Mammalian Artificial Chromosomes as a Synthetic Biology Tool for Transgene Expression

Charles Wang
Clemson University, charles.sj.wang@gmail.com

Follow this and additional works at: https://tigerprints.clemson.edu/all_theses

Recommended Citation
https://tigerprints.clemson.edu/all_theses/3233

This Thesis is brought to you for free and open access by the Theses at TigerPrints. It has been accepted for inclusion in All Theses by an authorized administrator of TigerPrints. For more information, please contact kokeefe@clemson.edu.
MAMMALIAN ARTIFICIAL CHROMOSOMES AS A SYNTHETIC BIOLOGY TOOL FOR TRANSGENE EXPRESSION

A Thesis
Presented to
the Graduate School of
Clemson University

In Partial Fulfillment
of the Requirements for the Degree
Master of Science
Chemical Engineering

by
Charles Shu-Jun Wang
December 2019

Accepted by:
Marc R. Birtwistle, Committee Co-Chair
Mark A. Blenner, Committee Co-Chair
Sarah Harcum
Mammalian artificial chromosomes, or MACs, have been studied as a potential avenue for hosting large numbers of transgenes in mammalian cells. MACs have several advantages over viral-based methods for transgene expression, including a lack of limits on loading capacity, which bypasses issues associated with integration into the genome. One area of research in which MACs can be applied is the biomanufacturing of protein-based therapeutics, where reported genome instability in Chinese hamster ovary (CHO) cells can lead to reduced product titer. MACs can potentially aid in solving this issue by providing alternate hosting sites for transgenes for integration of protein-based therapeutic production. However, some hurdles exist in the path of utilizing MACs as a biology tool, including the acquisition of sufficient mass and concentration of a MAC, the molecular cloning of a transgene into a MAC, and delivery of the cloned MAC to target mammalian cells. To address this, improvements were made at the steps of transformation of the MAC into \textit{E. coli}, isolation of positive colonies, and subsequent kit purification to generate sufficient masses and concentrations for downstream applications. Using Gibson Assembly, a selectable marker, glutamine synthetase (GS), was successfully cloned onto the MAC, yielding the construct MAC-GS. MAC-GS was subsequently electroporated into suspension CHO cells, and selection by removal of L-glutamine demonstrated the functionality of GS. These results represent a positive step forward for the implementation of MACs as a useful synthetic biology tool.
ACKNOWLEDGMENTS

I have so many people that I need to thank for helping me get through these last two and a half years. First and foremost, I have to thank my advisors, Marc Birtwistle and Mark Blenner. They took a huge chance on me when I started the program here at Clemson University, they’ve been supportive of me throughout the entire process of learning how to think critically and do good science (including the occasional critique whenever I didn’t know what I was talking about), and their financial support over the course of the last 8 months really helped keep me afloat enough to make it to the finish line. My other committee member, Sarah Harcum, has also been super-helpful, especially with regards to her knowledge of CHO cell culture and bioprocessing.

So many people contributed to the research that I did herein: Chris Saski, who was nice enough to let us use his PFGE machine; Alex von Kriegsheim, who ran our histone samples through mass spec free of charge; all of the AMBIC mentors that helped out on my project, including (but not limited to) Brita Anderson, Chris Tan, John Follit, and Ronan Kelly; Natalay Kouprina, Vladimir Noskov, Jung Hyun Kim, and the rest of the NIH folks, who were instrumental in helping us nail down how to handle artificial chromosomes; all of the graduate students in the Blenner and Birtwistle labs, whose general help got me through a lot of tough days; Caitlin and Scott Anglin, who basically crash-coursed me in molecular biology during my first year and a half of doing lab work; and, finally, Lu Xiaoming, Baylee Westbury, and Ben Usry, who put in tireless effort and many hours of work to help me collect all of these results.
Lastly, I have to thank my family. My sister, Bella Wang, has been my ultimate role model for my entire life; she and her husband, Carl Engle-Laird, have been incredibly supportive of me through the trials and tribulations of graduate school. And, last but not least, I absolutely would not be here without the support and opportunities that my parents, Shi Yuling and Wang Jianshe, have given me.
# TABLE OF CONTENTS

<table>
<thead>
<tr>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>TITLE PAGE ....................................................................................................................</td>
</tr>
<tr>
<td>ABSTRACT .....................................................................................................................</td>
</tr>
<tr>
<td>ACKNOWLEDGMENTS ..............................................................................................</td>
</tr>
<tr>
<td>LIST OF TABLES .........................................................................................................</td>
</tr>
<tr>
<td>LIST OF FIGURES ......................................................................................................</td>
</tr>
</tbody>
</table>

## CHAPTER

I. INTRODUCTION ........................................................................................................ 1

II. BACKGROUND AND SELECTED LITERATURE REVIEW .................................................. 5

   2.1 What Are MACs? ........................................................................................................ 5
   2.2 Chinese Hamster Ovary (CHO) Cells ....................................................................... 11
   2.3 Current Transgene Delivery Methods .................................................................... 13
   2.4 Summary ................................................................................................................ 20

III. PURIFICATION OF AN ARTIFICIAL CHROMOSOME IN HIGH YIELD .................. 22

   3.1 Rationale for Selection of Type of MAC .................................................................. 22
   3.2 Transformation of MAC into *E. coli* with Electroporation .................................. 23
   3.3 Purification of MAC with PureLink HiPure Plasmid Midiprep Kit ....................... 27
   3.4 MAC Characterization with Sanger Sequencing and Pulsed-Field Gel Electrophoresis .......................................................... 31
   3.5 Summary ................................................................................................................ 33

IV. MOLECULAR CLONING OF TRANSGENES ONTO AN ARTIFICIAL CHROMOSOME ......................................................... 35

   4.1 Gibson Assembly as a Cloning Method ................................................................. 35
Table of Contents (Continued)

4.2 Gibson Assembly of GS Expression Cassette into MAC ................................................................. 38
4.3 Verification of Gibson Assembly Success .......................................................... 42
4.4 Summary ........................................................................................................................................ 47

V. DELIVERY OF TRANSGENES ON AN ARTIFICIAL CHROMOSOME TO MAMMALIAN CELLS ................. 48

5.1 Verification of Positive Control Constructs
   NBS1-GFP and VHL-GFP ............................................................................................................. 48
5.2 Electroporation of NBS1-GFP into CHOZN GS<sup>-/-</sup> .................................................................. 51
5.3 Electroporation of CHOZN GS<sup>-/-</sup> with MAC-GS ............................................................... 55
5.4 Summary ........................................................................................................................................ 56

VI. SUMMARY OF EXPERIMENTAL RESULTS AND FUTURE DIRECTIONS ........................................ 57

6.1 Summary of Experimental Results ................................................................................................. 57
6.2 Future Directions ............................................................................................................................. 58

APPENDICES ......................................................................................................................................... 60

A: Histone-Mediated Transfection as a Delivery Method ................................................................. 61

A.1 Expression of Histones in \textit{E. coli} ............................................................................................. 61
A.2 Purification of Histones with Fast Protein Liquid Chromatography (FPLC) ................................. 64
A.3 Identification of Histones using Peptide Mass Fingerprinting ..................................................... 67
A.4 Ultrafiltration of Histones and Quantification of Histone Concentration ........................................ 67
A.5 Experimental Approach to Histone-Mediated Transfection ......................................................... 70
A.6 Histone-mediated Transfection in HEK 293 .................................................................................. 71
A.7 Electroporation and Histone-Mediated Transfection in CHOZN GS<sup>-/-</sup> ........................................ 75
A.8 Summary ......................................................................................................................................... 82

REFERENCES ......................................................................................................................................... 83
# LIST OF TABLES

<table>
<thead>
<tr>
<th>Table</th>
<th>Description</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>3.1</td>
<td>DNA mass and concentration requirements for molecular cloning and delivery</td>
<td>27</td>
</tr>
<tr>
<td>3.2</td>
<td>Modifications to the PureLink HiPure Plasmid Midiprep Kit protocol</td>
<td>29</td>
</tr>
<tr>
<td>A.1</td>
<td>Accession numbers from mass spectrometry and their corresponding identities from UniProt</td>
<td>67</td>
</tr>
</tbody>
</table>
LIST OF FIGURES

