Platform technology for the enhanced manufacture and utilisation of synthetic mRNA

Author: Edward Curry

Supervisor: Dr Adam Brown

A thesis submitted in fulfilment of the requirements

for the degree of Doctor of Philosophy

Chemical & Biological Engineering

December 2023

Declaration of Originality

In accordance with the regulations of the University of Sheffield, I declare that all work contained within this thesis is entirely my own and has not been submitted for the attainment of any other degree or qualification. Whereby others have contributed to the work constituting publications mentioned in this thesis, their role is clearly explained. Any material referenced from other work is attributed to the corresponding author.

Abstract

The use of synthetic messenger RNA (mRNA) in the fields of medicine and biotechnology is becoming increasingly prevalent, due to the well characterised advantages of the molecule as a means of transferring genetic information. The rapid expansion of synthetic mRNA related applications necessitates the development of new technologies for the manufacturing of the molecule in a more cost efficient and less resource intensive manner. Presented in this thesis are two approaches for improving the manufacture of synthetic mRNA, in addition to a novel usage of the synthetic mRNA molecule itself.

Firstly, a new method is described for identifying RNA polymerases that could be used in existing *in vitro* transcription processes. Current manufacturing methods suffer from poor product related impurity profiles, as a result of off-target activity from T7 RNA polymerase. Using a novel cell free coupled transcription-translation system, a library of 10 new RNA polymerases is introduced, with diverse yield profiles. The nature of active polymerase and promoter pairs discovered in this study allow for conclusions to be drawn on which RNA polymerases are optimal targets for future characterisation.

Secondly, a new manufacturing platform for the production of synthetic mRNA is introduced, with *E. coli* being used as a cellular chassis for the accumulation of recombinant mRNA. The use of a cell as a factory for mRNA production reduces raw material costs, due to the ability of *E. coli* to synthesise the components of the transcription reaction endogenously from inexpensive media. Coordinated mRNA, DNA, cell and media engineering, primarily focussed on disrupting interactions between synthetic mRNA molecules and host cell RNA degradation machinery, increases product yields 40-fold compared to standard ‘unengineered’ *E. coli* expression systems.

Finally, synthetic mRNA is used as a means of identifying rate limiting bottlenecks in the production of monoclonal antibodies (mAb) in Chinese Hamster Ovary (CHO) cells. This novel biotechnology application is achieved through transfection of a panel of mRNA 'control nodes', encoding transcription factors regulating pathways known to be implicated in the production of recombinant proteins. Control nodes which impact the titre of the monoclonal antibody being produced allow for identification of genes critical to the production of that mAb. The use of chemical modulators targeting control node implicated genes is shown to be a viable strategy for increasing mAb titre. Taken together, the work presented in this thesis provides a basis for the manufacturing of the next-generation of mRNA therapeutics, and shows the applicability of synthetic mRNA technology outside the field of medicine

Acknowledgements

I would like to thank firstly UK Research & Innovation and AztraZeneca for providing funding for this project.

A massive thanks goes to the supervisor of this project, Dr. Adam Brown, for his continued support over the last 4 years. Without his guidance and encouragement no part of this thesis would have been possible. I am also appreciative of my industrial supervisor, Dr. Martyn Hulley for his input throughout. To all those who have contributed to the completion of this thesis, I am deeply grateful.

Essential to finishing any PhD project is a supportive lab group, in both a professional and personal capacity. I am lucky to have been part of the Adam Brown & David James lab group. There are far too many people to name individually here – thank you.

Most important of all through this PhD however was the support of my family, especially Molly and Archie.

Contents

List of Figures

List of Tables

List of Abbreviations

A	Adenine
AAV	Adeno-associated virus
ARCA	Anti-Reverse Cap Analog
ARE	AU rich element
ATTR	transthyretin amyloidosis
bp	Base pair
CAR	Chimeric Antigen Receptor
CGE	Capillary Gel Electrophoresis
CHO	Chinese Hamster Ovary
CLD	Cell line development
cLuc	Cypridinia Luciferase
CMV	Cytomegalovirus
Covid	Coronavirus Disease
CRISPR	Clustered Regularly Interspace Short Palindromic Repeat
CSE	Conserved Sequence Elements
Da	Dalton
DC	Dendritic Cell
ddPCR	Digital Droplet PCR
DdRp	DNA-dependent RNA polymerase
DNA	Deoxyribonucleic acid
dsRNA	Double stranded RNA
DTE	Difficult to express
DTT	Dithiothreitol
EDTA	Ethylenediaminetetraacetic acid
ER	Endoplasmic reticulum
ETE	East to express
G	Guanine
GFP	Green Fluorescent Protein
GMP	Good Manufacturing Process
GOI	Gene of interest
GSC	Glioblastoma Stem Cells
GTP	Guanosine triphosphate
HDR	Homology Directed Repair
HEK	Human Embryonic Kidney
HIV	Human Immunodeficiency Virus
HPLC	High Performance Liquid Chromatography
IgG	Immunoglobulin G
iPSC	Induced Pluripotent Stem Cell
IRES	Internal Ribosome Entry Site
IVT	*In vitro* transcription
LB	Lysogeny Broth
LIF	Laser Induced Fluorescence
LNP	Lipid Nanoparticle
m1Ψ	N1-methylpseudouridine

m7G	Methyl-7-guanosine
mAb	Monoclonal Antibodies
miRNA	MicroRNA
MN	Meganucleases
mRNA	Messenger RNA
NHEJ	Non-homologous End Joining
nsP1-4	Non-structural Proteins 1-4
nt	Nucleotide
NTP	Nucleoside Triphosphate
OD	Optical Density
ORF	Open Reading Frame
Ori	Origin of replication
PBS	Phosphate Buffered Saline
PCC	Propionyl-CoA Carboxylase
PCR	Polymerase chain reaction
PH1	primary hyperoxaluria type I
PKR	RNA-dependent Protein Kinase
Poly(A)	Polyadenosine
RdRp	RNA-dependent RNA polymerase
RIG-I	Retinoic acid-inducible gene I
RNA	Ribonucleic acid
RNA-LP	RNA-Lipid Particle
RNAP	RNA Polymerase
RNase	Ribonuclease
r-protein	Recombinant protein
rRNA	Ribosomal RNA
RSV	Respiratory Syncytial Infection
RT-ddPCR	Reverse Transcription Digital Droplet PCR
saRNA	Self-amplifying ribonucleic acid
SARS-CoV2	Severe Acute Respiratory Syndrome Coronavirus 2
SDS	Sodium Dodecyl Sulfate
SDS-PAGE	SDS Polyacrylamide gel electrophoresis
SFV	Semliki forest virus
SINV	Sindbis virus
siRNA	Short Interfering RNA
snRNA	Small Nuclear RNA
SOC	Super optimal broth with catabolite repression
TALEN	Transcription Activator-like Effector Nucleases
TE	Tris-EDTA
TLR	Toll-like receptor
tRNA	Transfer RNA
UTR	Untranslated Region
VEE	Venezuelan equine encephalitis virus
VEGF-A	vascular endothelial growth factor A
ZFN	Zinc-finger Nucleases
Ψ	Pseudouridine

Chapter 1 – Introduction

The work presented in this study aims to provide novel tools to advance the manufacture of synthetic mRNA, in addition to introducing a new utilisation of synthetic mRNA in the field of biopharmaceutical production. This introductory chapter provides a contextualisation of the tools introduced in this thesis, within the field of synthetic mRNA. Much of the research interest in synthetic mRNA focuses on the use of mRNA as a therapeutic molecule. An overview of mRNA therapeutics is given, in the context of their initial development, current approved therapeutics, and future directions for the field. Additionally, current non-therapeutic usages of synthetic mRNA are discussed. A commentary is also given on the design of synthetic mRNA, and efforts to improve its functionality. Finally, a summary of methods for producing synthetic mRNA is given, both in terms of well-established *in vitro* transcription based manufacturing, and novel methods for the production of diverse RNA species.

1.1 Synthetic mRNA as a functional molecule in medicine and biotechnology

1.1.1 - Synthetic mRNA as a therapeutic molecule

Messenger mRNA (mRNA) is one of the fundamental components of the central dogma of molecular biology, facilitating the flow of genetic information from DNA into functional protein (Crick, 1970). Since its discovery in 1961, mRNA has been the focus of extensive research, both in terms of its fundamental role in all domains of life, and more recently its functionality in the field of biopharmaceuticals and biotechnology (Brenner et al., 1961). Much of the recent attention on synthetic mRNA has stemmed from its authorisation as a vaccine in the Covid-19 pandemic, the culmination of decades of pre-clinical development (Baden et al., 2021; Chung et al., 2020; Sahin et al., 2014).

The concept of injecting nucleic acids directly into cells was first demonstrated by Wolff et al. in 1990, when both mRNA and plasmid DNA was used to stimulate reporter gene expression in mouse skeletal muscle (Wolff et al., 1990). At this time, DNA was believed to be the more promising therapeutic format, as a consequence of its increased stability over mRNA. Indeed, a number of DNA-based therapies are now approved, utilising a variety of delivery formats (Aiuti et al., 2009; Ginn et al., 2018; Kulkarni et al., 2021; Opalinska and Gewirtz, 2002). In recent years however, interest in mRNA-based therapies has risen exponentially, a result of several key advantages existing when using mRNA as a therapeutic agent over more traditional DNA based vectors.

Firstly, there is no need for the mRNA to enter the nucleus, as transcription is able to proceed immediately upon access to the cytoplasm (Zou et al., 2010). This is in contrast to DNA based technologies, whereby entry into the nucleus is required, facilitated by nuclear envelope breakdown (Tavernier et al., 2011). It has also been observed that RNA transfection is more efficient than DNA transfection in quiescent cells (Zou et al., 2010). Secondly, unlike DNA based gene therapeutics, there is no risk of genome integration when using mRNA. Integration into the genome can lead to insertional mutagenesis, potentially disrupting tumour suppressor genes, activating proto-oncogenes, or causing abhorrent splicing of gene products (Cesana et al., 2012; Mitchell et al., 2004). The safety of DNA based gene therapies has previously been called into question, with cases of patients contracting Leukaemia, or, in one example, producing a fatal immune response to the treatment (Hacein-Bey-Abina et al., 2008; Yarborough and Sharp, 2009). Furthermore, the expression of proteins produced by mRNA delivered to the cell is transient, leading to higher levels of control of protein dosage (Wang et al., 2013). Finally, the production of *in vitro* transcribed (IVT) mRNA

is cost effective and simple compared to complex DNA delivery formats such as adeno-associated virus, and is easily scalable (Pascolo, 2017).

Whilst the use of mRNA for transferring genetic information to target cells has long been theorised as a means of therapeutic intervention, several critical discoveries over the last few decades have been key to unlocking the potential of this approach. The presence of a eukaryotic mRNA cap structure was first identified in 1975, with the first commercial m^7GpppG cap analogues coming to market in 1983 (Muthukrishnan et al., 1975; Sahin et al., 2014). Developments in capping technology now allow for the production of low-immunogenicity transcripts, which are effectively all capped in the correct orientation (Grudzien-Nogalska et al., 2007; Henderson et al., 2021). Methods for introducing mRNA more reliably into cells were also pioneered in the 1970s and 80s, culminating in the using of cationic lipids as a means of RNA transfection (Malone et al., 1989). Current mRNA delivery methods still rely on lipid-based nanoparticles, with novel chemistries that allow for efficient cellular uptake (Gote et al., 2023; Hou et al., 2021). Despite several attempts to use mRNA in the fields of protein replacement and cancer immunotherapy in the 1990s and early 2000s, progress was hindered by the immunogenicity of such treatments (Alexopoulou et al., 2001; Heil et al., 2004; Heiser et al., 2002; Zhou et al., 1999). The use of modified nucleotides as a means of ablating toll-like receptor (TLR) activity was therefore a critical discovery; one that earned the authors the 2023 Nobel prize for medicine (Karikó et al., 2005). Indeed, Covid-19 mRNA vaccines were N1-methyl pseudouridine modified in order to prevent immune activation (Morais et al., 2021). Such advances have been critical in the approval of mRNA vaccines against Covid-19, and underpin the development of future

mRNA therapeutics. Further details on the design of functional mRNA is given in section 1.2.

