DNA Template Process Development for the mRNA Technology Platform

Cynthia Hitti

Department of Bioengineering

Graduate Program in Biological and Biomedical Engineering

McGill University, Montreal, Canada

July 2024

A thesis submitted to McGill University in partial fulfillment of the requirements of the degree of Master of Engineering, Biological and Biomedical Engineering

Table of Content

Abstract

English Version

Continued growth in the biotechnology industry is pushing the vaccine and therapeutic manufacturing landscape towards alleviating the burden of long and complex process steps, while reducing costs of goods and ensuring quicker responses to disease outbreaks. Recently, the SARS-CoV-2 (Severe Acute Respiratory Syndrome Coronavirus 2) pandemic has highlighted messenger RNA (mRNA) technology's promise due to its relatively fast process development, its scalability, flexibility, and safety profile. Facing increasing popularity and demand, current research focuses on improving the unit steps of mRNA manufacturing for maximal product yield and quality. However, the rapid development of this technology remains limited by its upstream process. The primary manufacturing step involves the production of the DNA template for subsequent transcription. The template yield, quality, and purity are critical as they have been shown to impact the attributes of the mRNA produced. DNA template generation currently remains the most time-consuming and costly step in the mRNA manufacturing chain as bacterial fermentation, which can take several days to weeks, is currently the industry standard method. To reduce the burden of DNA template preparation times, synthetic production strategies should be explored, while maintaining final mRNA product yields, quality, and purity.

This study aims to develop, optimize, and compare DNA template production and purification methods for use in *In Vitro* Transcription (IVT) to maximize mRNA yield. We first establish a bacterial fermentation method for DNA template preparation as a baseline production method. We then explore the use of Polymerase Chain Reaction (PCR) as an alternative synthetic approach to the conventional *Escherichia coli* (*E. coli*) fermentation-based DNA template production technique. This is first done with reporter construct DNA sequences, and then validated using longer DNA sequences. Here, we demonstrate the promising potential of PCR for DNA template generation for use in the mRNA technology to render the overall manufacturing time shorter. We systemically compare these two methods based on manufacturing time, associated resources/costs, yield, and quality of the end-product. For both DNA template production methods, we also explore and compare different lab-scale purification strategies while implementing a Tangential Flow Filtration (TFF) scale-down technique with the use of Amicon columns. In summary, this work contributes to the optimization of production and purification methods for DNA templates suited for mRNA production with the overall goal of improving the mRNA technology platform.

Version Française

La croissance continue de l'industrie biotechnologique pousse le secteur de la production de vaccins et de produits thérapeutiques à alléger le fardeau des étapes de processus longues et complexes, tout en réduisant les coûts des marchandises et en garantissant des réponses plus rapides aux épidémies. Récemment, la pandémie de SARS-CoV-2 a mis en évidence les promesses de la technologie de l'ARN messager (ARNm) en raison de son processus de développement relativement rapide, de son évolutivité, de sa flexibilité et de son profil de sécurité. Face à la popularité et à la demande croissante, la recherche actuelle se focalise sur l'amélioration des étapes individuelles de la fabrication de l'ARNm afin d'obtenir un rendement et une qualité maximaux du produit. Cependant, le développement rapide de cette technologie reste limité par son processus en amont. La première étape de la fabrication implique la production de la matrice d'ADN pour la transcription ultérieure. Le rendement, la qualité et la pureté de la matrice sont essentiels, car des études ont démontré leur impact sur les attributs de l'ARNm obtenu. La production de matrices d'ADN demeure actuellement l'étape la plus longue et la plus coûteuse de la chaîne de production de l'ARNm. En effet, la méthode standard utilisée en industrie repose sur la fermentation bactérienne, un processus pouvant s'étendre sur plusieurs jours, voire plusieurs semaines. Pour réduire le temps de génération de matrices d'ADN, des stratégies de production synthétique devraient être explorées, tout en maintenant les rendements, la qualité et la pureté du produit ARNm final.

Cette étude vise à développer, optimiser et comparer des méthodes de production et de purification de matrices d'ADN pour la transcription in vitro (IVT) afin de maximiser le rendement en ARNm. Nous commençons par établir la fermentation bactérienne pour la préparation de matrice d'ADN comme méthode de production de base. Nous explorons ensuite l'utilisation de l'amplification en chaîne par polymérase (PCR) comme approche synthétique alternative à la technique conventionnelle de production de matrices d'ADN basée sur la fermentation d'*E. coli*. Cette méthode est d'abord appliquée à des séquences d'ADN rapporteurs, puis validée à l'aide de séquences d'ADN plus longues. Nous démontrons ici le potentiel prometteur de la PCR pour la génération de matrices d'ADN à utiliser dans la technologie de l'ARNm afin de réduire le temps de production. Nous comparons systématiquement ces deux méthodes sur la base du temps de fabrication, des ressources/coûts associés, du rendement et de la qualité du produit final. Pour les deux méthodes de production de matrice d'ADN, nous explorons et comparons également

différentes stratégies de purification à l'échelle du laboratoire tout en mettant en œuvre une technique à échelle réduite de filtration à flux tangentiel (TFF) avec l'utilisation de colonnes Amicon. En résumé, ce travail contribue à l'optimisation des méthodes de production et de purification des matrices d'ADN adaptées à la production d'ARNm dans le but général d'améliorer la plateforme technologique de l'ARNm.

Acknowledgments

First and foremost, I would like to acknowledge and express my sincere gratitude to my supervisor, Dr. Amine Kamen, for giving me the opportunity to be part of the establishing team of the recently launched mRNA manufacturing project within the Viral Vectors and Vaccines Bioprocessing Group. I am deeply thankful for all the knowledge and skills I have acquired through this research project, and to have been a part of a lab environment fostering collaboration. The invaluable experiences of this Master's degree have not only been enriching but also transformative, enhancing my autonomy, critical thinking, and problem-solving abilities.

I would also like to convey my deepest appreciation to Dr. Ayyappasamy Sudalaiyadum Perumal for always being available to discuss research progress despite his busy schedule and offer valuable insights and suggestions whenever needed. An additional special thanks to the National Research Council (NRC) of Canada in Ottawa, particularly Dr. Hiva Azizi, for providing valuable insights early on in the project and for providing plasmid DNA sequences. I would also like to thank Dr. Caroline Wagner for generously accepting to review my thesis.

Last but not least, I would like to thank the rest of the mRNA team at the Kamen Lab; Julia Puppin Chaves Fulber and Maryam Youssef. Words fall short in conveying the immense pleasure it was working alongside these two brilliant researchers. Their collaborative spirit and insightful ideas have been instrumental for my research. A special note of thanks to Julia Fulber for consistently being available to brainstorm and exchange ideas, and for all the encouragement and belief in me during the most stressful times of the project, I truly appreciated it. And to Maryam Youssef, who is not only my lab partner, but also one of my closest friends. The shared experiences, unwavering support, the laughter, and all the much-needed distractions have made this journey all the more memorable. Thank you for always being there for me through both the good and challenging times, professionally and in my personal life. Talking openly about both our respective projects was invaluable in building the overall vision of the work and in overcoming research challenges. Your friendship and constant emotional support have been treasured.

Finally, I would like to acknowledge my funding for this project through FRQNT (Fonds de recherche du Québec, Nature et technologies).

Contribution of Authors

The present work was conducted in partial fulfillment of the requirements of the degree of Master of Engineering, Biological and Biomedical Engineering (BBME). This thesis was written and edited in its entirety by myself, Cynthia Hitti. Data shown in the Appendix Section was collected via collaborations: Appendix A was done in close collaboration with Master's student Maryam Youssef from the Kamen Lab, and the plasmid DNA sequencing results in Appendix 1B (i) and 2B (i) were generously provided and analyzed by Professor Ioannis Ragoussis' Lab (McGill Genome Center), and more particularly by Dr. Spyridon Oikonomopoulos and research assistant Mira Lotfi, in the context of a growing collaboration for quality control purposes of the DNA and mRNA.

Additional Contributions details:

- Maryam Youssef (MY): Co-author on Appendix A tables and Experimental Insights.
- Julia Puppin Chaves Fulber (JPCF) and Dr. Ayyappasamy Sudalaiyadum Perumal (ASP): Experimental Insights.
- Ioannis Ragoussis' Lab (McGill Genome Center): Generated plasmid DNA sequencing and corresponding poly(A) tail data analysis and results and developed an in-house pipeline to calculate the poly(A) length from the raw nanopore signal data.
- Prof. Amine Kamen (AK): Project Supervision and Funding Acquisition.

List of Abbreviations

A	Adenosine
AEX	Anion-Exchange Chromatography
APC	Antigen-presenting Cell
aPCR	Assembly Polymerase Chain Reaction
BCA	Bicinchoninic acid assay
C	Cytosine
CAI	Codon Adaptation Index
CBR2opt	Click beetle red codon optimized mutant luciferase
CDS	Coding sequence
CE	Capillary Electrophoresis
CFREF	Canada First Research Excellence Fund
D2R	DNA to RNA
DF	Diafiltration
DNA	Deoxyribonucleic Acid
DNA Pol	DNA Polymerase
dPCR	Digital PCR
dsRNA	double-stranded Ribonucleic Acid
E. coli	*Escherichia Coli*
EF	Endotoxin Free
eGFP	enhanced Green Fluorescent Protein
eIF4E	Eukaryotic Translation Initiation Factor
ELISA	Enzyme-Linked Immunosorbent Assay
G	Guanine
GMP	Good Manufacturing Practice
GOI	Gene of interest
HCP	Host Cell Protein
HF	High-Fidelity
HIC	Hydrophobic Interaction Chromatography
HPLC	High-Performance Liquid Chromatography
HTS	High throughput sequencing
IEC	Ion-Exchange Chromatography
IEX	Ion-Exchange Chromatography
IFN-I	Type I Interferon
IP-RP-HPLC	Ion Pair Reversed-Phase High-Performance Liquid Chromatography
IRES	Internal Ribosomal Entry Site
IVT	In Vitro Transcription
kb	kilobases
KDS	Potassium dodecyl sulfate
LAL	Lyophilized Amebocyte Lysate
LB	Lysogeny Broth
LC/MS	Liquid chromatography/ Mass spectrometry
LNP	Lipid Nanoparticle
m1ψ	N1-methylpseudouridine

m5C	5-methylcytidine
m5U	5-methyl uridine
m6A	N6-methyladenosine
m7G	methyl-7 guanine
MCB	Master Cell Bank
MOPS	3-(N-morpholino)propanesulfonic acid
mRNA	messenger Ribonucleic Acid
MWCO	Molecular weight cut-off
NGS	Next-Generation Sequencing
nLuc	nanoLuciferase
NTPs	Nucleotide Triphosphates
OD	Optical Density
ORF	Open Reading Frame
PABP	Poly(A)-Binding Protein
PCR	Polymerase Chain Reaction
pDNA	plasmid DNA
PDT	Plasmid DNA Template
PEG	Polyethylene Glycol
PEI	Polyethylenimine
PIC	Preinitiation Complex
PLS	Partial Least Squares regression
poly(A)	polyadenylation
qPCR	Quantitative Polymerase Chain Reaction
RNA	Ribonucleic Acid
RNAP	RNA Polymerase
RNase	Ribonuclease
RP-HPLC	Reversed-Phase High-Performance Liquid Chromatography
RQC	Ribosome-mediated quality control pathway
RT-PCR	Reverse transcription polymerase chain reaction
s2U	2-thiouridine
SARS-CoV-2	Severe Acute Respiratory Syndrome Coronavirus 2
SDS	Sodium dodecyl sulfate
SDT	Synthetic DNA Template
SEC	Size-Exclusion Chromatography
T	Thymine
TFF	Tangential Flow Filtration
TLRs	Toll-like receptors
tRNA	Transfer Ribonucleic Acid
U	Uridine
UC	Ultracentrifugation
UF	Ultrafiltration
uORF	upstream Open Reading Frame
UTRs	Untranslated Regions
ψ	pseudouridine

List of Figures and Tables

List of Figures:

List of Tables:

1. Introduction

1.1. Emergence of the mRNA Technology for Vaccine and Therapeutics Manufacturing

Messenger RNA (mRNA) technologies have the potential to revolutionize areas of medicine, such as the prophylaxis of infectious diseases, and gene and cell therapy applications [1]. Particularly, *in vitro* transcribed mRNA has come into focus over the last few years as a new drug class to deliver genetic information [2]. The clinical efficacy of mRNA vaccines was successfully demonstrated during the COVID-19 pandemic when two mRNA-based vaccines were made in record time, the Pfizer-BioNTech (BNT162b2) and the Moderna (mRNA-1273) COVID-19 vaccines [3-5]. They have both risen to be lead candidates as they could be developed at high speed compared to traditional vaccine platforms, providing quick responses to disease outbreaks while proving a high degree of efficacy and safety [6]. Since then, various mRNA vaccines and therapeutics candidates have entered the clinical stage of development. For example, currently undergoing Phase 3 clinical trial (NCT05330975) is mRNA-1345, a vaccine candidate targeting respiratory syncytial virus [7]. Additionally, mRNA treatments for rare monogenic disorders [8] and personalized cancer vaccines are currently being explored [9,10]. Compared to DNA based-vaccines and therapies, mRNA technology is considered more effective and safer since the mRNA, encoding an antigen or a therapeutic protein of interest, does not need to enter the nucleus to be functional. It is translated instantly once in the cytoplasm, either stimulating an immune response or supplying a mutated or underexpressed protein [2]. Unlike DNA therapeutics, random genome integration is impossible, eliminating risks of insertional mutagenesis typically associated with viral vector and plasmid DNA (pDNA) methods [11]. Given its fast, scalable, cell-free manufacturing process, and its global success endorsed by its role in the COVID-19 pandemic, the mRNA technology holds immense potential to revolutionize the field.

As the interest in mRNA technology is rising, it is becoming critical to address manufacturing challenges for the establishment of a well-defined and cost-effective process. This is imperative for future global commercialization purposes of this new drug class. The mRNA production platform typically includes the following steps: (1) DNA template production and purification, (2) *in vitro* transcription (IVT), (3) mRNA purification, and (4) lipid nanoparticle (LNP) encapsulation, formulation, and storage. mRNA stability is one of the main issues faced by this technology, as it is susceptible to degradation by nuclease enzymes, which imposes ultracold

storage requirements that under-resourced countries cannot meet, limiting global expansion of the mRNA platform [12-14]. Moreover, process optimization to increase mRNA yield is important to counter the high production costs and scarcity associated with good manufacturing practice (GMP) grade enzymes [15]. Further progress in synthetic approaches for the production of the DNA template or even long mRNA molecules could also alleviate the currently required downstream purification process. Moreover, the rapid rise of mRNA therapeutics has resulted in the creation of a gap in the regulatory framework, such that the current guidelines either do not apply, lack globally accepted definitions, or simply omit RNA therapeutics [16]. In fact, the approval of mRNA based COVID-vaccines in the emergency context of the pandemic was incredibly fast, leading to uncertainty in the regulatory guidelines and classification of RNA-based therapies, hindering commercialization. However, this rapid approval process highlighted a clear contrast with the typical mean duration for the development and approval of a novel therapeutic, which is approximately 10-15 years [17]. Overall, standardization of mRNA manufacturing could facilitate regulatory approval, enabling consistent product yield and quality. To this end, determining and optimizing strategies to increase both the DNA template yield and purity as well as the mRNA yield and purity is crucial for allowing increased access to this technology.

1.2. Research Impact, Motivation & Aims

Following the rise and the growing success of the mRNA technology platform during the pandemic, McGill University has been awarded the largest research grant in its history, Canada First Research Excellence Fund (CFREF) in the value of $165 million, to launch the DNA to RNA (D2R) initiative. This research initiative aims to restore biopharmaceutical production capacity in Canada, while ensuring equitable treatment accessibility. The research described in this thesis is based in this D2R initiative, and particularly, situates itself within the Biomanufacturing stream with the goal of establishing RNA manufacturing capacity within McGill.

This research project focuses on the upstream manufacturing, predominantly on the establishment, development, and optimization of the DNA template production and purification aspects to increase production yields and maximize mRNA yield and quality through IVT, while considering the possible reduction of the cost- and time-effectiveness of the overall process. As such, we set out to compare two methods for the production of template DNA: (1) bacterial amplification and

(2) enzymatic production of the DNA. We also evaluate the possible purification methods for each of these production techniques, and their impact on the subsequently produced mRNA.

Our motivation comes from the fact that innovations in mRNA-based vaccines and therapeutics' development hold the promise to address current manufacturing challenges and to offer a more cost and time effective production method. *Figure 1,* showing the overall mRNA manufacturing steps and associated process timelines, depicts that the DNA template generation is the most time-consuming step. In fact, bacterial fermentation, considered the gold standard in both lab and industry scales for the production of plasmid DNA template (PDT), can take several days to weeks by itself, compared to the IVT reaction which may only take a few hours. As such, Polymerase Chain Reaction (PCR) is explored in this study as an alternative synthetic approach for DNA template production to mitigate the burden of this pDNA generation time.

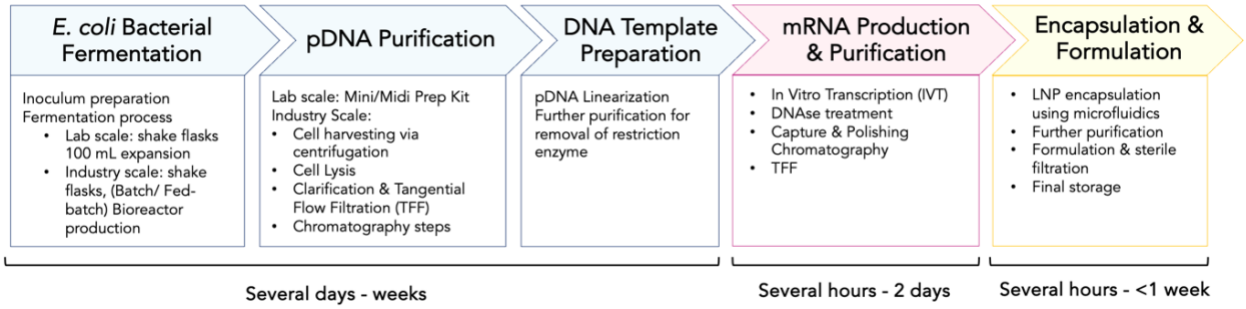


Figure 1: mRNA technology platform manufacturing process timeline [18-20].

These research goals are reinforced by our production process comparison of PCR and *E. coli*, as demonstrated in *Appendix A Table A,* which shows that, for example, PCR generated templates do not require linearization and the associated purification steps. This type of advantage of PCR generated templates would contribute to reducing the manufacturing process time. *Appendix A Table B* compares the fixed and variable costs associated with the *E. coli*-based and PCR generated DNA templates. The final evaluation of materials demonstrates that the price of 1 purified, and linearized microgram (µg) of DNA template amplified by PCR or *E. coli* fermentation is within 1-2 $ in each case. For both methodologies, the enzymes used contribute significantly to the overall cost of production (polymerase for PCR, restriction enzyme pDNA linearization for *E. coli*), as also shown in *Appendix A Figure A*. Therefore, this work has the potential to build upon the current state of the art and decrease overall production time and associated costs by optimizing both the production and purification processes of the template DNA for IVT.

2. Literature Review

2.1. mRNA Structure

To construct mRNA vaccines or therapeutics, the insertion of the encoded antigen or protein of interest in a DNA template that would later be used for mRNA transcription *in vitro* is required [15]. Upon delivery to the cytoplasm, the mRNA is recognized by cytosolic or endosomal receptors, leading to the activation of type I interferon (IFN-I) pathway and the production of pro-inflammatory cytokines and chemokines [15]. These lead to the activation of antigen-presenting cell (APC), stimulating a strong adaptative response [21].

mRNA is a single-stranded negatively charged molecule that can have a significantly large size (300–5,000 kDa, ~1–15 kilobases (kb)) and is used as a template during translation [22-24]. It includes the following elements, 5' Cap, 3' polyadenylation (poly(A)) tail, and an open reading frame (ORF) flanked by untranslated regions (UTRs) both at the 5' and 3' ends [25], as shown in *Figure 2*. Descriptions of each element are provided below. Each of these elements play a critical role in the function and stability of the mRNA such that modifications to them can not only improve the mRNA's translational efficiency and intracellular stability, but also improve its therapeutic functionality while reducing its immunogenicity [26].

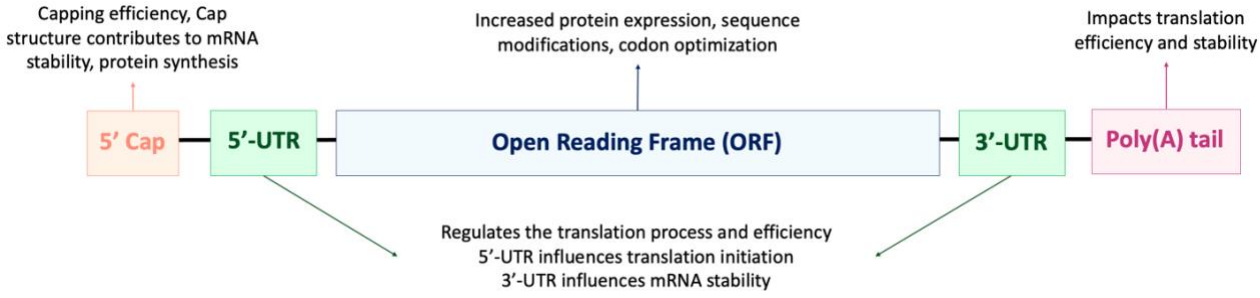


Figure 2: mRNA structural elements and their respective roles in the function and stability of the mRNA, including the 5' Cap, the 3' poly(A) tail, and the open reading frame (ORF) flanked by untranslated regions (UTRs) both at the 5' and 3' ends, adapted from [26].

First, the 5' Cap structure protects the mRNA from exonuclease degradation, avoids the innate immune response, and plays a key role in translation initiation by binding to a eukaryotic translation initiation factor (eIF4E) [27]. Different 5' Cap structures exist, mainly Cap 0 ($m^7GpppNp$) and Cap 1 (m^7GpppN_1mp). Cap 0, considered the simplest structure that effectively

allows gene expression and prevents mRNA degradation, features a methyl-7 guanine (m7G) nucleotide linked to the 5′ position of the RNA through a 5′ triphosphate bridge [27]. On the other hand, Cap 1 is generated when the first base of the mRNA is methylated at the 2’ position and is considered the most optimal structure for mRNA stability and expression, and it is recognized as self by the immune system, unlike Cap 0 which can activate an innate immune response [27]. Even though alternatives for translation machinery recruitment exists to the 5’ cap, such as cap-independent viral sequences and internal ribosomal entry site (IRES) sequences, high protein levels post-mRNA transfection are only achieved if 5’ capping of the mRNA is performed [28,29]. There exists two methods for capping: (1) post-translational capping using enzymes, such as the vaccinia virus capping enzyme, and (2) co-translational capping. Co-translational capping involves the incorporation of a Cap analog into the IVT reaction, such as the Cap dinucleotide ‘CleanCap®’ (from Trilink), thus achieving mRNA synthesis and capping in a single step [2].

The poly(A) tail regulates the stability and translational efficiency of the mRNA molecule, therefore functioning as a master regulator of gene expression in the cytoplasm [30,31]. It protects the mRNA from nuclease degradation by binding to the poly(A)-binding protein (PABP) [32], and PABPs interactions are affected by the length of the poly(A) tail, which is critical to maintain mRNA stability. Various studies have been conducted to establish the optimal poly(A) tail length, and have determined that 120 nt is considered optimal for mRNA vaccines, since no increase in expression can be observed beyond that length [33]. The poly(A) tail can be added post-transcriptionally by a polyadenylation step; however, standard practice in the biomanufacturing industry is to add the poly(A) tail sequence directly into the DNA template construct to reduce production steps [26]. Further poly(A) tail length optimizations can be performed using PCR-based methods, RNase (Ribonuclease) H/Oligo(dT) northern blot analysis, oligo(dT) selection, and even next-generation sequencing (NGS) protocols [34].

The 5’- and 3’-UTRs further contribute to the regulation of post-transcriptional gene expression. They can contain specific sequences that can bind to regulatory proteins, triggering the inhibition or initiation of translation. Moreover, translation efficiency is affected by the length and secondary structures of UTRs such that the migration of the 40S ribosomal unit can be inhibited by the presence of secondary structures with high structural stability, resulting in low translation

efficiency [32]. Particularly, both 5'- and 3'-UTRs regulate the translation process: the 5'-UTR influences translation initiation, while the 3'-UTR plays a significant role in mRNA stability [35].

Finally, the ORF provides the genetic information necessary to express the target protein or antigen of interest. As applied in this research, mRNAs encoding reporter proteins, such as fluorescence proteins or luciferase, are commonly used for effective demonstration of translation efficiency of the synthesized mRNAs. Even though the ORF is not one of the major elements affecting mRNA expression and stability, using modified nucleotide triphosphates (NTPs) such as N1-methylpseudouridine (m1ψ), has been found to reduce the immunogenicity of synthetic mRNA and to drive high levels of protein production, which is in part attributed to its ability to blunt TLR3 activation [36]. Other modified nucleotides include pseudouridine (ψ), 5-methylcytidine (m5C), 5-methyl uridine (m5U), or 2-thiouridine (s2U), and N6-methyladenosine (m6A), and also lead to the inactivation of toll-like receptors (TLRs) to inhibit the immunostimulatory effect of RNA and enhance translation [37-39]. For example, the Moderna COVID-19 spike-encoding mRNA vaccine (spikevax®, or mRNA-1273) completely replaces uridine with m1ψ during IVT [40].

2.2. Overall mRNA Technology Platform Manufacturing Process

The overall manufacturing process used in mRNA technology is depicted in *Figure 3*. This process can be easily translated for various applications, such as mRNA vaccines outside or inside the scope of infectious disease outbreaks and pandemic preparedness, and mRNA therapeutics and personalized medicine applications. In fact, the manufacturing process is sequence-independent, such that it is only primarily dictated by the nucleotides used, the length of the RNA, the capping chemistry, as well as the purification of the end-product, and its encapsulation/formulation. Flexibility of the mRNA technology platform is necessary to achieve various product scales depending on the end-application. As such, manufacturing process standardization is key for the flexibility of this platform, making it suitable for its various applications. For the purpose of this thesis, only the upstream manufacturing process, including the DNA production and purification as well as the IVT mRNA synthesis, will be discussed.

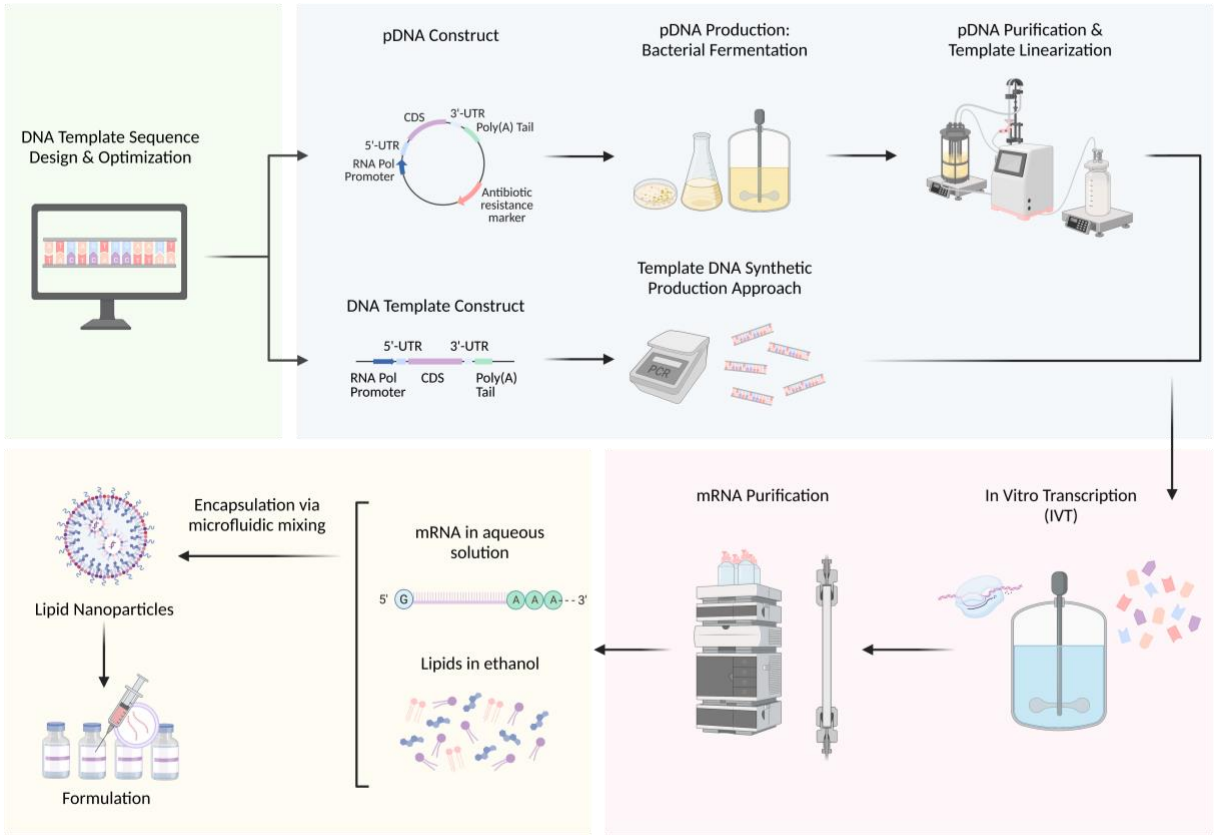


Figure 3: Manufacturing process for mRNA vaccines and therapeutics, from [41]:
(1) DNA template sequence design and optimization, (2) DNA template production and purification using either bacterial fermentation or synthetic approaches, (3) mRNA synthesis via IVT, (4) mRNA purification, and (5) Encapsulation for delivery, formulation, and storage.