<table>
<thead>
<tr>
<th>Figure</th>
<th>Description</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>2.1</td>
<td>Visual interpretation of the creation of ‘top-down’ artificial chromosomes</td>
<td>6</td>
</tr>
<tr>
<td>2.2</td>
<td>Visual depiction of the process by which the alphoid tetO MAC is constructed</td>
<td>8</td>
</tr>
<tr>
<td>2.3</td>
<td>The process of generating a SATAC</td>
<td>9</td>
</tr>
<tr>
<td>2.4</td>
<td>A diagram of how the ACE System works</td>
<td>10</td>
</tr>
<tr>
<td>2.5</td>
<td>Flow diagram of the mechanism of action of cationic lipid transfection</td>
<td>14</td>
</tr>
<tr>
<td>2.6</td>
<td>Diagram of MMCT</td>
<td>16</td>
</tr>
<tr>
<td>2.7</td>
<td>Diagram of PEG-mediated cell fusion</td>
<td>18</td>
</tr>
<tr>
<td>2.8</td>
<td>Histone-mediated transfection of DsRed2-H2A in HeLa cells</td>
<td>20</td>
</tr>
<tr>
<td>3.1</td>
<td>A sequence map of the MAC</td>
<td>24</td>
</tr>
<tr>
<td>3.2</td>
<td>A sequence map of RCA-Sat43</td>
<td>25</td>
</tr>
<tr>
<td>3.3</td>
<td>Chloramphenicol-containing agar plate of E. coli DH10β cells transformed with MAC via electroporation</td>
<td>26</td>
</tr>
<tr>
<td>3.4</td>
<td>Sanger sequencing results aligned with the MAC</td>
<td>31</td>
</tr>
<tr>
<td>3.5</td>
<td>A diagram of PFGE electrophoresis</td>
<td>32</td>
</tr>
<tr>
<td>3.6</td>
<td>A PFGE image of purified MAC samples, digested with SacI</td>
<td>33</td>
</tr>
<tr>
<td>4.1</td>
<td>Diagram of how Gibson Assembly works</td>
<td>36</td>
</tr>
<tr>
<td>4.2</td>
<td>A diagram of the process of inserting the GS expression into the MAC</td>
<td>37</td>
</tr>
<tr>
<td>Figure</td>
<td>Description</td>
<td>Page</td>
</tr>
<tr>
<td>--------</td>
<td>-----------------------------------------------------------------------------</td>
<td>------</td>
</tr>
<tr>
<td>4.3</td>
<td>Agarose gel of PCR products from pcDNA3.1_Zeo(+) GS</td>
<td>39</td>
</tr>
<tr>
<td>4.4</td>
<td>Plasmid map of pcDNA3.1_Zeo(+) GS</td>
<td>40</td>
</tr>
<tr>
<td>4.5</td>
<td>Agarose gel containing the purified insert and vector fragments for Gibson Assembly</td>
<td>41</td>
</tr>
<tr>
<td>4.6</td>
<td>Agar plate of <em>E. coli</em> DH10β transformed with the MAC-GS Gibson Assembly product</td>
<td>42</td>
</tr>
<tr>
<td>4.7</td>
<td>PCR verification of Gibson Assembly, using Gibson Assembly primers</td>
<td>43</td>
</tr>
<tr>
<td>4.8</td>
<td>The location of the verification primers on MAC-GS</td>
<td>44</td>
</tr>
<tr>
<td>4.9</td>
<td>Agarose gel of PCR with verification primers</td>
<td>44</td>
</tr>
<tr>
<td>4.10</td>
<td>PFGE analysis of MAC-GS</td>
<td>46</td>
</tr>
<tr>
<td>4.11</td>
<td>Sequencing reads aligned to the GS expression cassette</td>
<td>47</td>
</tr>
<tr>
<td>5.1</td>
<td>A plasmid map of VHL-GFP</td>
<td>49</td>
</tr>
<tr>
<td>5.2</td>
<td>PFGE image of VHL-GFP and NBS1-GFP</td>
<td>50</td>
</tr>
<tr>
<td>5.3</td>
<td>Fluorescence images of HT1080 fibrosarcoma cells transfected with X-tremeGENE9 (Roche)</td>
<td>51</td>
</tr>
<tr>
<td>5.4</td>
<td>Fluorescence overlay image of CHOZN GS-/- cells electroporated with NBS1-GFP</td>
<td>53</td>
</tr>
<tr>
<td>5.5</td>
<td>Fluorescence overlay image of CHOZN GS-/- cells independently electroporated with NBS1-GFP</td>
<td>53</td>
</tr>
<tr>
<td>5.6</td>
<td>Flow cytometry data of CHOZN GS/-/ cells electroporated with NBS1-GFP</td>
<td>55</td>
</tr>
<tr>
<td>5.18</td>
<td>MAC-GS viability and viable cell density under selection by removal of L-glutamine</td>
<td>56</td>
</tr>
<tr>
<td>Figure</td>
<td>Description</td>
<td>Page</td>
</tr>
<tr>
<td>--------</td>
<td>-----------------------------------------------------------------------------</td>
<td>------</td>
</tr>
<tr>
<td>A.1</td>
<td>SDS-PAGE of un-induced and induced <em>E. coli</em> cultures</td>
<td>63</td>
</tr>
<tr>
<td>A.2</td>
<td>FPLC chromatogram</td>
<td>65</td>
</tr>
<tr>
<td>A.3</td>
<td>FPLC chromatogram, zoomed in to elution step gradient</td>
<td>65</td>
</tr>
<tr>
<td>A.4</td>
<td>SDS-PAGE of FPLC fractions post-purification</td>
<td>66</td>
</tr>
<tr>
<td>A.5</td>
<td>SDS-PAGE of filtrates and retentates from ultrafiltration</td>
<td>69</td>
</tr>
<tr>
<td>A.6</td>
<td>Standard curve from BCA assay</td>
<td>70</td>
</tr>
<tr>
<td>A.7</td>
<td>Initial histone-mediated transfection experiments</td>
<td>73</td>
</tr>
<tr>
<td>A.8</td>
<td>HMT experiments with increased histone:DNA ratios</td>
<td>74</td>
</tr>
<tr>
<td>A.9</td>
<td>Experimental steps for each of the combinations of HMT and electroporation</td>
<td>75</td>
</tr>
<tr>
<td>A.10</td>
<td>Representative fluorescence images of combinations of electroporation and HMT</td>
<td>78</td>
</tr>
<tr>
<td>A.11</td>
<td>Representative fluorescence images of HMT/E with a 4:1 histone:DNA ratio</td>
<td>79</td>
</tr>
<tr>
<td>A.12</td>
<td>Representative fluorescence images of HMT/E with increased histone – DNA complexation times</td>
<td>81</td>
</tr>
</tbody>
</table>
CHAPTER ONE
INTRODUCTION

Synthetic biology represents the intersection of biology and engineering: designing and constructing circuits and networks of multiple transgenes in order to achieve various objectives, such as producing therapeutics or probing cellular biology in ways that other subfields of biology cannot. The ability to synthesize these transgene circuits has become easier and easier, as the cost of directly synthesizing DNA has fallen further and further. However, these circuits have zero use if the genes cannot be delivered to the target organisms in order to be used, and this delivery of multiple transgenes, especially into mammalian cells, remains a major hurdle for carrying out synthetic biology (Martella et al., 2016; Rennekamp, 2018).

An example of such a construct that is of interest in synthetic biology is the mammalian artificial chromosome (MAC). Also referred to in the literature as a human artificial chromosomes (HAC), MACs have been studied as a potential avenue for hosting large numbers of transgenes in mammalian cells. As the name implies, MACs are non-endogenous to a mammalian cell but contain elements recognizable by the cell for the purposes of DNA propagation through mitosis, like the centromere. MACs exhibit some features that are desirable for the purposes of synthetic biology. A MAC serves as a site for hosting transgenes that does not interfere with the function of the endogenous genome. Current viral vector-based methods for delivering transgenes to mammalian cells integrate them into the genome randomly and in locations that are accessed to different degrees by transcriptional machinery. Known as the ‘position effect,’ this
phenomenon may adversely affect transgene expression as a result. In contrast, a MAC serves as a well-defined, stable environment for hosting transgenes (Martella et al., 2016).

However, some challenges present in regards to utilization of MACs in synthetic biology. Acquiring enough of the MAC to carry out downstream processes is difficult: conventional plasmid preparation kits exhibit difficulties in extracting plasmids larger than 10 – 20 kb. Molecular cloning of transgenes onto a MAC is also a non-trivial challenge; site-specific recombination, while widely used in the literature, is limited in regards to cloning of multiple transgenes, and more conventional methods like restriction enzyme cloning or Gibson Assembly have not been extensively used in the literature. Furthermore, as previously mentioned, delivery of the MAC to mammalian cells is a significant roadblock for utilizing it in synthetic biology.

One such area that MACs could potentially address problems in is biomanufacturing of recombinant protein-based therapeutics, such as monoclonal antibodies (mAbs). The market for such therapeutics is large; world-wide sales of mAbs projected to reach $125 billion by 2020 (Ecker et al., 2015). Chinese hamster ovary (CHO) cells are the cell line most commonly used to produce protein-based therapeutics; 70% of those sales constitute therapeutics that were produced in CHO cells (Kunert and Reinhart, 2016). However, loss of mAb titer has been observed over time in CHO cell cultures, even with selection pressure. This loss is attributed to genomic instability resulting from rearrangement of genomic elements (Kunert and Reinhart, 2016; Wurm and Wurm, 2017). MACs can potentially offer a solution to this problem; hosting
transgenes on a separate chromosome with no known transposable elements that the CHO cell can recognize for sequence rearrangement. This would allow the CHO cell to maintain intact transgene cassettes inside the nucleus and thus retain acceptable protein titer for biomanufacturing purposes over the length of 60 generations, a typical duration for assessing clonal stability of transgenes in CHO cell cultures (Capella Roca et al., 2019).

This project was divided into three challenges in the expression of MACs in CHO cells: 1) obtaining sufficient MAC material for the delivery to CHO cells; 2) molecular cloning a reporter transgene into the MAC; and 3) delivering the cloned MAC to CHO cells. Currently, MACs are produced in *E. coli*, and the purification steps do not achieve high yields, due to the construct’s large size, low copy number, and instability of its centromeric repeats. Inserting transgenes into the MAC has typically been performed with site-specific recombination. However, using site-specific recombination to insert multiple transgenes into the MAC has some drawbacks, such as a lack of flexibility and cumbersome implementation. Delivery of the MAC into suspension CHO cells has not been extensively explored previously in the literature, so finding a suitable method of delivery was necessary.

As a summary of this thesis, Chapter 2 provides some background on the various aspects of this project and a brief review of literature directly relevant to this project, with a particular focus on MACs themselves, CHO cells, and various gene delivery methods. Chapter 3 addresses the efforts to obtain the MAC in high yield and purity, including transformation into *E. coli*; modifications of plasmid preparation procedure; and various
methods of verification and characterization of the MACs post-purification. Chapter 4 addresses molecular cloning of a transgene, including Gibson Assembly of the transgene into the MAC and subsequent verification of the sequence and size. Chapter 5 describes the studies to deliver the MAC to mammalian cells, including studies of electroporation of CHOZN GS\(^{+/−}\) with a positive control construct and the MAC loaded with a transgene as well as subsequent assessment of the electroporation results. Chapter VI summarizes the results of the project and future directions for the project. The appendix contains studies of histone-mediated transfection as a delivery method, including expression and purification of histones for HMT; pilot HMT studies in HEK 293 cells with HMT; and subsequent transfection studies with HMT and electroporation in CHOZN GS\(^{+/−}\).
CHAPTER TWO
BACKGROUND AND SELECTED LITERATURE REVIEW

This chapter details pertinent background and a review of selected literature about various topics: MACs themselves, the different types that exist, and the rationale behind choosing a particular type of MAC to work with; CHO cells as they pertain to the recombinant protein industry and their particular characteristics; and various methods used for transgene delivery to mammalian cells.

2.1 What Are MACs?

A mammalian artificial chromosome (MAC), as the name implies, is a DNA construct that is constructed synthetically and not the same as one of the constituent chromosomes of a mammalian cell. It contains a centromere, which allows the cell to split copies properly among daughter cells during mitosis. Because of this property, MACs have been considered as a promising way to host genes within a mammalian cell without some of the drawbacks of current methods for integrating transgenes into a cell’s genome (Kouprina et al., 2018).

Three types of artificial chromosomes exist in the literature: ‘top-down’ artificial chromosomes, ‘bottom-up’ (or de novo) artificial chromosomes, and satellite-DNA-based artificial chromosomes (SATACs).

2.1.1 ‘Top-Down’ MACs

“Top-down” formation of a MAC involves removing endogeneous genes from existing chromosomes that are already inside of a cell in order to isolate the centromere,
before inserting transgenes of interest, usually through site-specific recombination (SSR) techniques. Figure 2.1 shows this process. Endogenous genes were broken off by telomere-associated chromosome fragmentation (TACF), in which cells were transfected with plasmids that contained telomere sequences and were targeted to a particular chromosome. Through homologous recombination, that chromosome’s endogenous genes were swapped out with the telomeric plasmids, thus generating a miniature chromosome of telomeres, a centromere, and SSR acceptor sites (Farr et al., 1992; Farr et al., 1995; Kazuki et al., 2011; Vollrath et al., 1988).

![Figure 2.1](image_url). Visual interpretation of the creation of ‘top-down’ artificial chromosomes. Taken from Kazuki et al., 2011; image licensed under a Creative Commons Attribution-NonCommercial-No Derivative Works 3.0 Unported License: [http://creativecommons.org/licenses/by-nc-nd/3.0/](http://creativecommons.org/licenses/by-nc-nd/3.0/).

2.1.2 ‘Bottom-Up’ MACs

‘Bottom-up’ formation of a MAC relies on the fact that the centromere of a human chromosome consists of repeats of non-coding DNA, termed ‘satellite DNA.’ Alpha-satellite (or alphoid) DNA, the most common type of satellite DNA, consists of repeats of a 171 bp sequence arranged in a head-to-tail tandem repeat array. These arrays are organized further into higher-order repeats in order to form a functional centromere (Kouprina et al., 2018; Molina et al., 2017).

Previous groups constructed bottom-up MACs by making copies of these alphoid sequences through rolling-circle amplification (RCA). These longer fragments were
assembled into a target vector in *S. cerevisiae* through transformation-associated recombination (TAR). The process of TAR involves transforming linear fragments with homologous sequences at their ends and having *S. cerevisiae* cells join the fragments into a single construct (Kouprina and Larionov, 2008). This process yielded a partially-formed MAC of ~60 kb that can be transfected into mammalian cells by lipofectamine or other methods (Ebersole et al., 2005).

