

other TAL effector-based constructs for DNA targeting. *Nucleic Acids Res, 39*(12), e82. doi: 10.1093/nar/gkr218

Chapman, J. R., Taylor, M. R., & Boulton, S. J. (2012). Playing the end game: DNA double-strand break repair pathway choice. *Mol Cell, 47*(4), 497-510. doi: 10.1016/j.molcel.2012.07.029

Datta, P., Linhardt, R. J., & Sharfstein, S. T. (2013). An 'omics approach towards CHO cell engineering. *Biotechnol Bioeng, 110*(5), 1255-1271. doi: 10.1002/bit.24841

Doneanu, C. E., Xenopoulos, A., Fadgen, K., Murphy, J., Skilton, S. J., Prentice, H., . . . Chen, W. (2012). Analysis of host-cell proteins in biotherapeutic proteins by comprehensive online two-dimensional liquid chromatography/mass spectrometry. *mAbs, 4*(1), 24-44. doi: 10.4161/mabs.4.1.18748

Doshi, N., Demeule, B., & Yadav, S. (2015). Understanding Particle Formation: Solubility of Free Fatty Acids as Polysorbate 20 Degradation Byproducts in Therapeutic Monoclonal Antibody Formulations. *Mol Pharm, 12*(11), 3792-3804. doi: 10.1021/acs.molpharmaceut.5b00310

Duda, K., Lonowski, L. A., Kofoed-Nielsen, M., Ibarra, A., Delay, C. M., Kang, Q., . . . Frodin, M. (2014). High-efficiency genome editing via 2A-coupled co-expression of fluorescent proteins and zinc finger nucleases or CRISPR/Cas9 nickase pairs. *Nucleic Acids Res, 42*(10), e84. doi: 10.1093/nar/gku251

Fu, Y., Foden, J. A., Khayter, C., Maeder, M. L., Reyon, D., Joung, J. K., & Sander, J. D. (2013). High-frequency off-target mutagenesis induced by CRISPR-Cas

nucleases in human cells. *Nat Biotechnol, 31*(9), 822-826. doi: 10.1038/nbt.2623

Fu, Y., Sander, J. D., Reyon, D., Cascio, V. M., & Joung, J. K. (2014). Improving CRISPR-Cas nuclease specificity using truncated guide RNAs. *Nat Biotechnol, 32*(3), 279-284. doi: 10.1038/nbt.2808

Gaj, T., Gersbach, C. A., & Barbas, C. F., 3rd. (2013). ZFN, TALEN, and CRISPR/Cas-based methods for genome engineering. *Trends Biotechnol, 31*(7), 397-405. doi: 10.1016/j.tibtech.2013.04.004

Grav, L. M., Lee, J. S., Gerling, S., Kallehauge, T. B., Hansen, A. H., Kol, S., . . . Kildegaard, H. F. (2015). One-step generation of triple knockout CHO cell lines using CRISPR/Cas9 and fluorescent enrichment. *Biotechnol J, 10*(9), 1446-1456. doi: 10.1002/biot.201500027

Gstaiger, M., & Aebersold, R. (2009). Applying mass spectrometry-based proteomics to genetics, genomics and network biology. *Nat Rev Genet, 10*(9), 617-627. doi: 10.1038/nrg2633

Hockemeyer, D., Soldner, F., Beard, C., Gao, Q., Mitalipova, M., DeKelver, R. C., . . . Jaenisch, R. (2009). Efficient targeting of expressed and silent genes in human ESCs and iPSCs using zinc-finger nucleases. *Nat Biotechnol, 27*(9), 851-857. doi: 10.1038/nbt.1562

Hockemeyer, D., Wang, H., Kiani, S., Lai, C. S., Gao, Q., Cassady, J. P., . . . Jaenisch, R. (2011). Genetic engineering of human pluripotent cells using TALE nucleases. *Nat Biotechnol, 29*(8), 731-734. doi: 10.1038/nbt.1927

Holkers, M., Maggio, I., Liu, J., Janssen, J. M., Miselli, F., Mussolino, C., . . . Goncalves, M. A. (2013). Differential integrity of TALE nuclease genes following adenoviral and lentiviral vector gene transfer into human cells. *Nucleic Acids Res, 41*(5), e63. doi: 10.1093/nar/gks1446