2.2.1. Upstream Process: DNA Template Sequence Design & Optimization

The sequence design of the DNA template constitutes the first step in mRNA synthesis chain. The format of the DNA template depends on the DNA production method chosen, which will be thoroughly discussed in the section below, such that transcription templates later used in IVT can be pDNA, PCR product or synthetic double-stranded oligonucleotides [42,43]. The DNA template typically includes the following elements, which can be selected or modified accordingly to improve the translation and stability of the mRNA [44]: (1) Coding sequence (CDS), (2) 5'-UTR and 3'-UTR, (3) poly(A) tail, (4) RNA polymerase (RNAP) promoter sequence. Most of these elements have been discussed in Section 7.1 as they constitute the main structural elements required in the transcribed mRNA. Additionally, an antibiotic resistance marker sequence for

bacterial selection and restriction sites for DNA template linearization are required in the case where bacterial fermentation is chosen as the method of production for the template DNA.

The gene of interest (GOI) encoding for the therapeutic protein or antigen of interest is required for mRNA therapeutic production, as mentioned above. To reduce protein immunogenicity and increase protein expression, sequence optimization such as nucleotide and codon optimization can be done [45]. For example, prioritizing high guanine and cytosine (G and C) content compared to adenosine and thymine (A and T) increases ribosome association, mRNA stability, and therefore translation efficiency [46-48]. Optimization of the GC content in the GOI also alters RNA secondary structures which can interfere with gene expression, as demonstrated by Mauger et al. from Moderna, Inc. [23]. Moreover, codon degeneracy is used for sequence optimization as there exist multiple different choices in designing the coding sequence to produce the same desired protein [49]. There exists a positive correlation between the frequency of codons in a DNA sequence and the corresponding transfer ribonucleic acid (tRNA) in a species, and the concentration of tRNA determines the number of amino acids available for protein synthesis, affecting therefore the translation efficiency [50]. Thus, replacement of rare codons with more abundant and frequently used alternatives based on tRNA preferences has been shown to lead to a more controllable translation of the sequence to the desired protein [45,46,49]. On the other hand, non-optimal/rare codons can decrease mRNA half-life while increasing ribosomal pausing, which can trigger the ribosome-mediated quality control pathway (RQC) that mediates nascent polypeptide degradation, ribosome rescue, and mRNA decay [23,51,52]. Finally, all these optimizations can be assisted via bioinformatic tools. For example, key features such as the GC-content, optimal codon usage, codon adaptation index (CAI), and RNA structure prediction can be evaluated for mRNA engineering [46,53]. Deep learning techniques are also being explored for codon optimization to enhance protein expression, as demonstrated by Fu et al. [50].

Additionally, the DNA template for IVT should include the 5'-UTR and 3'-UTR sequences that will be later found in the corresponding mRNA sequence. As mentioned in Section 7.1, these elements do not directly encode proteins but play crucial roles in translation initiation and protein expression levels and affects the stability and half-life of the mRNA respectively [35,54]. In both research contexts and clinical trials, the most widely used UTRs are the β-globin UTRs [55]. The canonical cap-dependent scanning pathway of the eukaryotic mRNA translation starts with the

assembly of the 43S preinitiation complex (PIC) onto the mRNAs, where scanning is followed by the migration of the PIC along the 5'-UTR [56]. In order to minimize this scanning process, a shorter 5'-UTR with at least 20 nucleotides is recommended to maximize protein expression [46]. Moreover, studies have shown that secondary structure in the 5'-UTR tends to reduce translation efficiency and thus the protein output [57,58]. Thus, avoiding highly stable secondary structures near the 5'-end is important as they can disrupt ribosome loading and scanning. Additionally, to promote main CDS translation, potential upstream start codons such as AUG should be avoided to eliminate the presence of upstream open reading frame (uORF) that can trigger mRNA decay. Avoiding leaky scanning can be achieved by selecting a strong start codon with consensus sequence. For example, a study by Trepotec et al. showed that using a minimalistic 5'-UTR consisting of 14 nt combining the T7 promoter and the Kozak element consensus sequence (GCCACC) upstream of the start codon increased protein expression [59]. Finally, additional sequence elements can be introduced to the 5'-UTR depending on the therapeutic purpose to allow for selective translation. For example, in the context of cancer therapy, special 5'-UTR elements capable of translation under nutrient restriction may be needed for intratumor mRNAs injection [46]. On the other hand, 3'-UTRs are commonly derived from β-globin mRNAs and numerous 3'-UTRs, such as the hepatitis B virus (HBVpA) and the bovine growth hormone (BGHpA), have been studied for therapeutic mRNA applications [60]. 3'-UTRs contain many *cis*-regulatory elements that modulate mRNA localization and translation, the adenylate uridylate (AU-rich) element (ARE) being the most common one [46]. It has been shown by Holtkamp et al., that higher protein levels and prolonged persistence of the protein can be achieved by the introduction of two sequential β-globin 3'UTRs cloned head to tail between the CDS and the poly(A) tail [61]. Finally, 3'-UTR sequence optimization high throughput techniques have been discussed in literature such as a novel cell-based library screening process to identify 3'-UTRs that would lead to high protein expressions by Orlandini von Niessen et al. [62], and a massive parallel functional annotation assay for sequence optimization by Zhao et al. [63].

Other key elements for the DNA template include the poly(A) tail and the RNA pol promoter sequence. The poly(A) tail, as discussed in Section 7.1, regulates the stability and translation efficiency of the mRNA molecule, and can be added directly into the DNA template sequence to remove the post-transcriptional polyadenylation step. On the other hand, the RNA pol promoter sequence is key for transcription initiation during IVT as transcription begins when the RNA

polymerase binds to the promoter sequence, forming the initiation complex [64]. There exist various RNA polymerases from bacteriophages, such as T7, T3, or SP6 RNA polymerases, with T7 RNA polymerase (T7 RNAP) being considered the standard in both academic research and industry [15,65,66]. As such, the following T7 promoter sequence (5′-TAATACGACTCACTATA-3′) is the most commonly used one in the DNA template for future recognition of the T7 RNAP during IVT [59]. Moreover, there has been a lot of effort in engineering and optimizing T7 promoter sequences to increase mRNA abundance and further protein yield [67,68]. One example of an improved T7 promoter variant would be the "T7Max" (5'-AATTCTAATACGACTCACTATA GGGA-3'), demonstrated by Deich et al. [68]. Rapid amplification of cDNA 5′ ends coupled with deep sequencing (5′ RACE-Seq) was used by Conrad et al. for T7 promoter optimization and demonstrated that sequence motifs between the +4 and +8 nucleotide have a strong impact on transcriptional output [67]. Additionally, the high popularity of T7 RNAP in industry comes from its robustness and high fidelity at producing full-length RNA transcripts; however, it can also produce immunostimulatory byproducts such as smaller or loopback double-stranded RNA (dsRNA), which in turn affects protein expression [66,69]. In fact, dsRNA molecules are innate immune response activators capable of triggering Toll-like receptor 3 in the endosome and retinoic acid gene-I-like receptors such as MDA5 and RIG-I in the cytoplasm [69-71]. As these byproducts render the downstream purification process more complex, research has been conducted for the development of mutant T7 RNAP that would produce substantially less dsRNA during IVT compared with wild-type T7 RNAP, while maintaining RNA yield and purity. One such example includes the double-mutant T7 RNAP (G47A + 884G) identified by Dousis et al. [66].

Finally, if the DNA template is in the form of pDNA, the plasmid sequence should include an antibiotic resistance marker for bacterial selection, and a restriction site for DNA linearization, as mentioned above. The two most commonly used bacterial selection agents are ampicillin and kanamycin, allowing for the selection of the plasmids bearing their corresponding resistance genes, beta-lactamase and neomycin phosphotransferase II, respectively [72,73]. Despite the most available markers in the market, research in the field reported a novel selection marker *mfabI* (mutant *fabI*) for plasmid propagation in *E. coli* which further facilitates molecular manipulation of unstable sequences [74]. Moreover, to produce the desired mRNA transcripts of the correct length, pDNA should be linearized downstream of the insert by restriction digestion.

2.2.2. Upstream Process: DNA Template Production

After designing the DNA template encoding the antigen or protein of interest, the template needs to be generated, linearized if necessary, and purified to be used in the IVT. Currently, the biomanufacturing field is dominated by bacterial fermentation for the production of pDNA. Large scale production of pDNA, considered the base of various therapeutic applications, is one of the main challenges faced by the biopharmaceutical industry [75]. Since the advent of recombinant DNA technology in the mid 1970s, many genes have been successfully placed in vectors and expressed in foreign host cells [76]. Hundreds of biopharmaceutical companies are now using DNA in clinical development with various products licensed for commercial distribution [18]. Moreover, the rapid growth of the cell and gene therapy industry, particularly that of mRNA technology, has recently increased the demand for DNA in terms of both quality and scale [18]. As such, alternative methods to bacterial fermentation have also currently been proposed for DNA template manufacturing, such as synthetic and enzymatic approaches, which would alleviate these challenges during the upstream manufacturing process.

2.2.2.1. Industry Standard: Bacterial Fermentation Process

In both academia and in the biotechnology industry, plasmid generation is usually conducted via bacterial fermentation in the Gram-negative bacterium *E. coli* [75]. The bacterial fermentation process starts with the *E. coli* cell bank, from which the host competent cells are transformed with the previously designed pDNA to be amplified. For direct clinical applications in industry, GMP-grade plasmid DNA is mandatory and should be produced starting from a fully characterized GMP-produced Master Cell Bank (MCB) [77]. Following bacterial transformation and plating on LB (Lysogeny Broth) agar plates containing the corresponding antibiotic marker (such as ampicillin or kanamycin), single unique antibiotic resistant colonies are selected. Bacterial culture expansion typically follows the four main steps: creation of a master cell bank from well-characterized colony, inoculation, shake flasks expansion, and large-scale fermentation that can go up to 300-L scale in industry. Multiple strategies have been implemented to satisfy the increasing demand for high pDNA production yield and quality, such as the selection of high-producing strains and medium composition optimization in combination with different culture strategies, such as batch or fed-batch modes [78].

The choice of host strain has been shown to affect the quality of the purified pDNA [79]. Various *E. coli* strains have been routinely used for pDNA production at lab scale, such as DH5 [80], DH5α [81], DH10B [82], DH1 [83], SCS1-L [84], and JM108 [75]. At the industrial scale, the main *E. coli* strains used all belong to the *E. coli* K-12 class (DH5, DH5α, DH1, and JM108), and are characterized by mutations, ΔendA and ΔrecA, leading to efficient pDNA production with higher yield and quality [75]. These mutations, although strain and/or plasmid dependent, are known to minimize recombination, eliminate non-specific degradation by endonucleases, and improve plasmid stability [75]. On the other hand, the *E. coli* B strain BL21, has been suggested as a good alternative to K strains for plasmid production, due to better growth and metabolic characteristics [85]. However, the ΔendA and ΔrecA mutations need to be engineered to adapt BL21 for stable pDNA production, and proper growth using glucose as a carbon source and production strategy were developed. This was demonstrated by Phue et al. who achieved a 90% increase in volumetric yield compared to DH5α [85]. Moreover, the engineered strain VH33 Δ (recA deoR nupG) by Borja et al., with an elevated glucose concentrations cultivation, has shown higher yields compared to commonly used DH5α host; however, it produced lower percentage of supercoiled plasmid [86]. Gonçalves et al. demonstrated that *E. coli* GALG20, a *pgi*-gene knock-out, produced higher pDNA yields in bench-scale bioreactors when compared to DH5α [87]. Finally, Gram-negative bacterium *Vibrio Natriegens* has the potential to emerge as a novel and alternative host platform to *E. coli* for protein production in high yields [88,89]. With a doubling time of less than 10 min, this fastest growing free-living bacterium holds therefore great promise to also revolutionize plasmid DNA production, following further research [88,89].

Strain and vector engineering have been shown to improve plasmid stability, increase yield, enhance product safety, and even facilitate downstream processing [90]. For example, a high-copy number origin of replication is preferred as they can achieve higher pDNA yields, such as pUC-based plasmids, derived from ColE1, lacking the RNA one modulator (Rom) protein, and containing a point mutation in the RNA II sequence [91,92]. Moreover, Bower et al. reported that runaway R1-based plasmids have demonstrated higher copy numbers at higher temperatures [92]. On the other hand, various studies have been conducted regarding the effect of the bacterial fermentation cultivation strategy, as well as the growth medium used, on plasmid production. For example, Lopes et al. showed the advantage of using fed-batch culture mode rather than batch mode to achieve higher plasmid yields [78]. In their study, analyses were conducted via a novel

real-time monitoring system for pDNA production process using NIR spectroscopy combined with Partial Least Squares regression (PLS) modeling [78]. Moreover, coupling parameter control in fed-batch mode with high-copy pUC-based plasmids lead to an increase in plasmid yield with respect to cell mass – up to 1500 mg/L were achieved compared to 100-250 mg of pDNA/L for typical fermentation process [93]. Last year, Gotsmy et al. successfully designed and validated a novel three stage, growth-decoupled fed-batch process at the L-scale, using minimal medium based on genome-scale metabolic modelling [94]. In terms of bacterial growth medium, the most commonly used on is LB medium, with yeast extract as nitrogen source. According to Xu et al., glucose is considered the preferred carbon source for plasmid production [95]. They do however note that glycerol could also be used as a complementary carbon source to increase pDNA yields, due to its high specific pDNA productivity [95]. Finally, culture temperatures for *E. coli* are typically between 37 and 40^{o}C [96].

2.2.2.2. Alternative Synthetic DNA Approaches

As mentioned, the demand for DNA production has surged in recent years due to advancements in cell and gene therapies, but also in the context of the COVID-19 pandemic. However, bacterial fermentation for DNA template generation, although being the most commonly used method in academic research and industry, has proven to be challenging as it is both costly and time-consuming [18]. The fermentation process involves cloning and bacterial culture preparation steps, uses expensive reagents such as bacteria and antibiotics, and can take between several days to up to several weeks, without the guarantee that the correct bacterial clone will be obtained and isolated [42]. Additionally, this process is associated with concerns in terms of the GMP regulations for mRNA-based vaccines and therapeutics, due to the use of genetically modified antibiotic-resistant bacteria strains and the possibility of product biocontamination [97]. Therefore, alternative synthetic DNA approaches are currently being explored for DNA template generation to reduce the upstream process time and facilitate downstream purification [2]. Investments in research which develops synthetic approaches for manufacturing long DNA strands will support these efforts to reduce the time limitations imposed by bacterial fermentation approaches. Moreover, another benefit of the use of synthetic DNA template over pDNA for mRNA production is that the template does not need to be enzymatically linearized before IVT, thus removing this costly and time-consuming step in the overall upstream process.

PCR is an *in vitro* enzymatic synthesis method of specific DNA sequences which involves the use of primers to flank and select the region of interest to be amplified through multiple amplification cycles [98]. Several studies have successfully transcribed mRNA from a synthetic DNA template. In 2022, Mey et al. reported the use of a cost- and time-efficient synthetic DNA template (SDT) assembly process for use in IVT, demonstrating that the resulting SDT-mRNA is of quality comparable with that of the PDT-mRNA, in terms of yield, purity, and integrity [42]. In their study, they perform assembly PCR (aPCR) using synthetic pre-designed oligonucleotides as starting material, and then further amplified by PCR, effectively obtaining up to several micrograms of template DNA [42]. Another study, by Hu et al. in 2019, demonstrated the use of overlap extension PCR for synthetic DNA assembly, later transcribed into sgRNA *in vitro*, as a highly applicable and less time-consuming process [99]. Finally, in 2023, Wei et al. developed a universal integrated platform with control systems for on-demand mRNA preparation, which successfully integrates the use of a PCR module for DNA template amplification [100].

In the context of pandemic preparedness, rapid nucleic acid scale-up manufacture proposals and research are becoming more and more prevalent in both industry and academia. For example, Touchlight Ltd., a UK biotechnology company, has developed a novel, proprietary synthetic DNA vector, called "doggybone" or dbDNA™, and dual enzymatic manufacturing process for rapid DNA generation [18]. This technology allows for multi-gram DNA production in weeks, uniquely positioning itself for rapid and large-scale DNA production [18]. Another UK synthetic biology company, Evonetix Ltd., developed a DNA synthesis process in a benchtop device, called the Binary Assembly® process [101]. It uses a novel silicon chip to control the synthesis of double-stranded gene-length DNA and error-containing sequences removed, at many thousands of independent thermally controlled reaction sites [101].

2.2.3. Upstream Process: DNA Template Purification

The produced DNA template needs to be purified and linearized accordingly before it can be used in the IVT reaction, ensuring proper quality of the produced mRNA. The purification process varies depending on the method of DNA production employed due to the different process related impurities which will be generated depending on the manufacturing technique employed. Following bacterial fermentation, complex purification steps are required to obtain a high-purity plasmid free of host DNA, RNA, proteins, and endotoxins.

After cell harvesting by centrifugation, pDNA purification from bacterial cells generally starts with a lysis step. Cell harvesting can also be done via filtration; for example, tangential flow filtration (micro-filtration/diafiltration) is typically used to separate bacterial cells from the supernatant [102]. The methods used for cell disruption can typically be classified into two main categories: chemical (detergents, enzymes, etc.) and physio mechanical (shear, heat, freeze-thawing, etc.) [103]. The most commonly used lysis method is alkaline lysis, first reported by Birnboim and Doly in 1979 [104], which uses sodium hydroxide (NaOH at pH~12) accompanied by a detergent solution such as sodium dodecyl sulfate (SDS) and Triton® X-100 to disrupt the bacterial cell membrane [43,102,104]. This step releases intracellular components such as the plasmid DNA, RNA, and chromosomal DNA and proteins from bacteria cells resulting in a highly viscous solution [105]. During lysis, the alkaline environment denatures genomic DNA, such that some of the supercoiled DNA is converted to alternative forms including denatured supercoiled, multimeric, open circle, and linear [43]. Optimization of the lysis incubation time is necessary in order not to create irremovable genomic DNA fragments or damage the plasmid, while keeping a pH controlled in the 12-12.5 window to ensure plasmid quality [102,106]. Separation of cellular components from the highly viscous precipitate usually requires pre-filtration or centrifugation followed by clearing filtration, which can be costly and time-consuming [107]. Additionally, it is important to gently mix the cell lysate and neutralizing agent in low shear forces to avoid bacterial chromosomal DNA contamination, as DNA is sensitive to shearing forces [107,108]. To achieve complete and gentle mixing of large lysis volumes, batch mixing in a mechanically agitated vessel and/or in-line static mixers have been used [109,110]. Other cell lysis methods for pDNA purification have also been employed such as boiling cell lysis, first introduced by Holmes et al. which utilizes a lysozyme digestion to break down the cell wall followed by a heating step [111], and mechanical lysis, which is more common for protein purification due to the fact that high shear degrades nucleic acids [112]. The initial step in eliminating host cell contaminants during pDNA manufacturing involves precipitation/flocculation. The obtained lysate is then neutralized with a neutralization buffer before clarification, such as sodium or acidic potassium acetate with or without surfactant, RNAse, or $CaCl_2$ [105,108]. RNAse enzymes used to degrade the RNA prior to pDNA isolation can be added to the neutralization buffer; however, RNAse-free purification methods should be prioritized, as the ultimate process goal is to produce mRNA. Neutralization prompts the precipitation of detergent-solubilized proteins including high molecular weight

genomic DNA [102]. Isopropanol, polyethylene glycol (PEG), polyethylenimine (PEI), compaction agents and chaotropic salts have been used in DNA precipitation [105,113,114]. Post-clarification using depth filtration to remove precipitates, tangential flow filtration (TFF) can be done for concentration and for the removal of contaminating proteins and RNA [115]. Performing a concentration prior to chromatography steps allows for reduction in column loading time, while removing impurities such as small proteins, small-sized genomic DNA, and RNA [115]. Chromatography methods are based on different principles for obtaining high-purity plasmid: size-exclusion chromatography (SEC), ion-exchange chromatography (IEC), hydrophobic interaction chromatography (HIC), affinity chromatography, and reversed-phase high-performance liquid chromatography (RP-HPLC) [116-119]. As no single type of chromatography can effectively remove all residual contaminants, an optimal multi-step process is needed. For example, Guerrero-German et al. conducted pDNA purification using hollow-fiber tangential filtration, frontal IEC and HIC [120]. Another example would be a two-step chromatography purification: anion exchange (AEX) to concentrate pDNA and remove impurities such as host cell proteins (HCP) and genomic DNA, followed by HIC chromatography to capture the supercoiled pDNA, the most stable isoform, which was presented by Sartorius [121]. Moreover, multimodal chromatography has successfully been used for supercoiled pDNA purification with a two-ligand process (Capto™ adhere and PPA HyperCell™), leading to higher purity compared to traditional chromatography methods [122]. Finally, the obtained pDNA can be concentrated by ultra-centrifugation (UC) and sterile filtered through a 0.22-μm filter membrane [105].

Before the DNA template can be used in the IVT reaction, the purified plasmid needs to be linearized and further purified for the removal of the restriction enzymes. pDNA linearization should preferably be done with a restriction enzyme that leaves blunt or 5'overhangs at the 3'end of the DNA template rather than 3'overhangs, as they are favorable for proper run-off transcription by T7 or SP6 RNA polymerase and for reduction of unwanted products [123]. Examples of such restriction enzymes include SpeI, HindIII, NotI, SapI, and EcoRI [124]. The gold standard technique for removal of restriction enzymes from the final DNA template is phenol chloroform extraction, which is no longer preferred for large-scale clinical operations [125]. As such, other purification methods have been used such as chromatography and/or UF/DF (ultrafiltration/ diafiltration) [126]. For example, the use of sepharose-based strong AEX chromatography for the

preparation of linear DNA template for IVT has been proven successful, giving comparable results with phenol chloroform extraction [125].

Finally, for synthetically produced DNA templates, purification should only be carried out once as they are already in linear form, which reduces the DNA downstream processing time. DNA templates produced using PCR should be purified from short primers, dNTPs, enzymes, short-failed PCR products, and salts, typically using clean-up columns [127].

2.2.4. Process Analytics: Plasmid DNA Industry Quality Attributes

The identity, integrity, and purity of the pDNA, which later serves as the template in the IVT reaction, are crucial for ensuring the quality of the resulting mRNA product. *Table 1* shows the pDNA quality attributes in the context of template DNA production for use in IVT, as well as the associated analytical tools and acceptability criteria. This reveals the importance of analytical tools for quality control purposes.

	Attribute Details	Analytical Tools	Acceptability Criteria
Content	Yield/Concentration	UV spectrophotometry, Picogreen™ assay	
	Size	Gel electrophoresis	
Purity	Purity	UV spectrophotometry $A_{260/280}$ (Nanodrop)	Ratio 1.8 – 2.0
	Appearance	Visual inspection, USP <1>	Colorless, clear, no particle
	Plasmid conformation/form	Capillary electrophoresis (CE) or High-Performance Liquid Chromatography (HPLC)	> 90% supercoiled
	Homogeneity	AEX-HPLC	
Identity	Sequence	DNA sequencing Sanger/Next-generation sequencing (NGS), PCR	Conforms to ref. sequence
	Poly(A) tail length and integrity	DNA sequencing, Fragment analysis, analytical LC-MS, RP-HPLC, UP-HPLC	Correct sequence, single peak
Process Residuals	Residual host cell genomic DNA	Quantitative PCR (qPCR)	< 50 µg/mL
	Residual host cell protein (HCP)	Enzyme-linked immunosorbent assay (ELISA)	< 10 µg/mL
	Residual protein	SDS-PAGE or Bicinchoninic acid assay (BCA)	
	Residual host RNA	HPLC, Ribogreen™ assay, or agarose gel electrophoresis	< 2%
	Antibiotic residual (if applicable)	ELISA	
Safety	Endotoxins	Lyophilized Amebocyte Lysate (LAL) test, USP <85>	< 10 EU/mg pDNA
	Bioburden	Membrane filtration, USP <61>	< 1 CFU/mg, sterile
	Sterility	USP <71>	

Table 1: Analytical methods and process specifications for main pDNA product attributes and process parameters, adapted from: [128-132]

2.2.5. In Vitro Transcription: mRNA Synthesis

The purified linear DNA template encoding the antigen or therapeutic protein of interest obtained is then used to synthesize mRNA through IVT. This relatively straightforward and fast enzymatic reaction involves these key components: the linear DNA template, RNA Polymerase, NTPs, Magnesium ($MgCl_2$), and a reaction buffer [15]. Even though T7 RNAP is the most widely used RNAP in both academia and industry, one major downside of using it for mRNA synthesis is the resulting nonspecific synthesis of dsRNA byproducts through erroneous 3'extension, and research has been conducted to reduce such artifacts, as discussed in Section 7.2.1 [133]. Other elements can be added to the IVT reaction mix such as urea and spermidine. The addition of chaotropic agents, such as urea at a concentration of 1 M, can be used to decrease dsRNA formation, while the addition of spermidine at a concentration of 1 to 3 mM has been carefully chosen as to improve the overall efficiency of the RNA polymerase, prevent RNA synthesis inhibition, while not reaching its inhibitory effect at high concentrations [134-136]. Additionally, other modifications can be introduced in the IVT, such as: the use of modified NTPs to increase levels of protein production and reduce immunogenicity of synthetic mRNA, and the addition of CleanCap® reagent to allow for the 5' capping to be done co-transcriptionally instead of post-transcriptionally, both discussed in Section 7.1 [36-38,137]. Finally, Mg^{2+} is required as a cofactor for T7 RNAP, such that the ideal range of between 50 and 60 mM of free Mg^{2+} has been shown to be optimal [138,139].

The IVT reaction at large-scale is typically done in batch or fed-batch modes with fed-batch IVT having been shown to achieve higher mRNA yields [140]. Typical mRNA yields follow a baseline of 5 g/L, while higher yields of 12 g/L have been reached with Bayesian optimization [134,141]. Following IVT, the reaction mixture should be purified in order to obtain the desired mRNA product free of process related impurities, rendering the mRNA suitable for later patient administration post-formulation and encapsulation. Such impurities include the residual DNA templates, enzymes, NTPs, as well as immunogenic dsRNA and aborted mRNA products [142]. Conventional methods for purifying mRNA in a laboratory setting include lithium chloride (LiCl) precipitation and phenol-chloroform extraction and ethanol precipitation [25,32,143]. At larger scale, various downstream processing steps exist for the purification of clinical-grade mRNA, including multiple chromatography methods such as oligo(dT) affinity chromatography, ion-

exchange chromatography (IEX), ion-pair reverse-phase chromatography, as well as the use of TFF before finishing with sterile filtration [144-146].