Once the partially-formed MAC was transfected into mammalian cells, the cells recognized the presence of centromeric DNA and synthesized centromeric protein A (CENP-A). Interactions with the centromere protein B (CENP-B) box, an operator sequence for CENP-B, helped to establish chromatin formation and amplify the existing centromeric repeats to form a fully functional MAC (Okamoto et al., 2007). An SSR acceptor site was inserted via homologous recombination for gene delivery purposes (Kouprina et al., 2018).

A newer version of the bottom-up MAC, the alphoidtetO MAC, incorporates the tetracycline operator (tetO) into alphoid DNA sequences for conditional inactivation when the tetracycline repressor protein is present (Nakano et al., 2008). Figure 2.3 shows the process by which the alphoidtetO MAC was constructed.
2.1.3 Satellite-DNA-based Artificial Chromosomes (SATAcs)

SATAcs are MACs that contains elements of both top-down and bottom-up construction (Lindenbaum et al., 2004; Katona, 2015). Foreign DNA is integrated into ribosomal RNA gene sites on existing chromosomes, which induces large-scale amplification of the p arms. During this large scale amplification of the p arms, the centromere can be duplicated. When cell division occurs, a chromosome with a duplicate centromere undergoes breakage in the region between the two centromeres, producing a SATAC.

Figure 2.2. Visual depiction of the process by which the alphoid^{tetO} MAC was constructed. Taken from Kouprina et al., 2018 and reproduced with permission.
The SATAC was the basis for the Artificial Chromosome Expression (ACE) System, an attempt to establish a MAC-based platform to host transgenes in order to produce recombinant proteins. As seen in Figure 2.5, the ACE System, as a whole, has three components: the Platform ACE, a SATAC with a promoter, an SSR acceptor site \((attP)\), and a selectable marker for resistance to an antibiotic (puromycin). The ACE Integrase is a lambda integrase that has a point mutation to remove the requirement for host co-factors. The ACE Targeting Vector has an SSR donor site \((attB)\), the transgene
cassette, and a promoterless selectable marker. The recombination reaction places the ACE Targeting Vector’s selectable marker downstream of the Platform ACE’s promoter; thus, the Loaded ACE has a new selectable marker that can be used for subsequent selection to verify a successful recombination (Lindenbaum et al., 2004; Tóth et al., 2013).

Figure 2.4. A diagram of how the ACE System works. From Tóth et al., 2013; image licensed under a Creative Commons Attribution 4.0 International License: https://creativecommons.org/licenses/by/4.0/

2.1.4 Summary

The three different types of MACs all have interesting properties. Top-down MACs are retrofitted chromosomes; the existing genes on a chromosome are removed to
isolate the centromere so that new genes can be inserted. In that way, they are very
different from bottom-up chromosomes, which are built de novo and assembled piece by
piece. SATACs have elements of both top-down and bottom-up construction, in that they
are derived from existing chromosomes but exhibit elements of de novo construction in
how DNA sequences are amplified.

2.2 Chinese Hamster Ovary (CHO) Cells

2.2.1 CHO Cell Types as Related to the Biotechnology Industry

Many different CHO cell lines exist, some with useful genetic modifications. CHO-K1 is a clone of the original CHO line derived from Theodore Puck’s laboratory, from which many of these other CHO lines were derived (Wurm and Wurm, 2017). CHO-DXB11 was derived in 1980 from mutagenization of CHO-K1 cells with ethyl methanesulfonate, which deleted one locus of the gene for dihydrofolate reductase (DHFR) and generated a missense mutation in the other locus, thus making a \( DHFR^{+/−} \) variant (Urlaub and Chasin, 1980). CHO-DG44 was a later attempt at generating a \( DHFR^{+/−} \) variant that utilized gamma radiation to delete both alleles, as some CHO-
DXB11 cells demonstrated the ability to revert to the \( DHFR^{+/+} \) phenotype (Urlaub et al.,
1983). This mutation really opened the door for usage of CHO cells in bioprocessing;
Genentech succeeded in producing tissue plasminogen activator in CHO cells, and the
strategy used therein served as a model for the entire industry as a whole (Wurm and
Hacker, 2011).

Auxotrophic selection for CHO cells typically revolves around the genes that
code for two enzymes required for anabolism in CHO cells: DHFR and glutamine
synthetase (GS). DHFR converts dihydrofolate to tetrahydrofolate, which is an important cofactor in the biosynthesis of purines and thymidine (Urlaub and Chasin, 1980). With \textit{DHFR} knocked out, or with DHFR inhibited by methotrexate (MTX), CHO cells require glycine, hypoxanthine, and thymidine in the medium or a functional copy of \textit{DHFR} in order to grow (Wuest et al., 2012). GS ligates ammonia to L-glutamate to form L-glutamine and is the biosynthetic pathway for L-glutamine in CHO cells. With \textit{GS} knocked out, or with GS inhibited by methionine sulfoximine (MSX), CHO cells require either L-glutamine or a functional copy of \textit{GS} in order to grow (“CHOZN® Platform Technical Bulletin”).

2.2.2 The CHOZN® GS\textsuperscript{−/−} Cell Line

Developed by Sigma Aldrich Fine Chemicals (SAFC), the CHOZN® GS\textsuperscript{−/−} cell line was derived from a CHO-K1 adherent cell line that was adapted to suspension growth. Zinc finger nucleases were then used to alter a section of exon 6 of \textit{GS}. Exon 6 corresponds to the substrate-binding domain of GS, in order to render the protein non-functional. This made the CHO cells into glutamine auxotrophs. L-glutamine removal is a common route of selection in the biopharmaceutical industry because it is unstable at physiological pH in liquid cell culture medium and is thus usually added at time of use. Implementing L-glutamine selection is rather easy as a result, since addition of it is already a separate step in medium preparation. MSX selection is undesirable because it is toxic, it is an additional chemical that adds on cost to a bioprocess, and it requires subsequent removal during downstream purification. Knocking out \textit{GS} sidesteps all of the issues associated with MSX selection (“CHOZN® Platform Technical Bulletin”).
2.3 Current Transgene Delivery Methods

Existing methods in the literature for delivering genes to mammalian cells include cationic lipids, electroporation, microcell-mediated chromosome transfer (MMCT), polyethylene glycol (PEG)-mediated cell fusion, and histone-mediated transfection (HMT).

2.3.1 Cationic Lipids

Cationic lipid transfection has been widely used to transfect mammalian cells, especially of the adherent variety. The basic procedure involves three steps. First, the cationic lipid molecule interacts with the negatively-charged backbone of the DNA payload, forming an electrostatic complex. That complex then interacts further with the negatively-charged phospholipid membrane of the cell, entering it through an endocytotic pathway. The DNA payload then disengages from the cationic lipid and diffuses through the cytoplasm to the nucleus for gene expression (Chesnoy and Huang, 2000).

While cationic lipid transfection works fairly well for adherent cell lines, transfection efficiency is not as high for suspension cells. One given reason for this is that the cationic lipid complex does not attach to the surface of suspension cells very well, leading to less uptake of the lipid complexes (Basiouni et al., 2018).
2.3.2 Electroporation

Electroporation is another widely-used method for delivering genes to mammalian cells. As the name implies, electroporation involves subjecting cells to a strong electric field. The basic procedure of electroporation involves mixing cells with a DNA payload in suspension and loading the mix into a special cuvette that contains metal plates. The cuvette and content are then subjected to a high voltage. The prevailing notion is that the electric field weakens the cell membrane and creates pores, allowing DNA to diffuse into the cell. However, much about electroporation’s mechanism of action is up for debate (Escoffre et al., 2008).

One of the disadvantages of electroporation is the low resultant cell viability; altering a cell’s membrane can lead to cell death for many cells (Shigekawa and Dower,
1988). Despite low cell viability following electroporation, this method is the preferred gene delivery method for the CHOZN GS⁻ cell line ("CHOZN® Platform Technical Bulletin").

2.3.3 Microcell-mediated Chromosome Transfer (MMCT)

In terms of transferring MACs between cells, MMCT is the most commonly used method. MMCT involves 1) micronucleation of a donor cell line’s chromosomes, 2) separation of micronuclei from the cytoplasm, and 3) fusion with recipient cell lines, as shown in Figure 2.7. The first step begins by applying colcemid, a small molecule drug that arrests cells in the metaphase by inhibiting formation of the mitotic spindle. By inhibiting mitosis for long enough, nuclear envelopes will form around the individual chromosomes of the cell, creating micronuclei that contain between one and a few chromosomes inside of each cell. These micronuclei are important in order to isolate chromosomes. To separate the micronuclei from the cytoplasm, cells are then centrifuged at high speeds in the presence of cytochalasin B, which is used to separate the micronuclei from the cytoplasm of the cells. The micronuclei are subsequently filtered in order to collect micronuclei containing a single chromosome. To fuse the micronuclei with other cell lines, various cell fusion methods, including the use of Sendai virus, phytohemagglutinin-P (PHA-P), and polyethylene glycol (PEG), are then used to fuse micronuclei with recipient cell lines (Ege and Ringertz, 1974; Fournier and Riddle, 1977; Meaburn et al. 2005; Hiratsuka et al., 2015; Suzuki et al., 2016).
MMCT has some disadvantages. One such disadvantage is low efficiency for the transfer of micronuclei from donor cells to recipient cells. Another major disadvantage is the limited number of donor cell lines that can undergo MMCT. CHO and mouse A9
cells are the most commonly used donor lines, CHO and A9 cells undergo repeated hyperploidization, creating more chromosomes. This produces many micronuclei, increasing the chances of micronuclear fusion success with recipient cell lines (Liskovykh et al., 2016).

2.3.4 Polyethylene glycol (PEG)-mediated Cell Fusion

The methods for transgene delivery mentioned above are conventional for MACs. However, only a limited number of other methods exist in the literature. One novel method, described in Brown et al. (2016), used *S. cerevisiae* as the carrier for large DNA molecules by fusing the yeast cells to mammalian cells using polyethylene glycol (PEG) and dimethyl sulfoxide (DMSO). TAR was then used to join linear fragments of DNA into a yeast centromeric plasmid (YCp), as shown in Figure 2.8. Two fluorescent proteins, mCherry and GFP, were utilized with organism-specific promoters to evaluate efficiency of cell fusion and expression in mammalian cells, respectively. The flexibility of the YCp design via TAR allowed the authors to produce constructs of varying sizes, up to 1.1 Mb, which is approximately the size of a fully-formed MAC in a mammalian cell and about 18 times bigger than the 60 kb precursor MAC construct.
While this article presented a novel idea for getting around the barriers presented by more direct methods of gene delivery, implementing it would be a challenge in a bioprocessing context, because fungal contamination is seen as undesirable in mammalian cell cultures. Fungi can cause product spoilage by producing unwanted byproducts and ruining products, and said unwanted byproducts can also pose health hazards (Sandle, 2019).

2.3.5 Histones and Histone-mediated Transfection (HMT)

Histone-mediated transfection as a method for delivering transgenes to cells has been a subject of study since the late 1990s. Histones, as the proteins that condense DNA into chromatin, have a propensity for binding DNA via electrostatic interactions. Histones have also been found to contain nuclear localization signals (NLS), which facilitate transport from the cytoplasm to the nucleus. As such, histones have some
potential as carriers of transgenes. The linker histone (H1), the core histones that make up the nucleosome (H2A, H2B, H3, and H4), and other histone-like proteins have been explored at length for their propensities for transgene delivery (Han et al., 2019).