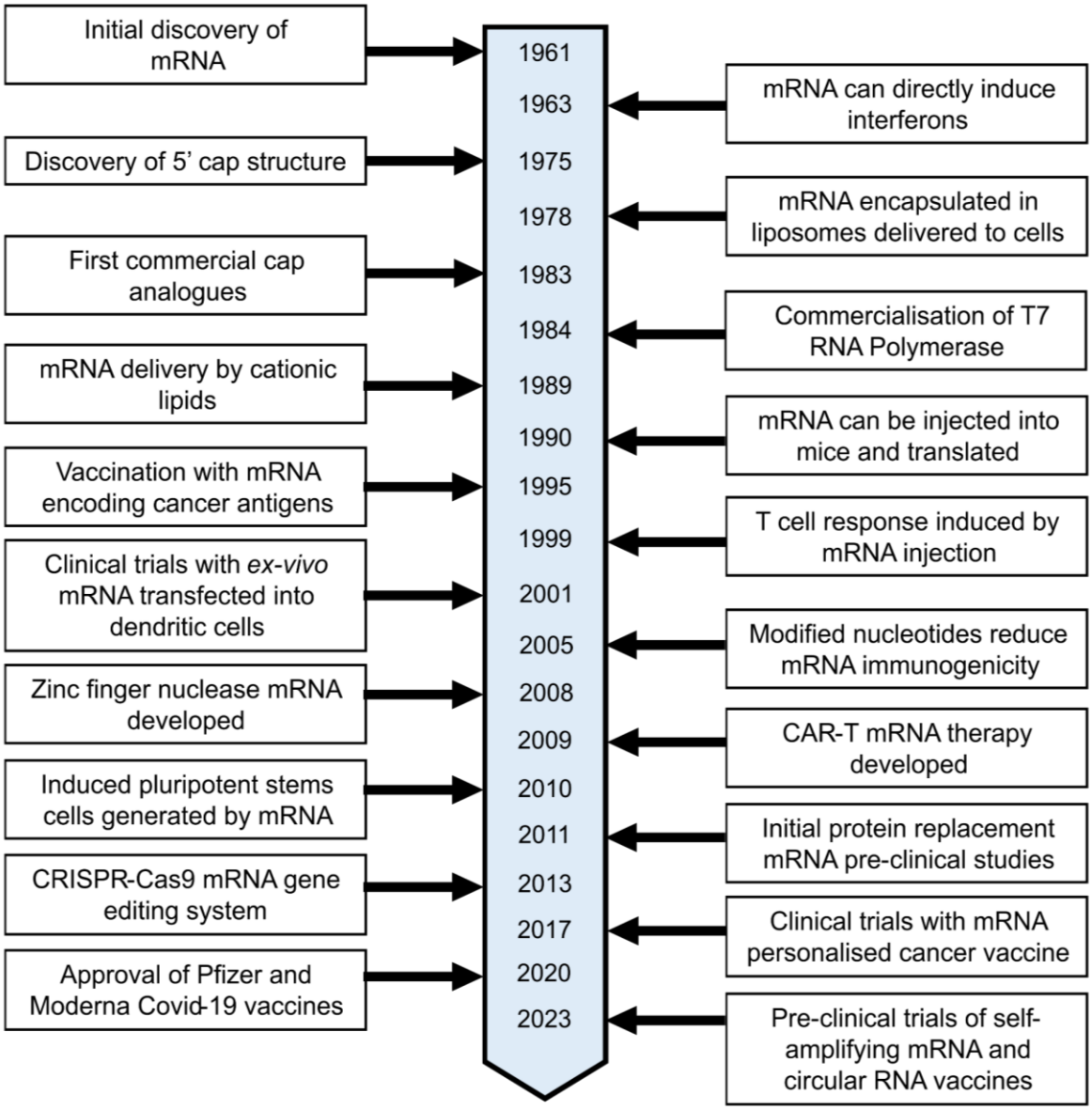


Figure 1.1 – A summary of the key discoveries and developments in the field of mRNA therapeutics. The timeline begins with the initial discovery of mRNA by Brenner in 1961, and follows the development and commercialisation of critical enabling technology such as cap analogues and modified nucleotides, up to the present day where new mRNA formats are entering pre-clinical trials. Information adapted from (Sahin et al., 2014; Vavilis et al., 2023a).

1.1.1 - Synthetic mRNA as a vaccination against infection diseases

The area in which therapeutic mRNA has had the largest impact is vaccination against infectious disease, including approved immunisations against Covid-19 and clinical trials of mRNA-based vaccinations against Malaria and Zika Virus (Baden et al., 2021; Erasmus et al., 2020; Ganley et al., 2023; Medina-Magües et al., 2021). The mRNA vaccine is injected intramuscularly, and taken up by muscle and dendritic cells in the vicinity. mRNA that enters the cell escapes from the lipid nanoparticle, and is translated by the endogenous cellular machinery into antigens that are displayed to T and B cells (Kwon et al., 2022). In the cases of Pfizer/BioNTechs approved BNT162b2 vaccine, this antigen is the SARS-CoV2 spike glycoprotein, featuring two proline substitutions which keep the protein in a prefusion conformation (Walls et al., 2020).

At the time of writing, there are numerous mRNA vaccines in Phase 1 or 2 trials against infectious disease other than Covid-19, as shown in Table 1.1 (Gote et al., 2023). Three of these have progressed to Phase 3, the final stage before market approval. Firstly, there is Moderna's Influenza virus candidate, mRNA-1010, which encodes hemagglutinin from four separate influenza strains (Lee et al., 2023a). Recently disclosed interim Phase 3 results show an improved immune response from mRNA-1010 compared to the currently marketed vaccine Fluarix. Secondly is mRNA-1345, Moderna's candidate vaccine for respiratory syncytial infection (RSV), encoding a RSV glycoprotein. The vaccine has proved effective in Phase 1 and 2 trials, with the expected completion of phase 3 being in November 2024. Finally, is another Moderna candidate, mRNA-1647, a vaccine against cytomegalovirus (CMV). The vaccine contains 6 mRNA species, 5 of which encode a membrane bound complex, with the 6th encoding CMV glycoprotein. mRNA-1647 already has been shown to be more

effective than previous candidates (gB/MF59 vaccine – 50% efficacy), with phase 3 trials expected to conclude in late 2025 (Hu et al., 2023).

Table 1.1 (overleaf) – mRNA vaccines against infectious diseases in clinical trials. Adapted from (Vavilis et al., 2023a).

Vaccine Name	Infectious disease targeted	Clinical Trial Phase
Autologous dendritic cells electroporated with WT1 mRNA	Leukemia	2
BG505 MD39.3 mRNA, BG505 MD39.3 gp151 mRNA, and BG505 MD39.3 gp151 CD4KO mRNA	HIV	1
BNT111	Melanoma	2
BNT113	Head and Neck Cancer	1-2
Core-g28v2 and eOD-GT8	HIV	1
DC-006 vaccine loaded with amplified cancer stem cell mRNA	Ovarian Cancer	1-2
Dendritic cells loaded with mRNA	Prostate Cancer	1-2
eOD-GT8 60mer mRNA	HIV	1
Human CMV pp65-LAMP mRNA-pulsed autologous dendritic cells	Glioblastoma	2
Individualized Cancer RNA Immunotherapy (IVAC®) vaccines: IVAC_W_bre1_uID and IVAC_W_bre1_uID/IVAC_M_uID	Personalised Cancer Vaccine	1
Influenza vaccines (mRNA-1020, mRNA-1030, and mRNA-1010)	Influenza	1-2
Langerhans-type dendritic cells with mRNA	Melanoma	1
Lipo-MERIT	Melanoma	1
mRNA -1215	Nipah Virus	1
mRNA-1189	Epstein-Barr virus	1
mRNA-1345	RSV	2-3
mRNA-1345 and mRNA-1273.214	RSV	3
mRNA-1647	Cytomegalovirus	2
mRNA-1653	Metapneumovirus and Parainfluenza	1
mRNA-1893	Zika Virus	1
mRNA-4157	Melanoma	2
National Cancer Institute (NCI)-4650	Personalised Cancer Vaccine	1-2
RNA-lipid particle (RNA-LP) vaccines	Adult Glioblastoma	1
Self-amplifying ribonucleic acid (saRNA) vaccines (PF-07852352, PF-07836391, PF-07836394, PF-07836395, PF-07836396, and PF-07867246)	Influenza	1
W_ova1 vaccine	Ovarian Cancer	1
WT1 mRNA-loaded autologous monocyte-derived dendritic cells	Glioma	1-2

1.1.2 - Synthetic mRNA as a cancer immunotherapy

Cancer immunotherapies using mRNA have now been in development for over 25 years (Beck et al., 2021; Boczkowski et al., 1996). Cancer vaccinations function by inducing a T-cell response in a tumour specific manner, through encoding tumour associated antigens or antigen receptors (Beck et al., 2021). One approach for achieving this is ex-*vivo* delivery, where dendritic cells, an antigen presenting cell type, are loaded with tumour antigen by mRNA delivery, and then reintroduced back into the patient (Perez and De Palma, 2019). A second ex-*vivo* approach is to introduce chimeric antigen receptor (CAR) mRNA to T cells. CAR-T cells are able to redirect patient T cells to target and kill cancer cells upon reintroduction to the patient (Lin et al., 2021a). *In vivo* therapy has also been shown to be feasible, by delivery of CAR or T cell antigen receptor mRNA directly to the patient. The main challenge with the approach is ensuring the mRNA reaches the T cells at required levels. Antibody or small molecular ligand targeting has been used to enhance this approach (Cafri et al., 2020; Chung et al., 2020).

A multitude of candidate mRNA cancer immunotherapies have entered clinical trials over the last 15 years (Beck et al., 2021). Many recent trials are in combination with antibodies or chemotherapy drugs for improved overall efficacy (Kong et al., 2023). 23 mRNA cancer vaccines have entered clinical trials in the last 5 years, with the field dominated by BioNTech and Moderna. A wide variety of cancers are being targeted, including breast cancer, melanoma, and glioblastoma (Hong et al., 2020; Piper et al., 2021; Sahin et al., 2020). One such candidate therapy has entered phase 3 trials, a potential vaccine for glioblastoma. Glioblastoma stem cells (GSC) are targeted by dendritic cells loaded with GSC mRNA, in combination with survivin, an anti-apoptotic

peptide, and human telomerase reverse transcriptase. Progression free survival rates in patients have been improved 3-fold in early trials compared to the current treatment, bevacizumab (Piper et al., 2021).

1.1.3 - Synthetic mRNA as a protein replacement therapy

Synthetic mRNA has been posited as a potential means of protein replacement therapy, whereby a deficient or absent protein in a patient can be supplemented from a recombinant source. This has traditionally been achieved through administering purified protein, however this is costly to produce, and the recombinant protein is rapidly degraded within the cell (Vavilis et al., 2023a). The use of nucleic acids as a means of encoding replacement proteins has consequently been investigated extensively.

DNA can be delivered through either viral or non-viral delivery methods. Viral delivery methods, such as adeno-associated viruses, use a viral capsid to carry the DNA to the target cell, where it can integrate into the genome and be expressed (Bulcha et al., 2021). As with all integrative methods, the risk of oncogenesis is high, and in some cases, neutralising antibodies emerging during prior viral infection can reduce the effectiveness of such treatments (Vavilis et al., 2023a). Non-viral delivery of plasmid DNA has an improved safety profile, however transfection efficiency is low, reducing the dosage of the therapy (Ramamoorth and Narvekar, 2015). The negligible risk of insertional mutagenesis when using mRNA, combined with higher transfection efficiencies due to there being no requirement to enter the nucleus, make it an

attractive molecule for protein replacement therapy. This is not without challenges, however, due to the reduced stability of mRNA when compared with DNA.

Several active phase 1 and 2 trials have proceeded using mRNA encased in LNPs as a protein replacement. Moderna are running trials with candidate drug mRNA-3927, for the treatment of propionic acidemia. mRNA-3927 encodes two subunits of the mitochondrial enzyme propionyl-CoA carboxylase (PCC) (Attarwala et al., 2023). Early data from studies involving small numbers of patients showed a reduction in serious symptoms after biweekly dosing. Several other phase 1 or 2 have concluded without significant positive results. AstraZeneca's candidate AZD8601, encoding vascular endothelial growth factor A (VEGF-A), was in development as a treatment for heart failure and type II diabetes (Anttila et al., 2023). Despite promising early results, the drug was removed from AstraZeneca's pipeline during phase 2 trials in 2022. Translate Bio have recruited patients for 2 mRNA candidates. The first, MRT5201, encodes ornithine transcarbamlyase for the treatment of ornithine transcarbamlyase deficiency, however trials were withdrawn before any patients were injected (Vavilis et al., 2023a). Trials did proceed for the second candidate drug MRT5005, encoding cystic fibrosis transmembrane regulator protein for the treatment of cystic fibrosis, however no beneficial effect was seen from the treatment (Rowe et al., 2023).

1.1.4 - Synthetic mRNA for genome engineering and reprogramming

Genome editing is a method whereby programmable nucleases are delivered to cells in order to precisely alter the DNA sequence, in order to treat genetic diseases. These nucleases can be grouped into four classes: meganucleases (MNs), zinc-finger nucleases (ZFNs), transcription activator-like effector nucleases (TALENs) and

clustered regularly interspaced short palindromic repeat (CRISPR) associated nucleases (Zhang et al., 2019b). Genome editing nucleases specifically recognise target DNA sequences, and cause double strand breaks which can then be repaired by non-homologous end joining (NHEJ), or homology directed repair (HDR), causing the insertion of a donor DNA sequence (Ceccaldi et al., 2016). MNs, TALENs and ZFNs rely on protein-DNA interactions for nuclease targeting, whilst the CRISPR-associated method uses a guide RNA sequence to target specific regions of the genome (Gaj et al., 2013). Whilst DNA or protein based transfer of genome editing nucleases have been the predominant methods of delivery, the use of mRNA for transient nuclease expression is becoming more prevalent (Zhang et al., 2019b).

Clinical trials have proceeded using mRNA as a method of genome engineering. In one example, Sangamo Therapeutics attempted to treat HIV (human immunodeficiency virus) through the deletion of the HIV-1 co-receptor CCR5 by ex-*vivo* delivery of ZFN encoding mRNA (DiGiusto et al., 2016). Long terms trials are still ongoing for this genome editing mRNA, in addition to a version targeting T-cells, with initial results being promising in terms of reducing the HIV viral load long term (Zeidan et al., 2021). CRISPR-Cas9 components delivered through lipid nanoparticle encapsulated mRNA, developed by Intellia Therapuetics, are also in clinical trials for the treatment of a number of liver diseases, such as primary hyperoxaluria type I (PH1), and transthyretin amyloidosis (ATTR). A single administration of LNP-mRNA encoding CRISPR-Cas9 components targeting the transthyretin gene was able to achieve a knockdown in protein levels of >97% after 12 months in mice (Finn et al., 2018). Numerous other mRNA based genome engineering strategies are in the preclinical stage, targeting diseases such as glycogen storage disease, acute

intermittent porphyria, galactosemia, and muscular dystrophy (Cao et al., 2021; Jiang et al., 2018a; Roseman et al., 2018; Stadelmann et al., 2022).