2.2.6. Process Analytics: mRNA Product Main Quality Attributes

The identity, content, purity, and safety of the mRNA obtained through the IVT reaction are important measures of the mRNA quality and are affected by the quality of the template DNA used. *Table 2* shows the mRNA product attributes and their associated analytical tools.

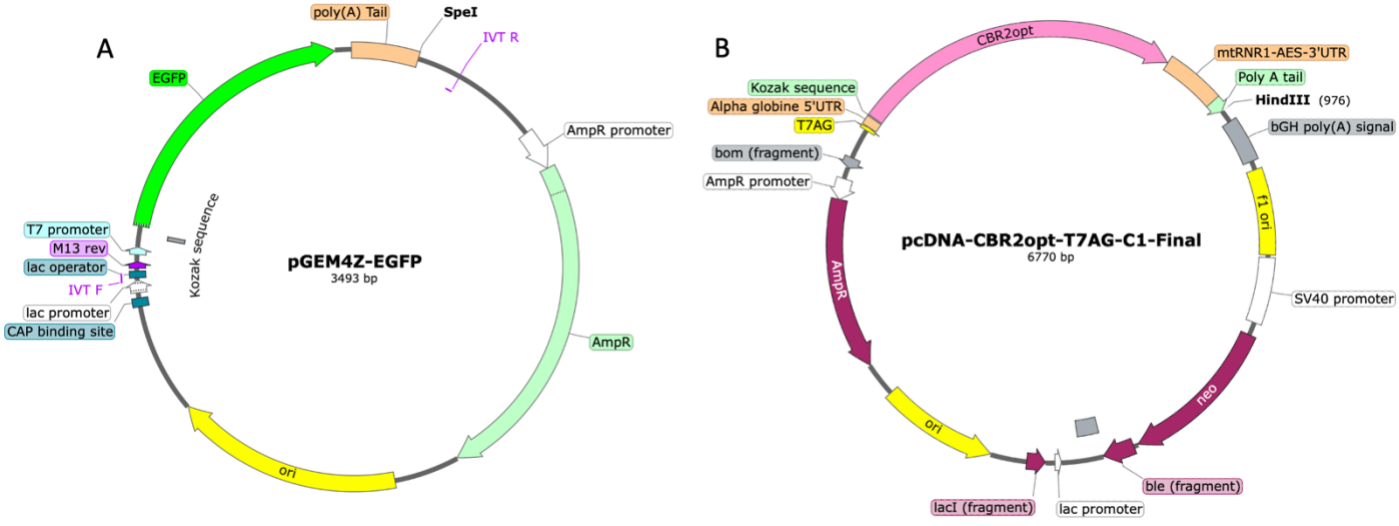
	Attributes/Process Parameters	Analytical Tool	Acceptability Criteria
Identity	Sequence confirmation	High throughput sequencing (HTS), Sanger Sequencing, Reverse transcription polymerase chain reaction (RT-PCR)	Conforms to reference sequence
	Integrity, Length	RNA electrophoresis, Bioanalyzer	Correct length, single peak
Content	mRNA concentration	Ribogreen™ Assay, Digital PCR (dPCR), UV spectrophotometry, qPCR	>1.5 g/L, 95–70%
Purity	5’ capping efficiency	Ion pair reversed-phase (IP-RP-) HPLC, Liquid chromatography- mass spectrometry (LC/MS), ribozyme assay with CE	>50–85%
	3’ poly(A) tail length	RNA electrophoresis, analytical LC-MS, LC-UV/MS, IP-RP-HPLC	100–120 bp
	RNA purity, shorter RNA	RNA electrophoresis, analytical RP-HPLC, RP-UPLC, IEX-HPLC, western blot oligonucleotide mapping	>50%
	Product related impurities - dsRNA	Immunoblot, dot blot, ELISA, analytical LC	<1 ng/µg RNA
	Process related impurities - residual DNA template	qPCR, fluorescence-based assays	<330 ng/mg
	Process related impurities - residual enzymes, HCP	NanoOrange, ELISA	<300-500 ng/mg RNA
	Product related impurities - aggregate quantitation	SEC-HPLC	
Safety	Endotoxin	USP <85>	
	Bioburden	USP <61>, <62>, <1115>	

Table 2: Analytical methods and process specifications for main mRNA-based product attributes and process parameters, adapted from: [129,147]

3. Methodology

3.1. Plasmid DNA Constructs

pGEM4Z-EGFP (peGFP) was purchased on Addgene (Plasmid #183475) (USA) and pcDNA-LUC-CBR2opt-T7AG-C1 (pnLuc), 230907-BA45-Genscript (pBA45), and 230907-Reference-Genscript (pRefspike) were generously provided by the National Research Council of Canada (NRC, Canada). These plasmids encode for enhanced green fluorescent protein (eGFP), nanoLuciferase (nLuc), and nucleic acid components of Pfizer's bivalent vaccine (BA4/5 SARS-CoV-2 Spike and Alpha reference sequences), respectively. The plasmid sequences are shown in *Figure 4*. peGFP is 3493 bp long with the AmpR sequence encoding for β-lactamase conferring resistance to ampicillin, carbenicillin, and related antibiotics. Other elements of peGFP include a 720 bp eGFP coding sequence, a 173 nt poly(A) tail, the T7 promoter (TAATACGACTCAC TATAGG), and an SpeI restriction site. pnLuc is 6770 bp long with the same AmpR sequence as peGFP, conferring similar resistance to ampicillin. It also contains a 1632 bp sequence encoding for nLuc and more specifically CBR2opt (click beetle red codon optimized mutant luciferase), a 68 nt segmented poly(A) tail, the CleanCap® AG compatible T7 promoter (TAATACGACTCAC TATAAG), and a HindIII unique restriction site. pBA45 and pRefspike are 6913 bp and 6928 bp respectively, and both contain the AmpR sequence. Their respective SARS-CoV-2 spike protein coding sequences lengths are 3807 bp and 3822 bp, each with a 110 nt poly(A) tail.



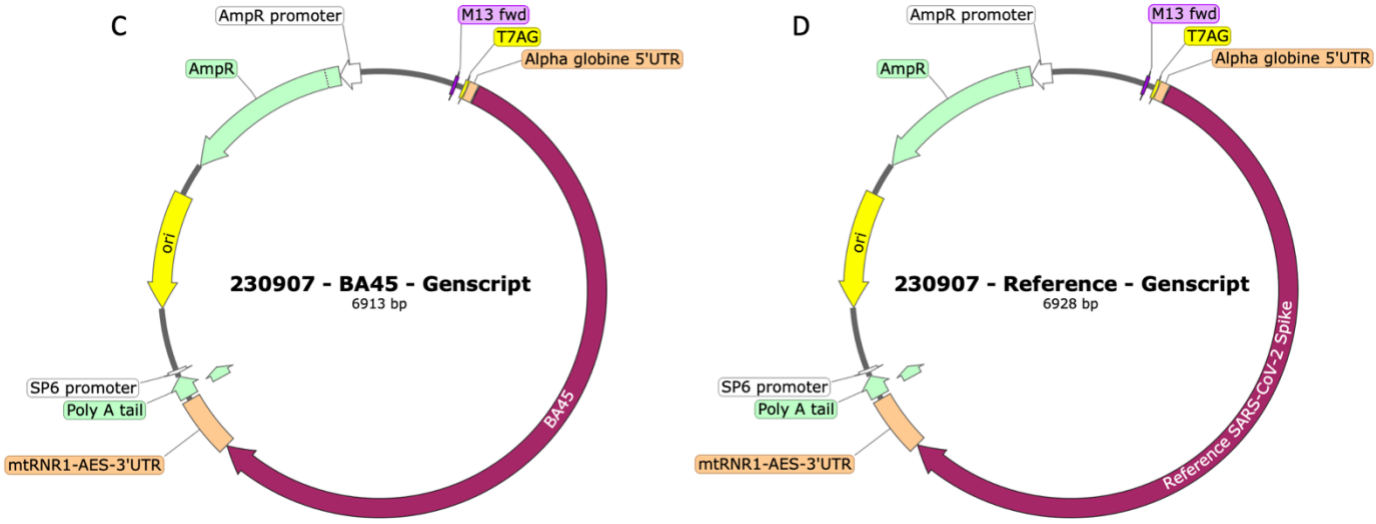


Figure 4: Plasmid maps showing the relative positions of the major elements of the plasmid sequences, such as the origin of replication, the antibiotic resistance gene, the coding sequence, etc., in addition to the plasmid name and length, pictures taken from SnapGene software: (A) pGEM4Z-EGFP (peGFP) was purchased on Addgene (Plasmid #183475), (B) pcDNA-LUC-CBR2opt-T7AG-C1 (pnLuc) provided by the National Research Council of Canada (NRC), (C) 230907-BA45-Genscript (pBA45) provided by the NRC, and (D) 230907-Reference-Genscript (pRefspike) provided by the NRC.

eGFP and nLuc coding sequences have been chosen for this study as they can later serve as rapid reporters of gene expression. As such, eGFP and nLuc plasmids were used throughout the entirety of the project, while the two SARS-CoV-2 Spike protein encoding plasmids were only used for validation purposes of the PCR production process for longer transcripts.

3.2. Template DNA Production

3.2.1. Method 1: Plasmid DNA Production using Bacterial Fermentation

3.2.1.1. Bacterial Fermentation

pDNA was generated using bacterial fermentation in *E. coli*. First, three different bacterial strains were compared in their ability to produce pDNA at small-scale: (1) MAX Efficiency® DH10Bac™ Competent Cells (Life Technologies), (2) NEB® Stable Competent *E. coli* , and (3) *Vibrio Natriegens* (ATCC14048), a gram-negative marine bacterium. To perform such transformations, LB agar plates were prepared with ampicillin (100 µg/mL) for appropriate selection of the plasmids for both *E. coli* strains. For the *Vibrio Natriegens* strain, LBv2 medium was used instead of LB for plate preparation and bacteria inoculation, with the following v2 salt

composition: 204 mM NaCl, 4.2 mM KCl, 23.14 mM $MgCl_2$. Transformations were performed following the protocol provided by the manufacturer, with the heat-shock step done for: (1) 45 seconds at 42°C for DH10Ba competent cells, (2) 30 seconds at 42°C for NEB® Stable Competent, and (3) 45 seconds at 42°C for *Vibrio Natriegens*. Following bacterial transformation and plating on prepared LB or LBv2 agar plates containing the antibiotic marker respectively, *E. coli* strains were incubated overnight at 37ºC and *Vibrio Natriegens* at 30ºC. This was then followed by colony selection from agar plates using a sterile inoculation loop, with at least five different colonies picked for each plate. Inoculation was done in 5 mL LB medium containing 5 µL ampicillin (1000x). The bacterial cultures were then incubated for approximately 8 to 10 hours in a shaking incubator (225 rpm) at 37°C. Visual observation was carried out to ensure that bacterial growth occurred. Growth was verified by the clarity of the media, with turbid media representing potential bacterial growth. OD_{600} (optical density at 600 nm) measurements were taken off-line for culture growth monitoring. Bacterial growth rate was computed from the OD_{600} readings using the following formula:

$$\mu = \frac{\ln(OD_2) - \ln(OD_1)}{t_2 - t_1}$$

Where: µ = bacterial growth rate (per hour); $OD_t = OD_{600}$ reading at time t; $t_2 - t_1$ is the time interval between readings. The doubling time was then computed using the following formula:

$$t_d = \frac{\ln(2)}{\mu}$$

Larger-scale pDNA production was then done using the selected *E. coli* strain, DH10Bac. OD_{600} measurements were taken to adequately expand the cultures from the 5 mL inoculums. After approximately 5 hours in the shaking incubator, each 5 mL inoculum was expanded into a 100 mL flask, which were left to grow in the shaking incubator at 37°C and 225 rpm for approximately 20 hours.

3.2.1.2. Plasmid DNA Isolation and Purification

pDNA was isolated and purified from bacterial cells using the QIAprep® Spin Miniprep kit for culture volumes of 5 mL and NucleoBond Xtra Midi EF (Midi prep kit for endotoxin-free plasmid DNA) for culture volumes of 100 mL. Purification of pDNA started with a lysis step after cell

harvesting by centrifugation at 4,500-6,000 x g for ≥ 10 min at 4°C and resuspension in buffer RES-EF and RNase A. The bacterial cells were lysed though the common alkaline lysis method by a sodium hydroxide/sodium dodecyl sulfate (NaOH/SDS) treatment, such that the release of pDNA, RNA, and chromosomal DNA and proteins from bacteria cells results in a highly viscous solution. After lysis, the samples were cleared from cell debris and precipitated to ensure high plasmid purity, which, in the case of large culture volumes, was achieved using the NucleoBond® Xtra Column Filter and Column system. Before clarification and loading onto the equilibrated column, the lysate was neutralized by a neutralization buffer containing potassium acetate causing SDS to precipitate as potassium dodecyl sulfate (KDS), pulling down chromosomal DNA, proteins, and other cell debris. The lysate was then loaded onto the column and two washing steps were performed for endotoxins removal before pDNA is eluted in endotoxin-free buffer (H_2O-EF). Quantification of the pDNA was performed via UV absorption measurements at 260 nm as described below in Section 8.5. Plasmid length and quality were checked using agarose gel electrophoresis.

3.2.2. Method 2: DNA template production using Polymerase Chain Reaction

3.2.2.1. Oligonucleotide Primer Pairs Design

Specific oligonucleotide primer pairs to each PCR templates were generated using Primer-BLAST, a National Center for Biotechnology Information (NCBI) online primer design platform.

PCR Template	Primer Sequences	Template strand	Length	Tm	GC %	Self-Comple-mentarity	Self 3' Comple-mentarity
eGFP plasmid DNA (pGEM4Z-EGFP)	F Primer: 5' TTGTGTGGAATTGTGAGCGG 3'	Plus	20	59.05	50.00	4.00	1.00
	R Primer: 5' GTCGGGGCTGGCTTAACTAT 3'	Minus	20	59.53	55.00	4.00	2.00
nLuc plasmid DNA (pcDNA-LUC-CBR2opt-T7AG-C1)	F Primer: 5' TGTACGGGCCAGATATACGC 3'	Plus	20	59.40	55.00	4.00	2.00
	R Primer: 5' GCAACTAGAAGGCACAGTCG 3'	Minus	20	58.93	55.00	4.00	2.00
BA4/5 SARS-CoV-2 Spike protein plasmid DNA (230907-BA45-Genscript)	F Primer: 5' ACGTTGTAAAACGACGGCCA 3'	Plus	20	60.81	50.00	8.00	2.00
	R Primer: 5' CCGATGAGCTCTAGCATTTAGG 3'	Minus	22	58.42	50.00	6.00	2.00
Reference SARS-CoV-2 Spike protein plasmid DNA (230907-Reference-Genscript)	F Primer: 5' ACGTTGTAAAACGACGGCCA 3'	Plus	20	60.81	50.00	8.00	2.00
	R Primer: 5' CCGATGAGCTCTAGCATTTAGG 3'	Minus	22	58.42	50.00	6.00	2.00

Figure 5: Characteristics of the designed oligonucleotide primer pairs for each PCR template, generated using Primer-BLAST.

Figure 5 shows the specific primer pairs designed for each PCR DNA template and their characteristics. The obtained PCR amplicons should contain the following elements to be used as a template for mRNA production: the 5'-UTR and 3'-UTR flanking the coding sequence, the

poly(A) tail sequence, and the T7 promoter sequence. As such, the forward primer (F primer) was designed to be a few bp before the 5'-UTR and the reverse primer (R primer) a few bp after the poly(A) tail sequence, such that all required elements will be present within the amplicon after PCR amplification.

3.2.2.2. PCR Reaction Conditions

Each PCR amplification reaction was set up according to the manufacturer's suggested protocols. Four different DNA Polymerases (DNA Pols) were used for comparison purposes: Q5 Hot Start High-Fidelity (HF) DNA Pol (#M0494) acquired from New England Biolabs (NEB), Phusion HF DNA Pol (#F553) acquired from Thermo Scientific, Platinum™ Taq DNA Pol HF (#11304011) acquired from Invitrogen, and recombinant Taq DNA Pol (#10342053) also purchased from Invitrogen.

Reaction and cycling conditions are shown in *Figure 6* for all four DNA Pols mentioned above. Reaction conditions ranged from: 1X reaction buffer (either separate in the PCR kit or provided with the enzyme mix), 200 µM each dNTPs, 0.2-0.5 µM forward and reverse primers, 1-2 ng (0.04 ng/µL) plasmid template and polymerase. Master mix formulations were prepared each time for ease of pipetting and accurate mixing of reagents. Cycling conditions varied with the DNA Pol used according to manufacturer's recommendation, and all reactions were cycled for 35 cycles. Amplification product yield was determined by measuring the concentration of unpurified PCR products using Quant-iT™ PicoGreen™ dsDNA Assay (Life Technologies, UK). PCR products were initially diluted in 1X TE Buffer (typically following a 20x dilution) and then serial diluted in a 96-well plate. Analysis was conducted based on the standard curve generated. Quality of the PCR amplicons was checked using DNA agarose gel electrophoresis.

A

Q5® Hot Start High-Fidelity 2X Master Mix			
Component	25 µL Reaction	50 µL Reaction	Final Concentration
Q5 High-Fidelity 2X Master Mix	12.5 µL	25 µL	1X
10 µM Forward Primer	1.25 µL	2.5 µL	0.5 µM
10 µM Reverse Primer	1.25 µL	2.5 µL	0.5 µM
Template DNA	Selected 1 ng	Selected 2 ng	< 1,000 ng
Nuclease-Free Water	to 25 µL	to 50 µL	

Thermocycling Conditions Q5 HF DNA Polymerase:				
Step		Temp	Time (manufacturer)	Time (selected)
Initial Denaturation		98 °C	30 seconds	30 seconds
PCR Cycles (35 Cycles)	Denaturation	98 °C	5-10 seconds	10 seconds
	Annealing	*50-72 °C	10-30 seconds	30 seconds
	Extension	72 °C	20-30 seconds/kb	1 minute
Final extension		72 °C	2 minutes	2 minutes
Hold		4-10 °C		4 °C infinite hold

* Annealing temperature range was provided by manufacturer. Annealing temperature (T_a) screening was performed to determine ideal T_a for each plasmid DNA: nLuc and eGFP

B

Phusion High-Fidelity DNA Polymerase PCR Kit			
Component	25 µL Reaction	50 µL Reaction	Final Concentration
H_2O	to 25 µL	to 50 µL	
5X Phusion HF Buffer	5 µL	10 µL	1X
10 mM dNTPs	0.5 µL	1 µL	200 µM each
10 µM Forward Primer	1.25 µL	2.5 µL	0.5 µM
10 µM Reverse Primer	1.25 µL	2.5 µL	0.5 µM
Template DNA	Selected 1 ng	Selected 2 ng	
Phusion DNA Polymerase	to 25 µL	to 50 µL	0.02 U/µL

Thermocycling Conditions Phusion HF DNA Polymerase:				
Step		Temp	Time (manufacturer)	Time (selected)
Initial Denaturation		98 °C	30 seconds	30 seconds
PCR Cycles (35 Cycles)	Denaturation	98 °C	5-10 seconds	10 seconds
	Annealing	*X °C	10-30 seconds	30 seconds
	Extension	72 °C	15-30 seconds/kb	1 minute
Final extension		72 °C	5-10 minutes	7 minutes
Hold		4 °C		infinite hold

* Annealing temperature (T_a) screening was performed to determine ideal T_a for each plasmid DNA: nLuc and eGFP

C

Platinum Taq High-Fidelity DNA Polymerase PCR Kit			
Component	25 µL Reaction	50 µL Reaction	Final Concentration
H_2O	to 25 µL	to 50 µL	
10X HF PCR Buffer	2.5 µL	5 µL	1X
50 mM $MgSO_4$	1 µL	2 µL	2.0 mM
10 mM dNTPs	0.5 µL	1 µL	0.2 mM each
10 µM Forward Primer	0.5 µL	1 µL	0.2 µM
10 µM Reverse Primer	0.5 µL	1 µL	0.2 µM
Template DNA	Selected 1 ng	Selected 2 ng	< 500 ng
Platinum Taq DNA Pol (5 U/µL)	0.1 µL	0.2 µL	1 U/rxn

Thermocycling Conditions Platinum Taq HF DNA Polymerase:				
Step		Temp	Time (manufacturer)	Time (selected)
Initial Denaturation		94 °C	30seconds-2minutes	1 minute
PCR Cycles (35 Cycles)	Denaturation	94 °C	15 seconds	15 seconds
	Annealing	*~55°C	30 seconds	30 seconds
	Extension	68 °C	1 minute/kb	2 minutes
Hold		4 °C		infinite hold

* Depending on melting temperature of primers (T_m). Annealing temperature (T_a) screening was performed to determine ideal T_a for each plasmid DNA: nLuc and eGFP

D

Taq DNA Polymerase PCR Kit			
Component	25 µL Reaction	50 µL Reaction	Final Concentration
Autoclaved, distilled water	to 25 µL	to 50 µL	
10X PCR Buffer, -Mg	2.5 µL	5 µL	1X
50 mM $MgCl_2$	0.75 µL	1.5 µL	1.5 mM
10 mM dNTPs Mix	0.5 µL	1 µL	0.2 mM each
10 µM Forward Primer	1.25 µL	2.5 µL	0.5 µM
10 µM Reverse Primer	1.25 µL	2.5 µL	0.5 µM
Template DNA	Selected 1 ng	Selected 2 ng	1-500 ng
Taq DNA Polymerase (5 U/µL)	0.1 µL	0.2 µL	1.0-2.5 U/µL

Thermocycling Conditions Taq DNA Polymerase:				
Step		Temp	Time (manufacturer)	Time (selected)
Initial Denaturation		94 °C	3 minutes	3 minutes
PCR Cycles (35 Cycles)	Denaturation	94 °C	45 seconds	45 seconds
	Annealing	*~55 °C	30 seconds	30 seconds
	Extension	72 °C	90 seconds/kb	3 minutes
Final extension		72 °C	10 minutes	10 minutes
Hold		4 °C		infinite hold

* Depending on melting temperature of primers (T_m). Annealing temperature (T_a) screening was performed to determine ideal T_a for each plasmid DNA: nLuc and eGFP

Figure 6: PCR Components and Thermocycling conditions for: (A) Q5 Hot Start HF 2X Master Mix DNA Pol, (B) Phusion HF DNA Pol, (C) Platinum Taq HF DNA Pol, and (D) Taq DNA Pol.

3.3. Template DNA Purification

3.3.1. Plasmid DNA Linearization

Before it can be used in the IVT, the DNA plasmids, generated using *E. coli* and purified using either the QIAprep® Spin Miniprep kit or the NucleoBond Xtra Midi EF kit, should be linearized and further purified for the removal of the restriction enzymes. peGFP was linearized using SpeI-HF® (purchased from NEB) and pnLuc was linearized using HindIII-HF® (NEB), as both restriction sites, shown in the corresponding sequence maps in *Figure 7,* are positioned right after the poly(A) tail. Moreover, for the use of T7 RNA polymerase during IVT, a 5'overhang is preferable to ensure polymerase stability and reduce artifacts, which is achieved after linearization using endonucleases such as SpeI and HindIII [123,124]. SpeI-HF® and HindIII-HF® were chosen

as HF as they have the same specificity as their native forms with significantly reduced star activity, and function in rCutSmart™ Buffer. Linearization was done overnight in the incubator at 37°C, followed by heat inactivation at 80°C for 20 minutes.

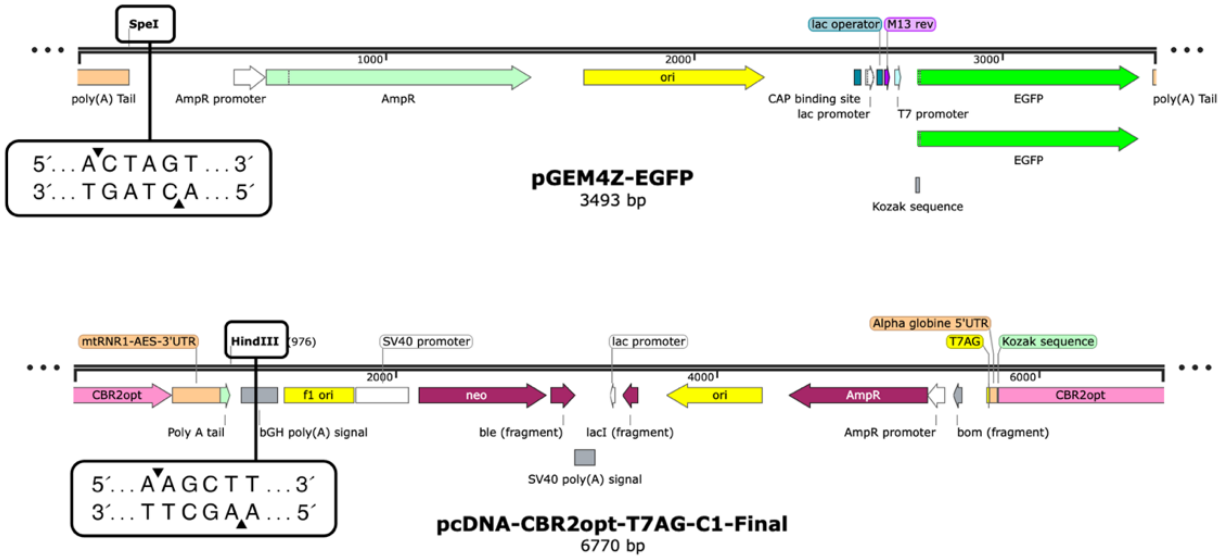


Figure 7: peGFP and pnLuc restriction sites, showing the 5'overhang preferred for IVT, sequence maps taken from SnapGene.

3.3.2. Template DNA Purification prior to IVT

Removal of the enzyme is then required to isolate the final IVT DNA template. On the other hand, for DNA templates generated using PCR, no further linearization step is required, rendering the overall manufacturing of IVT templates shorter. Post-PCR purification is needed for the removal of the unused nucleotides, primers, enzyme, and buffer components.

Three different purification methods were successfully compared, using both DNA constructs (eGFP and nLuc) to demonstrate their ability to be used for several sequence lengths. The first two purification methods are lab-scale and rely on the use of a commercially available kits (NucleoSpin® Gel and PCR Clean-up Mini kit purchased from Macherey-Nagel). These first two methods are (1) DNA extraction from agarose gel and (2) PCR Clean-up purification methods. DNA gel extraction allows for the isolation and purification of linear DNA fragments based on size. Following agarose gel electrophoresis, the DNA band of interest was cut from the gel under UV light and purified following the NucleoSpin® manufacturer's protocol. The second lab-scale method, PCR-clean-up, is not only suitable for PCR clean-up but also suitable for DNA

concentration and removal of salts, enzymes, etc. from enzymatic reactions. As such, it was completed using a modified version of the NucleoSpin® manufacturer's protocol which enhances DNA recovery by performing three elution steps and heating the elution buffer (5 mM Tris/HCl, pH 8.5) and column to 70^{o}C.

Finally, the third purification method we examined is a scalable in-house optimized TFF scale down technique that uses MilliporeSigma™ Amicon™ ultra-0.5 centrifugal filter units with high-recovery Ultracel™-PL regenerated cellulose membranes with a molecular weight cut-off (MWCO) of at least 30 kDa. A MWCO of 100 kDa was chosen for the removal of restriction enzyme from the *E. coli* produced DNA, allowing for the buffer (10 mM Tris-HCl, pH 8.0) and restriction enzymes to flow through while retaining and concentrating the template DNA. For PCR products, a MWCO of 30 kDa was chosen, as recommended by the manufacturer.

3.4. *In Vitro* Transcription (IVT) & mRNA Purification

mRNA was produced from the previously purified linear DNA templates using the MEGAscript™ T7 Transcription Kit (Thermo Fisher Scientific, Waltham, MA, USA), with IVT experiments conducted without co-transcriptional 5'capping. The obtained mRNA was purified using a standard phenol-chloroform purification protocol. Quantification of purified mRNA was done using Quant-it Ribogreen RNA Assay Kit (from Invitrogen). mRNA sequence length, purity, and quality were evaluated by formaldehyde gel electrophoresis and subject to analysis using the Agilent 2100 Bioanalyzer.