One such study on HMT explored the capabilities of core histones H2A and H2B to transfect mammalian cells in order to deliver transgenes. Fusion proteins were constructed from combinations of histone H2B and an NLS from the simian virus 40 large tumor antigen, with green fluorescent protein (GFP) included. These proteins were used as carrier molecules for a DNA construct that expressed a fusion protein, histone H2A connected to the fluorescent protein DsRed2. The H2B-containing fusion proteins were shown to be capable of delivering the plasmid payload to the nucleus, as measured by expression of DsRed2-H2A, with LipofectAMINE 2000 as a positive control. Including the NLS on the H2B-GFP fusion protein and histone H2A appeared to have an additive effect on the percentage of cells expressing DsRed2-H2A, as seen in Figure 2.9 (Wagstaff et al., 2007).

One possible limitation of histone-mediated transfection is that it acts similarly to cationic lipids in its mode of action. Like cationic lipids, histones also associate with DNA largely through electrostatic interactions. The similarity of the mechanism of action for attracting DNA could imply that histones also share the same pitfalls as cationic lipids when it comes to transfecting suspension cells.
2.4 Summary

Chapter 2 covers much of the important background information concerning MACs, CHO cells, and gene delivery methods. Three types of MACs exist in the literature: top-down chromosomes, in which existing chromosomes have their endogenous genes removed in order to isolate the centromere and load new transgenes; bottom-up chromosomes, which are constructed de novo and use the target mammalian cell’s machinery to build up the centromere; and SATACs, which employ elements of both top-down and bottom-up MACs in construction and have been investigated as platforms for recombinant protein production.

A brief history of CHO cells is also covered in Chapter 2; the various strains of CHO cells are described along with common methods of selection for the purposes of
recombinant protein production. Particular attention is paid to the CHOZN GS⁻/⁻ line, as the target cell line for much of the work in this thesis.

In addition, some transgene delivery methods in use for mammalian cells are discussed. Cationic lipid transfection is widely used among adherent cells, but it has some efficiency issues with suspension cell lines. Electroporation is also widely used but has the side effect of decreased cell viability post-electroporation. MMCT is a technique specific to MACs that is commonly used to transfer MACs between mammalian cell lines, but it also exhibits efficiency issues along with a limited number of cell lines that can support the technique. PEG-mediated cell fusion is interesting, especially in its ability to transfer very large DNA constructs, but it has issues when applied to real-world situations. HMT has some interesting properties as a transfection agent, like histones’ affinities for DNA and nuclear localization properties, and additive effects on transfection efficiencies have been shown, but it potentially runs into some of the same drawbacks that cationic lipid transfection does because of the similarity in the mechanism of action.
CHAPTER THREE
PURIFICATION OF AN ARTIFICIAL CHROMOSOME IN HIGH YIELD

The first major challenge to overcome in order to utilize the MAC as a synthetic biology tool was to obtain enough of it in high yield for use in downstream applications. As mentioned previously, conventional plasmid preparation kits have difficulties with isolating plasmids larger than 10 – 20 kb in large quantities, whereas the MAC is sized in the area of 60 kb. This chapter describes the MAC itself and its characteristics as well as the work conducted to improve the isolation and purification of the MAC.

3.1 Rationale for Selection of Type of MAC

When considering whether to utilize top-down or bottom-up MACs for this project, the main factor in the decision-making process was how to insert transgenes onto the MAC. The majority of groups that have published on MACs use SSR after delivering them to cells in order to load transgenes onto MACs, and several groups have published on integration of multiple genes onto a MAC (Yamaguchi et al., 2011; Tóth et al., 2014; Suzuki et al., 2014; Yoshimura et al., 2015; Lee et al., 2018). However, loading multiple transgenes through SSR has some limitations. Efforts to implement sequential integration of transgenes onto a MAC have required an iterative implementation that tacks on additional time as the number of transgenes increases; simultaneous integration of multiple transgenes requires a different integrase for each transgene. Conventional cloning methods, like Gibson Assembly, offer more flexibility and less cumbersome ways to load multiple transgenes onto a MAC but have not been extensively explored in
the literature. Cloning was hypothesized to be easier with bottom-up MACs because they are, at some point, purified before delivery to mammalian cells, which offers an opportunity to utilize conventional cloning methods in order to introduce transgenes onto the MAC. In contrast, top-down MACs and SATACs are constructed and loaded with transgenes exclusively inside a cellular environment, ruling out the opportunity to use conventional cloning methods for loading transgenes.

3.2 Transformation of MAC into E. Coli with Electroporation

As mentioned previously, bottom-up MACs were selected as the MAC for this project. A copy of the MAC was donated by Dr. Natalay Kouprina of the National Institutes of Health. The MAC, as previously described in Nakano et al. (2008), contained an ~50 kb α-satellite DNA cloned into a 10,209 bp vector (referred to in Nakano et al. as ‘RCA-Sat43’). This MAC also contained other useful gene cassettes for compatibility with different organism types. A sequence map of the MAC with α-satellite repeats is shown in Figure 3.1.
RCA-Sat43, shown in Figure 3.2, contains a bacterial artificial chromosome (BAC) cassette that includes the sopABC genes and origin of replication (ori2) that allow bacteria to partition a plasmid as if it were a chromosome, as well as a selectable marker that confers resistance to chloramphenicol (CamR). It also has a yeast artificial chromosome cassette (YAC) that includes the CEN6/ARSH4 gene, a fusion of the centromere from chromosome VI of S. cerevisiae fused to an autonomously replicating
sequence associated with histone H4. These two cassettes were present in order to facilitate RCA and TAR, as mentioned in Chapter 2. In addition, a blasticidin resistance cassette with an SV40 promoter and poly(A) signal were included, for selection in mammalian cells. The α-satellite repeats were inserted at the MluI restriction enzyme site during TAR to generate the MAC seen in Figure 3.1.

Figure 3.2. A sequence map of RCA-Sat43. Image generated by SnapGene software (from GSL Biotech; available at snapgene.com).
In order to transform this into *E. coli*, due to the large size of the DNA construct in question, conventional heat shock methods were deemed nonviable for transformation. As such, higher-efficiency techniques, like electroporation, were used to deliver the MAC to *E. coli*. 1 ng of the MAC was transformed into 15 µL 10-beta Electrocompetent *E. coli* cells (New England Biolabs) in a 1 mm gap cuvette at 1800 V, 25 µF, and 200 Ω.

Colonies were subsequently streak-plated onto agar plates containing chloramphenicol for selection, and colonies were observed, as seen in Figure 3.3. Of note, the outgrowth step and incubation steps were carried out at 30°C instead of 37°C because the array of centromeric repeats is unstable, and repeats can be lost during the transformation process if carried out at typical conditions. The cells were shaken at lower-than-normal speeds during the outgrowth step, as well, for this reason.

Figure 3.3. Chloramphenicol-containing agar plate of *E. coli* DH10β cells transformed with MAC via electroporation.
importantly, two key steps following successful transformation were also affected by this possibility of losing repeats during transformation. Post-plating, instead of picking a single colony, multiple colonies were picked for making glycerol stocks. In addition, when growing cultures from glycerol stocks, instead of simply stabbing the glycerol, an entire stock would be used for each culture. The idea behind this was to hedge bets on how the centromeric repeats were maintained across colonies by selecting multiple colonies and using a larger sample size of bacterial colonies in order to recover MAC of the correct size.

3.3 Purification of MAC with PureLink HiPure Plasmid Midiprep Kit

Once E. coli cells had been successfully transformed with the MAC, extracting and purifying it via plasmid preparation was necessary for subsequent characterization and molecular cloning. Table 3.1 lists the DNA mass and concentration requirements for the molecular cloning and delivery methods used later on. While molecular cloning itself did not require high DNA masses and concentrations, delivery did require high amounts and concentrations of DNA.

<table>
<thead>
<tr>
<th>Application</th>
<th>DNA Mass Required</th>
<th>DNA Concentration Required</th>
</tr>
</thead>
<tbody>
<tr>
<td>Molecular Cloning</td>
<td>250 ng</td>
<td>50 ng/µL</td>
</tr>
<tr>
<td>Delivery</td>
<td>30 – 50 µg</td>
<td>600 – 1000 ng/µL</td>
</tr>
</tbody>
</table>

Purification of the MAC, however, was a non-trivial challenge. Most kits for plasmid preparation have an upper limit on the size of plasmid that can be purified; that limit is often in the range of 10 – 20 kb, and yields drop as the size of the plasmid...
increases. Furthermore, the MAC’s origin of replication for *E. coli* is derived from the F plasmid (Kononenko et al., 2005), which is inherently single-copy in *E. coli*.

From Nakano et al. (2008), the Large Construct Kit (QIAGEN) was the kit of choice for extracting the MAC from *E. coli*. The kit itself is technically challenging to carry out, though; its associated protocol features two isopropanol extractions, a 1 h exonuclease digestion step, and multiple high speed spins of 30 mins to 1 h, all of which adds up to a protocol runtime of 7 h or more. Yields from the Large Construct Kit, from a 500 mL culture, were on the order of 9 µg at a concentration of 45 ng/µL, for downstream applications.

A search of literature on MACs yielded a kit that had been used in another paper by Logsdon et al. (2019) to extract MACs from bacteria, the PureLink HiPure Plasmid Midiprep Kit (Thermo Fisher Scientific). The kit itself had additional protocol modifications for large constructs like bacterial artificial chromosomes (BACs), and it required much less culture for extraction (100 mL vs. the Large Construct Kit’s 500 mL). Initial midipreps from this kit yielded plasmid masses of around 4 µg at 40 ng/µL, which was still not enough to meet the DNA demands of electroporation.

Subsequent optimization of the stock protocol was carried out to increase yields. A table detailing the original protocol (“PureLink™ HiPure Plasmid DNA Purification Kits”) and major changes made can be found below, in Table 3.2.
Table 3.2. Modifications to the PureLink HiPure Plasmid Midiprep Kit protocol.

<table>
<thead>
<tr>
<th>Original Step</th>
<th>Modified Step (If Changed)</th>
<th>Reasoning</th>
</tr>
</thead>
<tbody>
<tr>
<td>Prepare a 20-h bacterial culture of 100 mL, grown to an OD_{600} of 5.0.</td>
<td>Grow 100 mL bacterial culture to OD_{600} of 0.2 (as measured on Nanodrop pedestal).</td>
<td>An OD_{600} of 0.2 on the Nanodrop pedestal roughly corresponds to the switch from exponential to stationary phase. This maximizes MAC yield.</td>
</tr>
<tr>
<td>Apply 10 mL Equilibration Buffer to the midiprep column. Allow solution to drain by gravity flow.</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Centrifuge bacterial cells at 9000 x g for 15 mins. Remove supernatant.</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Add 8 mL Resuspension Buffer w/ RNase A to the pellet and resuspend cells until homogeneous.</td>
<td>Do not resuspend cells by vortexing. Pipette up and down instead.</td>
<td>This avoids the potential of shearing the MAC.</td>
</tr>
<tr>
<td>Add 8 mL Lysis Buffer and immediately mix by inversion until mixture is homogeneous.</td>
<td>Do not vortex. Incubate at room temperature for 5 mins.</td>
<td></td>
</tr>
<tr>
<td>Add 8 mL Precipitation Buffer and mix immediately by inversion until mixture is homogeneous.</td>
<td>Do not vortex.</td>
<td></td>
</tr>
<tr>
<td>Centrifuge lysate at 12,000 x g at room temperature for 5 mins.</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Load the supernatant onto the equilibrated column. Allow the solution to flow through by gravity.</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Wash the column with 2 x 10 mL Wash Buffer. Discard the flow-through.</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Place a sterile centrifuge tube under the column.</td>
<td>Make sure this tube is a 15 mL tube and not a 50 mL tube.</td>
<td>This allows the MAC to be dissolved in a smaller amount of water later on.</td>
</tr>
</tbody>
</table>
Table 3.2 (continued).