Synthetic mRNA has also been used in the context of the reprogramming cell fate, whereby external stimuli can be used to alter cellular phenotype. This was first achieved with mRNA encoding Yamananka stem cell factors, allowing for the reprogramming of multiple human cell types to induced pluripotent stem cell (iPSCs) (Takahashi and Yamanaka, 2006; Warren et al., 2010). More recently this strategy has been used to generate antigen specific $CD8^+$ T cells, hepatocyte like cells, and tumour-associated macrophages (Garcia-Llorens et al., 2023; de Mey et al., 2023; Zhang et al., 2019a). Reprogramming cell fates using synthetic mRNA has proven especially effective, due to high transfection efficiencies and no risk of genomic integration. Complex mixtures of mRNAs can be simultaneously delivered, such as in the case of $CD8^+$ T cell generation where a mix of four mRNAs are used – toll-like receptor 4, CD70, interferon gamma, and interleukin-10 receptor α (de Mey et al., 2023).

1.1.5 - Other functions of synthetic mRNA in biotechnology

Whilst the vast majority of synthetic mRNA research is focused in the therapeutic field, a number of applications have been developed using the molecule in a biotechnology or fundamental research context. Firstly, synthetic mRNA devices have been developed that bind to endogenous proteins and allow for distinguishing between different cell types. In one example, mRNA is engineered to bind in its 5' untranslated region (UTR) region to LIN28A, a protein marker of iPSCs. Binding of LIN28A prevents translation of encoded GFP, meaning that in fully differentiated cells a fluorescent signal is seen post-mRNA transfection, whereas cells with the iPSC phenotype

produce no signal (Kawasaki et al., 2017). In addition, synthetic mRNA devices are described in the literature in the context of gene circuits, that selectively activate translation based upon levels of miRNA that are present endogenously (Kameda et al., 2023; Tan et al., 2023). Such mRNA devices have function in the field of biopharmaceutical production in mammalian cells, whereby the gene circuit allows for upregulation of protein expression at defined points in a manufacturing process, such as a temperature shift in fed-batch culture (Donaldson et al., 2022).

In addition to the gene circuits described above, other uses of synthetic mRNA in the field of biopharmaceutical manufacturing are described in the literature, in the context of Chinese hamster ovary (CHO) monoclonal antibody (mAb) production. Large amounts mRNA can be transfected in to CHO cells, to replicate conditions that are present in high productivity manufacturing scenarios (Bydlinski et al., 2020; Coats et al., 2020). Using mRNA transfection as a means of mimicking a high productivity scenario allows for the detection of rate limiting steps in the production of a mAb, without needing to go through the costly and time consuming process of generating stable cell lines for initial screening experiments. Chapter 5 in this thesis describes a further method for identifying rate limiting steps in CHO cell mAb production. The focus of this chapter however, is on the elucidation of rate limiting steps in stable cell lines that have already been generated, where product titre may be lower than anticipated.

1.2 - The structure & optimisation of functional synthetic mRNA

Conventional linear synthetic mRNA has several key features, essential for the functionality of the molecule *in vivo* (Gote et al., 2023; Sahin et al., 2014). Firstly, the 5' cap structure is essential for translation in eukaryotic cells, and a range of cap analogs have been developed in order to improve capping efficiency and mRNA performance (Kocmik et al., 2018; Su et al., 2011). The 5' and 3' untranslated regions (UTRs) are also critical to mRNA function, with secondary structure in these sequences conferring stability to the molecule by preventing degradation (Asrani et al., 2018; Leppek et al., 2018, 2022). The coding region of the mRNA has also been engineered to improve function, with codon optimisation algorithms being employed to increase both translation rate and stability (Presnyak et al., 2015; Zhang et al., 2023). Finally, synthetic mRNA features a poly-A tail region, normally comprising 120-150 nucleotides, which functions as a platform for poly-A binding proteins, and prevents degradation of the mRNA molecule (Tavernier et al., 2011). The design of synthetic mRNA is crucial to its function *in vivo.* Multiple strategies for improving the design of synthetic mRNA have been described in the literature, with a view to improving the translatability, stability, or immunogenicity of the molecule.

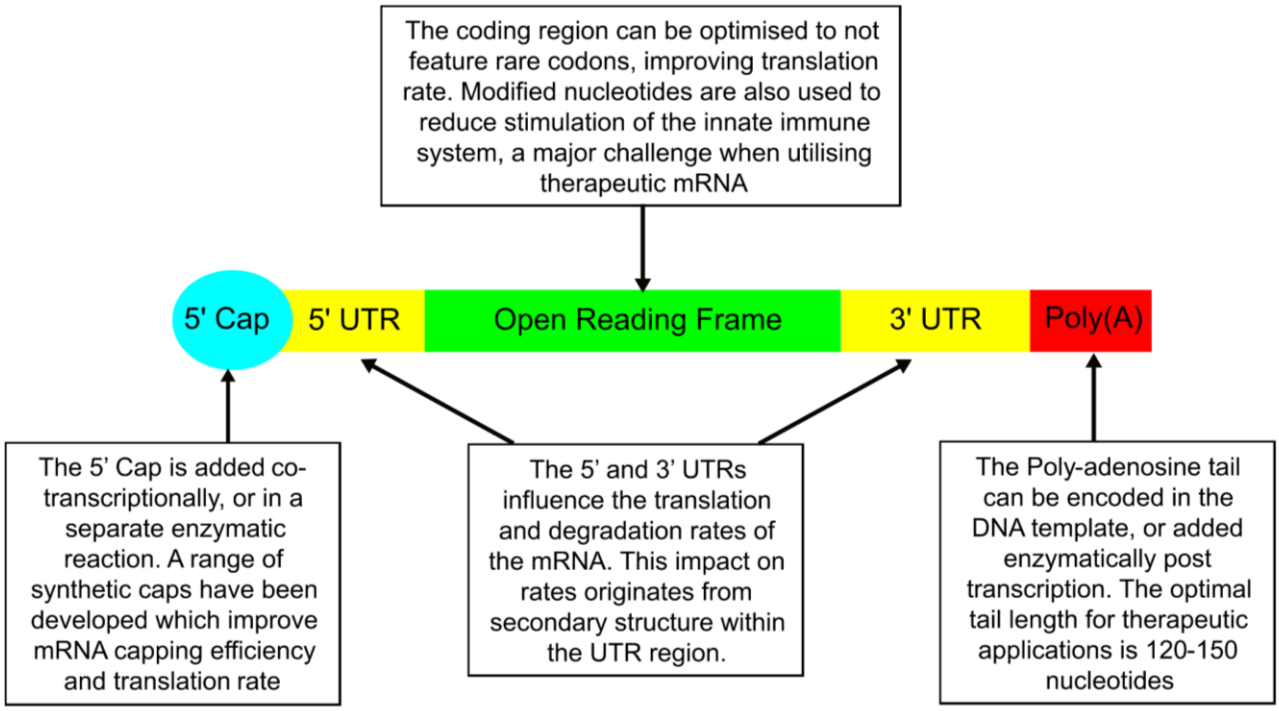


Figure 1.2 – Features critical to the function of synthetic mRNA *in vivo*. The 5' cap, 5' and 3' untranslated regions, open reading frame and poly(A) tail all play a role in allowing the mRNA molecule to be translated upon introduction to the cell. As described in section 1.2.1, development of each of these features allows for increased product dosage due to increased rate of translation, and enhanced mRNA stability.

1.2.1 - Features of Conventional Linear Synthetic mRNA

1.2.1.1 - The 5'-Cap Structure

A 5' cap is found on nearly all eukaryotic mRNA, with the naturally occurring cap having a m^7GpppNmpNNm structure (also known as Cap-0) (Furuichi, 2015). The cap is essential to a variety of mRNA processing events, including the initiation of translation. The cap interacts with the initiation factor eIF4E – a process which is rate limiting and essential for efficient protein synthesis (Carroll and Borden, 2013). Therefore, when producing functional synthetic mRNA, a cap structure must be incorporated into the mature molecule. Two strategies exist for the addition of the 5'

cap. Firstly, the cap may be added co-transcriptionally by providing the cap dinucleotide in the transcription reaction, alongside the four nucleoside triphosphate (NTPs) (Su et al., 2011). One of the major challenges when using a co-transcriptional capping is that the RNA polymerase is able to incorporate the cap structure in the reverse orientation, producing Gpppm^7GpNpN instead of m^7GpppGpNpN. This occurs as nucleophilic attack of the α-phosphate of the transcript is equally likely to be initiated by either guanosine at the ends of the cap structure (Kocmik et al., 2018). The result of this is a pool of capped transcripts, of which approximately half are capped in the reverse orientation, rendering them unable to interact with eIF4E, significantly reducing the rate of protein synthesis (Rhoads, 2016).

To address the issue presented by reverse cap incorporation, anti-reverse cap analogs (ARCAs) have been developed that feature m^7-guanosine modifications including a 2'-O-methyl and 3'-O-methyl group that ensures correct cap orientation (Jemielity et al., 2003; Shanmugasundaram et al., 2022; Stepinski et al., 2001). Various additional modifications have been made to ARCAs, such as the introduction of methylene groups between phosphate moieties, with a view to shielding the transcript from Dcp1/Dcp2 decapping enzymes (Grudzien-Nogalska et al., 2007). Improvements to translation efficiency have also been observed when incorporating benzyl or 4-methoxy-benzyl at the N2 position of the cap analog (Kocmik et al., 2018). A further consideration that must be made when co-transcriptionally capping the mRNA is the ratio of cap analog to GTPs in the transcription reaction. The cap analog and guanosine compete for the first position at the 5' end of the mRNA, resulting in the presence of a population of uncapped, and therefore translationally inactive mRNA. It is therefore necessary to use an excess of cap analog in proportion to guanosine, with a ratio of 4:1 being commonly utilised (Hadas et al., 2019; Loomis et

al., 2018). When using such ratios, 20% of transcripts will remain uncapped, reducing overall product quality. More recently, novel technologies such as CleanCap, developed by TriLink, has been shown to achieve in excess of 98% capping efficiency, whilst adding a cap-1 structure co-transcriptionally, as shown in Figure 1.3 (Henderson et al., 2021; Senthilvelan et al., 2023). Cap-1 is a commonly utilised modified cap structure, featuring methylation of the 2'-O of the ribose on the nucleotide which immediately follows the cap (Ghosh and Lima, 2010). CleanCap is a trimeric molecule, which binds to the +1 and +2 nucleotides of the transcription template through complementary base pairing, followed by incorporation of the complementary NTP at the +3 position (Henderson et al., 2021).

The alternative strategy to co-transcriptional capping is to cap the transcript enzymatically, post-transcription. This reaction proceeds utilising a Vaccinia virus derived capping enzyme, adding a cap-1 structure to the mRNA (Zarghampoor et al., 2019). The Vaccinia virus capping enzyme is a heterodimeric complex of a 97 kDa D1 protein and 31 kDa D12 protein (Fuchs et al., 2016). All the catalytic activity of the enzyme is localised to the D1 subunit, which features RNA triphosphatase, guanylyltransferase and methyltransferase activity. D12 acts as an allosteric stimulator of the D1 methyltransferase activity (Kyrieleis et al., 2014). More recently, a Faustovirus capping enzyme system has become commercially available, and displays activity at a broad range of temperatures, on mRNA substrates with higher levels of secondary structure (Chan et al., 2023). These enzyme complexes can be readily overexpressed and purified, or purchased from a commercial vendor, however the amounts of enzyme required to produce high levels of capped mRNA render the process costly (Fuchs et al., 2016). Furthermore, post transcriptional capping requires

an extra step in the mRNA synthesis process when compared with co-transcriptional capping, increasing manufacturing timescales.

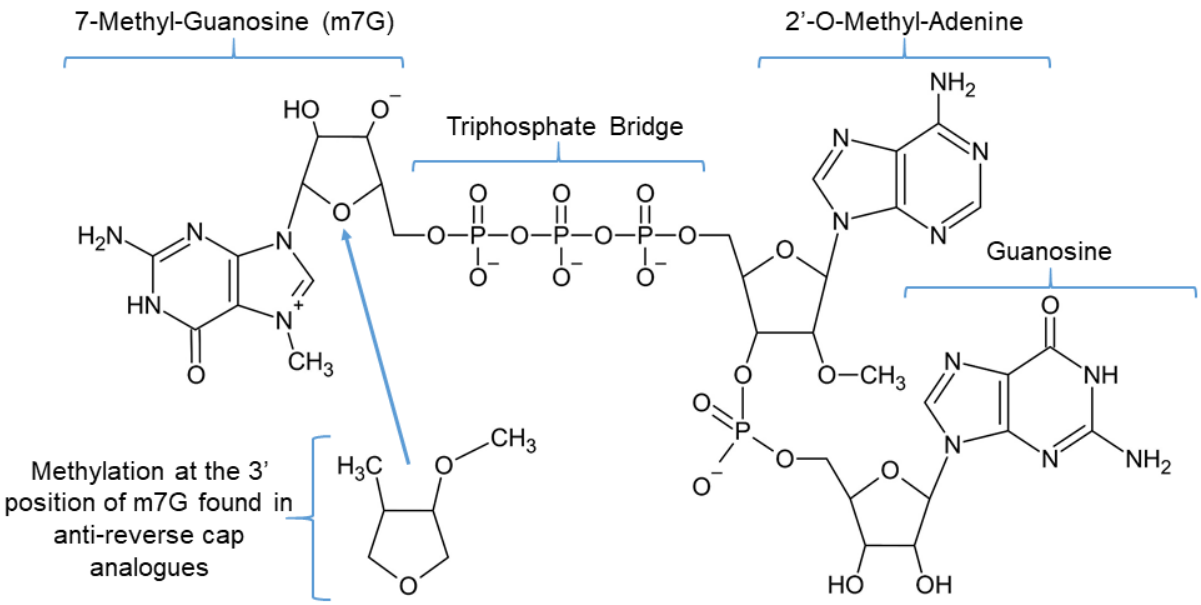


Figure 1.3 – The structure of the Trilink CleanCap AG reagent, commonly used in mRNA manufacture, achieving capping efficiencies >98%. The 7-methylguanosine structure is attached through a 5' O to 5' O linked triphosphate bridge to the first adenine residue. This adenine residue is 2'-O methylated, forming a cap-1 structure that allows for the evasion of the cellular innate immune response, and is followed by a guanosine residue. The Clean cap reagent binds to the first and second transcribed nucleotides on the DNA template, acting as the site therein of transcription initiation for the addition of the third nucleotide. Also shown is the site of methylation found in anti-reverse cap analogues (ARCAs), which ensure incorporation of the cap analogue in the correct orientation for translation.