3.5. Analytics

- ***NanoDrop*:** pDNA and linear purified DNA were quantified by UV absorption at 260 nm using NanoDrop® UV spectrophotometer (Thermo Scientific). DNA purity was also verified using UV spectroscopy. An $A_{260/280}$ ratio between 1.80-1.90 and a ratio $A_{260/A230}$ between 2.0-2.2 was used as an indication of pure pDNA. If the ratio of absorbance at 260 nm and 280 nm was lower than these expected, this was used as an indication of the presence of protein, phenol or other contaminants that absorb strongly at or near 280 nm. If the 260/230 ratio is lower than expected, this was used as an indication of the presence of contaminants which absorb at 230 nm such as EDTA, phenol, carbohydrate, guanidine HCL, etc.

- ***DNA Agarose Gel Electrophoresis*:** DNA length and quality were verified using 1% agarose (UltraPure™ Agarose, Invitrogen) gel electrophoresis and QuickLoad Purple 1kb Plus DNA ladder (ThermoFisher). The gels were run at 90 V for 50 minutes. Invitrogen SYBR® Safe DNA gel stain was used for visualizing DNA on the agarose gel. Gel imaging was conducted using the ChemiDoc MP imaging system and associated Image Lab software.

- ***Picogreen Assay:*** dsDNA concentration was determined using the Quant-iT™ PicoGreen™ Kit (Invitrogen). Samples were serially diluted in a 1:2 ratio in a 96-well microplate and fluorescence intensity measurements were taken with the Agilent BioTek Synergy HTX MultiMode Microplate Reader using fluorescent excitation and emission wavelengths (480/520 nm) set on the associated Gen 5 software. Plate readings were done following a 15-minute incubation at room temperature, protected from light. A dsDNA standard curve was generated using the provided Lambda DNA standard in the Quant-iT™ PicoGreen™ Kit, ranging from 500 ng/mL to 40 ng/mL. Initial sample dilution is done in 1X TE Buffer, and Quant-iT™ PicoGreen™ dsDNA Reagent is used as the dsDNA dye, following a 200-fold dilution.

- ***Ribogreen Assay:*** mRNA concentration was determined using the Quant-iT™ RiboGreen™ Kit (Invitrogen). Samples were serially diluted in a 1:6 ratio in a 96-well microplate and fluorescence intensity measurements were taken with the Agilent BioTek Synergy HTX MultiMode Microplate Reader using standard fluorescein wavelengths (excitation ~480 nm, emission ~520 nm) set on the associated Gen 5 software. Plate readings were done following a 2-5 minutes incubation at room temperature, protected from light. An RNA standard curve was generated using the provided Ribosomal RNA standard in the Quant-iT™ RiboGreen™ Kit, ranging from 1000 ng/mL to 20 ng/mL. Initial sample dilution is done in 1X TE Buffer, and Quant-iT™ RiboGreen™ RNA Reagent is used as the RNA dye, following a 200-fold dilution.

- ***RNA Formaldehyde Gel Electrophoresis*:** mRNA length and quality were verified using formaldehyde gel electrophoresis and RiboRuler High Range RNA Ladder (ThermoFisher). The gel setup was washed with RNAse zap and MilliQ water prior to use. Gels were run in 1X 3-(N-morpholino)propanesulfonic acid (MOPS) buffer at 110 V for 50-60 minutes. After adding loading buffer (0.6 g agarose in 50 mL MilliQ water, 4 mL formaldehyde and 6 mL

MOPS 10X buffer), the mRNA samples were heated at 65-70 °C for 5 min then chilled on ice to denature the secondary structure. Gel imaging was conducted using the ChemiDoc MP imaging system.

- ***Bioanalyzer Analysis***: mRNA characterization and quality were assessed using the Agilent 2100 Bioanalyzer, at the Department of Pharmacology and Therapeutics, McGill University.

4. Results

4.1. Template DNA Production

4.1.1. Baseline Template DNA Production Method: *E. coli* Bacterial Fermentation

As *E. coli* fermentation is considered the gold-standard method for DNA template production, the first step was to establish this process in the Kamen Lab for later comparison purposes. Standard growth conditions in LB medium for MAX Efficiency® DH10Bac™ Competent Cells for 100 mL culture expansion volumes were used. First, to establish a reference analytical method, as well as to compare the results of known analytical assays, Nanodrop and Picogreen analytical tools were compared in their ability to measure pDNA content. *Figure 8* illustrates these results, where peGFP was used as the benchmark between both assays. It was shown that relative to the Picogreen measurements, the Nanodrop pDNA concentration measurements are consistently higher, by around a factor of 1.6 on average. For future analytical considerations, Nanodrop measurements were used for pDNA content.

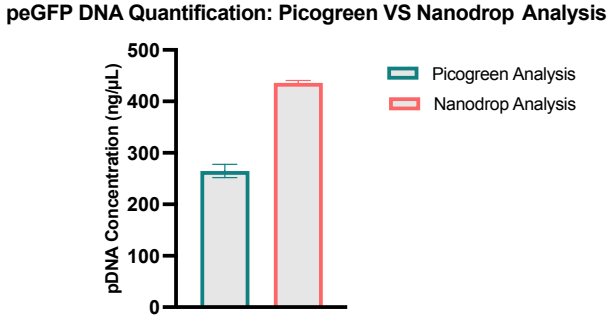


Figure 8: Comparison of analytical methods for pDNA content quantification, for peGFP bacterial fermentation productions. These results were measured post-extraction and Midi Prep purification.

Different bacterial strains were first compared in their ability to produce pDNA at small-scale to then select our benchmark strain for production scale-up. *Figure 9(A)* shows the transformation plates of the three bacterial strains which were assessed, (a) DH10Bac *E. coli*, (b) NEB Stable competent *E. coli*, (c) *Vibrio Natriegens* ATCC14048. pnLuc transformation was hypothesized successful for all three strains, as bacterial colonies grew in each culture plate. However, it was observed that the *E. coli* NEB Stable competent colonies were smaller and less distinctive than the *E. coli* DH10Bac ones, with some satellite colonies appearing on the corner of the plate. On the other hand, the *Vibrio Natriegens* colonies appeared overall larger than both *E. coli* strains. *Figure 9(B)* illustrates the DH10Bac and NEB stable *E. coli* strains OD_{600} growth curves. However, this

figure also demonstrates that no growth was observed for the *Vibrio Natriegens* strain as the corresponding inoculation step consistently failed despite colony growth on plates. We thus hypothesize that its transformation was not as efficient as previously observed. It was shown that the DH10Bac strain reached an OD of 1 more rapidly than the NEB Stable strain. Moreover, it can be observed that the DH10Bac strain reached a doubling time of around 2 hours from time t = 6h to time t = 8h. Similar growth trends are observed for the NEB strain with a time delay of 2 hours. For example, this strain only reached a doubling time of 2 hours during the interval between time t = 8h to time t = 9h45min (*Fig. 9(C)*).

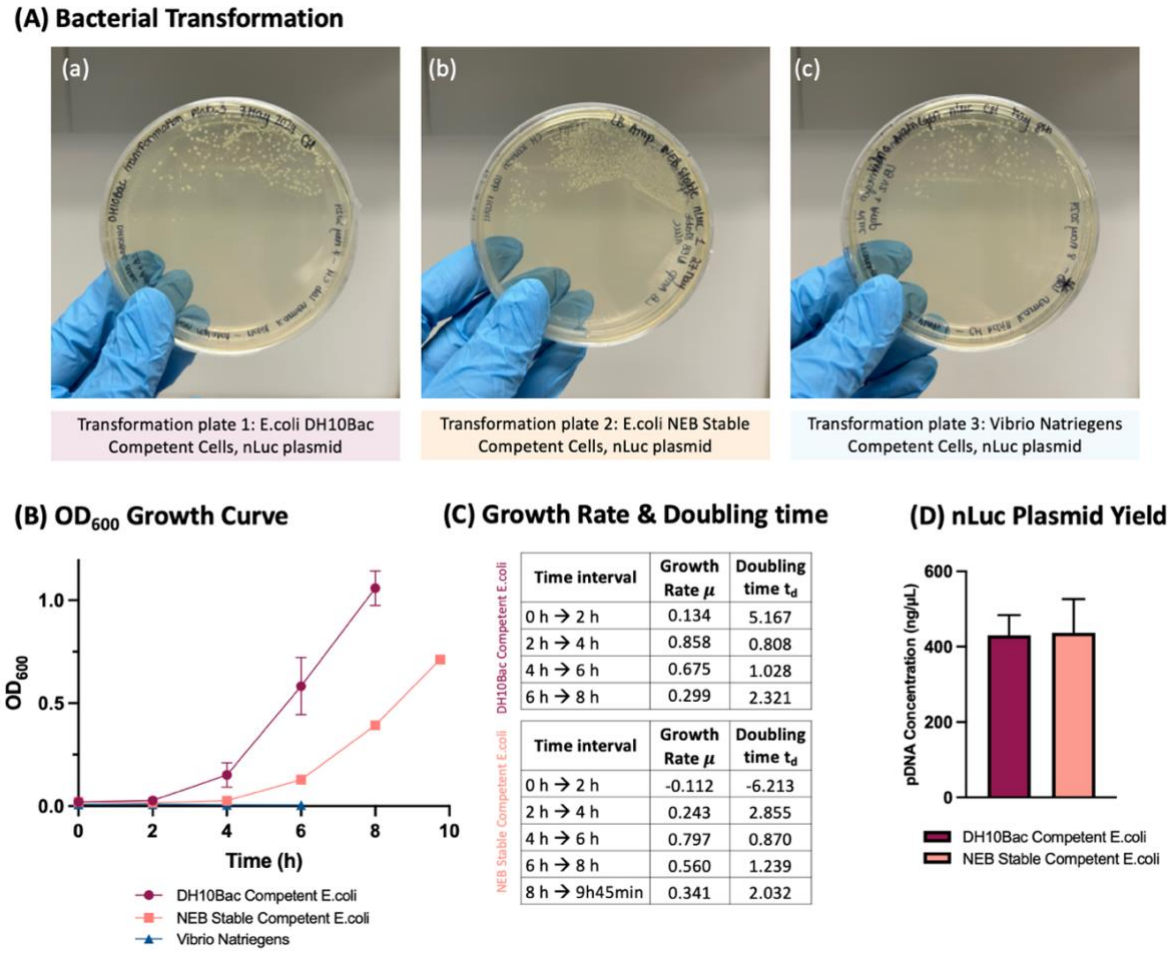


Figure 9: Comparison of Different Bacterial Strains for pDNA Production
(A) nLuc bacterial culture plates with ampicillin showing transformed colonies for three different bacterial strains: (a) *E. coli* DH10Bac competent cells, (b) *E. coli* NEB Stable competent cells, and (c) *Vibrio Natriegens* ATCC 14048. (B) OD_{600} growth curve with off-line measurements taken at different time points for 6 mL inoculation cultures at 37ºC (*E. coli* cultures in LB and *Vibrio Natriegens* cultures in LBv2 medium). (C) *E. coli* growth rates and doubling times. (D) nLuc pDNA yield measured using Nanodrop with a Mini prep purification elution volume of 50 µL, for two different *E. coli* competent cells: DH10Bac and NEB Stable competent cells.

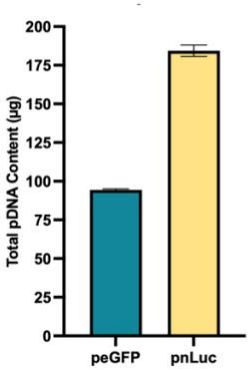
Finally, the nLuc plasmid titers (estimated via alkaline lysis/Mini prep followed by UV absorption at 260 nm using NanoDrop® UV spectrophotometer) were determined and shown in *Figure 9(D)*. No significant differences were observed between strains in terms of yield. Mini Prep plasmid extraction and purification were conducted starting with cultures of an OD of 0.8, for ease of comparison between the two *E. coli* strains.

The DH10Bac strain was then selected for scale-up production due to its ease of transformation, rapid growth, and sufficient plasmid yield. For our two reporter plasmids, i.e., peGFP and pnLuc, the biomass (estimated by the OD_{600} measurement) and the plasmid titers (estimated via alkaline lysis/Midi prep followed by UV absorption at 260 nm) were determined. These results are summarised in *Figure 10(A)* and *(B)*, respectively. Plasmid extraction and purification were conducted once the cultures were close to an OD_{600} of 1, as shown in *Figure 10(A)*. Starting with approximately 30 ng of pDNA, using this bacterial fermentation approach, around 95 µg of peGFP and around 185 µg of pnLuc were produced, serving as benchmark values for later comparison purposes. For quality control, the plasmid purity and integrity were evaluated using Nanodrop absorbance ratios and agarose gel electrophoresis. The $A_{260/280}$ and $A_{260/230}$ ratios obtained were 1.90 and 2.25 respectively for peGFP, and 1.96 and 2.31 respectively for pnLuc, confirming purity of pDNA post bacterial extraction and purification. Moreover, *Figure 10(C)* demonstrates the integrity of both reporter plasmids: as expected, different DNA conformations are shown on the agarose gel, with the desired supercoiled confirmation shown as the thickest band obtained for lines 2 to 5.

(A) Bacterial Fermentation

	OD_{600} measurements (5 mL inoculation at 37°C for 6h)	OD_{600} measurements (100 mL inoculation at 37°C for 18h)
eGFP Colony 1	0.060	1.16
eGFP Colony 2	0.051	1.216
nLuc Colony 1	0.059	1.02
nLuc Colony 2	0.059	0.942

(B) Plasmid Yield



(C) Plasmid Integrity

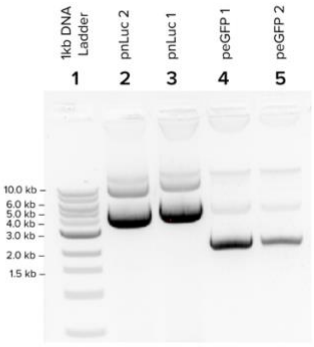


Figure 10: DH10Bac Bacterial Fermentation for pDNA Production
(A) OD_{600} values and biomass conversions taken at two different times of cultivation for *E. coli* cultures in LB medium, for 2 selected colonies per pDNA transformation plates i.e., (1) 6h 5 mL inoculation culture at 37°C, and (2) 18h 100 mL expansion culture at 37°C. (B) pDNA content measured using Nanodrop (UV

absorption at 260 nm) for both reporter plasmids, i.e., peGFP and pnLuc, with a Midi prep purification elution volume of 200 µL. (C) Agarose gel electrophoresis of pnLuc and peGFP produced in MAX Efficiency® DH10Bac™ Competent Cells and purified using the Midi Prep kit described in the methodology section 8.2.1.2.

4.1.2. Establishment of PCR as an Alternative Approach for DNA Production

In this section, we establish PCR as a potential alternative approach for the generation of DNA templates for IVT. Four different DNA Pol were compared in their ability to produce the desired sequence in terms of yield and sequence identity, as mentioned in the methodology section: Q5 Hot Start HF DNA Pol, Phusion HF DNA Pol, Platinum Taq HF DNA Pol, and recombinant Taq DNA Pol. PCR reactions were conducted using the designed primers listed in *Figure 5* Section 8.2.2.1. and the conditions recorded in *Figure 6* Section 8.2.2.2. Following the results displayed in *Figure 11*, 1 ng of starting DNA material was selected to be used for each 25 µL PCR reaction, for both reporter constructs, as no significant difference in the yield was observed when higher amounts of starting template were used.

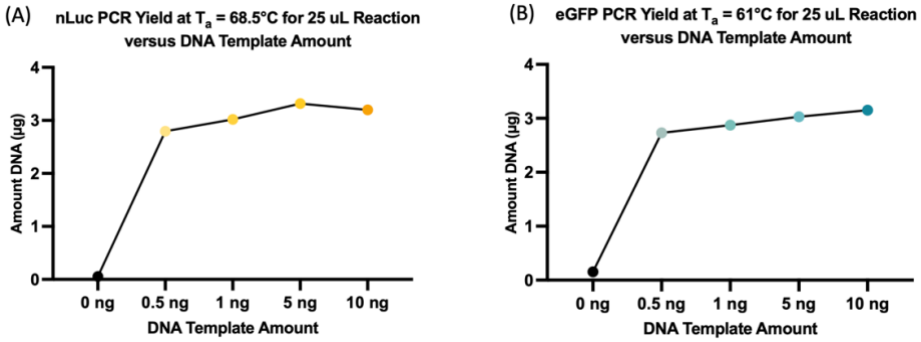


Figure 11: PCR DNA content yield versus starting DNA template amount, for nLuc and eGFP templates, measured using Picogreen.

After setting up the PCR conditions, an annealing temperature (T_a) screening was first conducted using both reporter sequences, i.e. nLuc and eGFP, for each of the four DNA Pol, to establish the best T_a for each polymerase. This step was critical in order to establish the optimal PCR conditions for each DNA Pol, prior to comparing relative polymerase performance.

The ideal T_a was determined by selecting the temperature condition satisfying both a high DNA yield and the correct sequence identity via agarose gel electrophoresis. The T_a screening was first completed for the synthesis of nLuc and eGFP amplicons using Q5 Hot Start HF DNA Pol, as shown in *Figure 12*. *Figure 12(1a), (1b)* and *(1c)* demonstrate that both reporter sequences, nLuc

and eGFP, were produced at the correct length. In the case of nLuc, the thickest band observed on the agarose gels is between the 2 kb and 3 kb ladder markers, confirming that the correct length (2172 bp) was obtained. We hypothesize that the other faint bands observed correspond to the initial circular pDNA supplied to the PCR reaction as it takes on different conformations migrating differently on the gel. Moreover, a non-specific PCR band was observed for nLuc amplicon production at around 300 bp for T_a values lower than 68ºC. As a result, the T_a screening for the nLuc template was consistently performed at higher T_a values than those used for the eGFP DNA template to reduce non-specific bands generated. In the case of the eGFP amplicon, no non-specific PCR bands were observed, and the desired PCR product (1193 bp) is between the 1 kb and 1.2 kb ladder markers, as expected. *Figure 12(2a)* and *(2b)* show that starting with 1 ng of pDNA for a 25 µL PCR reaction, we are able to produce around 3.3 µg of the nLuc DNA template and around 3.12 µg of the eGFP template using Q5 Hot Start HF DNA Pol. Following these results, the ideal T_a selected were 68.5ºC and 61ºC for PCR production of nLuc and eGFP sequences, respectively.

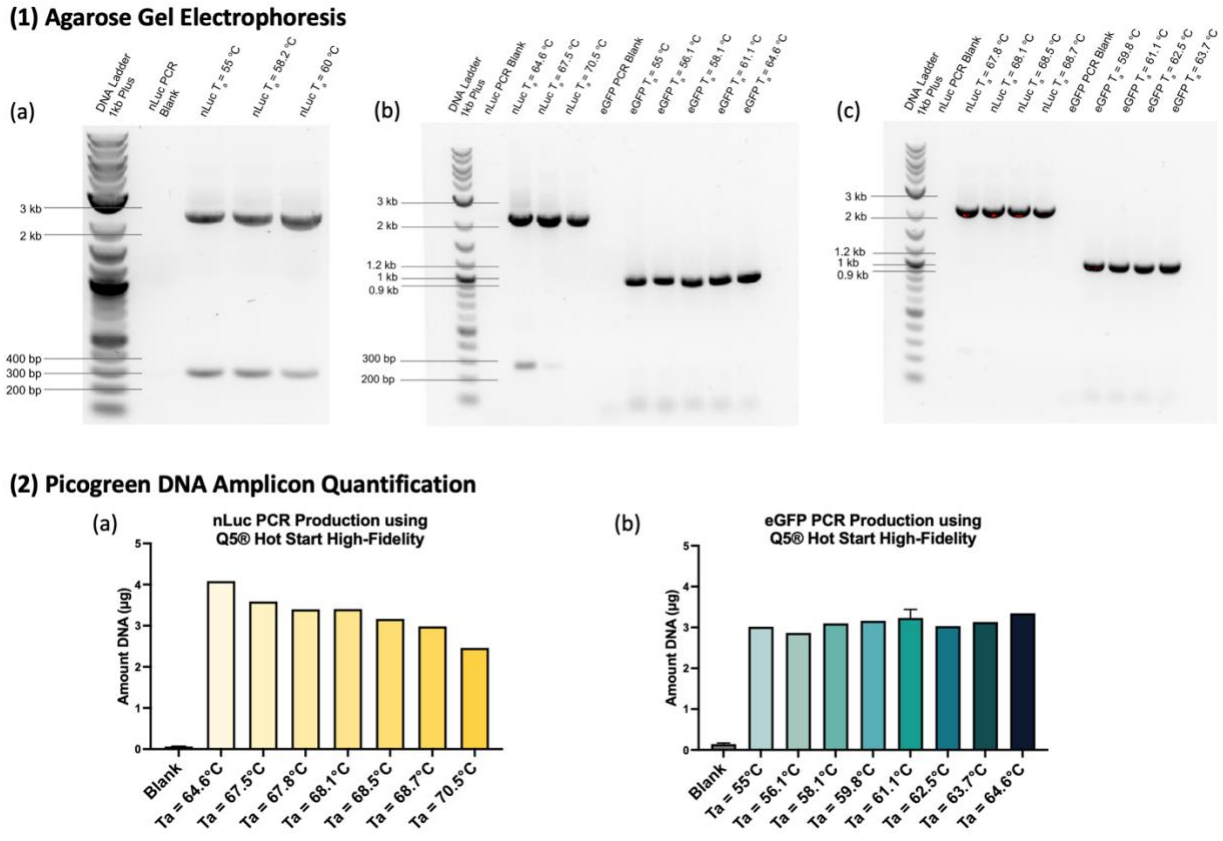


Figure 12: Annealing Temperature Screening for Q5 Hot Start HF DNA Pol.
(1a) Agarose gel of nLuc PCR amplicon for three different T_a: 55ºC; 58.2ºC; 60ºC. (1b) Agarose gel of nLuc amplicon for three different T_a: 64.6ºC; 67.5ºC; 70.5ºC, and eGFP amplicon for five different T_a:

55ºC; 56.1ºC; 58.1ºC; 61.1ºC; 64.6ºC. (1c) Agarose gel of nLuc amplicon for four different T_a: 67.8ºC; 68.1ºC; 68.5ºC; 68.7ºC, and eGFP amplicon for four different T_a: 59.8ºC; 61.1ºC; 62.5ºC; 63.7ºC. (2a) nLuc PCR production using Q5 Hot Start HF DNA pol, DNA content measured using Picogreen for T_a ranging from 64.6ºC to 70.5ºC. (2b) eGFP PCR production using Q5 Hot Start HF DNA pol, DNA content measured using Picogreen for T_a ranging from 55ºC to 64.6 ºC.

Similarly, a T_a screening was conducted for the three other DNA Pol mentioned above. As shown in *Figure 13(A), (C)* and *(E),* nLuc DNA template was successfully produced at an average of 2.28 µg, 3.63 µg, and 3.47 µg, using Phusion HF DNA Pol, Platinum Taq HF DNA Pol, and recombinant Taq DNA Pol, respectively. Regarding the eGFP DNA template, as shown in *Figure 13(B), (D)* and *(F),* average productions of 3.22 µg, 1.91 µg, and 1.43 µg were obtained using Phusion HF DNA Pol, Platinum Taq HF DNA Pol, and recombinant Taq DNA Pol, respectively. These measurements were taken using Picogreen dsDNA analysis for 25 µL PCR reactions starting with 1 ng of pDNA template encoding the sequence of interest, as nanodrop (UV absorbance) is not reliable for the quantification of unpurified PCR products.

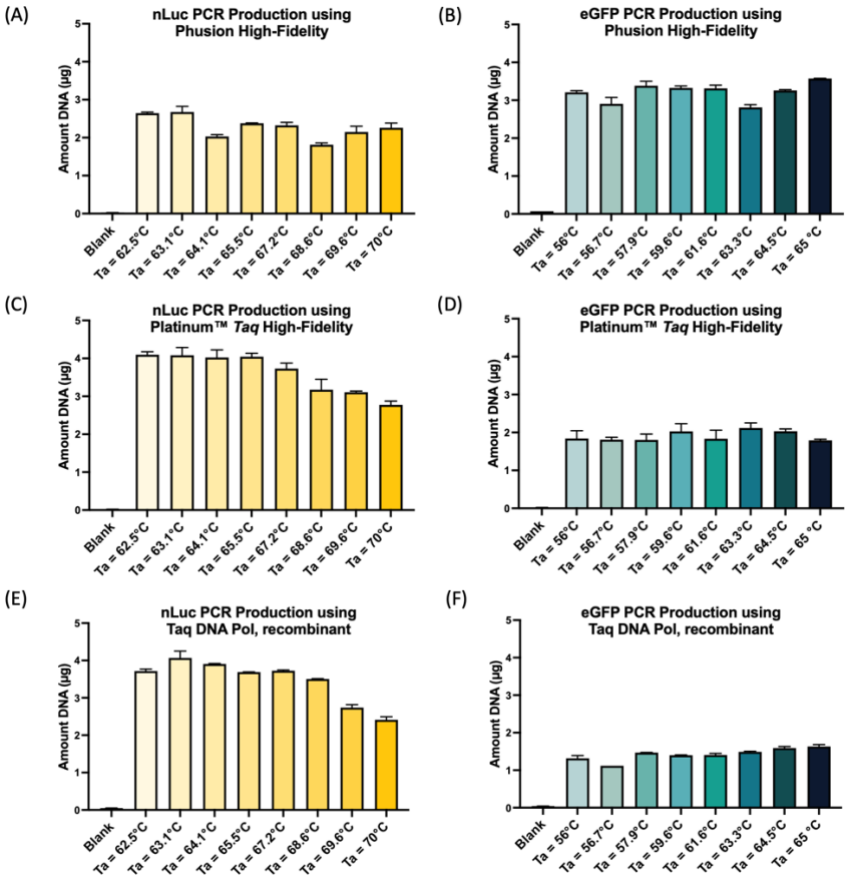


Figure 13: DNA content measured using Picogreen for a range of annealing temperatures for both reporter sequences, i.e. nLuc and eGFP, produced using 25 µL PCR reactions starting with 1 ng

of pDNA template, with three different DNA Pol: Phusion HF DNA Pol, Platinum Taq HF DNA Pol, and recombinant Taq DNA Pol respectively

Moreover, *Figure 14* illustrates that both reporter sequences were produced at the correct length. Non-specific undesired bands were observed on the agarose gels for nLuc PCR products of Phusion HF DNA Pol and recombinant Taq DNA Pol at around 200-300 bp for T_a values of 65.5°C and lower, and around 200-300 bp and 500-600 bp for T_a values of 68.6°C and lower, respectively. In the case of the eGFP amplicon, no non-specific PCR bands were observed, as shown in *Figure 14 (d), (e)* and *(f),* for the three different DNA Pol used. Following these results, the ideal T_a selected for nLuc and eGFP template PCR productions under the pre-defined conditions were: for (1) Phusion HF DNA Pol: 67°C and 61°C respectively, (2) Platinum Taq HF DNA Pol: 65°C and 60°C respectively, and (3) Taq DNA Pol: 70°C and 63°C respectively.