Hsu, P. D., Lander, E. S., & Zhang, F. (2014). Development and applications of CRISPR-Cas9 for genome engineering. *Cell, 157*(6), 1262-1278. doi: 10.1016/j.cell.2014.05.010

Huang, Y. M., Hu, W., Rustandi, E., Chang, K., Yusuf-Makagiansar, H., & Ryll, T. (2010). Maximizing productivity of CHO cell-based fed-batch culture using chemically defined media conditions and typical manufacturing equipment. *Biotechnol Prog, 26*(5), 1400-1410.

Ilko, D., Braun, A., Germershaus, O., Meinel, L., & Holzgrabe, U. (2015). Fatty acid composition analysis in polysorbate 80 with high performance liquid chromatography coupled to charged aerosol detection. *Eur J Pharm Biopharm, 94*, 569-574. doi: 10.1016/j.ejpb.2014.11.018

Janas, M. M., Wang, B., Harris, A. S., Aguiar, M., Shaffer, J. M., Subrahmanyam, Y. V., . . . Novina, C. D. (2012). Alternative RISC assembly: binding and repression of microRNA-mRNA duplexes by human Ago proteins. *RNA, 18*(11), 2041-2055. doi: 10.1261/rna.035675.112

Jayapal, K. P., Wlaschin, K. F., & Hu, W. S. (2007). Recombinant Protein Therapeutics from CHO Cells — 20 Years and Counting. *Chemical Engineering Progress, 103*, 40-47.

Jinek, M., Chylinski, K., Fonfara, I., Hauer, M., Doudna, J. A., & Charpentier, E. (2012). A programmable dual-RNA-guided DNA endonuclease in adaptive bacterial immunity. *Science, 337*(6096), 816-821. doi: 10.1126/science.1225829

Joung, J. K., & Sander, J. D. (2013). TALENs: a widely applicable technology for targeted genome editing. *Nat Rev Mol Cell Biol, 14*(1), 49-55. doi: 10.1038/nrm3486

Kerwin, B. A. (2008). Polysorbates 20 and 80 used in the formulation of protein biotherapeutics: structure and degradation pathways. *J Pharm Sci, 97*(8), 2924-2935. doi: 10.1002/jps.21190

Kim, H., & Kim, J. S. (2014). A guide to genome engineering with programmable nucleases. *Nature Reviews Genetics, 15*(5), 321-334. doi: 10.1038/nrg3686

Kim, S. H., & Lee, G. M. (2007). Down-regulation of lactate dehydrogenase-A by siRNAs for reduced lactic acid formation of Chinese hamster ovary cells producing thrombopoietin. *Appl Microbiol Biotechnol, 74*(1), 152-159. doi: 10.1007/s00253-006-0654-5

Kim, Y. G., & Chandrasegaran, S. (1994). Chimeric restriction endonuclease. *Proc Natl Acad Sci USA, 91*, 883-887.

Kishore, R. S., Kiese, S., Fischer, S., Pappenberger, A., Grauschopf, U., & Mahler, H. C. (2011). The degradation of polysorbates 20 and 80 and its potential impact on the stability of biotherapeutics. *Pharm Res, 28*(5), 1194-1210. doi: 10.1007/s11095-011-0385-x

Kishore, R. S., Pappenberger, A., Dauphin, I. B., Ross, A., Buergi, B., Staempfli, A., & Mahler, H. C. (2011). Degradation of polysorbates 20 and 80: studies on thermal autoxidation and hydrolysis. *J Pharm Sci, 100*(2), 721-731. doi: 10.1002/jps.22290

Kremkow, B., & Lee, K. H. (2013). Next-generation sequencing technologies and their impact on CHO cell-based biomanufacturing. *Pharm Bioprocess, 1*(5), 455-465.