1.2.1.2 - 5’ and 3’ Untranslated regions

5’ and 3’ UTR regions are known to be key determinants of the translatability and stability of mRNA molecules *in vivo*. Not contributing to the protein coding sequence, the 5’ and 3’ UTRs can feature extensive canonical and non-canonical base pairing and complex secondary structure, able to influence protein interactions and ribosome recruitment (Leppek et al., 2018). The average human 5’ UTR ranges from between 100-200 nucleotides, with the 3’ UTR being considerably longer at approx. 800 nucleotides (Pesole et al., 2001). The length of the 3’ UTR is critically important, with shorter 3’ UTRs having reduced translational efficiency, and longer 3’ UTRs often being degraded more rapidly (Orlandini von Niessen et al., 2019). Utilising UTRs in the production of synthetic mRNA has focused on incorporating sequences that increase the translatability and stability of the molecule. For this purpose, the UTRs of α and β-globin mRNAs have often been used due to the well-established high stability and translation rate of such transcripts (Ross and Sullivan, 1985). Indeed, the Pfizer/BioNTech SARS-nCov2 vaccine utilises the human α-globin 5’ UTR, whilst the Moderna Covid vaccine features 110 nucleotides of the human α-globin 3’ UTR (Xia, 2021). Further enhancements to stability have been observed when two consecutive human beta-globin 3’ UTR sequences have been utilised in the production of antigen encoding mRNA (Holtkamp et al., 2006).

Various additional 5’ and 3’ UTR sequences have now been screened for their ability *in vivo* to modulate stability and translation efficiency (Kim et al., 2022b). Recently, a cell culture based systematic evolution of ligands by exponential enrichment (SELEX) method has been used to select 3’ UTR motifs displaying favourable properties (Orlandini von Niessen et al., 2019). Several of the best performing sequences in this study originated from mitochondrial mRNAs featuring a low number of predicted

miRNA binding sites – a feature believed to increase translation rate (Jackson et al., 2010). Furthermore, attempts to construct a minimal 5' UTR have yielded a 14 nucleotide sequence containing a consensus Kozak sequence and T7 promoter, that has shown higher expression levels than the human α-globin in multiple cell lines (Trepotec et al., 2019a). Deep learning algorithms now allow for optimisation of UTR sequences that mimic sequence patterns seen in endogenous human UTRs (Gong et al., 2023). In certain cases, if the mRNA is being used for immune stimulation, for example, it may be advantageous for mRNA stability to be low, leading to the mRNA having a short half-life *in vivo*. For this purpose, AU rich elements (AREs) have been used which are able to recruit ARE binding proteins which upregulate degradation (Kramer and Carrington, 2014).

1.2.1.3 – mRNA Coding Sequence Optimisation

The degenerate nature of the genetic code provides a means to optimise the coding region of an mRNA sequence, both in terms of translation efficiency and stability, by substituting synonymous codons. Codon optimality is determined by the numbers of charged transfer RNA available, and the demand for each tRNA from translating ribosomes. Optimal codons are more readily available in the cytoplasmic pool, and can therefore be decoded faster reducing delay to translation elongation (Novoa and Ribas de Pouplana, 2012). Increased protein expression when using optimised codons has been widely reported in the production of recombinant biotherapeutics, with a >1000-fold increase reported in one case (Mauro, 2018). Additionally, it has been demonstrated that replacement of non-optimal codons with optimal codons leads to a significant increase in mRNA stability. mRNAs with less than 40% optimal codons

were shown to have a median half-life of 5.3 minutes, in contrast to 20.1 minutes when the optimal codons level was above 70% (Presnyak et al., 2015).

Whilst it is clear that in many cases there are obvious benefits to optimising codons based on cytoplasmic tRNA availability, careful consideration must be used when doing so for. The usage of certain codons in mRNA can lead to the formation of secondary structure, which has a profound effect on elongation rate. An example is RD114-TR retroviral envelope protein, which when codon optimised produces an inactive protein, as a result of insufficient glycosylation due to rapid elongation (Zucchelli et al., 2017). Furthermore, codon optimisation has been shown to impact proper protein folding, leading to inactive protein, or aggregation. One such example is MDR1 gene, whose protein product, P-glycoprotein, folds incorrectly after codon optimisation, and is unable to insert into the cell membrane (Kimchi-Sarfaty et al., 2004). Algorithms have been developed to overcome such issues, that can identify regions of slowly translated mRNA, and recommend codon replacements less likely to have deleterious impacts on the protein product in a process known as ‘codon harmonisation’ (Angov et al., 2008; Papamichail et al., 2018).

The most cutting edge mRNA design algorithms allow for the cooperative optimisation of both codon usage and mRNA stability through increased levels of inherent secondary structure, producing mRNA sequence that is both rapidly translated, and has an extended half-life (Leppek et al., 2022; Zhang et al., 2023). It is important that levels of secondary structure are balanced in order to not impede the cellular translational machinery, whilst still conferring stability to the molecule. The recently described LinearDesign algorithm, developed by Zhang and colleagues, has been shown to output mRNA sequences that increase antibody titre by 128-fold in mice (Zhang et al., 2023).

1.2.1.4 - Poly(A) Tail

The polyadenosine (Poly(A)) tail is an essential feature of eukaryotic mRNA, enhancing translation and mediating the prevention of mRNA degradation (Weill et al., 2012). The poly(A) tail acts as a platform for poly(A)-binding protein (PABP), with PABP able to interact with eIF4G at the 5' end, activating the translation of the mRNA (Machida et al., 2018). The optimal tail length for synthetic therapeutic mRNA has been shown to be between 120-150 nucleotides (Sahin et al., 2014). When producing IVT mRNA, the tail can be encoded in the DNA template, or added post-transcriptionally by poly(A) polymerase. Enzymatic tail addition by poly(A) polymerase allows for the addition of modified nucleotides, such as radiolabelled cordycepin (Martin and Keller, 1998). This method of tail addition, however, is limited in its utility as enzymatic polyadenylation produces a heterogeneous population of tail lengths. For this reason, encoding the poly(A) tail in the DNA template is preferred in the production of therapeutic mRNAs, as tail length can be more closely controlled (Sahin et al., 2014). One potential issue when using a template encoded poly(A) tail is that plasmids encoding extensive poly(A) regions are difficult to maintain in *E. coli*, due to rapid shortening of the tail through recombination. To overcome this, a segmented tail containing spacer sequences that do not reduce translation efficiency is used, which is more easily maintained in *E. coli* (Trepotec et al., 2019b).

1.2.1.5 - Modified Nucleotides

One of the most prominent issues when using mRNA for therapeutic applications is the upregulation of the innate immune system through Toll-like receptors (TLRs) (Karikó et al., 2008; Lee et al., 2023b; Takeda et al., 2003). TLRs are capable of recognising molecular patterns and motifs related to pathogens, coordinating the immune system in response (Karikó et al., 2005). A strategy for the reduction of such

immune upregulation is the incorporation of derivatives of naturally occurring modified nucleotides into the mRNA molecule. Modified nucleotides avoid upregulation of certain components of the innate immune system by mimicking endogenous mRNA (Freund et al., 2019). Eukaryotic mRNA contains a considerably larger number of modified nucleotides when compared to prokaryotic mRNA, therefore allowing the innate immune system to recognise endogenous mRNA based upon evolutionary origin (Karikó et al., 2005). Nucleotide modifications can cause steric clashes in the binding regions of innate immune system components, preventing recognition of the mRNA. For example, 2'-O-methylated RNA has been shown to cause steric exclusion from the Retinoic acid-inducible gene I (RIG-I) binding pocket, preventing downstream responses (Hyde and Diamond, 2015).

Examples of modified nucleotides include pseudouridine, 2-thiouridine and 5-methylcytidine (Hadas et al., 2019). The incorporation of the aforementioned modified nucleosides into mRNA has been shown to reduce, or completely inactivate, the activation of TLR3, TLR7 and TLR8 (Karikó et al., 2005). Furthermore, the use of pseudouridine (Ψ) has been shown to reduce activation of the RNA-dependent protein kinase (PKR), a translational repressor (Anderson et al., 2010). Retinoic acid inducible gene I (RIG-I) activity can also be ablated when N-6-methyladenosine or pseudouridine are utilised (Durbin et al., 2016). Approved mRNA vaccines, such as those against SARS-CoV-2, utilise N1-methylpseudouridine (m1Ψ) in place of uridine (Nance and Meier, 2021). m1Ψ elicits even less immunogenicity than Ψ, whilst also increasing that rate of translation (Kim et al., 2022a; Svitkin et al., 2017).

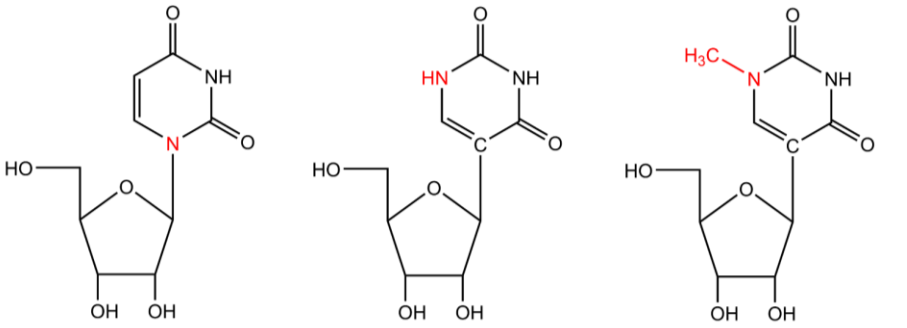


Figure 1.4 – The structure of uridine, and the two most commonly utilised modified nucleotides found in place of uridine in mRNA therapeutics, psuedouridine (Ψ) and N1-methyl-psuedouridine ($N^1\Psi$). Ψ was originally shown to dramatically reduce the levels of innate immune stimulation *in vivo*, with m1Ψ reducing the levels of immune activation further, whilst also increasing the rate of translation. It is for these reasons that approved mRNA therapeutics are currently fully m1Ψ modified.

1.2.2 Alternative synthetic mRNA product formats

1.2.2.1 Self-Amplifying mRNA

Conventional linear mRNA, used in current approved mRNA therapeutics, is non-replicative. The mRNA molecule is introduced to the target cell, causes translation of the antigen or therapeutic protein, as is degraded by endogenous RNA turnover pathways (Sahin et al., 2014). With conventional linear mRNA, target protein expression is directly proportional to the amount of mRNA that successfully enters the cell. In contrast to this are self-amplifying mRNA species (saRNA), which encode viral replication machinery, in addition to the open reading frame (ORF) of interest (Blakney et al., 2021). This viral replication machinery leads to intracellular amplification of the mRNA from a subgenomic promoter, meaning that a small amount of mRNA introduced to the cell leads to high levels of expression (Vogel et al., 2018).

The replication machinery from alphavirus genomes encoding an RNA-dependent RNA polymerase (RdRp) is commonly used for saRNA amplification. Examples include Venezuelan equine encephalitis virus (VEE), Sindbis virus (SINV), and Semliki forest virus (SFV) (Bloom et al., 2021). The replication machinery, comprising non-structural proteins 1-4, is encoded on a long transcript upstream of the target protein sequence. The encoded RdRp is able to drive transcription of the target protein mRNA through transcription from the corresponding alphavirus subgenomic promoter seqeunce (Rupp et al., 2015). A COVID-19 vaccine is in development using saRNA technology, however this is not without technical challenges (Maruggi et al., 2022). The considerable length of saRNA (approx. 9-12 kb) makes it challenging to manufacture, store and introduce efficiently to cells (Blakney et al., 2021). Despite this, saRNA remains a highly promising therapeutic modality, due to the considerably lower required dosage to elicit the same response as conventional mRNA.