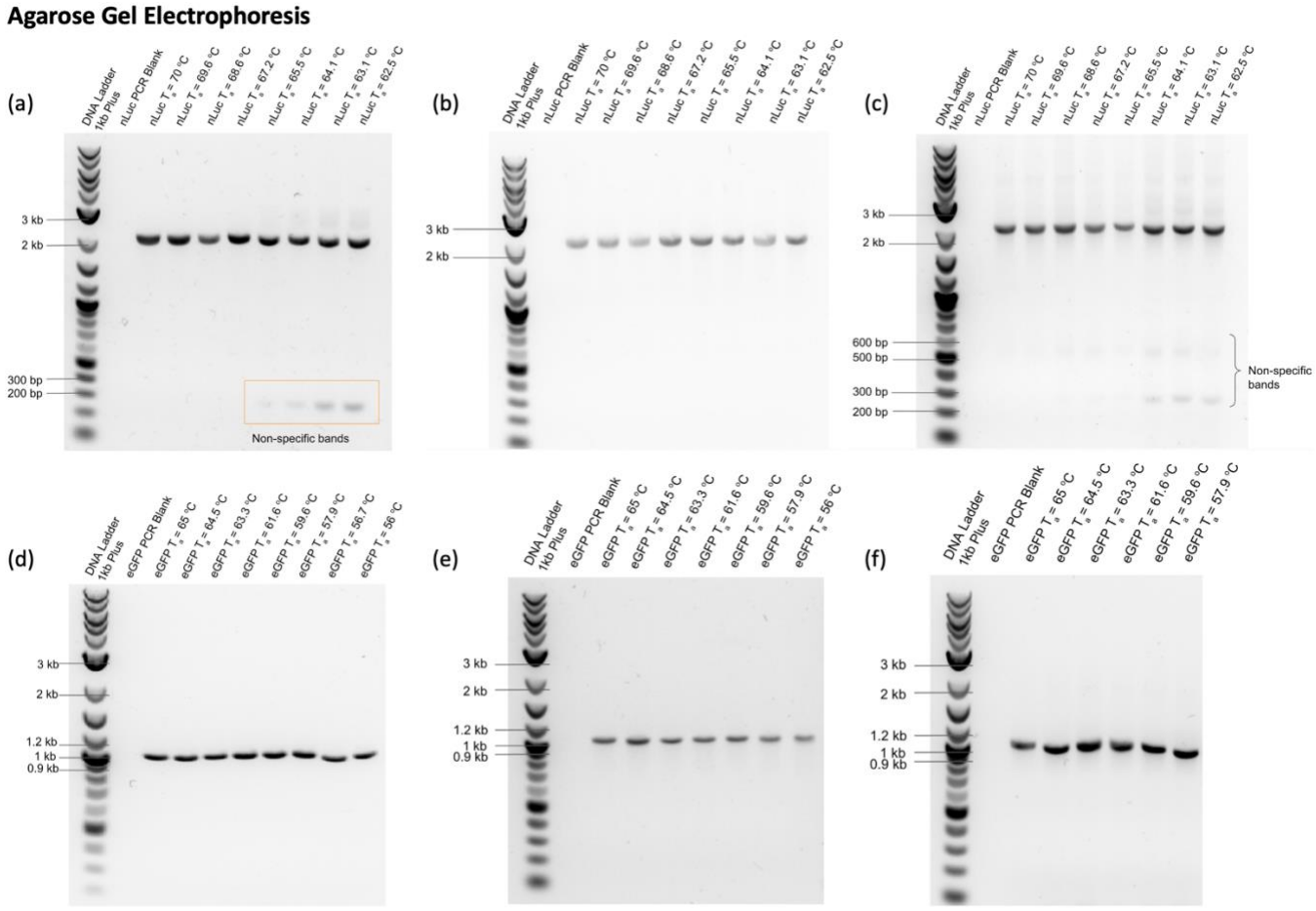


Figure 14: Agarose Gel Electrophoresis of PCR products for both reporter sequences, i.e. nLuc and eGFP, produced using PCR with three different DNA Pol: (a, d) Phusion HF DNA Pol, (b, e) Platinum Taq HF DNA Pol, and (c, f) recombinant Taq DNA Pol, respectively.

Following the selection of the ideal T_a for each polymerase, CDS and primer pair combination, we aimed to scale up the DNA production in order to produce sufficient material for subsequent mRNA transcription. To this end, for each polymerase and DNA sequence, the PCR yield was evaluated in both the 25 µL and 50 µL reaction scale. Overall DNA yields were compared between polymerases, as well as between reaction scales. *Figure 15* illustrates that as the reaction volume was doubled, the DNA yield increased for every polymerase and template evaluated. Specifically, for Q5 HF and Phusion HF polymerases, nearly twice the yield was observed for both the nLuc and eGFP amplicons, as expected theoretically. In the cases of Platinum Taq HF and Taq DNA Pol, performance of the polymerase was sequence dependent. Furthermore, despite the sequence length disparity, both nLuc and eGFP template yields were similar when produced in 50 µL reactions using Q5 HF and Phusion HF polymerases, with average production yields of 5.47 µg and 6.28 µg for nLuc respectively, and 5.28 µg and 6.70 µg for eGFP respectively (*Fig.15*).

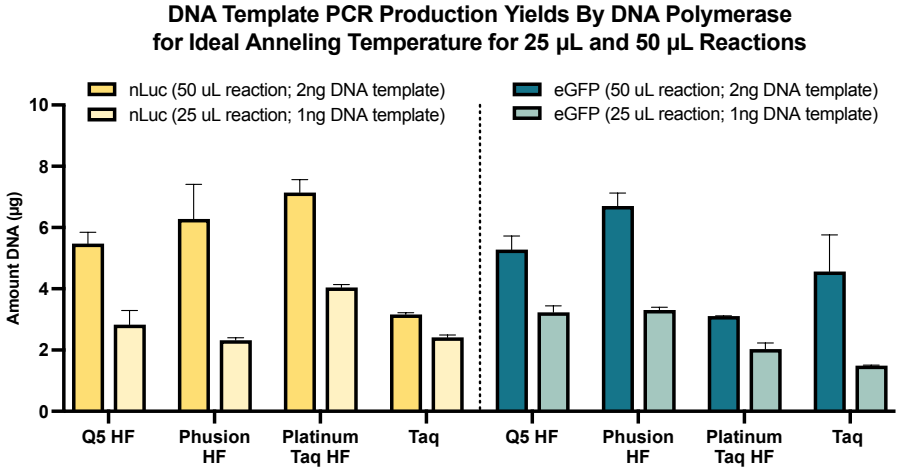


Figure 15: DNA Template PCR production yields by DNA Pol for ideal T_a previously determined, for 25 µL and 50 µL PCR reactions. Data measured using Picogreen analysis.

Q5 Hot Start HF DNA Pol was then selected for proof of concept of PCR production for longer DNA sequences: BA45 and Refspike with an amplicon size of respectively 4416 bp and 4431 bp, compared to the 2172 bp nLuc and 1193 bp eGFP amplicons, considering the chosen primers. *Figure 16(A)* demonstrates that under the same PCR conditions, the yield was slightly lower for longer DNA sequences. Specifically, under the highest production T_a (64°C), 2.63 µg and 2.34 µg were produced on average for BA45 and Refspike amplicons, compared to 3.09 µg and 3.23 µg for nLuc and eGFP, respectively. Moreover, correct sequence length was again verified using agarose gel electrophoresis, as shown in *Figure 16(B)*.

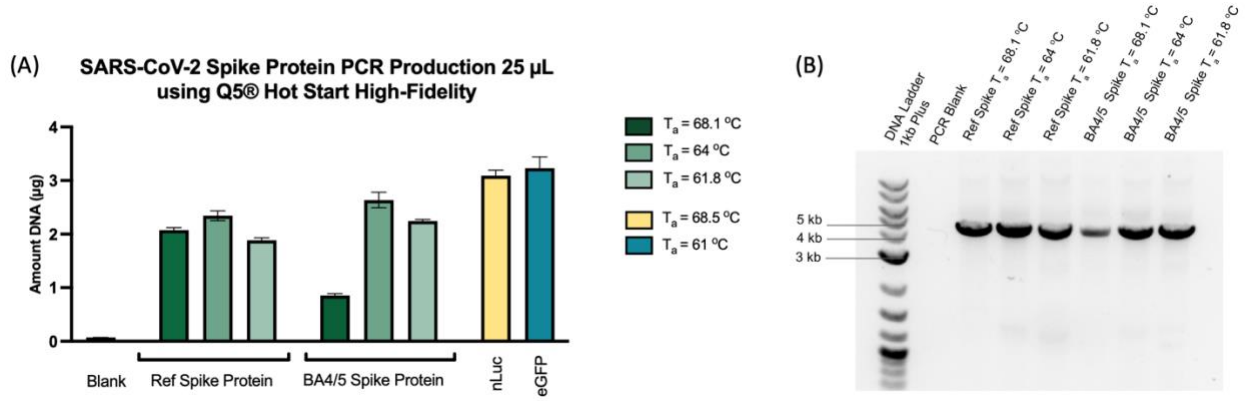


Figure 16: PCR Production for longer DNA templates encoding SARS-CoV-2 Spike protein. (A) DNA template PCR production yields for 25 µL reactions using Q5 Hot Start HF DNA Pol, measured using Picogreen, for different annealing temperatures. Three different T_a were used for the production of the SARS-CoV-2 protein GOI sequences: 68ºC, 64ºC, and 61.8ºC. (B) Agarose Gel Electrophoresis of PCR products encoding SARS-CoV-2 Spike protein at the three different T_a.

4.2. Template DNA Purification

4.2.1. Template DNA Preparation: pDNA origin

4.2.1.1. pDNA Linearization

After the GOI encoding pDNA has been generated using *E. coli* fermentation, and extracted and purified from bacterial cells, the DNA templates produced were linearized to fulfill the ultimate goal of mRNA transcription. As mentioned in Section 8.3.1, HindIII and SpeI restriction enzymes were used to linearize both reporter constructs, pnLuc and peGFP respectively. *Figure 17* shows that the linearization of both plasmid DNA constructs was successful.

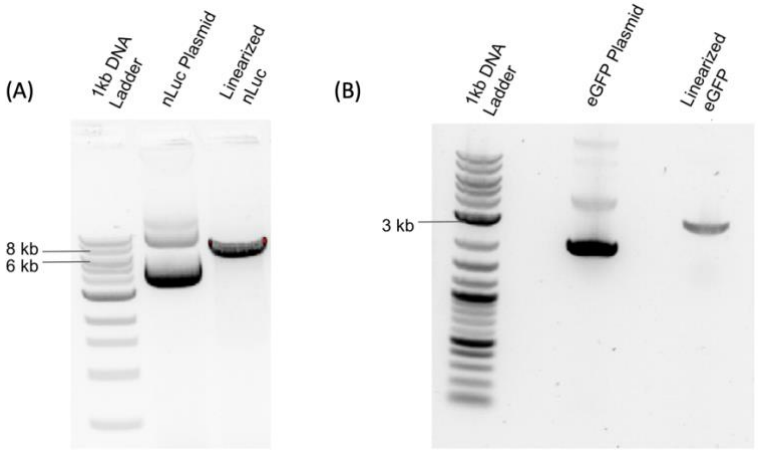


Figure 17: Agarose Gel Electrophoresis of both reporter constructs, (A) nLuc and (B) eGFP, in their plasmid and linearized forms.

The unique band on the DNA agarose gel for the linearized nLuc in *Figure 17(A)* appears to be between the 6 kb and 8 kb ladder markers, confirming that the correct length of the linear plasmid was obtained: 6.77 kb. Similarly, for eGFP in *Figure 17(B)*, the unique band for the linearized plasmid conformation is close to the 3 kb ladder marker, confirming identity with correct sequence length of 3.493 kb.

4.2.1.2. Further DNA Template Purification

Removal of the restriction enzymes is then required to isolate the final DNA template and proceed with IVT. Three different purification methods for the DNA template with pDNA origin were successfully compared: (1) DNA extraction from agarose gel, referred to in this section as gel extraction, (2) PCR clean-up, and (3) in-house TFF scale down method using Amicon Columns. Additional details regarding these methods are mentioned in Section 8.3.2.

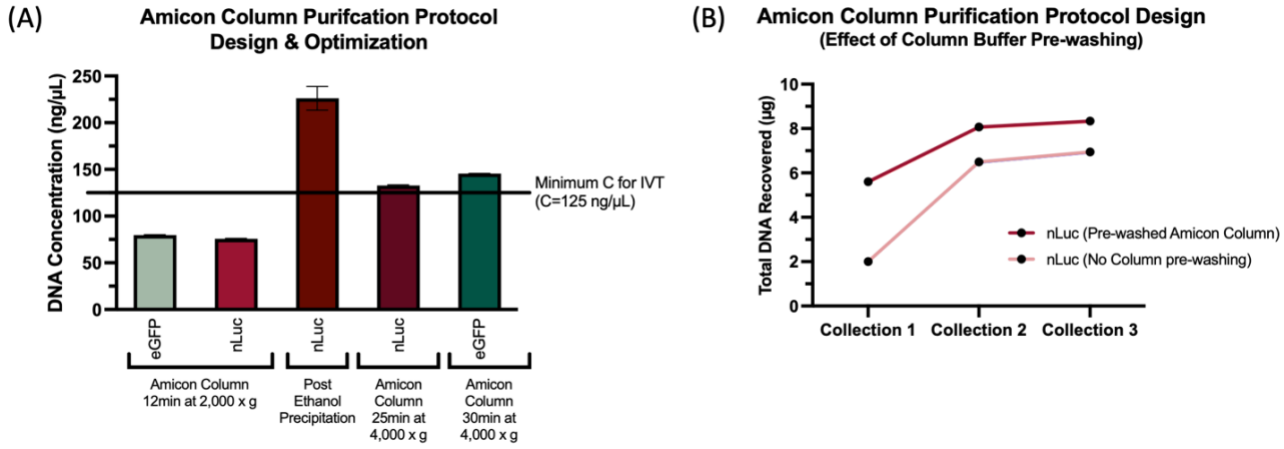


Figure 18: Amicon column-based purification protocol establishment and optimization.
(A) Amicon column purification protocol design and optimization, showing DNA concentration obtained via nanodrop measurements (in ng/µL) for nLuc and eGFP templates for various conditions: centrifugation speed and time. Values are compared to the minimum concentration required for using the obtained DNA template in IVT protocol in the lab. (B) Effect of column buffer pre-washing and sample collection number on total nLuc DNA recovered (in µg) using Amicon column purification at 4,000 x g for 30 min collection steps.

Establishment and optimization of an Amicon column-based purification protocol was done by evaluating different conditions such as centrifugation speed, time, and sample collection number. nLuc was selected for the optimization of the purification protocol, which was then validated using the eGFP sequence. First, it was clearly demonstrated that increasing the collection number to three increases DNA recovery (*Figure 18(B)).* We also demonstrate that pre-washing the column

with the Tris-HCl buffer contributed to an increased DNA recovery, as shown in *Figure 18(B)*. We began using a centrifugation speed of 2,000 x g for 12 min which, as shown in *Figure 18(A)*, did not result in a DNA template that was concentrated enough for the IVT protocol that we have established in the lab. Thus, in these cases, it would need to be followed by ethanol precipitation for DNA concentration to be used in IVT, which renders the purification process longer and decreases the post-purification recovery yield. Therefore, by increasing the centrifugation speed and time to 4,000 x g for 30 min for each sample collection step, we not only observed higher DNA recoveries of around 80 % as demonstrated below in *Figure 19*, but with concentrations high enough for IVT.

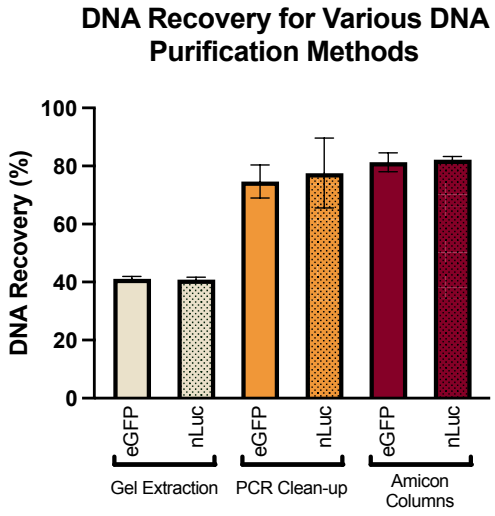


Figure 19: Histogram of the DNA template with pDNA origin purification recovery percentages (%) for both reporter constructs, i.e. nLuc and eGFP, for three different purification methods: gel extraction, PCR clean-up and Amicon column purification.
These results are based on nanodrop measurements of the post-purification yield versus the pre-purification yield to obtain the respective recovery %, starting with 10 µg pre-purification yield.

After establishing the Amicon column-based purification protocol, we systematically compared the three purification methods mentioned above for both DNA constructs, i.e. nLuc and eGFP, to demonstrate its ability to be used for several gene sequence lengths. *Figure 19* shows the DNA template recovery percentages comparison for the three purification methods for nLuc and eGFP, based on yield recoveries between post and pre-purification of linearized plasmids. Recovery percentages are similar across both linearized pDNA constructs, demonstrating no effect on sequence length for the three purification methods. Moreover, the average recovery percentages obtained for the DNA template encoding eGFP are 41.1%, 74.64%, and 81.3% using gel extraction, PCR clean-up and Amicon column purification methods, respectively. Similarly, for the nLuc template, the average recovery percentages obtained are 40.85%, 77.55%, and 82.2%. Overall, gel extraction purification demonstrated lower DNA recovery compared to the two other

methods, while the Amicon column-based purification showed similar or higher recoveries compared to PCR clean-up purification.

Purity was also evaluated using Nanodrop absorbance ratios for quality control purposes. For the eGFP template the $A_{260/280}$ and $A_{260/230}$ average ratios obtained across various experiments were around 1.88 and 1.23 for gel extraction, 1.88 and 1.97 for PCR clean-up, and 1.85 and 1.93 for Amicon column-based purification. Similarly, for nLuc template the following $A_{260/280}$ and $A_{260/230}$ ratios were obtained on average: 1.87 and 1.69, 1.90 and 2.06, and 1.93 and 2.14, for the three purification methods respectively. As such, PCR clean-up and Amicon column-based purifications performed better overall compared to gel extraction. This validates the use of Amicon column ultra-0.5 centrifugal filter units with high-recovery regenerated cellulose for DNA template purification before IVT, such that the MWCO chosen depends on the restriction enzyme used. Here, 100 kDa was chosen as SpeI-HF has a MW of 21.7 kDa and HindIII-HF a MW of 34.9 kDa.

4.2.2. Template DNA Purification: PCR generated DNA Template

To establish the reference analytical method for PCR generated DNA template, Nanodrop and Picogreen analytical assay results were compared in their ability to measure purified DNA content originally produced using PCR. It is critical to note that quantification of unpurified PCR products was not done using Nanodrop as primers and unused nucleotides can interfere with the absorbance reading.

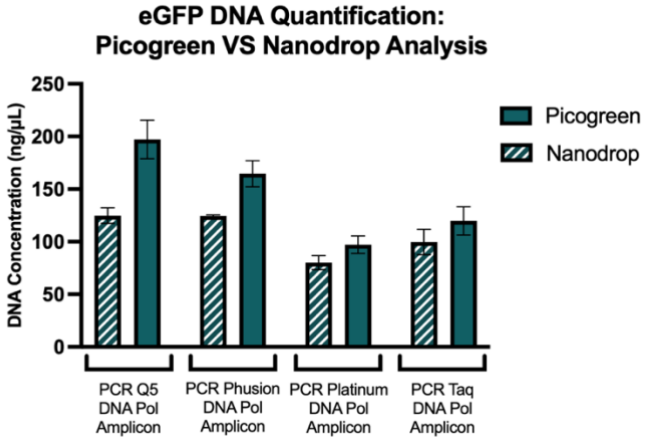


Figure 20: Comparison of analytical methods for PCR generated DNA content quantification. These results were measured post purification of the PCR products via PCR Clean-up. eGFP is used as the benchmark sequence for this analysis and has been produced via PCR using four different DNA Pol: Q5 HF, Phusion HF, Platinum Taq HF, and recombinant Taq.

Moreover, *Figure 20* illustrates the comparative results of using Picogreen or Nanodrop for the quantification of purified PCR products produced using four different DNA polymerases, using eGFP amplicons as the benchmark between both assays. It was shown that relative to the Picogreen concentration measurements, the Nanodrop readings are consistently lower, by around a factor of 0.76 on average. For future analytical considerations, Picogreen analysis was used for PCR generated DNA template, as unpurified PCR amplicons cannot be measured using Nanodrop.

For PCR generated DNA templates, only PCR clean-up and Amicon column-based purification methods were compared, as they outperformed gel extraction method for purification of linearized pDNA. First, PCR clean-up was used to purify nLuc and eGFP amplicons that were generated via PCR respectively, using four different DNA Pol, i.e., Q5 HF, Phusion HF, Platinum Taq HF, and recombinant Taq. *Figure 21* illustrates the PCR DNA template yields pre- and post- PCR clean-up purification, as well as the computed average recovery percentage for each DNA Pol for both sequences. Specifically, for amplicons produced using Q5 HF DNA Pol, PCR clean-up recoveries were overall higher than for other DNA Pol used, for both nLuc and eGFP.

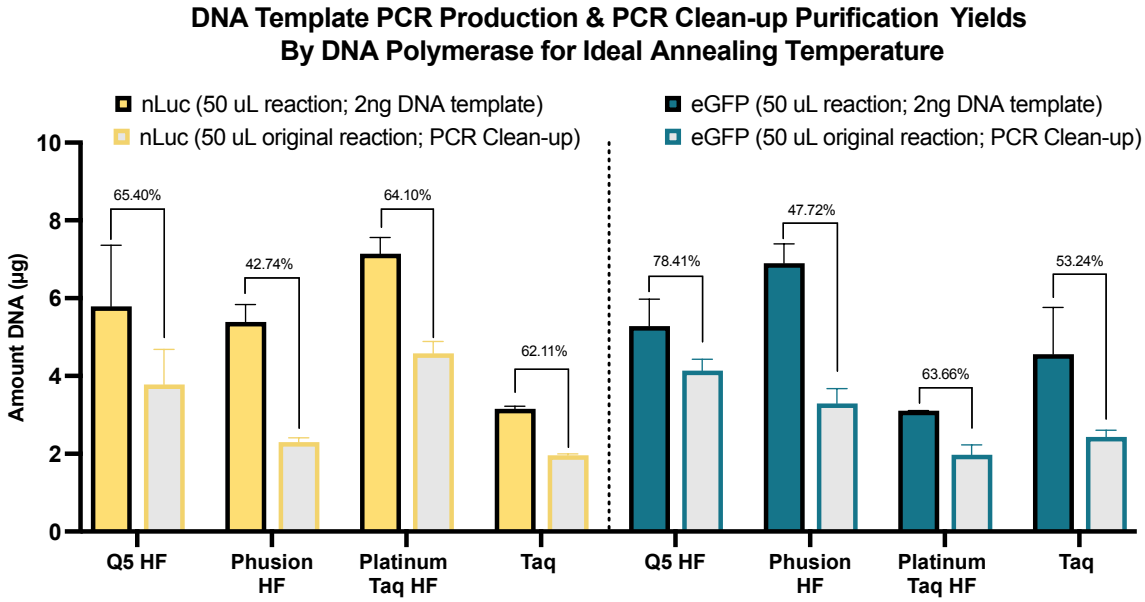


Figure 21: DNA Template PCR production and PCR clean-up purification yields by DNA Pol for ideal T_a previously determined, for 50 µL PCR reactions and for both reporter constructs. Average computed purification recovery percentages for each DNA Pol are shown in the plot.
Data measured using Picogreen analysis.

The highest purification recovery percentage, 78.41 %, was obtained for the eGFP sequence produced with Q5 HF, such that Q5 PCR DNA purification recovery was sequence dependent. In the cases of Phusion HF and Platinum Taq HF, performance of PCR clean-up purification was not sequence dependent, with averages of 42.74 % and 64.10 % respectively for nLuc and 47.72 % and 63.66 % respectively for eGFP amplicons. Additionally, Nanodrop data analysis was conducted post- PCR clean-up purification for DNA sequence purity, determined by the $A_{260/280}$ ratio. On average, the following $A_{260/280}$ ratios were obtained: 1.83, 1.82, 1.76, and 1.84 for nLuc produced using Q5 HF, Phusion HF, Platinum Taq HF, and recombinant Taq, respectively. Similarly, for eGFP DNA templates, the average ratios obtained were 1.84, 1.88, 1.84, and 1.87, respectively. The ratios were all within the desired range of 1.8 - 2.0, confirming purity.

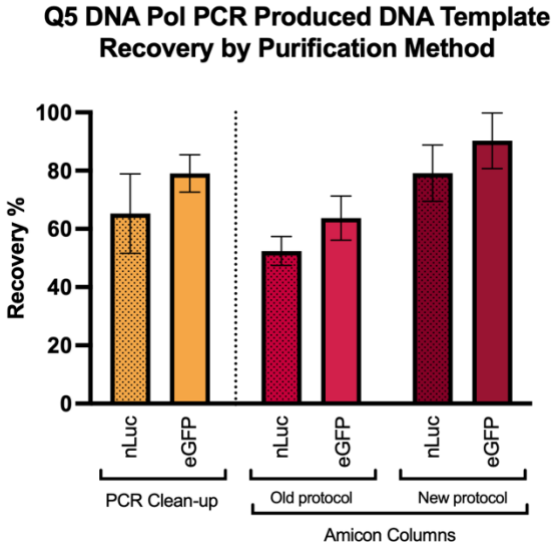


Figure 22: Histogram of the purification recovery percentages for both reporter DNA template produced using Q5 HF DNA Pol through PCR, for two different purification methods: PCR clean-up and Amicon column purification.
Amicon columns old protocol: buffer pre-washing followed by 3 concentrations, 4,000 x g for 30 min each, 100 kDa. Amicon columns new protocol: buffer pre-washing followed by 2 concentrations, 14,000 x g for 20 min each, 30 kDa. These results are based on Picogreen measurements of the post-purification yield versus the pre-purification yield to obtain the respective recovery %.

Amicon column-based purification was then conducted for both reporter sequences produced using Q5 HF DNA Pol. *Figure 22* shows the PCR produced nLuc and eGFP templates recovery percentages comparison for the two purification methods, based on yield recoveries between post and pre-purification, measured using Picogreen dsDNA analysis. Following the Amicon-based purification protocol designed for linearized pDNA template previously, the average recovery percentages obtained were 52.4 % and 63.7 %, for nLuc and eGFP templates respectively. Thus, optimization of the Amicon-based purification protocol of PCR products was conducted to increase these recoveries. By increasing the centrifugation speed from 4,000 x g for 30 min to 14,000 x g for 20 min for each concentration steps, we were able to increase the average recovery percentages to 79.1 % and 90.3 % for nLuc and eGFP respectively. Amicon column-based

purification performed overall similarly or better than the PCR clean-up purification of PCR DNA products based on yields. Additionally, it can be observed that for both purification methods, the purification recoveries were observed to be higher for the eGFP template than the nLuc template, a sequence dependency that was not observed with the purification of linearized pDNA template (*Fig. 19*).

4.2.3. Comparison of Purification Methods for Different Origin DNA templates

Overall, different purification methods were compared based on purification recovery and sample purity for two reporter constructs, nLuc and eGFP, produced using either *E. coli* fermentation or PCR. In summary, it was shown that the PCR clean-up and Amicon column-based purification protocols performed well for both types of DNA templates, with high recovery percentages – for both linearized pDNA template and PCR amplicons (*Fig. 23*). As such, the Amicon column-based purification was proven to be suitable for the preparation of DNA templates for use in mRNA transcription.

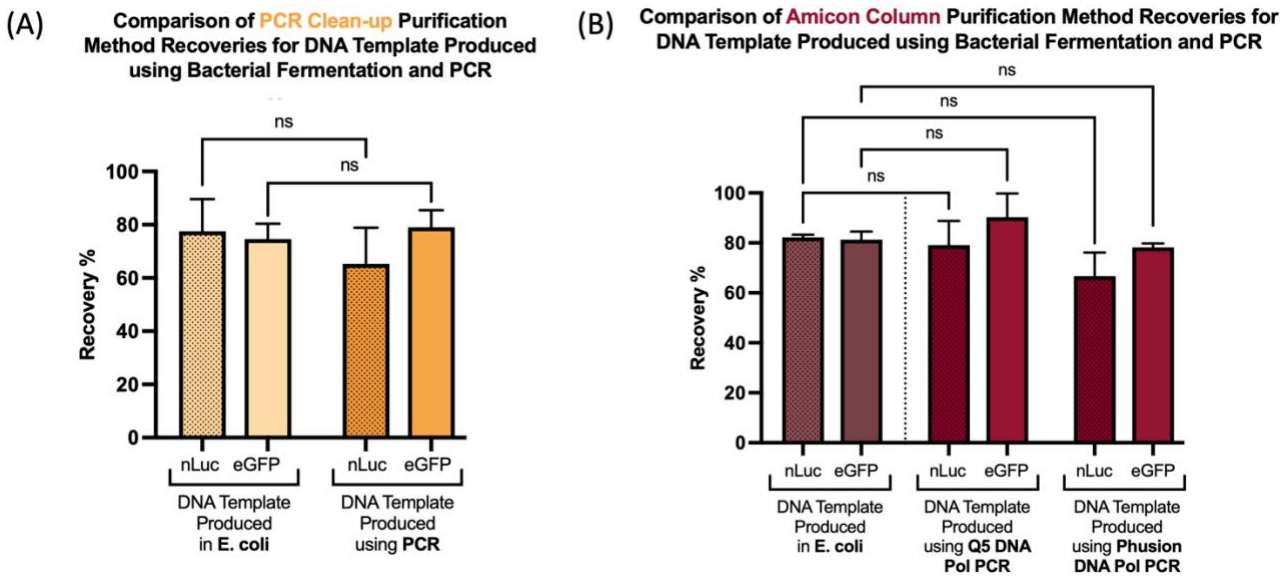


Figure 23: Summary of the purification recovery percentages by purification method for nLuc and eGFP DNA template sequences: (A) Comparison of PCR clean-up purification recovery percentages for DNA template produced using bacterial fermentation and PCR, (B) Comparison of Amicon based-column purification recovery percentages for DNA template produced using bacterial fermentation and PCR. Significance analysis performed using ordinary one-way ANOVA on GraphPad Prism software.