Kuscu, C., Arslan, S., Singh, R., Thorpe, J., & Adli, M. (2014). Genome-wide analysis reveals characteristics of off-target sites bound by the Cas9 endonuclease. *Nat Biotechnol, 32*(7), 677-683. doi: 10.1038/nbt.2916

Labrenz, S. R. (2014). Ester hydrolysis of polysorbate 80 in mAb drug product: evidence in support of the hypothesized risk after the observation of visible particulate in mAb formulations. *J Pharm Sci, 103*(8), 2268-2277. doi: 10.1002/jps.24054

Larsson, M., Vorrsjo, E., Talmud, P., Lookene, A., & Olivecrona, G. (2013). Apolipoproteins C-I and C-III inhibit lipoprotein lipase activity by displacement of the enzyme from lipid droplets. *J Biol Chem, 288*(47), 33997-34008. doi: 10.1074/jbc.M113.495366

Lee, J. S., Grav, L. M., Lewis, N. E., & Faustrup Kildegaard, H. (2015). CRISPR/Cas9-mediated genome engineering of CHO cell factories: Application and perspectives. *Biotechnol J, 10*(7), 979-994. doi: 10.1002/biot.201500082

Lee, J. S., Kallehauge, T. B., Pedersen, L. E., & Kildegaard, H. F. (2015). Site-specific integration in CHO cells mediated by CRISPR/Cas9 and homology-directed DNA repair pathway. *Sci Rep, 5*, 8572. doi: 10.1038/srep08572

Levy, N. E. (2014). Host cell protein impurities and protein-protein interactions in downstream purification of monoclonal antibodies. *Ph.D. dissertation, University of Delaware*.

Levy, N. E., Valente, K. N., Choe, L. H., Lee, K. H., & Lenhoff, A. M. (2014). Identification and characterization of host cell protein product-associated impurities in monoclonal antibody bioprocessing. *Biotechnol Bioeng, 111*(5), 904-912. doi: 10.1002/bit.25158

Liebler, D. C., & Zimmerman, L. J. (2013). Targeted quantitation of proteins by mass spectrometry. *Biochemistry, 52*(22), 3797-3806. doi: 10.1021/bi400110b

Lim, S. F., Chuan, K. H., Liu, S., Loh, S. O., Chung, B. Y., Ong, C. C., & Song, Z. (2006). RNAi suppression of Bax and Bak enhances viability in fed-batch cultures of CHO cells. *Metab Eng, 8*(6), 509-522. doi: 10.1016/j.ymben.2006.05.005

Lutz, E. P., Merkel, M., Kako, Y., Melford, K., Radner, H., Breslow, J. L., . . . Goldberg, I. J. (2001). Heparin-binding defective lipoprotein lipase is unstable and causes abnormalities in lipid delivery to tissues. *J. Clin. Invest., 107*, 1183-1192.

Maggio, E. T. (2012). Polysorbates, Peroxides, Protein Aggregation, and Immunogenicity: A Growing Concern. *J Excip Food Chem, 3*(2), 45-53.

Makarova, K. S., Haft, D. H., Barrangou, R., Brouns, S. J., Charpentier, E., Horvath, P., . . . Koonin, E. V. (2011). Evolution and classification of the CRISPR-Cas systems. *Nat Rev Microbiol, 9*(6), 467-477. doi: 10.1038/nrmicro2577

Mali, P., Esvelt, K. M., & Church, G. M. (2013). Cas9 as a versatile tool for engineering biology. *Nat Methods, 10*(10), 957-963. doi: 10.1038/nmeth.2649

Mali, P., Yang, L., Esvelt, K. M., Aach, J., Guell, M., DiCarlo, J. E., . . . Church, G. M. (2013). RNA-guided human genome engineering via Cas9. *Science, 339*(6121), 823-826. doi: 10.1126/science.1232033

Marichal-Gallardo, P. A., & Alvarez, M. M. (2012). State-of-the-art in downstream processing of monoclonal antibodies: process trends in design and validation. *Biotechnol Prog, 28*(4), 899-916. doi: 10.1002/btpr.1567

McConathy, W. J., Gesquiere, J. C., Bass, S. H., Tartar, A., Fruchart, J. C., & Wang, C. S. (1992). Inhibition of LPL activity by synthetic peptides of apolipoprotein C-III. *J. Lipid. Res, 33*, 995-1003.