<table>
<thead>
<tr>
<th>Step</th>
<th>Description</th>
<th>Notes</th>
</tr>
</thead>
<tbody>
<tr>
<td>Add 5 mL Elution Buffer <strong>warmed to 50°C</strong> to elute DNA. Discard the column.</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Add 3.5 mL isopropanol to the elution tube. Mix well.</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Centrifuge at 12,000 x g for 30 mins at 4°C. Carefully remove and discard supernatant.</td>
<td>Incubate tube at room temperature for additional 30 – 60 mins before removing supernatant.</td>
<td>This helps increase the yields of precipitated MAC.</td>
</tr>
<tr>
<td>Resuspend the DNA pellet in 3 mL 70% ethanol.</td>
<td>Make sure the 70% ethanol is chilled.</td>
<td>This improves the removal of excess salts from the MAC solution.</td>
</tr>
<tr>
<td>Centrifuge at 12,000 x g for 5 mins at 4°C. Carefully remove and discard supernatant.</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Air-dry the DNA pellet for 10 mins.</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
| Resuspend the DNA pellet in 50 – 100 µL Tris-EDTA (TE) buffer. | Resuspend in warmed water instead. Pour into a 1.5 mL microcentrifuge tube, and incubate the solution at 50°C for 1 h. | - TE can pose problems for cloning applications, so eluting in water is a safer choice in that regard.  
- Pouring the MAC solution avoids potential shearing from pipetting.  
- Incubating the solution at 50°C helps the MAC solubilize; at lower temperatures, the MAC can form aggregates. |
| Centrifuge the elution tube at 12,000 x g for 1 min. | Pouring the DNA solution out of the tube can leave liquid along the sides of the tube. This increases overall yield by collecting the leftover liquid that contains DNA. | |
These changes to the stock protocol led to greatly improved plasmid yields, on the order of 100 µg at a concentration of 1 µg/µL, enough to work with for downstream applications like electroporation.

3.4 MAC Characterization with Sanger Sequencing and Pulsed-Field Gel Electrophoresis

One method employed for characterization of the MAC was Sanger sequencing. Samples were sent to Eurofins Genomics for sequencing, using the universal primers SV40pA-R, pBABE3, and pEXF. The resulting reads were aligned with the provided sequence data for the MAC, as shown in Figure 3.4. Approximately 1,700 base pairs were matched with the MAC’s sequence, indicating, in part, that the MAC did have the correct sequence.

![Figure 3.4. Sanger sequencing results aligned with the MAC. The sequencing reads are highlighted in yellow above the provided sequence data, and dark-red regions on the arrows represent regions of the sequencing reads that match the plasmid map’s sequence. Image generated by SnapGene software (from GSL Biotech; available at snapgene.com).](image)

Typical agarose gels (0.5 – 2.0%) run with a unidirectional electric field cannot properly resolve DNA fragments larger than 10 kb, so a CHEF DR-III (Bio-Rad) was used to visualize the MAC in its entirety. Pulsed-field gel electrophoresis (PFGE) can resolve larger DNA fragments by using electrodes positioned at angles other than 180° relative to the gel, as shown in Figure 3.5, and turned on at different intervals instead of running continuously. In this way, DNA fragments on the megabase scale can be
resolved on an agarose gel (Mawer and Leach, 2013). Digestions were performed with 
\textit{SacI}, a unique cutter yielding a linear fragment that would allow visualization of the MAC in its entirety. The results are shown in Figure 3.6. The midiprep is shown in comparison with a positive control sample gifted from Dr. Kouprina. As can be seen, the majority of samples were of the same size as the positive control sample, indicating that the plasmid prep procedures produced a MAC of the correct size.

Figure 3.5. A diagram of PFGE electrophoresis. Image from Mawer and Leach, 2013.
Figure 3.6. A PFGE image of purified MAC samples, digested with SacI. From left to right, Lane 1: DNA molecular size ladder. Lane 2: positive control (MAC provided by Dr. Kouprina). Lane 3: MAC.

3.5 Summary

The first major challenge to implementing the MAC as a synthetic biology tool was overcome through several adjustments at each step of the transformation and purification process. During the process of electroporation into *E. coli*, cells were shaken at low speeds and grown at 30°C; multiple colonies were picked post-electroporation, as opposed to single colonies; and entire glycerol stocks were used instead of simple stabs.
The purification process also required several changes, like picking the correct kit to use; extracting the MAC prior to stationary phase; eliminating vortexing steps; pipetting with cut tips; and warming the eluted DNA prior to quantification. These changes yielded purified MAC samples that were validated with Sanger sequencing and PFGE electrophoresis.
In order to actually deliver transgenes to mammalian cells using MACs, they need to first be integrated into the MACs themselves. Since conventional molecular cloning methods had not been extensively explored in the literature, it remained to be seen whether or not such a method could be used to integrate transgenes onto the MAC. This chapter covers the work that was done to insert a selectable marker (for nutritional selection via glutamine removal) into the MAC via Gibson Assembly and characterize the subsequent products.

4.1 **Gibson Assembly as a Cloning Method**

Molecular cloning of transgenes onto the MAC was carried out via Gibson Assembly. As seen in Figure 4.1, in Gibson Assembly, the DNA segments to be inserted into the target plasmid are generated with polymerase chain reaction (PCR), using primers that add overlapping sequences with the target plasmid to the DNA segment. The segments generated from PCR are then combined with the linearized target plasmid, either through PCR amplification or digestion with restriction enzymes, and three additional enzymes: an exonuclease, a polymerase, and a ligase. The exonuclease removes single bases from the 5’ end of the DNA segments, creating single-stranded overhangs that pair together through base-pair complementarity; the polymerase adds in nucleotides to fill in any gaps; the ligase joins the nicked ends of the DNA backbone of any segments to create a whole DNA segment (Gibson et al., 2009).
Gibson Assembly was used to replace the Blasticidin-S resistance cassette with a cassette for the glutamine synthetase (GS) gene. As seen in Figure 4.2, the MAC is linearized with *Sac*I and *Bam*HI, cutting out the Blasticidin-S resistance cassette, before Gibson Assembly is used to insert the GS expression cassette. As mentioned in Chapter 2, GS is used as a selectable marker in glutamine auxotrophs, like CHOZN GS<sup>-/-</sup>, and was a key transgene to investigate for delivery to mammalian cells. The GS expression cassette was derived from pcDNA3.1_Zeo(+) GS, a gift from Gyun Min Lee, and is controlled by an SV40 promoter and poly(A) termination signal.
PCR of the GS expression cassettes was carried out using the reaction setup and thermocycling protocol specified by NEB for their Q5 Polymerase. The forward primer was designed to anneal to the beginning of the SV40 promoter. Its 5’ overhang contains a SacI cut site and 20 bp homologous to the region upstream of the Blasticidin-S resistance cassette. The reverse primer was designed to anneal to the end of the SV40 poly(A) signal. Its 5’ overhang contains a BamHI cut site and 20 bp homologous to the region downstream of the Blasticidin-S resistance cassette.

For Gibson Assembly, 251 ng of the MAC was combined with 131.3 ng of the GS insert obtained from PCR, for a 1:8 vector:insert molar ratio, along with Gibson
Assembly Master Mix. The mixture was incubated and electroporated according to NEB’s associated protocols for Gibson Assembly.

4.2 Gibson Assembly of GS Expression Cassette into MAC

Figure 4.3 shows an agarose gel of the PCR products from pcDNA3.1_Zeo(+) - GS. Of note, two bands showed up in the gel. The Bleomycin resistance cassette (BleoR) is also controlled by an SV40 promoter and poly(A) termination signal, as shown in Figure 4.4. From the plasmid’s sequence data, the GS cassette was expected to be 1.6 kb, and the BleoR cassette was expected to be 1.1 kb, and the bands on the gel match these expected values.
Figure 4.3. Agarose gel of PCR products from pcDNA3.1_Zeo(+) GS. From left to right, Lane 1: 1 kb DNA molecular size ladder. Lane 2: No template control (NTC), a negative control. Lane 3: PCR products from pcDNA3.1_Zeo(+) GS, abbreviated GS. Both products, corresponding to GS and BleoR, are pointed out.
The band containing the GS expression cassette was cut out and recovered with a Monarch DNA Gel Extraction Kit (NEB). A subsequent agarose gel, Figure 4.5, showed a single clean band corresponding to the size of the GS expression cassette. The MAC double-digested with SacI and BamHI is also shown, to demonstrate that both pieces necessary for Gibson Assembly were present.
Post-Gibson Assembly, the products were electroporated into *E. coli* DH10β, plated, and incubated at 30°C. As can be seen in Figure 3.12, one colony, labeled “MAC-GS,” popped up on a plate, seen in Figure 4.6. A midiprep of this colony yielded ~100 µg at a concentration of ~1100 ng/µL.

In terms of increasing cloning efficiency, the best place to start would involve increasing the amount of vector DNA used. Since the MAC is much larger than typical plasmid vectors, mass units can be misleading, since the molecular weights of the MAC and typical plasmid vectors are so different. A given mass amount of the MAC represents a smaller molar amount than the same mass of a typical plasmid vector would. For example, for the modified Gibson Assembly protocol, 500 ng of the MAC is equivalent to 0.0135 pmol. In contrast, 500 ng of pUC19, a commonly-used cloning vector, is
equivalent to 0.302 pmol, a ~22-fold increase in molar amount. For a 1:8 vector:insert ratio, with that low a molar amount of the MAC, it ends up being the limiting reagent in the Gibson Assembly process. Using a higher vector mass could lead to higher colony counts from subsequent transformation into *E. coli*.

Figure 4.6. Agar plate of *E. coli* DH10β transformed with the MAC-GS Gibson Assembly product.

4.3 Verification of Gibson Assembly Success

To verify the presence of the GS cassette on the MAC, two different PCR verification steps were used. In one PCR, the primers used for Gibson assembly were used, and MAC-GS was compared to a no-template control (N.C.), a positive control (P.C.), pcDNA3.1 Zeo(+)-GS, and the original MAC, as seen in Figure 4.7. The positive control yielded 2 bands, as expected; the no-template control yielded nothing, as
expected, the MAC yielded a band of 1.8 kb, the size of the blasticidin cassette, and GS-MAC yielded a band that matched the larger band on the positive control lane.

As further verification of the success of Gibson Assembly, another PCR was performed on MAC-GS with a different set of verification primers. The forward primer had homology with the MAC, while the reverse primer had homology with the GS expression cassette, as seen in Figure 4.8. A band of 1.9 kb was expected from the PCR reaction; Figure 4.9 shows an agarose gel of the PCR results, confirming the expected size and the presence of the GS cassette on the MAC.
Figure 4.8. The location of the verification primers on MAC-GS. Image credit: Dr. Xiaoming Lu (Clemson University).