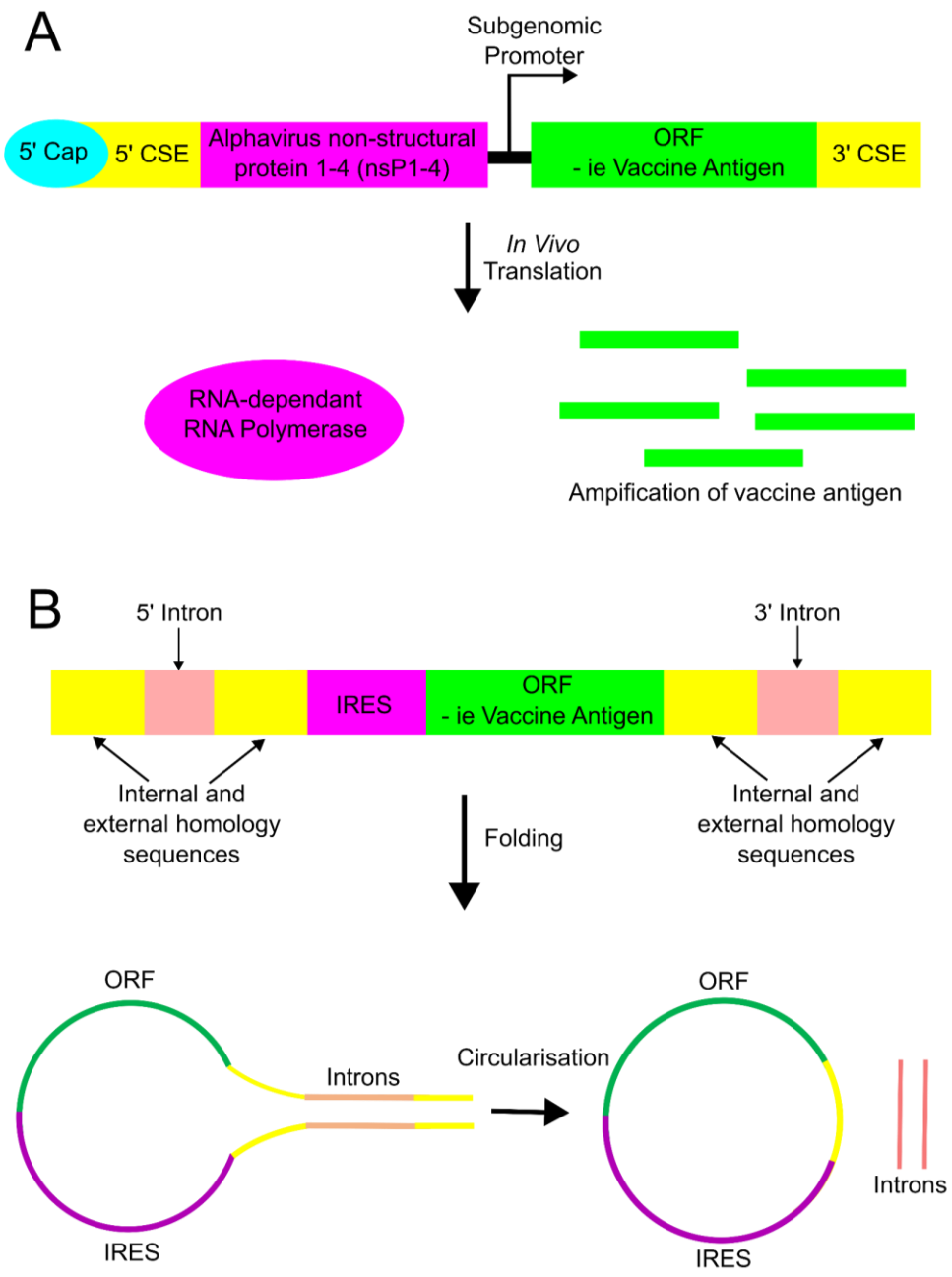


Figure 1.5 – The structure of novel mRNA formats. A) Self-amplifying mRNA is transfected as a single long mRNA transcript, encoding alphavirus non-structural proteins 1-4 (nsP1-4), and the ORF of interest. Conserved sequence elements (CSEs) are also encoded, which are essential for viral replication. nsP1-4 encodes a RNA dependent RNA polymerase (RdRp) which is able to amplify the ORF from a subgenomic promoter *in vivo*. B) One method for RNA circularisation is by the permuted intron-exon method, whereby intronic sequences can self-catalyse the covalent ligation of the RNA molecule into a closed circle. Shown here is a schematic of optimised sequences for this purpose, which feature regions of internal and external homology which bring the intron sequence together in a more efficient manner.

1.2.2.2 Circular mRNA

Circular RNA is drawing increasing interest as a potential therapeutic format. Circular RNA species occur in nature in eukaryotic cells as a result of back splicing, and have been shown to have regulatory function within the cell (Ebbesen et al., 2017; Salzman, 2016). The closed loop nature of circular RNAs confer a key advantage in the context of synthetic mRNA; they are completely resistant to exonuclease degradation due to not having exposed 5' and 3' termini, and are therefore more stable *in vivo* (Seephetdee et al., 2022; Wesselhoeft et al., 2018). Additionally, it has been shown that circular RNA elicits considerably less immune stimulation than linear mRNA, even without the use of modified nucleotides (Liu et al., 2022a; Wesselhoeft et al., 2019)

Several methods exist for producing synthetic circular RNA. Firstly, T4 RNA ligase can be used to circularise a linear precursor molecule *in vitro,* often using a DNA oligo to bridge the ligation junction to improve the efficiency of the reaction (Costello et al., 2020; Petkovic and Müller, 2015). Secondly, and most commonly utilised, is the permuted intron-exon method, which uses the activity of group I self-splicing introns to ligate the RNA molecule (Wesselhoeft et al., 2018). Such group I self-splicing introns are found in T4 bacteriophage thymidylate synthetase gene, or pre-tRNA from the cyanobacteria *Anabaena* (Rausch et al., 2021). This method only requires the presence of Mg^{2+} and GTP for the self-ligation of the molecule *in vitro* or *in vivo* (Meganck et al., 2021). It is also possible to use group II self-splicing introns, however their activity is less well characterised (Obi and Chen, 2021). 'Twister' ribozyme sequences have also been shown to function in a similar manner, circularising short RNA aptamers (Litke and Jaffrey, 2019).

The lack of 5'-cap structure on circular RNA means that translation must be driven through an internal ribosome entry site (IRES). Algorithms, such as DeepCIP, have

been developed in order to optimise the IRES sequence in a circular context (Zhou et al., 2023). The efficiency of circularisation by the permuted intron-exon method has also been improved, in order to yield a higher proportion of circular molecules from the input population of linear RNA. This is achieved through internal homology and spacer sequences that cause formation of a splicing bubble, whereby all required splicing machinery is in close proximity (Wesselhoeft et al., 2018; Zhu et al., 2022a). Utilising these optimised sequences, circular RNA vaccine candidates have now been described targeting SARS-CoV-2 (Qu et al., 2022), *Staphylococcus aureus* infection (Zhu et al., 2022a), and stimulating T cell responses in tissues for cancer immunotherapy (Amaya et al., 2023).

1.3 Methods for the production of synthetic mRNA

The overwhelming majority of synthetic mRNA is produced by *in vitro* transcription (IVT) – a well-established reaction, used for large scale manufacturing of mRNA of diverse size. A bacteriophage RNA polymerase is used in a cell free system to produce the mRNA in a run-off transcription reaction utilising a linearized plasmid or PCR product as the DNA template (Sahin et al., 2014). IVT is inherently a simple reaction – the RNAP polymerase is combined with a magnesium containing buffer, the four NTPs, the DNA transcription template and cap analog (Beckert and Masquida, 2011). Excluding the 5' cap structure, all features of the mature mRNA can be encoded in the DNA template. Modified nucleotides can be incorporated into the mRNA molecule by providing the nucleotide in the reaction mixture in place of the corresponding ribonucleotide triphosphate. On a laboratory scale, mRNA can be quickly purified from the reaction by DNase digestion and RNA precipitation, yielding a functional mRNA molecule (Pardi et al., 2018).

1.3.1 Large scale manufacturing of synthetic mRNA

One of the key advantages of using *in vitro* transcription for large scale manufacturing is the scalability of the process – any yield can theoretically be achieved by increasing the amount of input material (Baronti et al., 2018). Plasmid DNA is the predominant transcription template used on large scale, as a consequence of its ability to be propagated in *E. coli* (Walker et al., 2003). The plasmid must be linearized by a restriction enzyme to allow for run-off transcription, producing a population of mRNA species with homogenous 3' ends. This step is currently necessitated due to the low termination efficiency of the endogenous T7 terminator sequence (approx. 70%), however the development of synthetic high efficiency terminator sequence may allow

for IVT without the need for plasmid linearization (Calvopina-Chavez et al., 2022; Mairhofer et al., 2015).

On large scale, the IVT reaction proceeds in inert reaction vessels, or single use reaction bags. As with small scale IVT, the reaction only utilises a small number of components, however these components must be GMP-grade, considerably increasing the cost of such materials (Rosa et al., 2021). As with small scale mRNA production, T7 remains the dominant RNAP used for commercial manufacture. RNase inhibitor is typically added to the reaction to preserve the integrity of the transcribed mRNA from any trace RNase contamination (Ouranidis et al., 2022). The reaction proceeds over the course of 1-3 hours, making the actual production process incredibly quick compared to many other biopharmaceutical products such as DNA vector based vaccines or monoclonal antibodies (Agostinetto et al., 2022; McElwain et al., 2022).

Critical to the manufacturing of any synthetic mRNA is the purification of the molecule to reduce the immunogenicity and off-target effects of the final formulation. Several purification unit operations are required, normally comprising a chromatography step such as oligo-d(T) affinity chromatography, followed by tangential flow filtration (Rosa et al., 2021). The purified mRNA requires encapsulation for efficient delivery to the target cell. This is most commonly achieved using liposome-derived lipid nanoparticles (LNPs). These are formed through precisely controllable microfluidic methods, whereby mRNA in aqueous buffer is mixed with cationic lipids dissolved in ethanol, forming inverted micelles around the mRNA (Hou et al., 2021). LNP formulation methods continue to develop, enabling more efficient cellular uptake of the mRNA, or incorporation of different chemical modalities which allow subcellular targeting of the LNP (Hou et al., 2021; Tenchov et al., 2021).

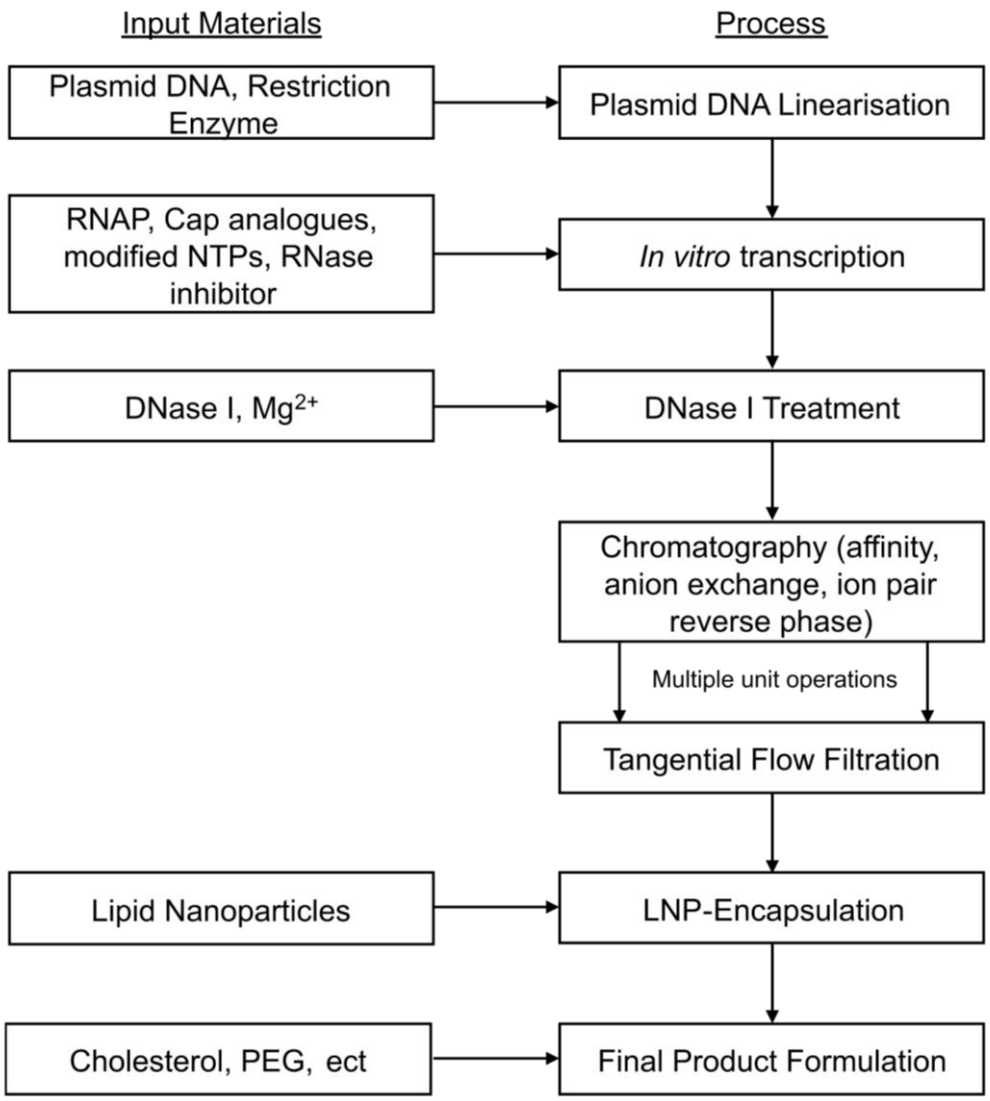


Figure 1.6 - A simplified process flow diagram for a typical mRNA manufacturing process. The production of mRNA by *in vitro* transcription (IVT) is a multistage process, involving the addition of large quantities of purified recombinant components such as cap analogues. Several purification unit operations are required for the final product formulation, due to the immunogenicity induced by contaminating RNA species such as double stranded RNA.