Menetret, J. F., Schaletzky, J., Clemons, W. M., Jr., Osborne, A. R., Skanland, S. S., Denison, C., . . . Akey, C. W. (2007). Ribosome binding of a single copy of the SecY complex: implications for protein translocation. *Mol Cell, 28*(6), 1083-1092. doi: 10.1016/j.molcel.2007.10.034

Moreno-Mateos, M. A., Vejnar, C. E., Beaudoin, J.-D., Fernandez, J. P., Mis, E. K., Khokha, M. K., & Giraldez, A. J. (2015). CRISPRscan: designing highly efficient sgRNAs for CRISPR-Cas9 targeting in vivo. *Nat Methods, 12*(10), 982-988.

Pennisi, E. (2013). The CRISPR Craze. *Science, 341*, 833-836. doi: 10.1126/science.341.6148.833

Prasad, K. N. (1980). Butyric acid: A small fatty acid with diverse biological functions. *Life Sciences, 27*, 1351-1358.

Ramirez, C. L., Foley, J. E., Wright, D. A., Muler-Lerch, F., Rahman, S. H., Cornu, T. I., . . . Joung, J. K. (2008). Unexpected failure rates for modular assembly of engineered zinc fingers. *Nat Methods, 5*(5), 374-375. doi: 10.1038/nmeth0508-374

Ran, F. A., Hsu, P. D., Lin, C. Y., Gootenberg, J. S., Konermann, S., Trevino, A. E., . . . Zhang, F. (2013). Double nicking by RNA-guided CRISPR Cas9 for enhanced genome editing specificity. *Cell, 154*(6), 1380-1389. doi: 10.1016/j.cell.2013.08.021

Ronda, C., Pedersen, L. E., Hansen, H. G., Kallehauge, T. B., Betenbaugh, M. J., Nielsen, A. T., & Kildegaard, H. F. (2014). Accelerating genome editing in CHO cells using CRISPR Cas9 and CRISPy, a web-based target finding tool. *Biotechnol Bioeng, 111*(8), 1604-1616. doi: 10.1002/bit.25233

Sander, J. D., & Joung, J. K. (2014). CRISPR-Cas systems for editing, regulating and targeting genomes. *Nat Biotechnol, 32*(4), 347-355. doi: 10.1038/nbt.2842

Schmid, G., Zilg, H., Eberhard, U., & Johannsen, R. (1991). Effect of free fatty acids and phospholipids on growth of and product formation by recombinant baby hamster kidney (rBHK) and Chinese hamster ovary (rCHO) cells in culture. *J. Biotechnology, 17*, 155-167.

Stein, A., & Kiesewetter, A. (2007). Cation exchange chromatography in antibody purification: pH screening for optimised binding and HCP removal. *J Chromatogr B Analyt Technol Biomed Life Sci, 848*(1), 151-158. doi: 10.1016/j.jchromb.2006.10.010

Sun, T., Li, C., Han, L., Jiang, H., Xie, Y., Zhang, B., . . . Zhu, J. (2015). Functional knockout of FUT8 in Chinese hamster ovary cells using CRISPR/Cas9 to produce a defucosylated antibody. *Engineering in Life Sciences, 15*(6), 660-666. doi: 10.1002/elsc.201400218

Thompson, L. H. (2012). Recognition, signaling, and repair of DNA double-strand breaks produced by ionizing radiation in mammalian cells: the molecular choreography. *Mutat Res, 751*(2), 158-246. doi: 10.1016/j.mrrev.2012.06.002

Tomlinson, A., Demeule, B., Lin, B., & Yadav, S. (2015). Polysorbate 20 Degradation in Biopharmaceutical Formulations: Quantification of Free Fatty Acids, Characterization of Particulates, and Insights into the Degradation Mechanism. *Mol Pharm, 12*(11), 3805-3815. doi: 10.1021/acs.molpharmaceut.5b00311

Urnov, F. D., Rebar, E. J., Holmes, M. C., Zhang, H. S., & Gregory, P. D. (2010). Genome editing with engineered zinc finger nucleases. *Nat Rev Genet, 11*(9), 636-646. doi: 10.1038/nrg2842

Valente, K. N. (2014). Optimization and application of proteomic methods for characterization of host cell protein impurities from Chinese hamster ovary cells. *Ph.D. dissertation, University of Delaware*.