Figure 4.9. Agarose gel of PCR with verification primers. From left to right, Lane 1: Negative control (N.C.), PCR without DNA template. Lane 2: 1 kb DNA molecular size ladder. Lane 3: MAC-GS, with the band’s size labeled.
To ensure that MAC-GS was the correct size and had not lost repeats somewhere along the way, PFGE was also performed on MAC-GS, as seen in Figure 4.10. MAC-GS and the original MAC were double-digested with *BamHI* and *SacI*, as well as single-digested with *EcoRV*, which cuts uniquely. The digests of MAC-GS and the original MAC are very close in size, and, combined with the *EcoRV* digests, show that MAC-GS did not lose repeats.
Samples were also sent off to GeneWIZ for Sanger sequencing. The resulting sequence data was aligned to the GS expression cassette; about 900 base pairs were successfully matched, showing that the Gibson assembly product did indeed contain part of the GS expression cassette.
4.4 Summary

The challenge of cloning transgenes onto a MAC via conventional cloning methods was met. Via Gibson Assembly, an expression cassette for the GS gene was cloned onto the MAC, and the subsequent products were confirmed to be consistent with correct assembly via restriction digest and PFGE, PCR, and Sanger sequencing.
CHAPTER FIVE

DELIVERY OF TRANSGENES ON AN ARTIFICIAL CHROMOSOME TO MAMMALIAN CELLS

Once a transgene had been successfully integrated onto the MAC, the other important objective to achieve was actually delivering these constructs to mammalian cells. As work in the literature by previous groups did not extensively cover delivery of the MAC to suspension CHO cells, finding a suitable method was necessary. Electroporation was the standard method for delivering DNA constructs to suspension CHO cells; that served as the starting point for studies of MAC delivery. This chapter describes the work conducted to characterize a positive control construct for MAC-GS, NBS1-GFP; deliver NBS1-GFP and MAC-GS to CHOZN GS⁻/⁻ cells via electroporation; assess electroporation success and efficiency; and determine transgene functionality via L-glutamine selection.

5.1 Verification of Positive Control Constructs NBS1-GFP and VHL-GFP

Finding an appropriate positive control for delivery studies on a MAC is difficult, due to its size. Few readily available DNA constructs approximate the expected size of 60 kb for a MAC. Dr. Jung Hyun Kim was kind enough to donate a pair of constructs to serve as positive controls, NBS1-GFP (~76 kb in size) and VHL-GFP (~46 kb). These constructs had been used in a previously-published paper along with MACs to correct genetic deficiencies in human cells (Kim et al., 2011); a sequence map of VHL-GFP is shown in Figure 5.1. These constructs had the desired qualities for a useful positive
control (namely, a size approximate to that of a MAC and a reporter gene that produces an FP).

Figure 5.1. A plasmid map of VHL-GFP. Image from Kim et al., 2011, reproduced with permission.

Post-maxiprep, PFGE was performed to visualize restriction digestes of NBS1-GFP and VHL-GFP, as seen in Figure 5.2. Both constructs were digested with NotI-HF and run against undigested versions. The digested version of VHL-GFP produced 3 fragments, analogous to the 3 NotI cut sites displayed in Figure 5.1. The three fragments on the gel correspond roughly to the three fragments predicted by the plasmid map: 33.3 kb, 7.8 kb, and 5.5 kb. The latter two fragments are difficult to predict because the ladder loaded only goes as low as 15 kb, but the 33.3 kb fragment appears to match up with the
predicted value. While an NBS1-GFP plasmid map was not available from the cited article, it was constructed in the same manner as VHL-GFP. *NBS1* is about 30 kb larger than *VHL*, and the larger fragment seen in the lane for NBS1-GFP appears to be about 30 kb larger than the largest fragment seen in the lane for VHL-GFP. From this, it can be inferred that the purified NBS1-GFP and VHL-GFP constructs were, indeed, the right size.

![PFGE image of VHL-GFP and NBS1-GFP](image)

**Figure 5.2.** PFGE image of VHL-GFP and NBS1-GFP. Lane 1: VHL-GFP. Lane 2: NBS1-GFP. Lane 3: DNA molecular size ladder.
To demonstrate the functionality of NBS1-GFP and VHL-GFP, Dr. Jung Hyun Kim transfected both constructs into HT1080 fibrosarcoma cells with X-tremeGENE 9 (Roche), a cationic lipid reagent similar to Lipofectamine 2000/3000 (Invitrogen). HT1080 cells were seeded into a 6-well plate 24 h prior to transfection such that the cells would be 70% confluent at the time of transfection. Transfection was carried out according to manufacturer’s instructions. Cells were imaged 24 h post-transfection and can be seen in Figure 5.3.

![Fluorescence images of HT1080 fibrosarcoma cells transfected with X-tremeGENE9 (Roche). Fluorescent cells are pointed out with white arrows. (a) VHL-GFP. (b) NBS1-GFP. Image credit: Dr. Jung Hyun Kim (National Institutes of Health).](image)

### 5.2 Electroporation of NBS1-GFP into CHOZN GS−/−

Once the size and functionality of NBS1-GFP and VHL-GFP had been verified, they were used as positive controls for electroporation. Since both constructs had the same functionality and only really differed in size, NBS1-GFP was used, on the basis that, if success was found with a 76 kb construct, success could be found with a smaller 60 kb MAC containing transgenes.
The protocol used for electroporation was provided in the CHOZN® Platform Technical Bulletin. In short, for each electroporation, 6.25 x 10^6 cells were isolated from a culture seeded 24 h beforehand, and 5 x 10^6 cells were combined with 30 – 50 µg of DNA. The cell-DNA mix was loaded into a 4 mm gap electroporation cuvette and electroporated with a Gene Pulser Xcell machine (Bio-Rad) using an exponential decay pulse at 300 V and 950 µF. The cells were transferred to 5 mL of growth medium in a suspension T25 flask (Greiner Bio-One) and observed 24 h later. For electroporation of NBS1-GFP, higher masses were used; the quoted range in the electroporation protocol was presumed to apply to smaller, high-copy plasmids, and an equivalent mass of a larger, low-copy plasmid would contain a lesser molar amount.

~100 µg NBS1-GFP was electroporated into CHOZN GS^-/- cells, and the cells were observed under fluorescence 24 h post-electroporation, as seen in Figure 5.4. To estimate efficiency, 100 µL of culture was diluted into 1 mL total of growth medium in a well of a 24-well plate, and 5 fields of view were imaged; one such image is shown in Figure 5.4. Of 914 counted cells, 18 fluoresced, yielding an electroporation efficiency of 1.97%. This number may seem low, but it makes sense when taking into account the relationship between the size of a DNA construct and the efficiency of electroporating it into cells. The scale of electroporation is also important to consider: 1.97% of 5 x 10^6 cells is 98,500 cells, which is not an insignificant amount and not prohibitive for the purposes of selection.
To further verify the efficiency estimates, this experiment was independently conducted by Baylee Westbury, an undergraduate member of the project team. Cells were diluted and plated as before, except that the cells were diluted in a 1:20 ratio for ease of counting. One such image is shown in Figure 5.5. Of 290 cells counted, 6 cells fluoresced, yielding an efficiency of 2.07%.
As a more quantitative method of evaluating efficiency, flow cytometry was carried out on CHOZn GS\(^{-/-}\) cells electroporated with NBS1-GFP. 18 h post-electroporation, cells were analyzed with an NL-2000 flow cytometer (Cytek) with a 488 nm laser. The events corresponding to the general cell population were subsequently gated off. The bulk of the events observed had a mean fluorescence intensity (MFI) of less than 10\(^4\); these events were taken to represent autofluorescence of the cells that did not take up NBS1-GFP. Any event exhibiting a higher MFI than 10\(^4\) was gated off separately to represent the cells that had taken up NBS1-GFP. Figure 5.6 shows histograms of the fluorescence intensities of the general cell population (at left) and the population exhibiting high fluorescence intensities (at right). Of the 4,567 events observed in the general cell population, 101 showed a fluorescence intensity higher than 10\(^4\), indicating an electroporation efficiency of 2.21\%, a value similar to the percentages obtained from manual cell counting.
Figure 5.6. Flow cytometry data of CHOZN GS-/- cells electroporated with NBS1-GFP. The y-axis is counts of flow cytometry events, and the x-axis is the mean fluorescence intensity of B3, the most sensitive blue laser detector channel on the NL-2000, on a biexponential scale. (a) The general cell population, with the P1 gate showing the biggest cluster of the general cell population. (b) The fluorescence signal of the cells gated by P1. Fluorescence intensity from the strongest laser channel is displayed on the x-axis.

5.3 Electroporation of CHOZN GS⁺⁻ Cells with MAC-GS

Now that an efficiency had been established for electroporation of large constructs into CHOZN GS⁺⁻, electroporating MAC-GS into CHOZN GS⁺⁻ was the next step. 39 µg MAC-GS was electroporated into CHOZN GS⁺⁻, as detailed in Section 5.2, and, 24 h post-electroporation, cells were scaled up from 5 mL of growth medium in a T25 suspension flask to 10 mL of growth medium without L-glutamine in a T75 suspension flask (Greiner Bio-One). Cells were counted at least twice weekly, and the medium was changed once a week, according to the CHOZN® Platform Technical Bulletin, until viabilities passed 90%. As seen in Figure 5.7, the cells took approximately 19 days to recover from selection; viable cell density at the end of selection was 6.8 x 10⁶ cells/mL.
5.4 Summary

The challenge of delivering MACs loaded with a transgene to suspension CHO cells was met. Positive control constructs for the MAC, NBS1-GFP and VHL-GFP, were obtained, and their size and function were verified through PFGE and cationic lipid transfection into HT1080 fibrosarcoma cells. Electroporation of NBS1-GFP into CHOZN GS\(^{-}\) produced a low number of successfully electroporated cells but enough cells to proceed with selection. Efficiencies were estimated from epifluorescence microscopy and flow cytometry. MAC-GS was subsequently electroporated into CHOZN GS\(^{-}\), and selection by removal of L-glutamine yielded a culture of viable cells.
CHAPTER SIX
SUMMARY OF EXPERIMENTAL RESULTS AND FUTURE DIRECTIONS

6.1 Summary of Experimental Results

Chapter 3 addresses the first technical challenge of purifying the MAC in high yield and characterizing it. The MAC sample gifted by Dr. Kouprina was electroporated into *E. coli* DH10β, and positive colonies were recovered from the subsequent agar plating. A PureLink® HiPure Plasmid Midiprep Kit was used to extract the MAC from *E. coli* culture, producing yields of up to 100 µg. For verifying the characteristics of the MAC recovered via plasmid prep, Sanger sequencing confirmed the sequence data of some regions of interest on the MAC for subsequent cloning, and PFGE confirmed the MAC’s size as equivalent to that of a positive control sample MAC.

Chapter 4 addresses the next technical challenge of cloning transgenes onto the MAC. Gibson Assembly was used to replace the Blasticidin-S resistance expression cassette with a cassette for expression of glutamine synthetase, and a colony was recovered from subsequent electroporation into *E. coli* DH10β. To verify the characteristics of the assembly product MAC-GS, PCR was used to demonstrate the presence of the GS expression cassette; PFGE was used to confirm the size of MAC-GS; and Sanger sequencing was used to confirm the sequence of the GS expression cassette itself.