4.3. mRNA Transcription using Prepared DNA Template: IVT Reaction

Finally, after the DNA has successfully been generated and purified, it can be used as a template for mRNA synthesis. Herein, we assess the effect of the chosen DNA template purification method on the overall mRNA yield and purity.

First, DNA templates originating from *E. coli* fermentation and purified using one of the three different methods discussed above were used as starting material for IVT reactions. Results depicted in *Figure 24* have been obtained following mRNA synthesis, phenol chloroform purification, and Picogreen quantification. Starting all IVT reactions with 1 µg of DNA template, it was shown that the total mRNA yield varied based on the purification method used to prepare the DNA template. *Figure 24(A)* demonstrates that higher IVT yields were obtained when using a DNA template that has been purified using PCR-clean up or Amicon columns as opposed to that obtained from gel extraction. Moreover, mRNA productions also gave higher yields for the eGFP template compared to the nLuc template, which is hypothesized to be due to the sequence size difference. Specifically, it was shown that similar (i.e. nLuc) or higher (i.e. eGFP) mRNA productions were obtained for the IVT reactions where the DNA template was purified using Amicon columns compared to PCR clean-up.

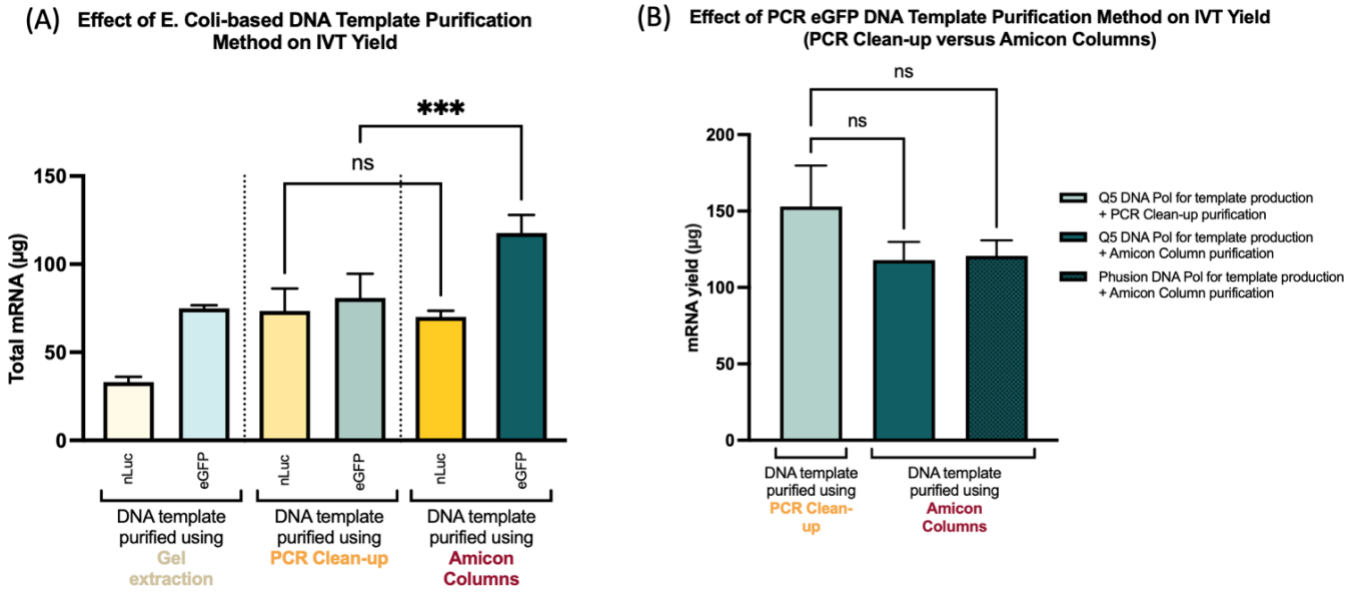


Figure 24: IVT yields depending on DNA template origin and purification method used.
(A) mRNA yields of IVT experiments using *E.coli*-based DNA template purified using Gel extraction, PCR Clean-up or Amicon columns, (B) mRNA yields of IVT experiments using PCR based DNA template

purified using PCR Clean-up or Amicon columns. IVT reactions were conducted starting with 1μg of DNA template. Data obtained using Picogreen and following phenol chloroform purification of the obtained mRNA. Significance analysis performed using ordinary one-way ANOVA on GraphPad Prism software.

For proof of concept, a similar comparison study was conducted for IVT reactions that used PCR generated and purified DNA templates instead of *E. coli*-based templates, specifically for the eGFP sequence. *Figure 24(B)* illustrates these results where no significant differences in yield were observed between IVT reactions that used PCR generated DNA template purified through PCR clean-up or Amicon columns. Overall, we can observe that the eGFP mRNA yield for PCR Clean-up DNA template was almost twice as high for PCR originating template rather than the *E. coli*-based one. Total mRNA yields obtained were on average 153.0 μg and 80.9 μg, respectively. On the other hand, Amicon column-based purification of the DNA template landed similar mRNA yields for both PCR and *E. coli*-based DNA templates, which further validates the use of Amicon columns for DNA purification. The total mRNA yields obtained were on average 117.7 μg for *E. coli* Amicon purified template, and 117.9 μg and 120.8 μg for Q5 and Phusion PCR Amicon prepared templates, respectively.

mRNA characterization was then conducted using formaldehyde gel electrophoresis, as shown in *Figure 25*. Purity and identity of the mRNA obtained through IVT reactions that used *E. coli*-based DNA templates were confirmed in *Figure 25(A)* and *(B)*, for both the eGFP and nLuc sequences respectively. Correct sequence lengths were obtained with the main band right below the 1 kb marker for eGFP mRNA and around 2 kb for the nLuc mRNA, as expected. More specifically, it can be seen that the mRNA produced with *E. coli*-based DNA template purified with gel extraction showed increased degradation (*Fig. 25(A)* lines 2-4 & *25(B)* lines 2-4) compared to when purified with PCR clean-up (*Fig. 25(A)* lines 5-7 & *25(B)* lines 5-7). Moreover, comparing the mRNA purity between PCR clean-up and Amicon purification DNA template, showed similar quality between the two experiments. Purity was evaluated using the Bioanalyzer profile, showing a clear and defined band and peak close to 2 kb marker, as anticipated (*Figure 25(C)* and *(D)*). Similarly, identity of the mRNA obtained through IVT reactions that used PCR-based DNA templates were confirmed in *Figure 25(E)* and *(F)*, with correct band lengths obtained in all cases.

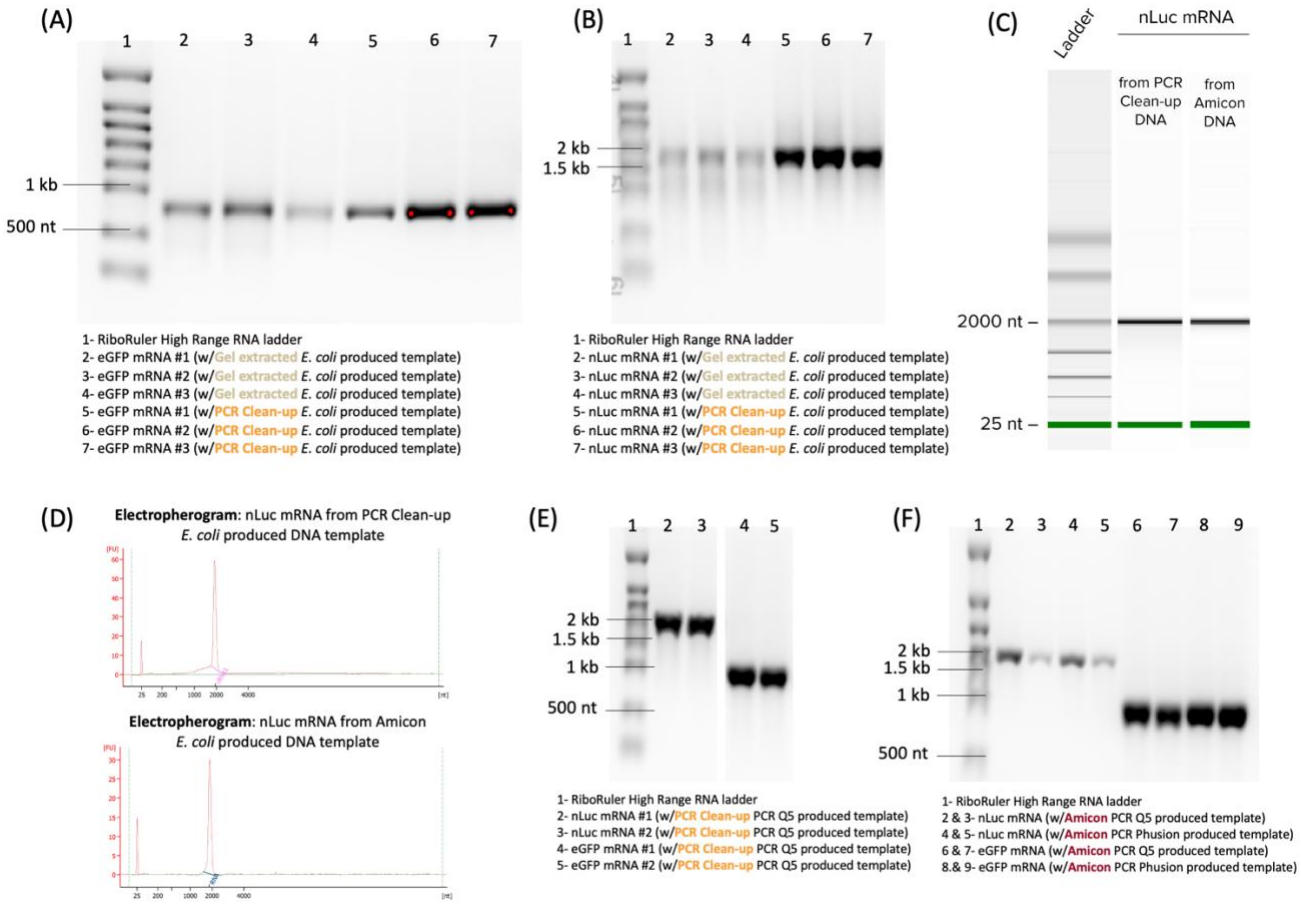


Figure 25: Characterization of nLuc and eGFP mRNA identity.
(A) mRNA formaldehyde gel electrophoresis of eGFP mRNA produced through IVT using *E. coli* generated DNA template purified using gel extraction or PCR Clean-up, (B) Same as (A) but for nLuc mRNA, (C) Bioanalyzer gel results of nLuc mRNA produced through IVT using *E. coli* generated DNA template purified using PCR Clean-up or Amicon columns, (D) Corresponding Bioanalyzer electropherograms, (E) Formaldehyde gel electrophoresis of nLuc and eGFP mRNA produced through IVT using Q5 PCR generated DNA template purified using PCR Clean-up, (F) mRNA gel electrophoresis of nLuc and eGFP mRNA produced through IVT using Q5 or Phusion PCR generated DNA template purified using Amicon columns.

5. Discussion and Conclusion

Owing to its cell-free production and fast reaction time, mRNA manufacturing is considered highly advantageous as compared to other more complex vaccine and therapeutic manufacturing processes [15]. However, plasmid production still represents the most time-consuming step of this process, where *E. coli* fermentation can take up to several days/weeks in comparison to the mRNA synthesis and its associated purification steps, which are generally completed within several hours to two days. More particularly, DNA linearization, which is needed when using pDNA sources as IVT DNA templates, constitutes one of the main cost contributors amongst the major operations for mRNA therapeutics and vaccines (*Appendix A Table B* and *Figure A*). Despite these hurdles, bacterial fermentation is still used as the gold standard technique for pDNA generation in both academic labs and industrial settings. As such, *E. coli* fermentation was first used in this study as our reference DNA template production method. DH10Bac was selected as our working *E. coli* strain due to its ease of transformation, fast growing, and good associated pDNA yields, as demonstrated. In fact, DH10Bac corresponds to a modified version of the DH10B strain, and both the DH5α and DH10B *E. coli* strains have similar phenotypes – *endA* and *recA* negative – preventing plasmid degradation, making them typical hosts used for pDNA production [148]. Furthermore, we demonstrated similar pDNA yields for the *E. coli* NEB Stable competent strain, which is suitable for high efficiency transformation, recommended to us by a representative of BIOVECTRA in the context of mRNA manufacturing [132]. Future studies in bacterial fermentation for plasmid production could examine alternative methods for transformation and bacterial growth of *Vibrio Natriegens* strains to alleviate bacterial culture time and take advantage of its fast-doubling time.

Despite the existing foundation of knowledge and regulatory approvals surrounding *E. coli*-based plasmid production, the ongoing surge of interest and research in cell-free, enzymatic or synthetic production of pDNA represents a promising avenue for future manufacturing of these molecules. This would have a strong impact, not only in the context of the mRNA manufacturing platform, but also in the context of DNA-based vaccines, as well as gene and cell therapy processes [149]. In this thesis, PCR was shown to be an attractive potential alternative method for DNA template generation for use in mRNA transcription. To our knowledge, no direct and systematic comparison of the DNA template origins for IVT was previously recorded. Therefore, we developed a straight-

forward and time-reductive workflow for the production of DNA template for IVT using PCR. A comparative table of our results is show in *Table 3*, which illustrates the advantages of using PCR over *E. coli*-based DNA templates from a manufacturing perspective.

	PCR Generated DNA Template	*E. coli* Generated DNA Template
Production Time	1-2 hours	3-7 days
Linearization Step	✗	✓
Production Yields *(nLuc example)*	Q5 DNA Pol: ~5-6 µg per 50 µL reaction Phusion DNA Pol: ~6 µg per 50 µL reaction → *Easily scaled out (parallelized)*	Mini Prep High Copy Plasmid: DH10Bac: ~20 µg per 50 µL collection NEB Stable Competent: ~20 µg per 50 µL collection
Purification Recoveries	PCR Clean-up: 65.3 % (nLuc); 79.0 % (eGFP) Amicon columns: 79.1 % (nLuc); 90.3 % (eGFP)	Gel extraction: 40.8 % (nLuc); PCR Clean-up: 77.55 % (nLuc); 74.6 % (eGFP) Amicon columns: 82.2 % (nLuc); 81.3 % (eGFP)
Comparable Cost (per µg of purified DNA):	$1.04 CAD	$ 1.752 CAD *(Linearization enzymes: most expensive)*

Table 3: Overall Comparative Table between PCR and *E. coli* generated DNA template. Average comparable costs obtained from *Appendix A Table B* Analysis.

We demonstrated that the choice of DNA Pol has an impact on the sequence synthesis yield. Each DNA Pol is characterized by a set of properties, such as its fidelity, 3'→ 5' exonuclease, 5'→ 3' exonuclease, etc. that also depend on the reaction conditions for various applications, such that Q5 DNA Pol can be used for long range PCR for example. We also demonstrate that careful considerations should be taken when establishing PCR reaction parameters, as they are dependent on various factors including but not limited to the DNA Pol used, the designed primers, and the sequence to be amplified (i.e., length, nucleotide composition, etc.). The use of defined enzymes could help in the standardization of the upstream process, from the PCR reaction to the purification step. As PCR generates amplicons containing all the sequence elements discussed in the literature review section of this thesis, no linearization step is required, making thus purification a unique step. We also compared different purification methods – two lab-scale and one TFF scale-down technique using Amicon columns – for different DNA template origins and demonstrated similar or higher purification recoveries for the established Amicon columns-based purification method when compared to PCR clean-up, while highlighting good yields on both ends. Additionally, sequence dependency was shown for purification of PCR products illustrated by higher recoveries for shorter sequences, a trend that was not observed in the purification of linear DNA template

originating from pDNA. Despite this result, sequence length does not constitute a fundamental limitation to the use of PCR for DNA template generation but should rather be taken into consideration when setting the production and purification workflow to maximize end-yields. In fact, as shown in *Appendix A Table A*, long range PCR can generate sequences of up to 20 kb, covering the typical length of therapeutic mRNAs. The Amicon column-based purification method was proven advantageous at lab-scale and could be implemented at large-scale using TFF. On the other hand, low recoveries were obtained when purification was carried using gel extraction, which can be due to multiple factors such as: long exposure to UV light, uncomplete melting of agarose plug, introduction of more salts and other components to the purification when the agarose plug was not small enough, etc. Moreover, PCR production yields can be scaled out by running multiple 50 µL reactions in parallel – with the only limitations being the number of PCR tubes a machine can hold, the number of PCR machines available, as well as the overall costs of enzymes. As shown in *Appendix A Table A*, DNA yields are highly dependent on the scale of operation. Around 5 µg/mL of culture for lab-scale production of high copy pDNA to around 2.1 g/L for a 300 L production, and around 50-200 ng/µL for PCR reactions which can be parallelized for higher yields. Finally, the resulting SDT-mRNA was of a comparable quality with that of the PDT-mRNA, as shown through the quality control conducted via formaldehyde gel electrophoresis. We also showed that the DNA template origin did not impact the mRNA synthesis yield, using the eGFP construct as our benchmark sequence for comparison purposes. In a therapeutic context, *in vitro* then *in vivo* studies should be conducted to evaluate respective protein expression and function. For example, de Mey et al. demonstrated that their produced SDT-mRNA was comparable to their produced PDT-mRNA, showing their strong potential for use for vaccine development and immunogenicity screening [42].

Having established a foundational protocol for both bacterial amplification of DNA templates as well as their synthetic production and purification, future work can focus on the establishment of additional analytical techniques to ensure the quality of these templates. Specifically, monitoring of the poly(A) tail remains a challenge within the field. Recombination of homopolymeric poly(A) stretches in plasmid vectors still constitutes a major issue in poly(A) tail conservation, such that it has been shown by Trepotec et al. that using a segmented poly(A), such as the one in our nLuc reporter sequence, helps reduce such recombinations [150]. Throughout the progression of our study, we examined several methodologies to assess the length of the poly(A) tail, as shown in

Appendix B. Overall sequence conservation, except for the poly(A) region, was observed for both nLuc and eGFP DNA templates produced via *E. coli* or PCR. Sequencing analysis was conducted using various methods such as Sanger, Nanopore and PacBio sequencing (*Appendix 1B*). However, we were met with conflicting results regarding poly(A) tail length (*Appendix 2B*), which would require additional optimization for future progress. Alternative methodologies can include synthetic approaches for the production of poly(A) tails, followed by their ligation to the DNA template, as described by Graham et al. from Elegen (DNA synthesis company based in California, U.S.) [151]. Moreover, other methods have been proposed such as the one described by Arbuthnot et al., where they developed a method that uses type IIS restriction enzymes to build a poly(A) tract via repeated restriction digestion followed by their ligation to lengthen the poly(A) tail and circular plasmid propagation [152]. Additionally, for platform technologies, the inclusion of unique restriction sites on either side of the poly(A) tail would facilitate the analysis of this DNA segment. The excised poly(A) tail would then be easily analyzed through the fragment analyzer, gel electrophoresis or other analytical technologies.

In our study, we used reporter constructs which were provided to us and unmodified by us to establish the project. The results obtained for those sequences could be generalized for different mRNA templates of similar length; efficiency of the established process should be validated for longer sequences. However, for the implementation of a singular flexible plasmid design, which can be applied to produce several mRNAs for various applications, the template can be designed to account for these necessary analytical assays. In this study, pDNA was used as the origin sequence for later PCR amplification. Though this represents a step to reduce current manufacturing times, the next step would be to implement the use of a fully synthetic DNA template. As such, DNA synthesizers could be used to generate pre-designed template fragments sequences to be assembled and later amplified via PCR following the developed workflow. Typically, DNA synthesizers can produce DNA sequences of around 100 bp [153]. For example, de Mey et al. successfully used a synthetic DNA template for IVT, such that the SDT was generated using three synthetic oligonucleotides designed to hybridize together during assembly PCR, to then be amplified by PCR before purification [42].

To conclude, the aims of this research project to establish and optimize the DNA template production and purification steps for use in IVT have been fulfilled. The major potential of PCR

as the method of choice for DNA synthesis was successfully demonstrated. As such, the use of enzymatically-produced synthetic DNA templates has the potential to overcome mRNA manufacturing concerns. In fact, pDNA production involving bacterial fermentation is not only subject to large-scale bioreactors batch failures but is also associated with a large footprint and large associated upstream costs and time. The purification process that follows is similarly slower and more expensive as it is done pre- and post-linearization of the sequence of interest. Thus, synthetic DNA approaches are highly strategic from a manufacturing standpoint due to their ease of generation, fast production, and scalability, all of which are favorable for rapid tech transfer.

6. Appendix

APPENDIX A: Evaluating DNA Template Generation Techniques: PCR versus Bacterial Fermentation, in terms of production process and associated costs.

APPENDIX A Table A: Production Process Comparison of PCR generated DNA template and *E. coli* generated DNA template (PDT).

	PCR Generated Template	*E. coli* Generated Template
Reaction volume/ Throughput	• 10-100 µL (fast and homogeneous cooling/heating is hard to achieve for higher volume) • 96 reactions can run in parallel	> 1000 L (bioreactor production) [1]
Starting DNA content	1 pg - 100 ng (depending on PCR reaction volume)	10 pg - 100 ng (depending on the amount of bacterial cells for transformation)
DNA yield	Several µg (around 50-200 ng/µL) Difficult to predict: depends on the size of the amplicon, design of the primers, starting amount of template and primers, amplification efficiency, reaction volume, numbers of PCR cycles, etc.	Lab Scale (Mini Prep): • High Copy Plasmid: ~5 µg/mL of culture • Low Copy Plasmid: ~2 µg/mL of culture High Yield Optimized Industry Scale: • 2.1 g/L for a 300 L production [2] • 1.58- 2.96 g/L in 6 L production [3] • 1-1.5 g/L for processes ~1000 L [4]
Maximum sequence length	• Standard PCR (less than 5 kb) [5] • Long range PCR: Sequences between 5 kb - 20 kb [5, 6]	Plasmid size can be 10s of kilobases long Examples in literature: • 16.8 kb plasmid successfully amplified in *E. coli* • 21.4 kb plasmid in *E. coli* • 205 kb plasmid transformed into *E. coli* and isolated [7]
Expected error rates	Error rate depends on PCR conditions such as DNA polymerase, amplicon size, number of PCR cycles, etc. Examples in literatures: [8] • *Taq* DNA polymerase I: 1×10^{-5} to 2×10^{-4} errors/base/doubling • Q5 substitution error rate: 5.3×10^{-7} ($\pm 0.9 \times 10^{-7}$) • Phusion substitution error rate: 3.9×10^{-6} ($\pm 0.7 \times 10^{-6}$) • Long range error rate (error per base pair): *Taq* blend: $8\text{-}14\times10^{-6}$ and *Pfu* fusion: 1.3×10^{-6} [9]	$1x10^{-9}$ errors per base pair errors per replication [10]
Amplification time required	• For standard PCR: 1-2 h • For long range PCR: ~ 6 h	• Several days-weeks
Downstream processing	• Required for removal of primers and loose nucleotides (reaction mixture components)	• Required for extraction from cells • Required to remove cell debris, host cell proteins, etc.
Linearization	Not Required	Required

Appendix A Table B: **Cost Analysis Comparison of PCR versus *E. coli* fermentation for DNA template generation**

This cost analysis does not include the cost of consumables, which may impact the result of the analysis. This estimate was completed using available listed prices of raw materials on common supplier websites. Prices for large scale and bulk orders may vary with these suppliers.

PCR Generated Template		***E. coli* Generated Template**	
Item/Reagent	**Anticipated Cost**	**Item/Reagent**	**Anticipated Cost**
Fixed Costs (FC)		**Fixed Costs (FC)**	
PCR Thermocycler	*Example:* • C1000 Touch™ Thermal Cycler with 96–Deep Well Reaction: List price $12,930 CAD	**Shaker Incubator**	*Example:* • Thomas Scientific: MaxQ Stackable Incubated Shaker: List Price ~$24,000 CAD
Micro-centrifuge	Approximately $4,000 CAD *Example*: • Centrifuge 5425/5425 R - Microcentrifuge: $4,407 CAD	**Dry Incubator**	Approximately $ 1,500 CAD *Example:* • Thermo Scientific Heratherm IMC18 Incubator: $1,524 CAD
		Large Centrifuge (Purification)	Approximately $50,000 CAD *Examples:* • Beckman Coulter Avanti J-E Centrifuge: $41,332 CAD [7] • Thermo Scientific™ Sorvall BIOS A Centrifuge: $61,700 CAD [8]
Variable Costs (VC)		**Variable Costs (VC)**	
DNA Polymerase	• Q5® Hot Start High-Fidelity DNA Polymerase 500 units (2 U/µL): $708.30 [1] o $ 0.7/50 µL reaction • Phusion™ High-Fidelity DNA Polymerase 500 units (2 U/µL): $692.00 [2] o $ 0.69/50 µL reaction • Taq DNA Pol recombinant 500 units (5 U/µL): $210 [3] o $ 0.525/50 µL reaction	**Bacterial Cells**	• 20-50 µL of competent cells for bacterial transformation *Examples:* • NEB® Stable Competent *E. coli* (20 x 0.05 ml): $510.00 CAD [9] o Assuming 30µL of cells used per prep: ~$15.30 CAD/Giga prep • One Shot™ TOP10 Chemically Competent *E. coli* (21 x 50 µL/tube): $680.00 CAD [10] o Assuming 30µL of cells used per prep: ~$19.43 CAD/Giga prep
Nucleotides (dNTPs)	• 10mM: 1mL stock, 842.00CAD [4] o $0.842 CAD /50µL reaction	**Medium**	For Plates: • Luria Broth Base (Miller's LB Broth), powder 2.5 kg: 522.00 CAD [11] o Price per plate: ~ $0.261 CAD • Agar powder 2.5 kg: 838.00 CAD [12] o Price per plate: ~ $0.201 CAD For Bacterial culture: • Medium estimated cost: $0.2-0.35 CAD/L [13] • Assuming 250 mL culture for a Giga Prep: ~$0.05-0.09 CAD/Giga Prep

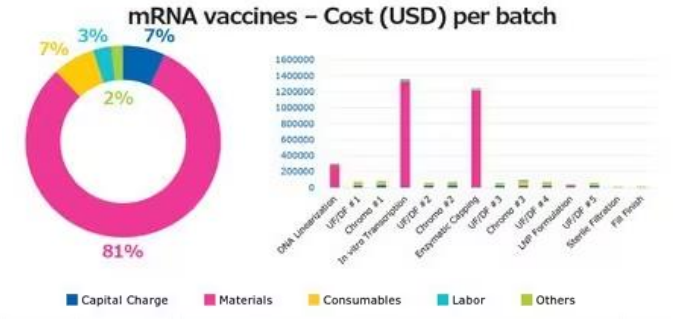
Primers (Forward and Reverse)	• 25 nmol: 3.60 CAD
	  ○ Cost per 50µL PCR reaction: ~$0.004 CAD/ 50µL reaction
		Antibiotic	For Plates:
			• Ampicillin powder(5g): $119.69CAD [14]
			  ○ Working concentration: 100µg/mL
			  ○ Price per plate: ~ $0.0239 CAD
			• Kanamycin powder(5g):$116.68CAD [15]
			  ○ Working concentration: 50µg/mL
			  ○ Price per plate: ~$0.0116 CAD
			For bacterial culture:
			• Ampicillin: Assuming 250 mL/Giga Prep, 25 µg required: $0.598 CAD
			• Kanamycin: $0.2917 CAD/Giga Prep

Reaction Buffer	*Examples:*
	• Taq DNA Polymerase PCR Buffer (10X) 2mL: $99.25 CAD [5]
	  ○ ~ $0.25/ 50µL reaction
	• Phusion HF Buffer Pack 6mL: 46.62 CAD [6]
	  ○ ~ $0.077/ 50µL reaction
		Linearization Enzyme	*Example:*
			• SpeI-HF: 2500 units, 440 CAD [16]
			  ○ 10 units of enzyme/ µg of DNA: ~$1.75CAD/ µg DNA
			  ○ Assuming 10 mg of DNA per Giga prep:[17] ~$17,500CAD/10mg DNA

Purification	Lab Scale:
	• QIAquick PCR Purification Kit (250 preps): 640 CAD
	  ○ $2.56 CAD/ prep
	• NucleoSpin® Gel and PCR Clean-Up (250 preps): 607.27 CAD
	  ○ $2.43 CAD/ prep
	Industrial scale purification was not explored for this analysis.
		Purification	Lab Scale:
			• 164$/ sample for each Giga Prep [18]
			Large Scale (Industrial scale) purification was not explored for this analysis.