Valente, K. N., Lenhoff, A. M., & Lee, K. H. (2015). Expression of difficult-to-remove host cell protein impurities during extended Chinese hamster ovary cell culture and their impact on continuous bioprocessing. *Biotechnol Bioeng, 112*(6), 1232-1242. doi: 10.1002/bit.25515

Valente, K. N., Schaefer, A. K., Kempton, H. R., Lenhoff, A. M., & Lee, K. H. (2014). Recovery of Chinese hamster ovary host cell proteins for proteomic analysis. *Biotechnol J, 9*(1), 87-99. doi: 10.1002/biot.201300190

Valton, J., Dupuy, A., Daboussi, F., Thomas, S., Maréchal, A., Macmaster, R., . . . Duchateau, P. (2012). Overcoming Transcription Activator-like Effector (TALE) DNA Binding Domain Sensitivity to Cytosine Methylation. *J Biol Chem, 287*(46), 38427-38432.

Walsh, G. (2012). New biopharmaceuticals: a review of new biologic drug approvals over the years, featuring highlights from 2010 and 2011. *BioPharm International, Process Development Forum*.

Walsh, G. (2014). Biopharmaceutical benchmarks 2014. *Nat Biotechnol, 32*(10), 992-1000.

Wei, C., Liu, J., Yu, Z., Zhang, B., Gao, G., & Jiao, R. (2013). TALEN or Cas9 - rapid, efficient and specific choices for genome modifications. *J Genet Genomics, 40*(6), 281-289. doi: 10.1016/j.jgg.2013.03.013

Wiedenheft, B., Sternberg, S. H., & Doudna, J. A. (2012). RNA-guided genetic silencing systems in bacteria and archaea. *Nature, 482*(7385), 331-338. doi: 10.1038/nature10886

Wu, P. (2013). Practical experiences integrating upstream and downstream processing. *Am Pharml Rev, 16,* 84-86.

Wu, S. C. (2009). RNA interference technology to improve recombinant protein production in Chinese hamster ovary cells. *Biotechnol Adv, 27*(4), 417-422. doi: 10.1016/j.biotechadv.2009.03.002

Wurm, F. M. (2004). Production of recombinant protein therapeutics in cultivated mammalian cells. *Nat Biotechnol, 22*(11), 1393-1398.

Xu, X., Nagarajan, H., Lewis, N. E., Pan, S., Cai, Z., Liu, X., . . . Wang, J. (2011). The genomic sequence of the Chinese hamster ovary (CHO)-K1 cell line. *Nat Biotechnol, 29*(8), 735-741. doi: 10.1038/nbt.1932