1.3.2 *In vitro* Transcription utilising T7 RNA Polymerase

As previously mentioned, the vast majority of *in vitro* transcription is catalysed by the DNA-dependent RNA polymerase from bacteriophage T7, a lytic phage of *E. coli* (Studier and Moffatt, 1986). T7 RNA Polymerase (T7 RNAP) is a member of the single subunit polymerase superfamily, which features representatives from many other bacteriophages, in addition to mitochondrial and chloroplast RNAPs (Sarcar and Miller, 2018). T7 RNAP is a 99 kDa protein, with a number of features that make it highly attractive for biotechnology applications such as *in vitro* transcription. Firstly, T7 RNAP is highly specific to its promoter, a 17 nucleotide double stranded DNA sequence, reducing the likelihood of off-target transcription (Shis and Bennett, 2014). Furthermore, T7 RNAP requires no additional transcription factors in order to initiate transcription. Finally, T7 RNAP is known to be a processive enzyme after initial transcription, with high fidelity, ensuring full length run-off products are accurately produced (Huang et al., 2000).

T7 RNAP has three functional domains, termed the ‘finger’, ‘palm’ and ‘thumb’ domain, which together enable the polymerase activity of the enzyme, in addition to four ‘accessory’ structural elements which have roles in recognising the promoter and melting the DNA duplex (Wang et al., 2018). The finger domain enables translocation of the polymerase along the DNA template by undergoing conformational changes (Yin and Steitz, 2004). The palm domain catalyses the polymerisation reaction, co-ordinating two magnesium ions proximal to two conserved aspartate residues in a cleft within the enzymes structure (Borkotoky and Murali, 2018). The thumb domain is flexible and is thought to prevent dissociation of the enzyme from the DNA template as it moves along it (Borkotoky and Murali, 2018).

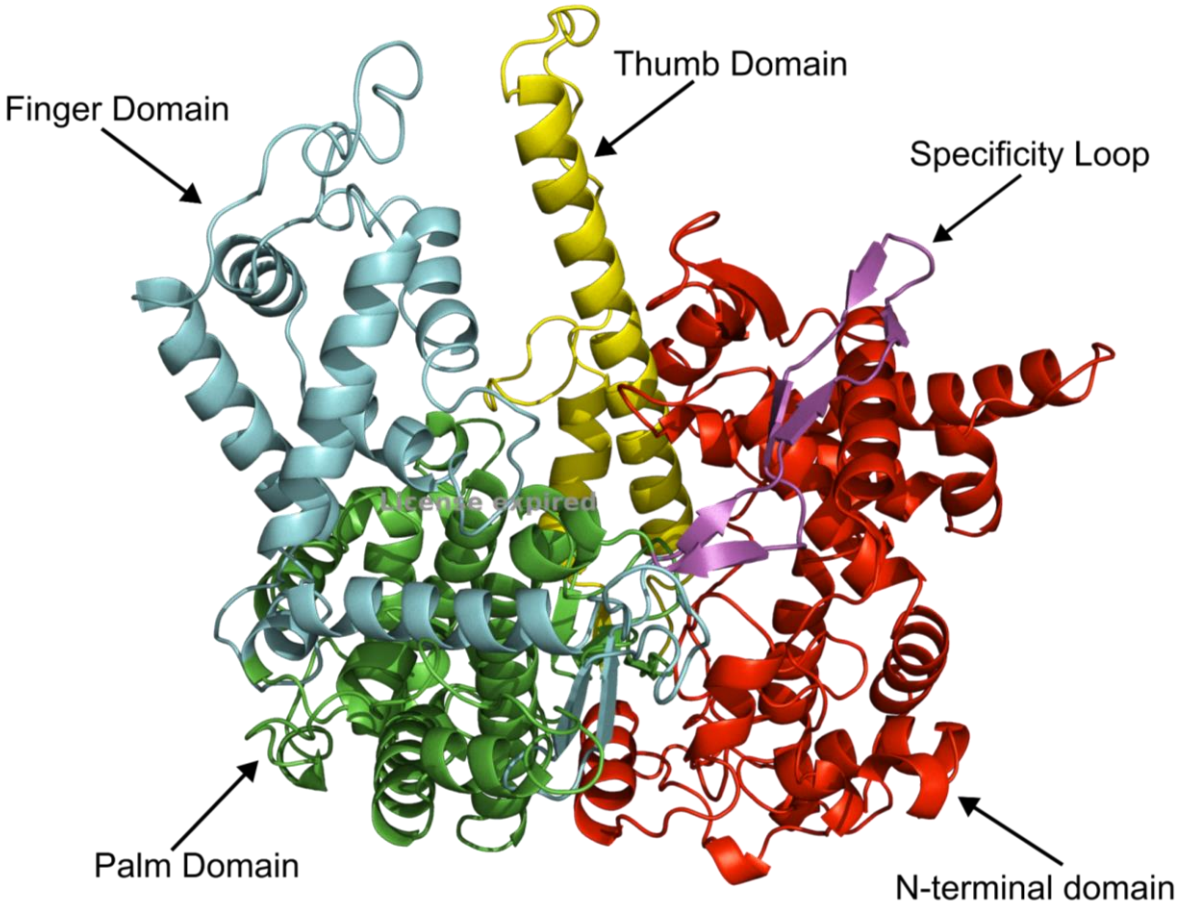


Figure 1.7 – The 3-dimensional structure of T7 RNA polymerase, as determined by X-ray crystallography (PDB deposition 1CEZ). Each of the domains is marked a different colour, with the N terminal domain red (1-325), thumb domain yellow (326-411), palm domain green (412-553 and 785-879), the finger domain blue (554-784), and the specificity loop magenta (740-769). α-helices are represented as spiral tubes, whilst β-strands are shown as arrows. The finger domain enables translocation of the polymerase. The palm domain catalyses the transcription reaction. The thumb domain stabilises the DNA template. The specificity loop is involved in promoter interactions. Image produced using PDB deposition 1CEZ.

1.3.3 Challenges when using T7 RNAP

Initial transcription is an unstable process – the polymerase must maintain a transcription bubble until a RNA-DNA hybrid can form of sufficient length to outcompete DNA reannealing (Gong et al., 2004). The initiation process comprises the addition of nucleoside monophosphates onto a first priming nucleoside until a RNA molecule of ~12 nucleotides is formed (Ramírez-Tapia and Martin, 2012). As a result of the unstable nature of initial transcript, a number of short abortive transcripts are released before the transition of the RNAP to elongation phase, most commonly between 2-6 nts (Ramírez-Tapia and Martin, 2012). The transfer of the polymerase to an elongation state occurs when the enzyme escapes its strong contacts with the promoter sequence.

In addition to the aforementioned abortive transcripts, T7 RNAP has been shown to facilitate the production of a number of RNA species after promoter release, which are shorter or longer than the template encoded product (Gholamalipour et al., 2018). Promoter-independent RNA synthesis by T7 can occur through several characterised mechanisms. Firstly, T7 can recognise certain RNA sequences as a replication template, self-replicating the RNA template in a manner analogous to viral RNA-dependent RNAPs (Konarska and Sharp, 1989). Furthermore, transcripts produced during the IVT reaction can ‘self-prime’, acting as a binding site for the RNAP, allowing for the extension of the RNA molecule (Gholamalipour et al., 2018). Two pathways are proposed for self-priming which initiates RNA-template mediated RNA synthesis. Cis self-priming occurs when regions proximal to the 3’ end of the RNA molecule exhibit complementarity to each other, leading to the formation of an intramolecular hairpin duplex region which T7 can bind and continue further transcription upon (Nacheva and Berzal-Herranz, 2003). Trans self-priming occurs when complementarity exists

between RNA transcripts, forming an intermolecular duplex upon which the RNAP can bind and initiate transcription (Gholamalipour et al., 2018). These pathways lead to production of long double stranded RNA molecules when using T7 RNAP. Such molecules are formed through promoter independent transcription from the 3' end of the DNA template, producing an antisense transcript with full complementarity to the expected RNA molecule, forming a double stranded duplex (Mu et al., 2018).

The number of off-target transcripts produced during IVT reactions using T7 RNAP is particular problem when producing mRNA therapeutics, where a homogenous sample is required to ensure a predictable response to the treatment. The scale of this issue was highlighted by Lu et al., whose RNA-seq data showed that when producing a 50nt RNA, only 0.99% of molecules were the correct length transcript (Lu et al., 2019). Double stranded RNA species are particularly immunogenic, due to their detection by cytosolic sensors RIG-I and MDA5, which subsequently induce the production of interferons (Loo et al., 2008; Yoneyama and Fujita, 2010). As a result, mRNA produced by IVT used for therapeutic purposes must undergo extensive purification to increase the uniformity of transcript lengths, and reduce immunogenicity derived from contaminating dsRNA. High performance liquid chromatography (HPLC) purified mRNA has been shown previously to not induce interferons, and be translated at higher levels, however the process is costly and time consuming, especially at large scales used in the pharmaceutical industry (Karikó et al., 2011).

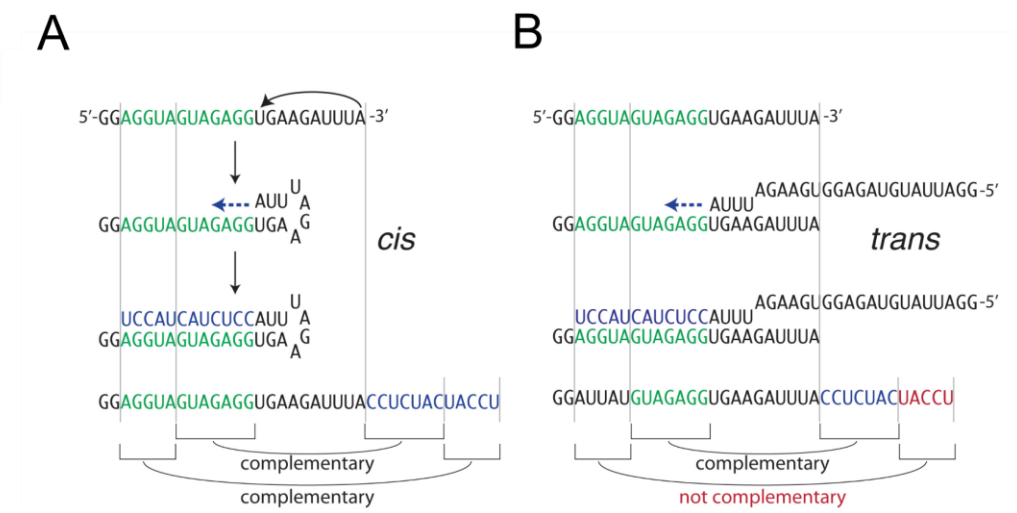


Figure 1.8 - Proposed pathways for self-priming which initiate RNA template mediated RNA synthesis. A – Cis self-priming. Regions at the 3' termini of the RNA molecule show complementarity. This allows for an intramolecular duplex to form after a hairpin loop. This duplex region can then act as a binding site for T7 RNAP, allowing for continued transcription of the same RNA molecule, leading to an extended product being formed. B – Trans self-priming. Complementarity exists between two separate RNA molecules, forming and intermolecular duplex. This duplex region, again, can act as a binding site for T7 RNAP. One the molecules in the duplex becomes extended. Adapted from (Gholamalipour et al., 2019).

1.3.4 Engineering T7 RNA polymerase for improved IVT-based manufacturing

The limitations of T7 RNA polymerase for use in IVT are well characterised. Consequently, several strategies have been pursued in order to improve product quality from the reaction, by restricting the production of off target RNA species. Firstly, attempts have been made to engineer T7 RNAP in order to improve its function. A P266L mutation in T7 RNAP has been shown to reduce the affinity of the polymerase to its promoter, allowing it to escape initial promoter contacts more easily by facilitating

the transition from initiation complex to elongation complex. This leads to reduced abortive transcript production, with the initial transcription complex being more stable, and the polymerase being more likely to proceed to elongation instead of releasing 2-6 nt transcripts (Guilleres et al., 2005; Ramírez-Tapia and Martin, 2012).

The C-terminal ‘foot’ domain has also been the focus of engineering efforts. When the T7 RNAP amino acid sequence 799-FAFA-883 was mutated to AAFA, 3’ homogeneity of product mRNA was significantly improved (Gardner et al., 1997). This was hypothesised to be due to weakening interactions with the initiating nucleotide, normally mediated by the FAFA sequences, meaning that initiation from non-promoter sequences was less likely. The addition of a C-terminal glycine residue (884G) was also shown to increase 3’ homogeneity (Dousis et al., 2023). As with the mutation described by Gardner et al., this mutation was predicted to reduce affinity to non-optimal templates such as transcribed mRNA, in this case by increasing the steric bulk of the foot domain. Dousis and colleagues also used a G47A mutation to improve 3’ homogeneity. In this case the mutation causes the RNAP to slightly favour being in the initiation complex state, in which promoter independent transcription is less likely. Whilst G47A by itself this would likely increase the amount of abortive transcripts, when combined with 884G, abortive transcript levels were not higher than wild type, and 3’ homogeneity was increased from 6-12%, to 90%. Despite this increase in product quality, IFN-β levels remain at 50% of that of the wild type enzyme, meaning that extensive purification would still be required in a therapeutic context.

Thermostable T7 RNAPs have also been developed, with a view to increasing reaction yield. A quadruple mutant featuring S633P, S430P, F849I and F880Y was identified through random mutagenesis, which showed activity at 48 ºC, enabling the use of a high-temperature IVT process (Boulain et al., 2013). More recently, commercially

available T7 thermostable variants from New England Biolabs have been shown to be transcriptionally active at 62 ºC (Wu et al., 2020). An impact of running the IVT reaction at a higher temperature is reduced production of anti-sense RNA by-products, as the RNA is unable to self-prime as effectively at higher temperatures. High temperature IVT does not, however, reduce the rate of production of 3' extended RNA products (Wu et al., 2020). Despite the potential benefits of increased transcription rate in mRNA manufacturing processes, thermostable T7 RNAPs are not widely used, due to the significant energy costs associated with raising the reaction vessel temperature. In addition, higher temperatures cause RNA degradation, reducing product integrity

1.3.5 The use of RNAPs other than T7 for mRNA manufacture

The activities of several other bacteriophage RNA polymerases have been characterised, with a view to using enzymes in IVT reactions with improved quality and yield profiles over T7 RNAP. The activities of the RNAP encoded by both T3 and SP6 bacteriophage were characterised in the 1980s, with both enzymes now being widely commercially available for use in IVT. Both T3 and SP6 produce a comparable profile of off target transcripts to T7 RNAP however, limiting any potential advantages to their use (Ling et al., 1989; Nam and Kang, 1988; Taylor and Mathews, 1993).