Chapter 5 addresses the last technical challenge of delivering the MAC with a transgene to mammalian cells and efforts to do so with electroporation. Electroporation
was carried out with different positive control constructs, NBS1-GFP and VHL-GFP; these constructs were initially run on a PFGE gel to confirm construct size and transfected into HT1080 fibrosarcoma cells for functional verification. Electroporation of NBS1-GFP into CHOZN GS-/- yielded low efficiencies but a not-insignificant number of transfected cells, a result validated by epifluorescence microscopy and flow cytometry. Electroporation of MAC-GS and subsequent nutritional selection yielded a viable culture of CHOZN GS-/- cells.

6.2 Future Directions

This project has shown some potential success in terms of cloning transgenes onto artificial chromosomes and delivering them to mammalian cells. However, much work remains to be done in this area, and many questions still remain unanswered:

1) Does a MAC remain as a separate chromosome or integrate into the genome? One way to do this is to carry out whole genome sequencing and align the results to the CHO reference genome. If a MAC has remained a separate construct, the sequencing reads corresponding to a MAC should align very poorly to the reference genome, as there is little homology between a MAC and the CHO genome. If a MAC has integrated, then there should be some reads containing partial homology to the reference genome.

2) Can a MAC remain as a separate chromosome over 60 generations, a typical duration for assessing clonal stability in CHO cells? One way to do this is by tracking the copy number of the GS gene, via quantitative PCR (qPCR), in the MAC-GS CHO strain in the presence or absence of selection pressure, with
CHOZN GS/- and CHO K1 (which has an intact GS gene) as benchmarks to compare GS gene expression to.

3) Can multiple transgenes (e.g. expression cassettes for IgG light and heavy chains) be inserted into a MAC? This would require further Gibson Assembly experiments. Measuring IgG titer would be the way to determine whether or not CHO cells with a MAC containing transgenes for IgG production can match the titer of a CHO cell strain with integrated transgenes for IgG production.

That being said, the foundation of viability for cloning a transgene onto a MAC and delivering it to mammalian cells (two tasks that were non-trivial and required much time and effort) has been laid down, paving the way for subsequent inquests into executing MACs as a usable technology in bioengineering.
Appendix A

Histone-Mediated Transfection as a Delivery Method

Histone-mediated transfection was of interest as a delivery method and a possible augmentation of electroporation. This appendix describes the work done to express and purify histones in *E. coli*; characterize the histones

A.1 Expression of Histones in *E. coli*

One particular downside to producing histones for use in HMT comes during the purification process. The fusion proteins produced in the study by Wagstaff et al. were purified under denaturing conditions, which adds additional steps and time for purification. Shim et al. (2012) improved the process of histone production by performing purification at high salt concentrations (2.0 M sodium chloride), using a polycistronic vector that would produce all four core histones at once, pET29a-YS14 (a gift from Jung-Hyun Min). At these conditions, histones dissociate from DNA and can remain in solution instead of dropping out into inclusion bodies. This simplified the process of histone purification, due to the non-denaturing conditions used. Non-denaturing conditions would greatly simplify the purification procedure, as opposed to having to denature proteins, isolate the resulting inclusion bodies, and then re-fold the proteins.

The vector pET29a-YS14 was purified from an *E. coli* DH5α stock with a PureYield Maxiprep Kit (Promega) and then transformed into *E. coli* BL21(DE3). A colony from the transformation was inoculated into 10 mL 2x YT medium and shaken at 170 rpm for 4 h at 37 °C. That culture was then seeded into 1 L of the same medium and
grown at the same conditions until the OD reached ~0.4. A sample of culture was taken at this point for further verification of expression. IPTG was added to the culture, to a final concentration of 0.4 mM, to induce expression. The culture was shaken for another 20 h.

Cells were pelleted at 4,500 x g for 10 mins at 4 °C and stored at -20 °C for future purification.

The cell pellet was re-suspended in 60 mL lysis buffer (20 mM Tris-Cl, pH 8.0; 2 M sodium chloride; 1 mM phenylmethanesulfonylefluoride (PMSF); 0.5 mM tris(2-carboxyethyl)phosphine (TCEP)). Lysis was carried out via sonication, using a Fisherbrand Model 120 Sonic Dismembrator (Fisher Scientific). The cell pellet was sonicated for 6 mins, at 70% power, on a 5 s off, 5 s on cycle. A sample of the cell lysate was saved for verification of expression.
Verification of expression was performed by running cell lysate samples on an 18% sodium dodecyl sulfate polyacrylamide gel electrophoresis (SDS-PAGE) gel, to separate the various proteins in the samples. Un-induced and induced *E. coli* samples were run on the gel to determine whether the histones had been expressed. The gel was subsequently stained with Coomassie Blue, which binds to proteins, and de-stained. Imaging of the gel, as well as all subsequent SDS-PAGE gels, was performed on a LI-COR Odyssey CLx. As seen in Figure A.1, bands that corresponded to the molecular weights of the histones and did not appear to show up in the un-induced culture were
pointed out on the gel. Size estimates were based on calculations of the expected weights from the vector’s sequence data. Histone H2A was expected to be ~17.9 kDa because of the associated S-tag, 6x His tag, and thrombin cleavage site. Histone H4 was expected to be ~13.2 kDa due to the additional associated thrombin cleavage site and 6x His tag. This gel showed at least preliminary evidence that the cells had properly expressed the histones.

A.2 Purification of Histones with Fast Protein Liquid Chromatography (FPLC)

The cell lysate was spun at 38,500 x g for 1 h at 4 °C in order to clarify the lysate. Such high speeds were necessary to remove the majority of the larger proteins, as the histones were smaller than 20 kDa. After centrifugation, the supernatant was collected and filtered through a 0.45 µm filter to remove particulates that might interfere with subsequent purification. Stock imidazole solution was added to a final concentration of 30 mM for subsequent chromatography.

The lysate was loaded onto a 5 mL HisTrap HP column (GE Healthcare) with an NGC Chromatography System (Bio-Rad) that had been equilibrated with Ni-buffer A (20 mM Tris-Cl, pH 8.0; 2 M sodium chloride; 0.5 mM TCEP). Two washes were performed: 10 CV of Ni-buffer A with 30 mM imidazole, and 10 CV of Ni-buffer A. Elution was performed using a step gradient of imidazole in Ni-buffer A: 60 mM, 110 mM, 140 mM, 170 mM, and 500 mM. The choices of 110 mM, 140 mM, and 170 mM were based on Shim et al., which had reported 110 mM and 170 mM as the ranges in which they saw roughly stoichiometric ratios of histones. The resulting chromatogram and a portion of it zoomed in on the elution steps are shown in Figures A.2 and A.3.
Figure A.2. FPLC chromatogram. Absorbance at 280 nm is shown as the orange trace. Relevant fractions that were run on SDS-PAGE are indicated.

Figure A.3. FPLC chromatogram, zoomed in to elution step gradient. Relevant fractions that were run on SDS-PAGE are indicated.

The fractions pointed out in the chromatograms and the lysate pre-purification were run on a 4 – 20% gradient SDS-PAGE to observe the size of the eluted proteins, as seen in Figure A.4. The input and flow-through were run in order to show that the column did actually pick up the proteins of interest. The wash fraction was shown in order to demonstrate that the undesired proteins were being washed out. Each of the fractions from the elution corresponded to a potential peak seen on the A280 trace (in orange).
The difference from the input appears to show bands in the 15 kDa range that are stronger in the input than the flow-through fraction. The bands from the various elution peaks show weak bands of other proteins at higher molecular weights, but the band from peak 5 shows the presence of two strong bands in the ranges of the expected values for histone sizes. It was slightly odd that the proteins eluted at such a high concentration of imidazole and not beforehand. However, this can potentially be explained by the fact that, while the stock Tris-Cl buffer was adjusted to pH 8.0 at room temperature, the FPLC was operated at 4 °C. The pH of Tris base is influenced by temperature; for instance, a 50 mM solution of Tris base that has a pH of 8.0 at 25 °C will, at 5 °C, have a pH of 8.58. This changes the ability of Tris base to act as a buffer and maintain the solution’s overall pH, which can subsequently impact the binding interactions involved in affinity chromatography.
A.3 Identification of Histones using Peptide Mass Fingerprinting

It was fairly likely that the band from peak 5 showed the presence of the histones in question, but, to confirm protein identity, a sample of the fraction was sent off to a collaborator, Alex von Kriegsheim, for identification via peptide mass fingerprinting. The unknown protein was cleaved into smaller peptide fragments using a protease. The peptides were then analyzed with mass spectrometry, and the resulting masses were run against a protein database to match them to expected peptide masses from known proteins. Known proteins that matched were returned as accession numbers. The accession numbers were then searched on UniProt to obtain the names of the proteins, and the four core histones were identified, as shown in Table A.1.

Table A.1. Accession numbers from mass spectrometry and their corresponding identities from UniProt.

<table>
<thead>
<tr>
<th>Accession Number</th>
<th>Identity</th>
</tr>
</thead>
<tbody>
<tr>
<td>A0A1L8G0S8</td>
<td><em>X. laevis</em> histone H4</td>
</tr>
<tr>
<td>A0A1L8G0S1</td>
<td><em>X. laevis</em> histone H3</td>
</tr>
<tr>
<td>Q92130</td>
<td><em>X. laevis</em> histone H2B</td>
</tr>
<tr>
<td>Q6INC9</td>
<td><em>X. laevis</em> histone H2A</td>
</tr>
</tbody>
</table>

A.4 Ultrafiltration of Histones and Quantification of Histone Concentration

In order to quantify the concentration of histones in the fraction, the Pierce Rapid Gold BCA Assay Kit (Thermo Fisher) was used. However, the kit is incompatible with high concentrations of imidazole and TCEP, so a portion of the histones was first buffer-exchanged to phosphate-buffered saline (PBS) via ultrafiltration, using an Amicon 10K molecular weight cutoff (MWCO) filter. 500 µL of the histones was spun at 14,000 x g
for 20 minutes, after which the volume of the retentate was measured before being brought back up to 500 µL with PBS. This process was repeated 3 times. In this way, the calculated concentrations of imidazole and TCEP were sufficiently diluted as to not impact the BCA assay’s results, as well as to avoid cytotoxicity in future cell culture experiments.

A potential concern was raised with regards to the size of the proteins. The histones themselves are around 15 kDa in size, so a 10K MWCO could potentially let proteins through. However, from Shim et al., the histones appeared to associate in solution to form nucleosomes, complexes of ~111 kDa, meaning that a 10K MWCO should be sufficient for buffer exchange. To test this, the filtrates and retentates after each of the spins were run on an SDS-PAGE to see if proteins were leaking through the filters. As can be seen in Figure A.5, it does not appear as though the 10K MWCO filter allowed any protein to leak through (or it did so in undetectable amounts).
A bicinchoninic acid (BCA) assay was subsequently run in order to quantify protein concentration. Protein standards of bovine serum albumin (BSA) in PBS were made, from 50 to 2000 µg/mL, in order to construct a standard curve, shown in Figure A.6, to judge the histone samples against. A linear regression of concentration vs. absorbance, in accordance with Beer’s Law, shows a linear trend with a correlation
coefficient of 0.9987. Using the standard curve, the histone sample was estimated to be 371.6 ng/µL.

Figure A.6. Standard curve from BCA assay.