Appendix A

SUPPLEMENTARY MATERIALS FOR CHAPTER 5

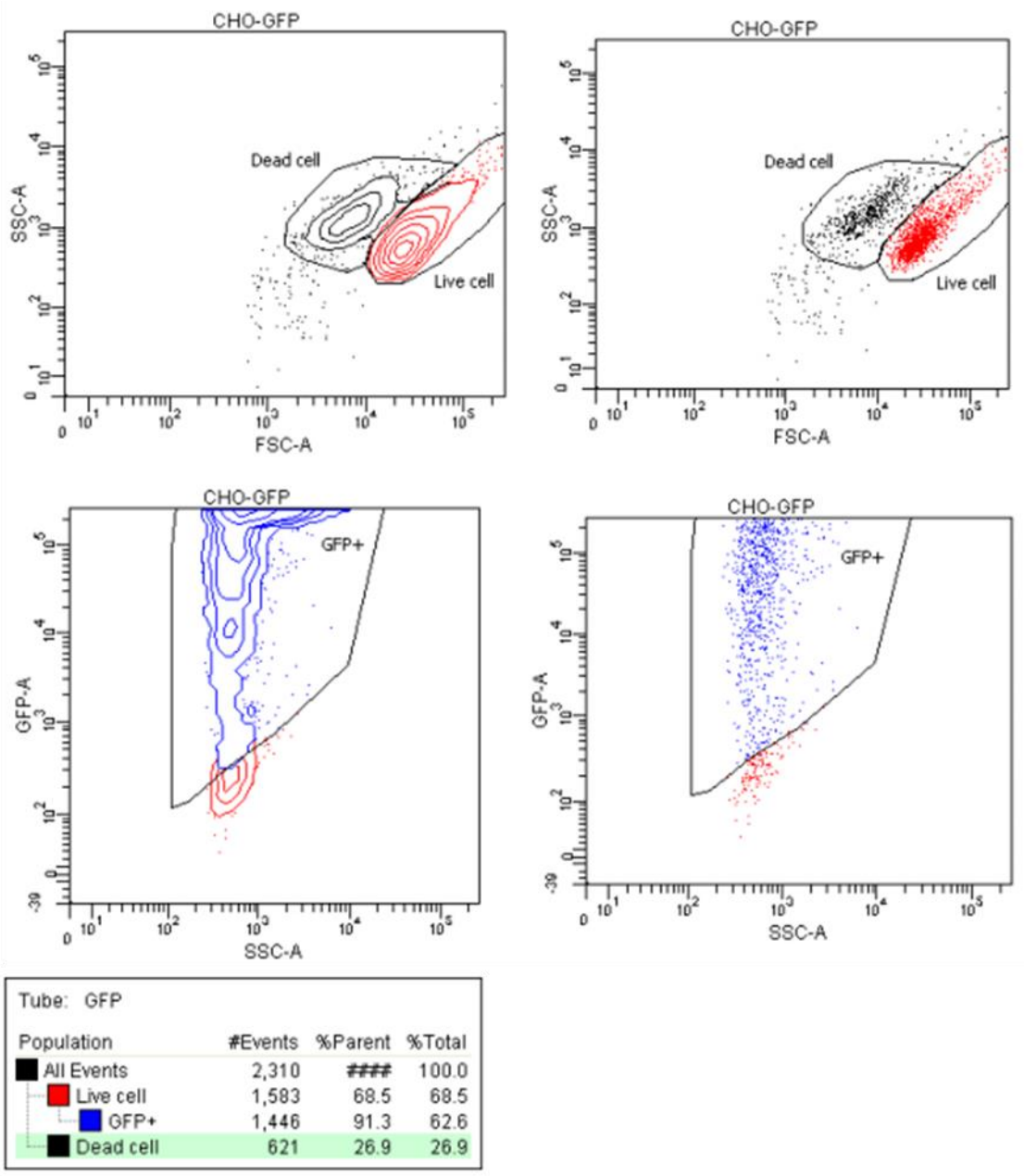
A.1 Preface

This appendix Chapter contains supplementary materials to support the work presented in Chapter 5.

A.2 Supplementary table and figures for Chapter 5

Table A.2: Sequencing primers.

Target Exon	Primer pair	Anneal Temp, ℃	PCR length, bp
E.5	GGATCCAGCTGGGCCTAACT	66	219
	CGCTCTGCAATCACACGAAT		
E.6	TGTGGACCAGCTGGTAAAGT	62	240
	TTGTAGGGCATCTGAGAGCG		
E.8	GCCTGAGGTCTCCACAAATA	58	182
	TTTTCTGAGTCTCTCCTGCT		



Tube: GFP

Population	#Events	%Parent	%Total
All Events	2,310	####	100.0
Live cell	1,583	68.5	68.5
GFP+	1,446	91.3	62.6
Dead cell	621	26.9	26.9

Figure A.1: Transfection efficiency of Lonza Nucleofector. The transfection efficiency of CHO-K1 cells is evaluated following transfection with pmaxGFP plasmid. Upon transfection with the pmaxGFP plasmid, the cells express GFP. The percentage of cells transfected with pmaxGFP is determined by a flow cytometer.

Appendix B

MATLAB SCRIPT FOR SGRNA SELECTION

B.1 Preface

The MATLAB script written for sgRNA selection is presented in this appendix. The selection of sgRNA obtained from this program was compared with the results obtain using CRISPy.