One P60-like viruses has had its RNAP characterised – Synechococcus phage Syn5 (Zhu et al., 2013a). The first characterised marine single subunit RNAP, Syn5 gave lower yields than T7, however was more efficient at low concentrations of ribonucleotides. Syn5 RNAP also shows higher levels of processivity than T7 RNAP, and more efficiently incorporates modified nucleotides. A further characterised single subunit RNAP was isolated from phiKMV-like phage KP34. The KP34 RNAP is of particular interest as it does not produce the self-primed 3' extended products which

are characteristic of T7 (Lu et al., 2019). The authors who characterised the polymerase showed that 38% of transcription products from a 50 nt template were the expected full length product – a vast improvement over the 0.99% seen with T7 RNAP (Lu et al., 2019). The activity of a psychrophilic RNAP, VSW-3, has also been described. VSW-3 is more efficient than T7 at low temperatures, and has reduced levels of dsRNA contamination (Xia et al., 2022). VSW-3 is now sold commercially as Ice-Lake RNAP, with recommended reaction temperatures of 4-25 ºC.

Despite several novel RNAPs being characterised, as described above, T7 RNAP continues to dominate the field of mRNA manufacture. This is primarily due to the poor yields observed with alternative RNAPs currently compared to T7. It is more economically viable to produce a high quantity of mRNA that needs extensive purification, than have a lower quantity with less downstream processing time (Kis et al., 2021). Despite this, the diverse characteristics displayed by novel characteristics do provide a rational for broadening investigations into new RNAPs – it is certainly possible that an RNAP exists in nature that balances the quality and yield profiles required to reduce the cost-per-dose in a manufacturing context. It is also true that as the diversity of RNA product formats, modified nucleotides, and cap structures increases, it may be found that T7 RNAP is suboptimal for the production of such next generation molecules. In such cases, it would be beneficial to have an extensive library of potential RNAPs with diverse characteristics.

1.4 Non-IVT Based methods for the production of synthetic RNA

The production of synthetic mRNA is dominated by IVT, a result of the simplicity and scalability of the reaction, as described previously. When producing other forms of RNA however, such as short aptamers, and dsRNA, a more diverse range of production technology exists. Described here is the use of chemical synthesis and cell-based RNA production systems, which have certain advantages when producing various RNA species

1.4.1 Chemical synthesis of synthetic RNA

The chemical synthesis of RNA is routinely used in the production of short RNA species, including short interfering RNA (siRNA) and microRNAs (miRNA) (Obika and Sekine, 2018). The fundamental technology used for chemical synthesis of RNA is almost identical to that used for the automated production of DNA oligos, developed in the 1980s (Josephson et al., 1984). Briefly, the process relies on the cyclic elongation of the RNA chain attached to a solid support. A stepwise process of deprotection, activation, coupling, oxidation and capping allows for the nucleotide chain to grow, with the addition of natural and modified bases being possible (Flamme et al., 2019). Potent chemicals such as 5-ethylthio-1H-tetrazole have been found to increase the efficiency of the main rate limiting step in the elongation process, coupling, whereby the next nucleotide is added to the growing chain (Ryczek et al., 2022).

As a technique, chemical synthesis is most useful when an entirely homogenous population of short RNA species (2-40 nt) is required, especially if the RNA contains a high proportion of modified nucleotides. When producing longer RNA species, the yield achieved rapidly drops off. This is a result of the coupling process not being 100%

efficient. Even with a coupling efficiency of 98%, a 50 nt RNA will only have a yield of 37%, due to the 2% loss of material with the addition of each base (Reese, 2005; Ryczek et al., 2022). It is possible to ligate short RNA chains to produce longer RNA molecules, however this increases the cost of the reaction, and is again not 100% efficient (Kiliszek et al., 2017). For these reasons, chemical synthesis of RNA is only possible for a small range of products, and is especially impractical for longer species such as mRNA.

1.4.2 Cell-based RNA production systems

Cellular production hosts present an attractive opportunity for the production of synthetic RNA – a result of the fact that cells endogenously encode the machinery required to transcribe RNA, or can be manipulated to encode orthologous transcriptional components. The main challenge when producing RNA in any living cell is the RNA degradation pathways present across all domains of life (Bandyra and Luisi, 2018; Court et al., 2013). This presents an issue in terms of yield, integrity and homogeneity of the final product. For this reason, the literature focuses on the production of inherently stable RNA species, such as dsRNA, and those with high levels of secondary structure. Various examples of such methods are described below.

1.4.2.1 Production of dsRNA in *E. coli*

Much of the literature surrounding the production of recombinant RNA *in vivo* has focussed on the use of *E. coli*; a result of its well understood RNA degradation pathways, and ease of genetic manipulation. Further to this, *E. coli* is simple and cheap to culture even on large bioreactor scale (McElwain et al., 2022; Terpe, 2006). The production of RNA in *E. coli* is analogous to that of recombinant protein production – T7 RNAP is encoded genomically within the cell, and is used to drive transcription of plasmid encoded RNA under the control of the T7 promoter, that has been transformed

into the cell. In the case of double stranded RNA, the *E. coli* strain HT115 is particularly well suited due to the deletion of the RNase III gene, the non-essential initiator of dsRNA degradation (Bento et al., 2020; Meng et al., 2020; Takiff et al., 1989). Cells harbouring the RNA encoding plasmid of interest are grown to exponential phase, and RNA production is induced. Cells can then be harvested, lysed, and the dsRNA purified away from other cellular components (Nwokeoji et al., 2016).

The design of the transcription templates is essential to high yield dsRNA production. Plasmid constructs can either be designed featuring convergent or divergent T7 promoters to allow for the transcription of the two RNA strands, or the two strands can be produced as one long transcript in a sense-loop-antisense fashion (Hough et al., 2022). With process optimisation, such as media conditions, fed batch yields of 182 mg/L have been achieved (Papić et al., 2018). One of the major areas of research utilising *E. coli*-produced dsRNA is the production of cost efficient RNAi based pesticides. The dsRNA produced can be applied to crops, and up-taken by pest species, where it is processed by Dicer into siRNA (Christiaens et al., 2018; Dalaisón-Fuentes et al., 2022). The siRNA is specific to the target pest species, reducing off target effects of the pesticide (Hough et al., 2022).

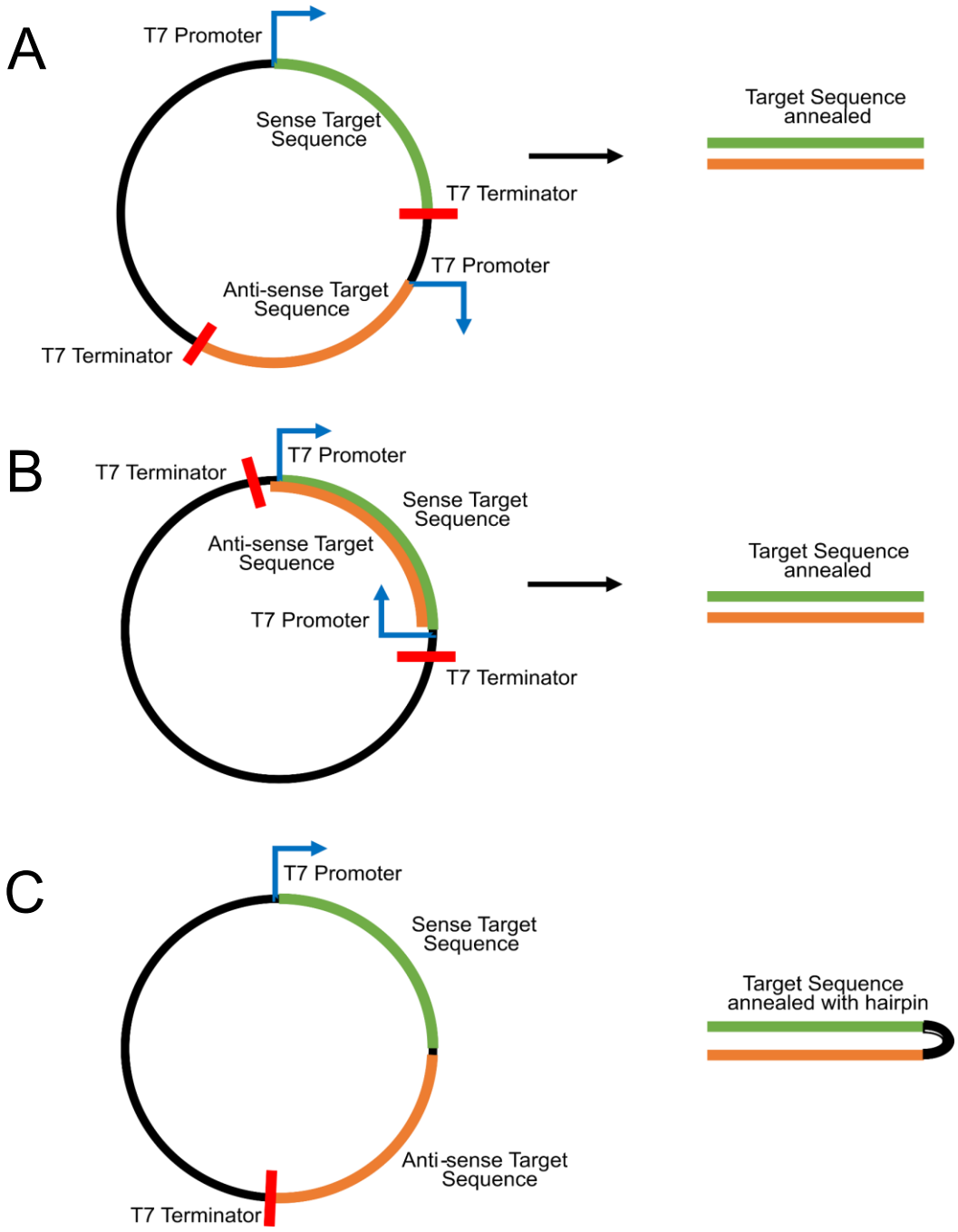


Figure 1.9 – Plasmid templates for dsRNA production *in vivo*. A) The two strands of the dsRNA molecule can be produced under separate T7 promoters and terminators on the same plasmid, and the two strands anneal *in situ*. B) The two RNA strands can be produced from the same region of DNA with two T7 terminators facing each other on the sense and antisense strand. As with (A), the two strands anneal after transcription. C) The dsRNA molecule can be produced as a single transcript, with a hair pin loop added to the sequence which allows for intramolecular contacts after transcription.

1.4.2.2 Producing RNA stabilised by scaffold sequences

Several strategies have been developed for the production of recombinant RNA other than dsRNA in bacterial cells, leading to milligram scale yields of short RNAs, such as microRNA and siRNA (Kaur et al., 2018). The first strategy involves hiding the RNA of interest within a stable RNA structure, often the anticodon stem-loop of tRNA, or sequences from 5S ribosomal RNA (Ponchon et al., 2009a). Both tRNA and rRNA are highly stable in the cell when compared with mRNA, exhibiting half lives in the range of days (Zhang et al., 2009). The tRNA scaffold sequence has the advantage of being smaller, at 60 nucleotides, compared to 80 in the rRNA structure (Ponchon and Dardel, 2011). A further advantage of using the tRNA scaffold is the naturally high abundance of tRNA in the cell, at approx. 130,000 copies in *E. coli* (Wei et al., 2019). The consequence of this is that the cell is already able to cope with the accurate processing of high levels of tRNA, and production of high yields of recombinant RNA shouldn't place a large burden on the cell. Yields of 10-50 mg per litre have been achieved using a tRNA scaffold, with recombinant RNA of up to 200 nucleotides in length (Ponchon and Dardel, 2011).

The production of recombinant RNA with scaffold sequences is simple in principle – the DNA sequence encoding the RNA of interest is inserted into a plasmid between two arms of the scaffold sequence, which when expressed will fold into a stable RNA structure (Nelissen et al., 2012). The vector requires a strong promoter that will recruit either the cells transcriptional machinery, or commonly T7 RNAP, in order to produce high yields of the recombinant RNA (Ponchon and Dardel, 2011). More sophisticated vectors have been designed for the rapid purification of the recombinant RNA, with sephadex or streptavidin aptamer sequences being placed between the scaffold sequence and the RNA of interest (Nelissen et al., 2015). Efficient strategies also exist

for the release of the RNA of interest from the stable scaffold post-purification. Hammerhead ribozyme sequences can be encoded either side of the recombinant RNA, which upon the addition of Mg^{2+} cleave the RNA of interest away from the any additional sequence (Nelissen et al., 2015).