TOTAL VC	**Comparable cost (per µg of purified DNA): $1.04 CAD** *Assumptions*: • All costs calculated per one 50µL reaction (unpurified): $4.15 CAD • Each 50µL reaction produces 5µg (C=100 ng/µl) • Purification recovery %: typically, 75−95 % (* written on the kit, optimistic): consider: 80%	**TOTAL VC**	**Comparable cost (per µg of purified DNA): $ 1.752 CAD** *Assumptions:* • All costs were calculated for bacterial culture for 1 Giga Prep • Assumed 1 Giga Prep will produce approximately 10 mg of DNA (as estimated by the manufacturer)

APPENDIX A Figure A: Taken from SigmaAldrich – Cost distribution among major unit operations for mRNA vaccines [1]



APPENDIX B: Evaluating Sequencing Methods for DNA Template Quality Control

APPENDIX 1B: Assessment of the DNA sequenced reads except for the poly(A) sequence.

(i) Plasmid DNA sequences produced using *E. coli* Fermentation:

Evaluation of plasmid sequence conservation through *E. coli* fermentation was done by Professor Ioannis Ragoussis' Lab using de-novo assembly of the sequenced reads into groups of similar molecules, using two different sequencing platforms: Nanopore and PacBio sequencing platform. The following results were obtained:

- Nanopore reads (most abundant assembled sequence), for peGFP and pnLuc:
    - Agreement with reference sequence, except the poly(A) region and nanopore adaptor overhang of reference.

*It would be recommended to cut the plasmid with a restriction site that is not right after the pol(A) tail, as is the case of SpeI and HindIII (for eGFP and nLuc respectively), for sequencing purposes: as the nanopore adaptor sequence was usually trimmed away and therefore, there is not an overhang of a constant sequence to define the boundaries.

- PacBio reads (most abundant assembled sequence), for peGFP and pnLuc:
    - Agreement with reference sequence, except the poly(A) region.

(ii) Amplicon DNA sequences produced using PCR:
Note: PCR initial DNA input was the pDNA sequence produced using E. coli fermentation

PCR error rates, based on the number of mismatches and gap/insertion errors, were determined by Sanger sequencing for both nLuc and eGFP amplicons produced using Q5 Hot Start HF. The average values for 3 different amplicons are reported in the table below. Similarly, the PCR error rates for both reporter sequences for a range of DNA Pol were also determined by Nanopore sequencing, such that the results are also reported in the table below.

Amplicon	DNA polymerase	Total # bp	# mismatches	# gap/insertion	Error Rate
nLuc	Q5 High-Fidelity	2172	1 (± 1)	1.3334 (± 0.5774)	1.1667×10^{-3} ($\pm 5.7735 \times 10^{-4}$)
nLuc	Q5 High-Fidelity	2172	0 (± 0)	0 (± 0)	0 (± 0)
eGFP	Q5 High-Fidelity	1193	0 (± 0)	1.6667 (± 0.5774)	1.1667×10^{-3} ($\pm 5.7735 \times 10^{-4}$)
eGFP	Q5 High-Fidelity	1193	0 (± 0)	0 (± 0)	0 (± 0)
nLuc	Phusion HF	2172	0 (± 0)	0 (± 0)	0 (± 0)
eGFP	Phusion HF	1193	0 (± 0)	0 (± 0)	0 (± 0)
nLuc	Platinum HF	2172	0 (± 0)	0 (± 0)	0 (± 0)
eGFP	Platinum HF	1193	0 (± 0)	0 (± 0)	0 (± 0)
nLuc	Taq Pol	2172	0 (± 0)	0 (± 0)	0 (± 0)
eGFP	Taq Pol	1193	0 (± 0)	0 (± 0)	0 (± 0)

*Sanger Sequencing was done with triplicates (in Blue)
*Nanopore Sequencing was done with triplicates (in White)

- Nanopore sequencing is known to have higher sensitivity compared to Sanger sequencing
- Nanopore sequencing outperforms Sanger sequencing in terms of length

Sanger sequencing of PCR amplicons was done at Genome Quebec to verify PCR accuracy. To sequence nLuc amplicon, the following primers were used: at the extremities of the amplicon, forward (5' TGTACGGGCCAGATATACGC 3') and reverse (5' GCAACTAGAAGGCACAGT CG 3'), and in the middle of the amplicon, forward (5' AAAGAAGTCGCGGAAGTGGC 3') and reverse (5' GGCCGAAGCCACAGGAGATA 3'). Similarly, for eGFP amplicon, the following primers were used: at the extremities of the amplicon, forward (5' TTGTGTGGAATTGTGAGC GG 3') and reverse (5' GTCGGGGCTG GCTTAACTAT 3'), and in the middle of the amplicon, forward (5' GGCACAAGCTGGAGTACAAC 3') and reverse (5' TTCTGCTTGTCGGCCATG AT 3'). To distinguish between Sanger sequencing errors and DNA Pol errors, 4 different primers were used, not only to ensure that the entirety of the amplicon was read, but to also ensure that both strands were sequenced. Error rate analysis includes bases confirmed on both strands. Nanopore sequencing was also conducted on PCR amplicons to confirm sequence accuracy, using Plasmidsaurus' services.

APPENDIX 2B: Assessment of the Poly(A) tail sequence.

(i) Poly(A) tail sequences of plasmid DNA produced using *E. coli* Fermentation:

Poly(A) Tail length of:	Expected Length	From Basecalled data of PacBio (median)	From Basecalled data of Nanopore (median)	From raw data of nanopore with Dorado v7 estimation (median)	From raw data of nanopore with in house estimation (median)
eGFP plasmid	173 bp	~ 61 bp	-	61 bp	57.09 bp
nLuc plasmid: first segment of poly(A)	24 bp	~ 32 bp	~ 28 bp	failed	32.36 bp
nLuc plasmid: second segment of poly(A)	34 bp	~ 71 bp	-	70 bp	69.3 bp
nLuc plasmid: full segmented poly(A)	68 bp**	-	-	119 bp	-

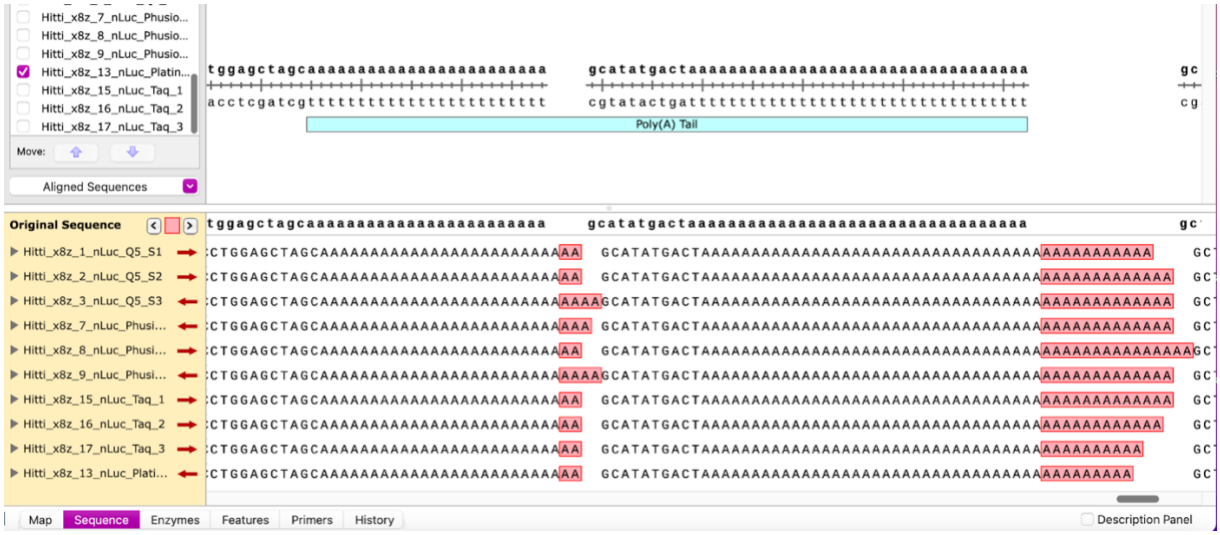
** nLuc segmented poly(A) sequence: AAAAAAAAAAAAAAAAAAAAAAAAGCATATGACTAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA

As the assembled plasmid sequences matched the expected one for both nLuc and eGFP plasmids, except for the poly(A) region, the poly(A) tail was treated separately as the assembly collapsed multiple reads into one representation that might or might not represent the true length. Data analysis and poly(A) distributions were conducted by Professor Ioannis Ragoussis' Lab (McGill Genome Center). Different sequencing methods were compared in their ability to sequence the poly(A) tail: (a) PacBio, (b) Nanopore, (c) Dorado v7.0.0 (new poly(A) tail estimation tool in the

nanopore basecaller that allows for the poly(A) tail length estimation from plasmids, given that we input the sequence before and after the region where we except the poly(A) to be), and (d) in-house estimation from nanopore date developed by Professor Ioannis Ragoussis' Lab.

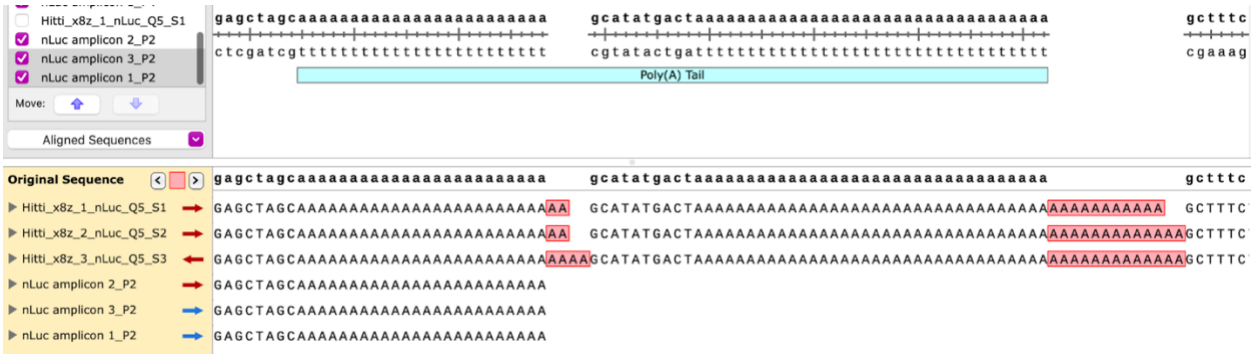
(ii) Poly(A) tail sequences of amplicon DNA produced using PCR:
Note: PCR initial DNA input was the pDNA sequence produced using E. coli fermentation

- Nanopore Sequencing results (Plasmidsaurus data, analyzed using SnapGene):
    - nLuc amplicon, produced using PCR with four different DNA Pol:
        - Reference sequence: Poly(A) tail: 68 bp, segmented
        - Nanopore results show insertions throughout the poly(A) tail – right after the first and second segmented regions.



    - eGFP amplicon, produced using PCR with four different DNA Pol:
        - Reference sequence: Poly(A) tail: 173 bp
        - Nanopore results show that poly(A) tail is shorter than expected, around 40 base pairs rather than 173 bp

- Sanger Sequencing results (Genome Quebec data, analyzed using SnapGene):
    - nLuc amplicon, produced using PCR with Q5 HF DNA Pol:
        - Reference sequence: Poly(A) tail: 68 bp, segmented
        - lines 1-3: nanopore sequencing; lines 4-6: sanger sequencing
        - Sanger sequencing results show that the poly(A) tail was not fully sequenced



    - eGFP amplicon, produced using PCR with Q5 HF DNA Pol:
        - Reference sequence: Poly(A) tail: 173 bp
        - lines 1-3: sanger sequencing; lines 4-5: Nanopore sequencing
        - Nanopore Sequencing stopped for poly(A) tail after 37 bp on average
        - Sanger Sequencing stopped for poly(A) tail after 62 bp on average

**Figures in 2B (ii) are screenshots taken from SnapGene when conducting 'Align to reference DNA sequence'.*

References for APPENDIX A

Table A

1. Chartrain, M., Bentley, L.K., Krulewicz, B.A., Listner, K.M., Sun, W.J. and Lee, C.B., Merck Sharp and Dohme LLC, 2011. Process for large scale production of plasmid DNA by E. coli fermentation. U.S. Patent 7,998,732. https://patents.google.com/patent/CA2554534C/en
2. 486. Rapid Deployment Plasmid Production Process, Combining Inducible High Yield Fed-Batch Fermentation Process with Novel Autolytic Plasmid DNA Purification", 2007, Molecular Therapy, vol. 15, pp. S187-S188. https://www.cell.com/molecular-therapy-family/molecular-therapy/fulltext/S1525-0016(16)44692-7
3. Carnes, A.E., Luke, J.M., Vincent, J.M., Schukar, A., Anderson, S., Hodgson, C.P. and Williams, J.A., 2011. Plasmid DNA fermentation strain and process-specific effects on vector yield, quality, and transgene expression. Biotechnology and bioengineering, 108(2), pp.354-363.
4. Chartrain, M., Bentley, L.K., Krulewicz, B.A., Listner, K.M., Sun, W.J. and Lee, C.B., Merck Sharp and Dohme LLC, 2011. Process for large scale production of plasmid DNA by E. coli fermentation. U.S. Patent 7,998,732. https://patents.google.com/patent/CA2554534C/en
5. Grunenwald, H., 2003. Optimization of polymerase chain reactions. PCR protocols, pp.89-99.
6. Long Range PCR, New England Biolabs. Available at : https://www.neb.com/en/applications/dna-amplification-pcr-and-qpcr/specialty-pcr/long-range-pcr
7. Islam, M.M., Odahara, M., Yoshizumi, T., Oikawa, K., Kimura, M., Su'etsugu, M. and Numata, K., 2019. Cell-penetrating peptide-mediated transformation of large plasmid DNA into Escherichia coli. ACS Synthetic Biology, 8(5), pp.1215-1218.
8. Potapov, V. and Ong, J.L., 2017. Examining sources of error in PCR by single-molecule sequencing. PloS one, 12(1), p.e0169774.
9. Hogrefe, H.H. and Borns, M.C., 2011. Long-range PCR with a DNA polymerase fusion. PCR Protocols, pp.17-23.
10. Lee, D.F., Lu, J., Chang, S., Loparo, J.J. and Xie, X.S., 2016. Mapping DNA polymerase errors by single-molecule sequencing. Nucleic acids research, 44(13), pp.e118-e118.

Table B

1. Q5® Hot Start High-Fidelity DNA Polymerase. New England Biolabs. Available online: https://www.neb.ca/m0493
2. Phusion™ High-Fidelity DNA Polymerase (2 U/µL). Thermo Scientific. Available online: https://www.thermofisher.com/order/catalog/product/F-530XL
3. Taq DNA Polymerase, recombinant (5 U/µL). Thermo Scientific. Available online: https://www.thermofisher.com/order/catalog/product/EP0401
4. dNTP Mix (10 mM ea). Invitrogen. Available online: https://www.thermofisher.com/order/catalog/product/18427013
5. *Taq* DNA Polymerase PCR Buffer (10X). Invitrogen. Available online : https://www.thermofisher.com/order/catalog/product/18067017?SID=srch-srp-18067017
6. Phusion HF Buffer Pack. Thermo Scientific. Available online: https://www.thermofisher.com/order/catalog/product/F518L?SID=srch-srp-F518L
7. Beckman Coulter Avanti J-E Centrifuge, 50/60 Hz, 200/208/240 V. Beckman Coulter 369001. Available online: https://www.fishersci.com/shop/products/avanti-je-208v-240v-50hz-60h/NC1587942
8. Thermo Scientific™ Sorvall BIOS A Centrifuges. Thermo Scientific™ 75007698. Available online: https://www.fishersci.com/shop/products/sorvall-bios-a-centrifuge-single-phase-1/75007698
9. NEB® Stable Competent *E. coli* (High Efficiency). New England Biolabs. Available online: https://www.neb.ca/detail.php?id=C3040
10. One Shot™ TOP10 Chemically Competent *E. coli.* Invitrogen. Available online: https://www.thermofisher.com/order/catalog/product/C404003
11. Luria Broth Base (Miller's LB Broth Base), powder. Invitrogen. Available online: https://www.thermofisher.com/order/catalog/product/12795027
12. Agar powder. Thermo Scientific Chemicals. Available online: https://www.thermofisher.com/order/catalog/product/A10752.0E

13. Cardoso, V.M., Campani, G., Santos, M.P., Silva, G.G., Pires, M.C., Gonçalves, V.M., Giordano, R.D.C., Sargo, C.R., Horta, A.C. and Zangirolami, T.C., 2020. Cost analysis based on bioreactor cultivation conditions: Production of a soluble recombinant protein using Escherichia coli BL21 (DE3). Biotechnology reports, 26, p.e00441.
14. MP Biomedicals™ Ampicillin, Sodium Salt. MP Biomedicals™ 0219452605. Available online: https://www.fishersci.ca/shop/products/mp-biomedicals-ampicillin-sodium-salt-3/ICN19452605
15. Kanamycin Sulfate (White Powder), Fisher BioReagents. Fisher BioReagents BP9065. Available online: https://www.fishersci.ca/shop/products/kanamycin-sulfate-white-powder-fisher-bioreagents/bp9065
16. SpeI-HF®. New England Biolabs. Available online: https://www.neb.ca/R3133
17. QIAGEN Plasmid Giga Kit. QIAGEN. Available online : https://www.qiagen.com/us/products/discovery-and-translational-research/dna-rna-purification/dna-purification/plasmid-dna/qiagen-plasmid-kits?catno=12191
18. Westman, David. “Plasmid Purification – Strategies and Costs.” *Biotage*, Biotage Sweden AB, 18 Jan. 2023. Available online: www.biotage.com/blog/plasmid-purification-strategies-and-costs

Figure A

1. Cost Modeling Vaccine Manufacturing: Estimate Production Costs for mRNA and Other Vaccine Modalities. Available online: www.sigmaaldrich.com/US/en/technical-documents/technical-article/pharmaceutical-and-biopharmaceutical-manufacturing/vaccine-manufacturing/cost-modeling-vaccine-manufacturing. Accessed 13 Dec. 2023.

References

1. Jackson, N.A.; Kester, K.E.; Casimiro, D.; Gurunathan, S.; DeRosa, F. The promise of mRNA vaccines: a biotech and industrial perspective. *npj Vaccines* **2020**, *5*, 11.

2. Sahin, U.; Karikó, K.; Türeci, Ö. mRNA-based therapeutics—developing a new class of drugs. *Nature reviews Drug discovery* **2014**, *13*, 759-780.

3. Jain, S.; Venkataraman, A.; Wechsler, M.E.; Peppas, N.A. Messenger RNA-based vaccines: Past, present, and future directions in the context of the COVID-19 pandemic. *Advanced drug delivery reviews* **2021**, *179*, 114000.

4. Anderson, E.J.; Rouphael, N.G.; Widge, A.T.; Jackson, L.A.; Roberts, P.C.; Makhene, M.; Chappell, J.D.; Denison, M.R.; Stevens, L.J.; Pruijssers, A.J. Safety and immunogenicity of SARS-CoV-2 mRNA-1273 vaccine in older adults. *New England Journal of Medicine* **2020**, *383*, 2427-2438.

5. Baden, L.R.; El Sahly, H.M.; Essink, B.; Kotloff, K.; Frey, S.; Novak, R.; Diemert, D.; Spector, S.A.; Rouphael, N.; Creech, C.B. Efficacy and safety of the mRNA-1273 SARS-CoV-2 vaccine. *New England journal of medicine* **2021**, *384*, 403-416.

6. Hogan, M.J.; Pardi, N. mRNA vaccines in the COVID-19 pandemic and beyond. *Annual review of medicine* **2022**, *73*, 17-39.

7. A Study of mRNA-1345 Vaccine Targeting Respiratory Syncytial Virus (RSV) in Adults ≥50 Years of Age.

8. Martini, P.G.; Guey, L.T. A new era for rare genetic diseases: messenger RNA therapy. *Human Gene Therapy* **2019**, *30*, 1180-1189.

9. Study of Personalized Tumor Vaccines (PCVs) and a PD-L1 Blocker in Patients With Pancreatic Cancer That Can be Treated With Surgery.

10. A Study of a Personalized Cancer Vaccine Targeting Shared Neoantigens.

11. Chaudhary, N.; Weissman, D.; Whitehead, K.A. mRNA vaccines for infectious diseases: principles, delivery and clinical translation. *Nature reviews Drug discovery* **2021**, *20*, 817-838.

12. Liang, Y.; Huang, L.; Liu, T. Development and delivery systems of mRNA vaccines. *Frontiers in Bioengineering and Biotechnology* **2021**, *9*, 718753.

13. Wadhwa, A.; Aljabbari, A.; Lokras, A.; Foged, C.; Thakur, A. Opportunities and challenges in the delivery of mRNA-based vaccines. *Pharmaceutics* **2020**, *12*, 102.

14. Uddin, M.N.; Roni, M.A. Challenges of storage and stability of mRNA-based COVID-19 vaccines. *Vaccines* **2021**, *9*, 1033.

15. Rosa, S.S.; Prazeres, D.M.; Azevedo, A.M.; Marques, M.P. mRNA vaccines manufacturing: Challenges and bottlenecks. *Vaccine* **2021**, *39*, 2190-2200.

16. Vervaeke, P.; Borgos, S.; Sanders, N.; Combes, F. Regulatory guidelines and preclinical tools to study the biodistribution of RNA therapeutics. *Advanced Drug Delivery Reviews* **2022**, *184*, 114236.

17. New Therapeutic Uses. *National Center for Advancing Translational Sciences* 2023.

18. Ohlson, J. Plasmid manufacture is the bottleneck of the genetic medicine revolution. *Drug discovery today* **2020**, *25*, 1891.

19. Skok, J.; Megušar, P.; Vodopivec, T.; Pregeljc, D.; Mencin, N.; Korenč, M.; Krušič, A.; Celjar, A.M.; Pavlin, N.; Krušič, J. Gram-scale mRNA production using a 250-mL single-Use bioreactor. *Chemie Ingenieur Technik* **2022**, *94*, 1928-1935.

20. Cai, Y.; Rodriguez, S.; Rameswaran, R.; Draghia-Akli, R.; Juba Jr, R.J.; Hebel, H. Production of pharmaceutical-grade plasmids at high concentration and high supercoiled percentage. *Vaccine* **2010**, *28*, 2046-2052.

21. Linares-Fernández, S.; Lacroix, C.; Exposito, J.-Y.; Verrier, B. Tailoring mRNA vaccine to balance innate/adaptive immune response. *Trends in molecular medicine* **2020**, *26*, 311-323.

22. Youn, H.; Chung, J.-K. Modified mRNA as an alternative to plasmid DNA (pDNA) for transcript replacement and vaccination therapy. *Expert opinion on biological therapy* **2015**, *15*, 1337-1348.

23. Mauger, D.M.; Cabral, B.J.; Presnyak, V.; Su, S.V.; Reid, D.W.; Goodman, B.; Link, K.; Khatwani, N.; Reynders, J.; Moore, M.J. mRNA structure regulates protein expression through changes in functional half-life. *Proceedings of the National Academy of Sciences* **2019**, *116*, 24075-24083.

24. Kowalski, P.S.; Rudra, A.; Miao, L.; Anderson, D.G. Delivering the messenger: advances in technologies for therapeutic mRNA delivery. *Molecular Therapy* **2019**, *27*, 710-728.

25. Pardi, N.; Hogan, M.J.; Porter, F.W.; Weissman, D. mRNA vaccines—a new era in vaccinology. *Nature reviews Drug discovery* **2018**, *17*, 261-279.

26. Kim, S.C.; Sekhon, S.S.; Shin, W.-R.; Ahn, G.; Cho, B.-K.; Ahn, J.-Y.; Kim, Y.-H. Modifications of mRNA vaccine structural elements for improving mRNA stability and translation efficiency. *Molecular & Cellular Toxicology* **2021**, 1-8.

27. Ramanathan, A.; Robb, G.B.; Chan, S.-H. mRNA capping: biological functions and applications. *Nucleic acids research* **2016**, *44*, 7511-7526.

28. Gallie, D.R. Cap-independent translation conferred by the 5′ leader of tobacco etch virus is eukaryotic initiation factor 4G dependent. *Journal of virology* **2001**, *75*, 12141-12152.

29. Bert, A.G.; Grépin, R.; Vadas, M.A.; Goodall, G.J. Assessing IRES activity in the HIF-1α and other cellular 5′ UTRs. *Rna* **2006**, *12*, 1074-1083.

30. Córdoba, K.M.; Jericó, D.; Sampedro, A.; Jiang, L.; Iraburu, M.J.; Martini, P.G.; Berraondo, P.; Avila, M.A.; Fontanellas, A. Messenger RNA as a personalized therapy: The moment of truth for rare metabolic diseases. *International Review of Cell and Molecular Biology* **2022**, *372*, 55-96.

31. Passmore, L.A.; Coller, J. Roles of mRNA poly (A) tails in regulation of eukaryotic gene expression. *Nature Reviews Molecular Cell Biology* **2022**, *23*, 93-106.

32. Kwon, H.; Kim, M.; Seo, Y.; Moon, Y.S.; Lee, H.J.; Lee, K.; Lee, H. Emergence of synthetic mRNA: In vitro synthesis of mRNA and its applications in regenerative medicine. *Biomaterials* **2018**, *156*, 172-193.

33. Linares-Fernández, S.; Moreno, J.; Lambert, E.; Mercier-Gouy, P.; Vachez, L.; Verrier, B.; Exposito, J.-Y. Combining an optimized mRNA template with a double purification process allows strong expression of in vitro transcribed mRNA. *Molecular Therapy-Nucleic Acids* **2021**, *26*, 945-956.

34. Jalkanen, A.L.; Coleman, S.J.; Wilusz, J. Determinants and implications of mRNA poly (A) tail size–does this protein make my tail look big? In Proceedings of the Seminars in cell & developmental biology, 2014; pp. 24-32.

35. Matoulkova, E.; Michalova, E.; Vojtesek, B.; Hrstka, R. The role of the 3'untranslated region in post-transcriptional regulation of protein expression in mammalian cells. *RNA biology* **2012**, *9*, 563-576.

36. Nance, K.D.; Meier, J.L. Modifications in an emergency: the role of N1-methylpseudouridine in COVID-19 vaccines. *ACS Central Science* **2021**, *7*, 748-756.

37. Karikó, K.; Muramatsu, H.; Welsh, F.A.; Ludwig, J.; Kato, H.; Akira, S.; Weissman, D. Incorporation of pseudouridine into mRNA yields superior nonimmunogenic vector with increased translational capacity and biological stability. *Molecular therapy* **2008**, *16*, 1833-1840.