B.2 MATLAB script for a complete list of possible sgRNA

```
clear all; clc;

Genomic_DNA = input('Input the full genomic DNA (CAPITAL letters of
A, T, C, G only): ', 's');

% Changing sense strand to antisense strand
Antisense_DNA = Genomic_DNA;
for i = 1:length(Antisense_DNA)
    if Antisense_DNA(i) == 'G'
        Antisense_DNA(i) = 'C';
    elseif Antisense_DNA(i) == 'C'
            Antisense_DNA(i) = 'G';
    elseif Antisense_DNA(i) == 'A'
            Antisense_DNA(i) = 'T';
    elseif Antisense_DNA(i) == 'T'
            Antisense_DNA(i) = 'A';
    end
end

% ---------------------------------------------------------------------
% ---------------------------------------------------------------------

% ***** Finding 'G(N)20 GG' in sense DNA... *****

% Finding location of GG and G
GG_sense = strfind(Genomic_DNA, 'GG');
for j = 1:length(GG_sense)
    start_sense_pos = GG_sense(j) - 21; % - 21 because location of
'GG' is first G in GG

    if start_sense_pos <= 0
        GG_sense(j) = 0;
    else GG_sense(j) = GG_sense(j);
        if Genomic_DNA(start_sense_pos) == 'G'  % Finding if sequence
ends with G
        GG_sense(j) = GG_sense(j);
        else GG_sense(j) = 0;
        end
    end
end

% Finding location of G
GG_sense(GG_sense==0) = []; % Deleting zero elements from array
G_sense = GG_sense - 21; % location of G with respect to GG

% Finding sgRNA sequence with 'G(N)20 GG'
for k = 1:length(G_sense)
```

```
    sgRNA_sense_1 = Genomic_DNA(G_sense(k):GG_sense(k)+1);
    sgRNA_sense(k,1) = cellstr(sgRNA_sense_1);
end

% Tabulating sgRNA with corresponding locations
for l = 1:length(sgRNA_sense)
    sgRNA_sense_2 = char(sgRNA_sense(l));
    Position(l) = strfind(Genomic_DNA, sgRNA_sense_2);
    End_Position(l) = Position(l) + 22;
end

Position_S = Position';
End_Position_S = End_Position';
Position_S = [Position_S, End_Position_S];

filename = 'Generic_sgRNA.xlsx';
xlswrite(filename, Position_S, 'sgRNA Positions', 'C3');
xlswrite(filename, sgRNA_sense, 'sgRNA Positions', 'B3');

% ------------------------------------------------------------------
% ------------------------------------------------------------------


% ***** Finding 'GG(N)20 G' in antisense DNA... *****

% Finding location of GG and G
GG_antisense = strfind(Antisense_DNA, 'GG');
for j = 1:length(GG_antisense)
    end_antisense_pos = GG_antisense(j) + 22;

    if end_antisense_pos >= length(Antisense_DNA)
        GG_antisense(j) = 0;
    else GG_antisense(j) = GG_antisense(j);
        if Antisense_DNA(end_antisense_pos) == 'G'  % Finding if
sequence ends with G
        GG_antisense(j) = GG_antisense(j);
        else GG_antisense(j) = 0;
        end
    end
end

% Finding location of G
GG_antisense(GG_antisense==0) = []; % Deleting zero elements from
array
G_antisense = GG_antisense + 22; % location of G with respect to GG

% Finding sgRNA sequence with 'GG(N)20 G'
for k = 1:length(G_antisense)
    sgRNA_antisense_1 =
Genomic_DNA(G_antisense(k):GG_antisense(k)+1);
```

```
    sgRNA_antisense(k,1) = cellstr(sgRNA_antisense_1);
end

% Tabulating sgRNA with corresponding locations
for m = 1:length(sgRNA_antisense)
    sgRNA_antisense_2 = char(sgRNA_antisense(m));
    Position(m) = strfind(Antisense_DNA, sgRNA_antisense_2);
    End_Position(m) = Position(m) + 22;
end

Position_AS = Position';
End_Position_S = End_Position';
Position_AS = [Position_AS, End_Position_AS];

filename = 'Generic_sgRNA.xlsx';
xlswrite(filename, Position_AS, 'sgRNA Positions', 'H3');
xlswrite(filename, sgRNA_antisense, 'sgRNA Positions', 'G3');
```