A.5 Experimental Approach to Histone-Mediated Transfection

In order to test the efficacy of histone-mediated transfection (HMT) in cell culture, pilot studies were first conducted with HEK 293, a commonly-used cell line that grows quickly, is easy to work with, and had been successfully transfected with reporter genes in the past. The DNA payload of choice for initial tests was mTFP1-C1, a gift from Robert Campbell and Michael Davidson (Ai et al., 2008). mTFP1-C1 is a 4.7 kb vector that expresses the FP mTFP1. FP genes were the reporter genes of choice for gene delivery experiments because of the quick turnaround in observing results. Nutritional or antibiotic selection can take anywhere from 1 to 3 weeks; in contrast, FPs can be seen
inside of cells via fluorescence microscopy within 24 h, a much shorter time period necessary for evaluation of gene delivery success and efficiency.

Initial DNA:cell and histone:DNA ratios were derived from Wagstaff et al. Originally, a 15x15 mm coverslip seeded to 80 – 90% confluency was transfected with 10 µg DNA. Scaling down to the growth area of a 96-well plate (0.32 cm²) yielded 1.42 µg DNA. As for the histone:DNA ratio, a simplified calculation of 1 histone per 208 bp DNA was used. This calculation did not assume that the DNA would definitively wrap around the histone, but it served as a simplifying assumption. Using this conversion, 1.42 µg mTFP1-C1 required 168.5 ng histones.

Wells of a 96-well plate were seeded at 10,000 cells/well and transfected 48 h later. A 2-fold change in either direction for both mTFP1-C1 and the histones were included. The negative control was set as histones with no DNA payload. For the positive control, lipofectamine, as an established transfection reagent, was used to transfect mTFP1-C1. Histones were mixed with mTFP1-C1 and brought up to 50 µL total volume with growth medium before being incubated at room temperature for 15 mins. The histone-DNA mix was incubated with cells for 1 h at 37 °C before adding 50 µL DMEM. Cells were incubated for 48 h post-transfection and observed with a Revolve microscope (Echo), using a 488 nm laser.

A.6 Histone-Mediated Transfection in HEK 293

Results of initial HMT experiments can be seen in Figure A.7. As expected, the negative control showed no fluorescence. The positive control showed many cells fluorescing, as expected. Looking at the well with the specified DNA and histone masses
(1.42 µg and 168.5 ng, respectively), only a few cells in the field of view showed fluorescence. The transillumination view is shown, as well, to demonstrate the number of cells in the field of view. Modifying the DNA mass while keeping histone mass constant did little to increase the number of transfected cells; however, increasing the number of histones without changing DNA mass led to a marked increase in the number of cells transfected. This result showed promise in terms of HMT, especially the trend of increasing the histone:DNA ratio, with the caveat of the cells essentially forming a monolayer. This was due to the original protocol’s stipulation that transfection occur at 80 – 90% confluency and then waiting 48 h post-transfection to observe efficiency.
Figure A.7. Initial histone-mediated transfection experiments. (a) Fluorescence image of HEK 293 cells transfected with 1.42 µg mTFP1-C1 and 166.5 ng histones. (b) Fluorescence image of HEK 293 cells transfected with 1.42 µg mTFP1-C1 and 333 ng histones.

The next question to explore, then, was how increasing histone mass even further would affect the number of cells transfected. Histone mass was increased by up to 16-fold (for a final mass of 2.66 µg histone), for the same DNA ranges tested in the previous experiment. While there was not a clear trend to be seen in the results, the trials with
higher histone masses (1.33 µg and 2.66 µg shown below) did show a marked increase in the number of fluorescent cells compared to the trials with lower histone masses (666 ng shown in Figure A.8). A possible explanation for this is that circular DNA merely associates with histones through electrostatic interactions to facilitate HMT instead of actually wrapping around the histones.

Figure A.8. HMT experiments with increased histone:DNA ratios. DNA mass for this set of images was maintained at 1.42 µg. (a) Transfection with 666 ng histones. (b) Transfection with 1.33 µg histones. (c) Transfection with 2.66 µg histones.
A.7 Electroporation and Histone-Mediated Transfection in CHOZN GS⁻/⁻

A.7.1 Initial Combinations of Electroporation and HMT

At this point, having established the ability of HMT to transfect mammalian cells, attention was turned to transfecting the CHOZN GS⁻/⁻ suspension cell line. To reiterate, electroporation was the standard method of gene delivery for CHOZN GS⁻/⁻; Sigma already had an electroporation protocol in place for it. HMT was of interest in the hopes that it would have an additive effect on efficiency when combined with electroporation. Some combinations of electroporation and histone-mediated transfection were explored: HMT alone, HMT followed up by electroporation of cells (HMT/E), and electroporation followed immediately by HMT (E/HMT). The workflows for each of these avenues are shown in Figure A.9.

Figure A.9. Experimental steps for each of the combinations of HMT and electroporation. Image credit: Dr. Xiaoming Lu (Clemson University).
The plasmid construct ColorfulCell was utilized as a positive control. ColorfulCell, a gift from Pierre Neveu (Sladitschek and Neveu, 2015), is a 14.7 kb construct that contains 6 unique FPs (Citrine, 3 copies of mCerulean, 3 copies of TagBFP, hmAzami-Green, iRFP670, and mCherry) and was considered a useful reporter gene construct due to its larger size than that of mTFP1-C1.

6.5 µg ColorfulCell were used in these experiments, due to limitations on the amount of MAC that could be extracted from *E. coli* cultures at the time that have since been remedied. On the basis of the previously-mentioned simplifying assumption of 1 histone per 208 bp DNA, 6.1 µg histones were used for these experiments.

Electroporation was carried out with a modified version of Sigma’s standard protocol for electroporating CHOZN GS⁻/⁻ cells. Cells from a CHOZN GS⁻/⁻ culture were initially counted to determine viability; electroporation requires culture viability of above 90%. Once viability was determined, for each sample to be electroporated, 6.25 x 10⁶ cells were isolated from the culture, centrifuged at 220 rcf for 5 mins, and re-suspended in 1 mL CD CHO Fusion medium (SAFC), supplemented with 5 mM L-glutamine. The DNA – histone mix was added to 800 µL cells, and 600 µL of that mixture was added to a 4 mm gap cuvette (VWR). The cuvette was electroporated with an exponential decay pulse at a capacitance of 950 µF. The standard voltage specified by Sigma was 300 V, but it had been optimized to 270 V, based on a protocol from another article (Panousis et al., 2005), for lower amounts of mTFP1 and ColorfulCell, on the basis of efficiency and cell viability post-electroporation. Post-electroporation, cells were incubated at room temperature for 5 mins to allow cell membranes to re-seal and then transferred to a T25
suspension flask (Greiner Bio-One) with 5 mL CD CHO Fusion medium, supplemented with 5 mM L-glutamine.

Cells were imaged on the Echo Revolve 24 h post-transfection. Samples were diluted 1:10 in growth medium for ease of cell counting, and 5 fields of view were captured on the Echo Revolve for calculating efficiency. Efficiencies were calculated from the average of counts from all 5 fields of view.

Figures A.10(a) and A.10(b) show HMT/E and its positive control (ColorfulCell only, without histones); roughly the same efficiency was estimated for each (6.6% for the sample vs. 4.5% for the positive control). Figures A.10(c) and A.10(d) show E/HMT and its positive control (ColorfulCell only, without histones); almost no cells lit up for either of the sample and the positive control (<0.25% efficiency for each). HMT alone (Figure A.10(e)) yielded no observable fluorescent cells. From these results, it appeared that following electroporation up with HMT had no real effect whatsoever. HMT itself appeared to have no discernible effect on electroporation as a gene delivery method. HMT alone appeared to have zero effect on cells in suspension culture.
Figure A.10. Representative fluorescence images of combinations of electroporation and HMT. All images at 10x objective, except 20x objective for (c). (a) HMT/E. (b) Positive control for HMT/E (DNA only). (c) E/HMT. (d) Positive control for E/HMT. (e) HMT.
A.7.2 Increasing Histone:DNA Ratio

A couple of suggestions were brought up for modifying the HMT/E procedure. One of those suggestions was increasing the histone:DNA ratio, in the hopes that more carrier histones would increase the efficiency of transfection. The histone ratio was increased to 4:1, for 24.4 µg histones to 6.5 µg ColorfulCell. Cells were otherwise transfected and observed post-transfection in the same manner as described for HMT/E in the previous set of experiments; representative fluorescence images can be seen in Figure A.11. Contrary to expectations, the efficiency of gene delivery actually dropped. Very few cells were to be found among the HMT/E cells (Figure A.11(a)), as compared to the positive control (Figure A.11(b)). Calculated efficiencies were 0.9%, compared to the positive control’s 5.2%.

Figure A.11. Representative fluorescence images of HMT/E with a 4:1 histone:DNA ratio. All images at 10x objective. (a) HMT/E. (b) Positive control for HMT/E (DNA only).

A.7.3 Increasing Histone – DNA Complexation Time

Another possible fix posited was increasing the time allowed for the histones to complex with the DNA payload. The initial complexation time of 15 mins was derived
from Wagstaff et al., but longer complexation times were hypothesized to possibly give more time for the histones to bind to the DNA via electrostatic interactions. Incubation times of 4 h, 8 h, and 24 h were tested. A couple of other steps in the method were modified as well. When the histone-DNA complexes were added to the cells and incubated for 1 h, the incubation occurred under shaking conditions in a 6-well plate at 100 rpm. This was hoped to aid in helping the cells maintain oxygen uptake. Post-electroporation, the cells in a T25 suspension flask were also incubated on an orbital shaker for the same reason.

Figure A.12 shows representative fluorescent images of each of the different timepoints, compared to the positive control of DNA only. As can be seen, across Figures A.12(b), (c), and (d), there really appears to be no difference in the number of fluorescent cells between the different time points.
Figure A.12. Representative fluorescence images of HMT/E with increased histone – DNA complexation times. All images at 10x objective. (a) Positive control for HMT/E (DNA only). (b) HMT/E with 4 h complexation time. (c) HMT/E with 8 h complexation time. (d) HMT/E with 24 h complexation time.

These experiments appeared to indicate that, while HMT worked reasonably well for an adherent cell line like HEK 293, that success apparently did not translate to suspension cells. It also did not appear to have any sort of additive effect when combined with electroporation; thus, the latter was used as the gene delivery mechanism of choice going forward.
A.8 Summary

In order to obtain histones for histone-mediated transfection studies, the plasmid pET29a-YS14 was transformed into *E. coli* BL21(DE3) and expressed via IPTG induction. The culture was lysed via sonication, and expression of histones was confirmed by running lysate of induced cells with lysate of un-induced cells on an SDS-PAGE gel. Induced cell lysate was subsequently centrifuged and purified by FPLC. Relevant fractions from the chromatogram were run on an SDS-PAGE gel, and one fraction corresponding to an elution peak at 500 mM imidazole produced bands of expected size with few other detectable bands. The identities of the proteins in the fraction were confirmed by mass spectrometry to be histones H2A, H2B, H3, and H4. The fraction was buffer-exchanged to PBS by ultrafiltration, SDS-PAGE was used to confirm that proteins were not lost in ultrafiltration, and a BCA assay was used to estimate a protein concentration of 371.6 ng/µL.

Pilot HMT studies in HEK 293 cells were carried out using mTFP1-C1 and purified histones. Some cells were successfully transfected, with a trend of increasing number of cells transfected as histone mass was increased. Decreasing DNA mass did not seem to show a discernible trend in terms of increasing or decreasing number of cells transfected. In CHOZN GS^-/, combinations of HMT and/or electroporation were explored, leading to the general observations that histones did not have an effect on electroporation efficiencies, while HMT alone did not produce any transfected cells.
REFERENCES