38. Fotin-Mleczek, M.; Duchardt, K.M.; Lorenz, C.; Pfeiffer, R.; Ojkic-Zrna, S.; Probst, J.; Kallen, K.-J. Messenger RNA-based vaccines with dual activity induce balanced TLR-7 dependent adaptive immune responses and provide antitumor activity. *Journal of immunotherapy* **2011**, *34*, 1-15.

39. Yang, L.; Tang, L.; Zhang, M.; Liu, C. Recent advances in the molecular design and delivery technology of mRNA for vaccination against infectious diseases. *Frontiers in immunology* **2022**, *13*.

40. Morais, P.; Adachi, H.; Yu, Y.-T. The critical contribution of pseudouridine to mRNA COVID-19 vaccines. *Frontiers in cell and developmental biology* **2021**, *9*, 3187.

41. Youssef, M.; Hitti, C.; Puppin Chaves Fulber, J.; Kamen, A.A. Enabling mRNA therapeutics: Current landscape and challenges in manufacturing. *Biomolecules* **2023**, *13*, 1497.

42. de Mey, W.; De Schrijver, P.; Autaers, D.; Pfitzer, L.; Fant, B.; Locy, H.; Esprit, A.; Lybaert, L.; Bogaert, C.; Verdonck, M. A synthetic DNA template for fast manufacturing of versatile single epitope mRNA. *Molecular Therapy-Nucleic Acids* **2022**, *29*, 943-954.

43. Prather, K.J.; Sagar, S.; Murphy, J.; Chartrain, M. Industrial scale production of plasmid DNA for vaccine and gene therapy: plasmid design, production, and purification. *Enzyme and microbial technology* **2003**, *33*, 865-883.

44. Deng, Z.; Tian, Y.; Song, J.; An, G.; Yang, P. mRNA vaccines: the dawn of a new era of cancer immunotherapy. *Frontiers in Immunology* **2022**, *13*.

45. Weng, Y.; Li, C.; Yang, T.; Hu, B.; Zhang, M.; Guo, S.; Xiao, H.; Liang, X.-J.; Huang, Y. The challenge and prospect of mRNA therapeutics landscape. *Biotechnology advances* **2020**, *40*, 107534.

46. Jia, L.; Qian, S.-B. Therapeutic mRNA engineering from head to tail. *Accounts of Chemical Research* **2021**, *54*, 4272-4282.

47. Mordstein, C.; Savisaar, R.; Young, R.S.; Bazile, J.; Talmane, L.; Luft, J.; Liss, M.; Taylor, M.S.; Hurst, L.D.; Kudla, G. Codon usage and splicing jointly influence mRNA localization. *Cell systems* **2020**, *10*, 351-362. e358.

48. Thess, A.; Grund, S.; Mui, B.L.; Hope, M.J.; Baumhof, P.; Fotin-Mleczek, M.; Schlake, T. Sequence-engineered mRNA without chemical nucleoside modifications enables an effective protein therapy in large animals. *Molecular Therapy* **2015**, *23*, 1456-1464.

49. McClellan, D.A. The codon-degeneracy model of molecular evolution. *Journal of Molecular Evolution* **2000**, *50*, 131-140.

50. Fu, H.; Liang, Y.; Zhong, X.; Pan, Z.; Huang, L.; Zhang, H.; Xu, Y.; Zhou, W.; Liu, Z. Codon optimization with deep learning to enhance protein expression. *Scientific reports* **2020**, *10*, 17617.

51. Simms, C.L.; Yan, L.L.; Zaher, H.S. Ribosome collision is critical for quality control during no-go decay. *Molecular cell* **2017**, *68*, 361-373. e365.

52. Presnyak, V.; Alhusaini, N.; Chen, Y.-H.; Martin, S.; Morris, N.; Kline, N.; Olson, S.; Weinberg, D.; Baker, K.E.; Graveley, B.R. Codon optimality is a major determinant of mRNA stability. *Cell* **2015**, *160*, 1111-1124.

53. Ikemura, T. Correlation between the abundance of Escherichia coli transfer RNAs and the occurrence of the respective codons in its protein genes: a proposal for a synonymous codon choice that is optimal for the E. coli translational system. *Journal of molecular biology* **1981**, *151*, 389-409.

54. Hinnebusch, A.G.; Ivanov, I.P.; Sonenberg, N. Translational control by 5′-untranslated regions of eukaryotic mRNAs. *Science* **2016**, *352*, 1413-1416.

55. Kariko, K.; Kuo, A.; Barnathan, E. Overexpression of urokinase receptor in mammalian cells following administration of the in vitro transcribed encoding mRNA. *Gene therapy* **1999**, *6*, 1092-1100.

56. Leppek, K.; Das, R.; Barna, M. Functional 5′ UTR mRNA structures in eukaryotic translation regulation and how to find them. *Nature reviews Molecular cell biology* **2018**, *19*, 158-174.

57. Ding, Y.; Tang, Y.; Kwok, C.K.; Zhang, Y.; Bevilacqua, P.C.; Assmann, S.M. In vivo genome-wide profiling of RNA secondary structure reveals novel regulatory features. *Nature* **2014**, *505*, 696-700.

58. Wan, Y.; Qu, K.; Zhang, Q.C.; Flynn, R.A.; Manor, O.; Ouyang, Z.; Zhang, J.; Spitale, R.C.; Snyder, M.P.; Segal, E. Landscape and variation of RNA secondary structure across the human transcriptome. *Nature* **2014**, *505*, 706-709.

59. Trepotec, Z.; Aneja, M.K.; Geiger, J.; Hasenpusch, G.; Plank, C.; Rudolph, C. Maximizing the translational yield of mRNA therapeutics by minimizing 5′-UTRs. *Tissue Engineering Part A* **2019**, *25*, 69-79.

60. Zinckgraf, J.W.; Silbart, L.K. Modulating gene expression using DNA vaccines with different 3′-UTRs influences antibody titer, seroconversion and cytokine profiles. *Vaccine* **2003**, *21*, 1640-1649.

61. Holtkamp, S.; Kreiter, S.; Selmi, A.; Simon, P.; Koslowski, M.; Huber, C.; Türeci, O.z.; Sahin, U. Modification of antigen-encoding RNA increases stability, translational efficacy, and T-cell stimulatory capacity of dendritic cells. *Blood* **2006**, *108*, 4009-4017.

62. von Niessen, A.G.O.; Poleganov, M.A.; Rechner, C.; Plaschke, A.; Kranz, L.M.; Fesser, S.; Diken, M.; Löwer, M.; Vallazza, B.; Beissert, T. Improving mRNA-based therapeutic gene delivery by expression-augmenting 3′ UTRs identified by cellular library screening. *Molecular Therapy* **2019**, *27*, 824-836.

63. Zhao, W.; Pollack, J.L.; Blagev, D.P.; Zaitlen, N.; McManus, M.T.; Erle, D.J. Massively parallel functional annotation of 3′ untranslated regions. *Nature biotechnology* **2014**, *32*, 387-391.

64. Brown, T.A. Assembly of the Transcription Initiation Complex. In *Genomes. 2nd edition*; Wiley-Liss: 2002.

65. Kwon, S.; Kwon, M.; Im, S.; Lee, K.; Lee, H. mRNA vaccines: the most recent clinical applications of synthetic mRNA. *Archives of pharmacal research* **2022**, *45*, 245-262.

66. Dousis, A.; Ravichandran, K.; Hobert, E.M.; Moore, M.J.; Rabideau, A.E. An engineered T7 RNA polymerase that produces mRNA free of immunostimulatory byproducts. *Nature Biotechnology* **2023**, *41*, 560-568.

67. Conrad, T.; Plumbom, I.; Alcobendas, M.; Vidal, R.; Sauer, S. Maximizing transcription of nucleic acids with efficient T7 promoters. *Communications Biology* **2020**, *3*, 439.

68. Deich, C.; Cash, B.; Sato, W.; Sharon, J.; Aufdembrink, L.; Gaut, N.J.; Heili, J.; Stokes, K.; Engelhart, A.E.; Adamala, K.P. T7Max transcription system. *Journal of Biological Engineering* **2023**, *17*, 4.

69. Mu, X.; Greenwald, E.; Ahmad, S.; Hur, S. An origin of the immunogenicity of in vitro transcribed RNA. *Nucleic Acids Research* **2018**, *46*, 5239-5249.

70. Durbin, A.F.; Wang, C.; Marcotrigiano, J.; Gehrke, L. RNAs containing modified nucleotides fail to trigger RIG-I conformational changes for innate immune signaling. *MBio* **2016**, *7*, e00833-00816.

71. Peisley, A.; Jo, M.H.; Lin, C.; Wu, B.; Orme-Johnson, M.; Walz, T.; Hohng, S.; Hur, S. Kinetic mechanism for viral dsRNA length discrimination by MDA5 filaments. *Proceedings of the National Academy of Sciences* **2012**, *109*, E3340-E3349.

72. Sutcliffe, J.G. Nucleotide sequence of the ampicillin resistance gene of Escherichia coli plasmid pBR322. *Proceedings of the National Academy of Sciences* **1978**, *75*, 3737-3741.

73. Hershfield, V.; Boyer, H.W.; Yanofsky, C.; Lovett, M.A.; Helinski, D.R. Plasmid ColE1 as a molecular vehicle for cloning and amplification of DNA. *Proceedings of the National Academy of Sciences* **1974**, *71*, 3455-3459.

74. Jang, C.-W.; Magnuson, T. A novel selection marker for efficient DNA cloning and recombineering in E. coli. *PloS one* **2013**, *8*, e57075.

75. Selas Castiñeiras, T.; Williams, S.G.; Hitchcock, A.G.; Smith, D.C. E. coli strain engineering for the production of advanced biopharmaceutical products. *FEMS microbiology letters* **2018**, *365*, fny162.

76. O'Mahony, K.; Freitag, R.; Hilbrig, F.; Mu¨ ler, P.; Schumacher, I. Strategies for high titre plasmid DNA production in Escherichia coli DH5α. *Process Biochemistry* **2007**, *42*, 1039-1049.

77. Schmeer, M.; Buchholz, T.; Schleef, M. Plasmid DNA manufacturing for indirect and direct clinical applications. *Human gene therapy* **2017**, *28*, 856-861.

78. Lopes, M.B.; Gonçalves, G.A.; Felício-Silva, D.; Prather, K.L.; Monteiro, G.A.; Prazeres, D.M.; Calado, C.R. In situ NIR spectroscopy monitoring of plasmid production processes: effect of producing strain, medium composition and the cultivation strategy. *Journal of Chemical Technology & Biotechnology* **2015**, *90*, 255-261.

79. Yau, S.Y.; Keshavarz-Moore, E.; Ward, J. Host strain influences on supercoiled plasmid DNA production in Escherichia coli: Implications for efficient design of large-scale processes. *Biotechnology and bioengineering* **2008**, *101*, 529-544.

80. Listner, K.; Bentley, L.; Okonkowski, J.; Kistler, C.; Wnek, R.; Caparoni, A.; Junker, B.; Robinson, D.; Salmon, P.; Chartrain, M. Development of a highly productive and scalable plasmid DNA production platform. *Biotechnology progress* **2006**, *22*, 1335-1345.

81. Lara, A.R.; Knabben, I.; Regestein, L.; Sassi, J.; Caspeta, L.; Ramírez, O.T.; Büchs, J. Comparison of oxygen enriched air vs. pressure cultivations to increase oxygen transfer and to scale-up plasmid DNA production fermentations. *Engineering in Life Sciences* **2011**, *11*, 382-386.

82. Lahijani, R.; Hulley, G.; Soriano, G.; Horn, N.A.; Marquet, M. High-yield production of pBR322-derived plasmids intended for human gene therapy by employing a temperature-controllable point mutation. *Human gene therapy* **1996**, *7*, 1971-1980.

83. Cooke, J.R.; McKie, E.A.; Ward, J.M.; Keshavarz-Moore, E. Impact of intrinsic DNA structure on processing of plasmids for gene therapy and DNA vaccines. *Journal of biotechnology* **2004**, *114*, 239-254.

84. Singer, A.; Eiteman, M.A.; Altman, E. DNA plasmid production in different host strains of Escherichia coli. *Journal of industrial microbiology and biotechnology* **2009**, *36*, 521-530.

85. Phue, J.N.; Lee, S.J.; Trinh, L.; Shiloach, J. Modified Escherichia coli B (BL21), a superior producer of plasmid DNA compared with Escherichia coli K (DH5α). *Biotechnology and bioengineering* **2008**, *101*, 831-836.

86. Borja, G.M.; Meza Mora, E.; Barrón, B.; Gosset, G.; Ramírez, O.T.; Lara, A.R. Engineering Escherichia coli to increase plasmid DNA production in high cell-density cultivations in batch mode. *Microbial Cell Factories* **2012**, *11*, 1-9.

87. Goncalves, G.A.; Prather, K.L.; Monteiro, G.A.; Carnes, A.E.; Prazeres, D.M. Plasmid DNA production with Escherichia coli GALG20, a pgi-gene knockout strain: Fermentation strategies and impact on downstream processing. *Journal of biotechnology* **2014**, *186*, 119-127.

88. Mojica, N.; Kersten, F.; Montserrat-Canals, M.; Huhn III, G.R.; Tislevoll, A.M.; Cordara, G.; Teter, K.; Krengel, U. Using Vibrio natriegens for High-Yield Production of Challenging Expression Targets and for Protein Perdeuteration. *Biochemistry* **2024**.

89. Xu, J.; Dong, F.; Wu, M.; Tao, R.; Yang, J.; Wu, M.; Jiang, Y.; Yang, S.; Yang, L. Vibrio natriegens as a pET-compatible expression host complementary to Escherichia coli. *Frontiers in Microbiology* **2021**, *12*, 627181.

90. Bower, D.M.; Prather, K.L. Engineering of bacterial strains and vectors for the production of plasmid DNA. *Applied microbiology and biotechnology* **2009**, *82*, 805-813.

91. Yanisch-Perron, C.; Vieira, J.; Messing, J. Improved M13 phage cloning vectors and host strains: nucleotide sequences of the M13mpl8 and pUC19 vectors. *Gene* **1985**, *33*, 103-119.

92. Bower, D.M.; Prather, K.L. Development of new plasmid DNA vaccine vectors with R1-based replicons. *Microbial cell factories* **2012**, *11*, 1-10.

93. Carnes, A.E.; Hodgson, C.P.; Williams, J.A. Inducible Escherichia coli fermentation for increased plasmid DNA production. *Biotechnology and applied biochemistry* **2006**, *45*, 155-166.

94. Gotsmy, M.; Strobl, F.; Weiß, F.; Gruber, P.; Kraus, B.; Mairhofer, J.; Zanghellini, J. Sulfate limitation increases specific plasmid DNA yield and productivity in E. coli fed-batch processes. *Microbial Cell Factories* **2023**, *22*, 242.

95. Xu, Z.-n.; Shen, W.-h.; Chen, H.; Cen, P.-l. Effects of medium composition on the production of plasmid DNA vector potentially for human gene therapy. *Journal of Zhejiang University-SCIENCE B* **2005**, *6*, 396-400.

96. Silva, F.; Passarinha, L.; Sousa, F.; Queiroz, J.A.; Domingues, F.C. Influence of growth conditions on plasmid DNA production. *Journal of microbiology and biotechnology* **2009**, *19*, 1408-1414.

97. Messenger RNA (mRNA)-Based, Personalized Cancer Vaccine Against Neoantigens Expressed by the Autologous Cancer.

98. Malke, H. HENRY A. ERLICH (Editor), PCR Technology: Principles and Applications for DNA Amplification. X+ 246 S. New York-Tokyo-Melbourne-Hong Kong 1989. McMillan Publishers (Stockton Press).£ 15.00. ISBN: 0-333-48948-9. **1990**.

99. Hu, Z.; Wang, L.; Shi, Z.; Jiang, J.; Li, X.; Chen, Y.; Li, K.; Luo, D. Customized one-step preparation of sgRNA transcription templates via overlapping PCR using short primers and its application in vitro and in vivo gene editing. *Cell & Bioscience* **2019**, *9*, 1-7.

100. Wei, H.; Rong, Z.; Liu, L.; Sang, Y.; Yang, J.; Wang, S. Streamlined and on-demand preparation of mRNA products on a universal integrated platform. *Microsystems & Nanoengineering* **2023**, *9*, 97.

101. Evonetix. Unlocking the promise of mRNA vaccines with gene synthesis technology. Available online: (accessed on April 2024).

102. MilliporeSigma. Cell Harvest, Lysis, Neutralization & Clarification of Plasmid DNA. Available online: https://www.sigmaaldrich.com/deepweb/assets/sigmaaldrich/marketing/global/documents/350/791/cell-harvest-lysis-an7049en-ms.pdf (accessed on April 2024).

103. Pethica, B. Lysis by physical and chemical methods. *Microbiology* **1958**, *18*, 473-480.

104. Bimboim, H.C.; Doly, J. A rapid alkaline extraction procedure for screening recombinant plasmid DNA. *Nucleic acids research* **1979**, *7*, 1513-1523.

105. Sun, B.; Yu, X.; Yin, Y.; Liu, X.; Wu, Y.; Chen, Y.; Zhang, X.; Jiang, C.; Kong, W. Large-scale purification of pharmaceutical-grade plasmid DNA using tangential flow filtration and multi-step chromatography. *Journal of bioscience and bioengineering* **2013**, *116*, 281-286.

106. Wright, J.L.; Jordan, M.; Wurm, F.M. Extraction of plasmid DNA using reactor scale alkaline lysis and selective precipitation for scalable transient transfection. *Cytotechnology* **2001**, *35*, 165-173.

107. Schmeer, M.; Schleef, M. Pharmaceutical grade large-scale plasmid DNA manufacturing process. *DNA Vaccines: Methods and Protocols* **2014**, 219-240.

108. Voss, C.; Flaschel, E. Method for producing extra-chromosomal nucleic acid molecules. **2010**.

109. Chamsart, S.; Karnjanasorn, T. Alkaline-cell lysis through in-line static mixer reactor for the production of plasmid DNA for gene therapy. *Biotechnology and bioengineering* **2007**, *96*, 471-482.

110. Shehadul Islam, M.; Aryasomayajula, A.; Selvaganapathy, P.R. A review on macroscale and microscale cell lysis methods. *Micromachines* **2017**, *8*, 83.

111. Holmes, D.S.; Quigley, M. A rapid boiling method for the preparation of bacterial plasmids. *Analytical biochemistry* **1981**, *114*, 193-197.

112. Lengsfeld, C.; Anchordoquy, T. Shear-induced degradation of plasmid DNA. *Journal of pharmaceutical sciences* **2002**, *91*, 1581-1589.

113. Nicoletti, V.; Condorelli, D. Optimized PEG method for rapid plasmid DNA purification: high yield from" midi-prep". *BioTechniques* **1993**, *14*, 532-536.

114. Cole, K. Purification of plasmid and high molecular mass DNA using PEG-salt two-phase extraction. *Biotechniques* **1991**, *11*, 18, 20, 22-14.

115. Eon-Duval, A.; MacDuff, R.H.; Fisher, C.A.; Harris, M.J.; Brook, C. Removal of RNA impurities by tangential flow filtration in an RNase-free plasmid DNA purification process. *Analytical Biochemistry* **2003**, *316*, 66-73.

116. Latulippe, D.R.; Zydney, A.L. Size exclusion chromatography of plasmid DNA isoforms. *Journal of Chromatography A* **2009**, *1216*, 6295-6302.

117. Eon-Duval, A.; Burke, G. Purification of pharmaceutical-grade plasmid DNA by anion-exchange chromatography in an RNase-free process. *Journal of chromatography B* **2004**, *804*, 327-335.

118. Bo, H.; Wang, J.; Chen, Q.; Shen, H.; Wu, F.; Shao, H.; Huang, S. Using a single hydrophobic-interaction chromatography to purify pharmaceutical-grade supercoiled plasmid DNA from other isoforms. *Pharmaceutical biology* **2013**, *51*, 42-48.

119. Sparks, R.; Elder, J. A simple and rapid procedure for the purification of plasmid DNA using reverse-phase C18 silica beads. *Analytical biochemistry* **1983**, *135*, 345-348.

120. Guerrero-Germán, P.; Montesinos-Cisneros, R.M.; Prazeres, D.M.F.; Tejeda-Mansir, A. Purification of plasmid DNA from Escherichia coli ferments using anion-exchange membrane and hydrophobic chromatography. *Biotechnology and applied biochemistry* **2011**, *58*, 68-74.

121. Sartorius. Plasmid DNA. Available online: https://www.biaseparations.com/solutions/pdna/ (accessed on March 2024).

122. Silva-Santos, A.R.; Rosa, S.S.; Prazeres, D.M.F.; Azevedo, A.M. Purification of plasmid DNA by multimodal chromatography. *DNA Vaccines: Methods and Protocols* **2021**, 193-205.

123. Breton, H.; G. Brett, R.; George, T. Minding your caps and tails – considerations for functional mRNA synthesis. 2021.

124. Bancel, S.; Issa, W.J.; Aunins, J.G.; Chakraborty, T. Manufacturing methods for production of RNA transcripts. **2018**.

125. Cui, T.; Fakhfakh, K.; Turney, H.; Gu¨er-Gane, G.; Toloczko, A.; Hulley, M.; Turner, R. Comprehensive studies on building a scalable downstream process for mRNAs to enable mRNA therapeutics. *Biotechnology Progress* **2023**, *39*, e3301.

126. MilliporeSigma. Enabling Capabilities and Solutions for all mRNA platforms: Process Development, Manufacturing and Formulation. Available online: https://www.sigmaaldrich.com/deepweb/assets/sigmaaldrich/product/documents/411/946/mrna-manufacturing-workflow-br8431en-ms.pdf (accessed on September 2023).

127. ThermoFisher. Purifying your PCR Product. Available online: https://www.thermofisher.com/ca/en/home/references/an-introduction-to-cloning-a-researchers-guide-to-cloning-dna/purifying-your-pcr-product.html (accessed on April 2024).

128. Crommelin, D.J.; Anchordoquy, T.J.; Volkin, D.B.; Jiskoot, W.; Mastrobattista, E. Addressing the cold reality of mRNA vaccine stability. *Journal of pharmaceutical sciences* **2021**, *110*, 997-1001.

129. USP. Analytical Procedures for mRNA Vaccine Quality (Draft Guidelines)- 2nd Edition. Available online: https://go.usp.org/l/323321/2022-02-08/6rxnz3/323321/16443533303H7od3V5/EA966W_mRNA_Vaccine_Chapter_2022_03_3.pdf (accessed on April 2024).

130. Gilar, M.; Doneanu, C.; Gaye, M.M. Liquid Chromatography Methods for Analysis of mRNA Poly (A) Tail Length and Heterogeneity. *Analytical Chemistry* **2023**, *95*, 14308-14316.

131. Gilar, M.; Gaye, M.M. Ion-Pair Reversed-Phase Liquid Chromatography Method for Analysis of mRNA Poly (A) Tail Heterogeneity.

132. BIOVECTRA. Guest Lecture for Course BIEN 602 (McGill University). **2024**.

133. Rong, M.; He, B.; McAllister, W.T.; Durbin, R.K. Promoter specificity determinants of T7 RNA polymerase. *Proceedings of the National Academy of Sciences* **1998**, *95*, 515-519.

134. Rosa, S.S.; Nunes, D.; Antunes, L.; Prazeres, D.M.; Marques, M.P.; Azevedo, A.M. Maximizing mRNA vaccine production with Bayesian optimization. *Biotechnology and Bioengineering* **2022**, *119*, 3127-3139.

135. Samnuan, K.; Blakney, A.K.; McKay, P.F.; Shattock, R.J. Design-of-experiments in vitro transcription yield optimization of self-amplifying RNA. *bioRxiv* **2021**, 2021.2001. 2008.425833.

136. Piao, X.; Yadav, V.; Wang, E.; Chang, W.; Tau, L.; Lindenmuth, B.E.; Wang, S.X. Double-stranded RNA reduction by chaotropic agents during in vitro transcription of messenger RNA. *Molecular Therapy-Nucleic Acids* **2022**, *29*, 618-624.

137. Henderson, J.M.; Ujita, A.; Hill, E.; Yousif-Rosales, S.; Smith, C.; Ko, N.; McReynolds, T.; Cabral, C.R.; Escamilla-Powers, J.R.; Houston, M.E. Cap 1 messenger RNA synthesis with co-transcriptional cleancap® analog by in vitro transcription. *Current protocols* **2021**, *1*, e39.

138. Young, J.S.; Ramirez, W.F.; Davis, R.H. Modeling and optimization of a batch process for in vitro RNA production. *Biotechnology and bioengineering* **1997**, *56*, 210-220.

139. Kern, J.A.; Davis, R.H. Application of Solution Equilibrium Analysis to inVitro RNA Transcription. *Biotechnology progress* **1997**, *13*, 747-756.

140. Pregeljc, D.; Skok, J.; Vodopivec, T.; Mencin, N.; Krušič, A.; Ličen, J.; Nemec, K.Š.; Štrancar, A.; Sekirnik, R. Increasing yield of in vitro transcription reaction with at-line high pressure liquid chromatography monitoring. *Biotechnology and Bioengineering* **2023**, *120*, 737-747.

141. Kis, Z.; Kontoravdi, C.; Shattock, R.; Shah, N. Resources, production scales and time required for producing RNA vaccines for the global pandemic demand. *Vaccines* **2020**, *9*, 3.

142. Zhang, J.; Liu, Y.; Li, C.; Xiao, Q.; Zhang, D.; Chen, Y.; Rosenecker, J.; Ding, X.; Guan, S. Recent Advances and Innovations in the Preparation and Purification of In Vitro-Transcribed-mRNA-Based Molecules. *Pharmaceutics* **2023**, *15*, 2182.

143. NewEnglandBiolabs. mRNA Purification (E2065). Available online: https://www.neb.com/en/protocols/2015/03/25/mrna-purification-e2065 (accessed on March 2024).

144. Webb, C.; Ip, S.; Bathula, N.V.; Popova, P.; Soriano, S.K.; Ly, H.H.; Eryilmaz, B.; Nguyen Huu, V.A.; Broadhead, R.; Rabel, M. Current status and future perspectives on MRNA drug manufacturing. *Molecular Pharmaceutics* **2022**, *19*, 1047-1058.

145. Lukavsky, P.J.; Puglisi, J.D. Large-scale preparation and purification of polyacrylamide-free RNA oligonucleotides. *Rna* **2004**, *10*, 889-893.

146. DeRosa, F.; Dias, A.; Karve, S.; Heartlein, M. Methods for purification of messenger RNA. US9850269B2, 2014.

147. Daniel, S.; Kis, Z.; Kontoravdi, C.; Shah, N. Quality by Design for enabling RNA platform production processes. *Trends in Biotechnology* **2022**, *40*, 1213-1228.

148. Nguyen, T.T.; Bui, L.M.; Byun, J.-Y.; Cho, B.-K.; Kim, S.C. Exploring the potential of a genome-reduced Escherichia coli strain for plasmid DNA production. *International Journal of Molecular Sciences* **2023**, *24*, 11749.

149. Wang, F.; Chen, Q.; Li, S.; Zhang, C.; Li, S.; Liu, M.; Mei, K.; Li, C.; Ma, L.; Yu, X. Linear DNA vaccine prepared by large-scale PCR provides protective immunity against H1N1 influenza virus infection in mice. *Veterinary Microbiology* **2017**, *205*, 124-130.

150. Trepotec, Z.; Geiger, J.; Plank, C.; Aneja, M.K.; Rudolph, C. Segmented poly (A) tails significantly reduce recombination of plasmid DNA without affecting mRNA translation efficiency or half-life. *Rna* **2019**, *25*, 507-518.

151. Graham, Z.; Kaur, S.; Meunier, M.; Dyer, R.; Unger, M. Enzymatic Addition of Poly-A Tail to Linear Synthetic DNA to Accelerate *in vitro* Transcription-mediated mRNA Synthesis. **2023**.

152. Arbuthnot, P.; Ely, A.; Bloom, K. A convenient method to generate and maintain poly (A)-encoding DNA sequences required for in vitro transcription of mRNA. *BioTechniques* **2019**, *66*, 37-39.

153. Hughes, R.A.; Ellington, A.D. Synthetic DNA synthesis and assembly: putting the synthetic in synthetic biology. *Cold Spring Harbor perspectives in biology* **2017**, *9*, a023812.