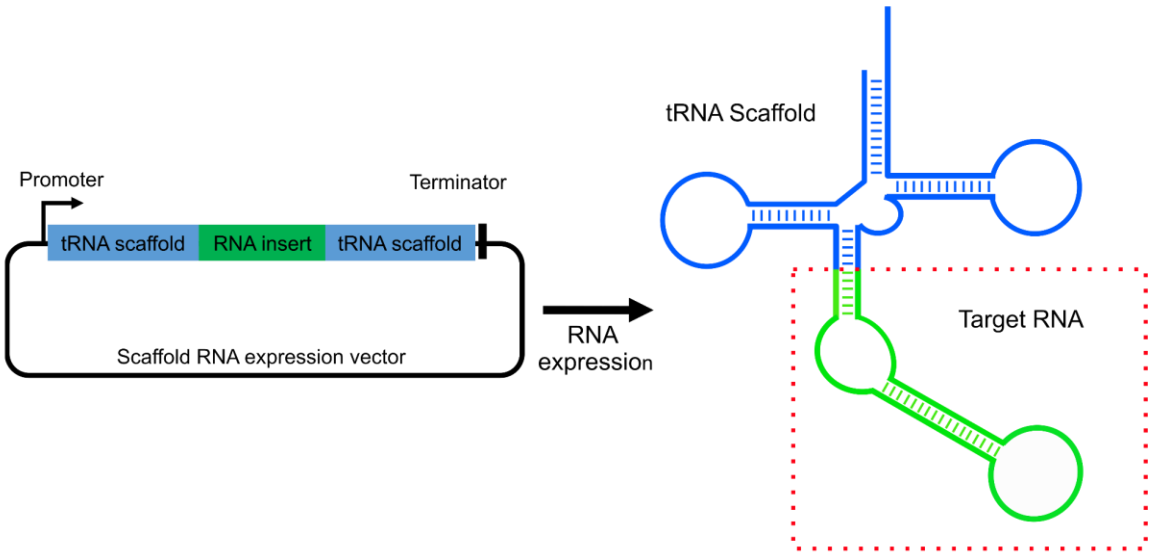


Figure 1.10 – Recombinant RNA expression within a tRNA scaffold. The RNA molecule of interest is inserted in the expression plasmid between the 5' and 3' arms of the tRNA molecule, commonly tRNA-lysine. Expression is driven from the plasmid by endogenous or T7 promoter, depending on the expression strain. After transcription, the recombinant RNA-tRNA chimera folds into a highly structured formation, increasing the stability of the molecule and allowing for increased levels of accumulation intracellularly.

1.4.2.3 Protecting RNA through RNA binding proteins

Short siRNA species have been successfully produced *in vivo*, by co-expressing a siRNA binding protein from plant viruses called p19. The natural role of p19 is to prevent silencing of siRNAs, by binding to the siRNA, sequestering it and preventing it being degraded (Ye et al., 2003). When co-expressed with the recombinant siRNA of interest, the RNA can accumulate in the cell to much higher levels than would be possible normally due to RNase III degradation (Kaur et al., 2018). The presence of the p19 binding protein also allows for simple downstream purification, by expressing a histidine tag at either termini of the protein. The RNA can then be released under denaturing conditions (Kikuchi and Umekage, 2018). Utilising this method, yields of 10 mg/L have been achieved in the production of siRNA under large scale fed batch fermentation conditions (Kaur et al., 2018). Whilst the co-expression of p19 is only suitable for the production of siRNAs, the underlying principles may be applicable to the production of mRNA. If the mRNA can be sequestered in a binding protein, such as poly(A) binding protein, then exonuclease degradation can be avoided, and downstream processing can be simplified.

1.4.2.4 Using alternative species to *E. coli* for synthetic RNA production

Several systems for RNA production in hosts other than *E. coli* are reported in the literature, utilising beneficial phenotypic features of different species. Firstly, *Rhodovulum sulfidophilum* has been used as a factory for the product of synthetic RNAs, with the bacteria then secreting the RNA extracellularly (Kikuchi and Umekage, 2018). A phototrophic marine organism, *R. sulfidophilum* produces no detectable nucleases, both intra- and extracellularly (Nagao et al., 2015). The unique ability of *R. sulfidophilum* to secrete RNA extracellularly allows for simple purification of the

recombinant RNA. *R. sulfidophilum* has been used for the successful production of short RNA aptamer sequences, such as the streptavidin aptamer, however reported yields are very low, with an extracellular concentration of 200 ng/L of recombinant RNA (Kikuchi and Umekage, 2018).

An efficient system for the production of a 160 nucleotide small nuclear RNA (snRNA) has recently been described in *Corynebacterium glutamicum,* with yields of 300 mg/L achieved after 24 hours of batch fermentation (Hashiro et al., 2019). *C. glutamicum* has traditionally been used for the production of amino acids and L-glutamate, showing suitability for use in large scale fermentation (Ikeda and Takeno, 2013). The snRNA target molecule, U1A-RNA, forms several structured hairpin loops, involved in the binding of U1A-RBD protein. In order to prevent the degradation of U1A-RNA, the authors use a RNase III knockout mutant, which is non-essential in *C. glutamicum* (Maeda et al., 2016). In order to drive efficient transcription of the snRNA, a strong promoter from *Corynephage* BFK20 is used. One proposal by the authors of this study has been to use the U1A-RNA as a new class of RNA scaffold, with the recombinant RNA of interest being encoded alongside the stable snRNA.

1.4.2.5 Applicability of cellular RNA production strategies to synthetic mRNA

Whilst the methods presented in this section are effective for the production of the forms of stable synthetic RNA described above, their applicability to mRNA manufacture appears to be more limited in their current form. Due to its length, and single stranded nature, mRNA is inherently more unstable. Cells must be able to tightly regulate the expression of mRNA within the cell, in order to respond to environmental stimuli. For this reason, highly efficient mRNA degradation pathways are present across all domains of life, initiated mainly by the ubiquitous essential enzyme RNase E (Bae et al., 2023; Börner et al., 2023). In order to make the production of mRNA in

a cellular chassis viable, strategies are needed for the stabilisation of the molecule *in vivo*¸ in order for higher titres of full length product to be accumulated.

1.5 Thesis Overview

As outlined in section 1.1, research interest in the applications of synthetic mRNA continues to grow year on year, both in the context of novel therapeutics, and other biotechnological applications. As the number of approved mRNA therapeutics increases, pressure will continue to grow on existing mRNA manufacturing infrastructure. It is essential that manufacturing processes are optimised, in order to both reduce the cost-per-dose of novel therapeutics and increase the robustness of manufacturing pipelines in times of emergency demand, such as the Covid-19 pandemic. Described in this thesis are new technologies for future optimised manufacturing of synthetic mRNA, both in terms of improving existing IVT processes through novel RNAP characterisation and introducing a new cellular mRNA production platform. Additionally, a new application for synthetic mRNA is described, in the context of improving titre of monoclonal antibodies in CHO cells.

The results in this thesis are presented in the form of three seperate publications. Chapter 2 gives general materials and methods used in the aforementioned studies, with more detailed methods being available in each publication chapter. Chapter 3 details work on characterising novel RNAPs for use in IVT reactions. For this first time, a synthetic biology approach is taken to characterise multiple diverse RNAPs simultaneously. A library of RNAPs was curated from genomic data to create a subset of enzymes for characterisation. A novel coupled transcription/translation method was introduced, allowing for rapid identification of active polymerase/promoter pairs. The

key outputs of this publication were the identification of several new active RNAPs, the description of a new method for characterising these RNAPs, and insight into which RNAPs may be optimal targets for characterisation in future studies.

Chapter 4 introduces a new platform for the production of synthetic mRNA in a cellular chassis. *E. coli* is co-opted as a production host for mRNA dues to its simplicity to culture. Methods are detailed for stabilising mRNA within the context of the cells, allowing for the considerable accumulation of molecule. A whole cell engineering approach is taken in order to optimise the mRNA production process, leading to the key output of a final system where mRNA can be readily overexpressed, extracted, and purified from *E. coli*. mRNA produced in *E. coli* is shown to be functional *in vivo* through transfection in mammalian cells.

Chapter 5 presents a novel application for synthetic mRNA, outside of current therapeutic uses. mRNA is used as a genetic ‘control node’ in order to overwhelm various processes within a monoclonal antibody producing CHO cell line, in order to elucidate which steps in the recombinant protein production pathway are rate limiting to titre. The information provided by this mRNA control node platform allows for the identification of bottlenecks in the production of diverse antibody products. Knowledge of such bottlenecks allows for the identification of specific chemical inhibitors and activators that can improve antibody titre.

Chapter 2 – Materials and Methods

2.1 General Methods and Recipes

2.1.1 Media preparation for *E. coli* culture

Lysogeny broth for routine *E. coli* culturing was prepared using 10 g/L NaCl, 10 g/L Tryptone and 5 g/L Yeast extract in distilled water. LB agar was prepared identically, with the addition of 15 g/L agar. Super optimal broth with catabolite repression (SOC) medium was prepared containing 20 g/L tryptone, 5 g/L yeast extract, 0.5 g/L NaCl, 3.6 g/L glucose, 2.47 g/L $MgSO_4$, 0.186 g/L KCl, and 0.952 g/L $MgCl_2$, dissolved in distilled water. All media was sterilised by autoclaving at 121 ºC for 15 minutes.

2.1.2 Antibiotics

Antibiotics were prepared in distilled water, or purchased as pre-dissolved solutions. Ampicillin stock solution was prepared at a concentration of 100 mg/ml, and used at a final concentration of 100 µg/ml. Kanamycin stock solution was prepared at a concentration of 50 mg/ml, and used at a final concentration of 50 µg/ml.

2.1.3 Transformation of *E. coli*

50 µL of *E. coli* cells were defrosted on ice for 5 minutes. For plasmid propagation and cloning procedures, high efficiency DH5α strains were used. For protein and RNA overexpression, BL21 strains were used. For re-transformation of plasmids, 1 ng of plasmid stock <50 ng/µL was used. For ligation procedures, 2-5 µL of the ligation reaction was used. Cells were incubated with DNA for 30 minutes on ice, before heat shocking at 42ºC for 30 seconds. Cells were then returned to ice for 5 minutes, before the addition of 950 µL of SOC medium. Cell suspensions were incubated for 1 hour at

37 ºC and 800 rpm, before plating 50 µL on LB agar plate containing the relevant antibiotic. Plates were incubated at 37 ºC for at least 16 hours.

2.1.4 Plasmid Amplification

Plasmids were routinely propagated for downstream usage in re-transformation, sequencing, *in vitro* transcription or coupled transcription-translation reactions. Colonies containing the plasmid of interested were used to grow 5 ml LB broth cultures overnight. Cells were harvested by pelleting at 8000 *g* for 10 minutes. Plasmid was purified from cells using the Qiagen Miniprep kit.

2.1.5 DNA Sequencing

DNA for sequence analysis was prepared at a concentration of 50 ng/µL, and mixed with 10 µM sequencing primer relevant to the plasmid of interest. Sequencing was carried out using the Genewiz premixed service.

2.1.6 Quantification of DNA and RNA by Nanodrop

DNA and RNA concentration and purity was determined using a Nanodrop 2000 (Thermo Fisher). Concentration was determined by absorbance at 260 nm. For DNA samples, purity was indicated by a A260 nm/ A280 nm ratio of 1.8-1.9. For RNA, an A260 nm/ A280 ratio of 1.9-2.1, and A260 nm/ A230 nm >2 indicated purity.

2.1.7 Quantification of protein by Bradford Assay

Protein concentration was determined using Bradford reagent solution (Bio-Rad), in 1 ml cuvettes containing 1 µL protein, 0.2 ml Bradford reagent, and 0.8 ml distilled water. After thorough mixing, A595 nm was determined, and concentration (mg/ml) calculated by multiplying the value by 15

2.1.8 Agarose Gel Electrophoresis

1% Agarose gels for the analysis of DNA and RNA were prepared by dissolving 1 g agarose in TAE buffer (40 mM Tris-HCl, 20 mM acetic acid, 1 mM EDTA) by heating. Gel visualisation was enabled through the addition of Ethidium Bromide solution to a 1:10000 dilution. New England Biolabs 1 kb-plus ladder was used as a standard for DNA gels, and Thermo Fisher Riboruler HR for RNA gels.

2.1.9 SDS Polyacrylamide gel electrophoresis

SDS Polyacrylamide gel electrophoresis (SDS-PAGE) was used for the analysis of protein expression levels. Precast 15 well NuPage 4-12% Tris-Glycine gels (Thermo Fisher) were used in the Xcell Mini-cell system (Thermo Fisher). Up to 15 µL of protein sample was mixed with 20 µL of Novex denaturing loading buffer (Thermo Fisher) and 4 µL of Novex reducing buffer (Thermo Fisher), before heating at 70 ºC. Gels were secured in the running tank containing Novex Tris-Glycine running buffer (Thermo Fisher), and samples loaded alongside 5 µL PageRuler prestained protein ladder. Gels were run at 225 V for 40 minutes. Gels were stained in 100 ml Imperial protein stain (Thermo Fisher), before destaining in distilled water for >8 hours.

2.1.10 Ethanol Precipitation of DNA or RNA

DNA or RNA was desalted by ethanol precipitation. DNA or RNA was mixed with 1/10 volume 3 M Sodium Acetate pH 5.2, and 3 volumes ethanol (>98% purity). Samples were incubated at -20 ºC for at least 1 hours, before centrifuging at 15000 x*g* for 30 minutes at 4 ºC. Supernatant was removed, and DNA/RNA pellets were washed with ice cold 70% ethanol. Pellets were dried for 10 minutes at room temperature, before resuspension of DNA/RNA in Tris-EDTA (TE) buffer or H_2O.

2.2 Vector Construction

2.2.1 Primer Design

Primers were designed manually using Snapgene, or automatically using Primer3. NEB Tm calculator was used to ensure differences in Tm <5 ºC between primer pairs for PCR. Primers were purchased from Integrated DNA technologies, as 100 µM stock solutions in TE buffer, diluted to a working concentration of 10 µM

2.2.2 Polymerase chain reaction (PCR)

PCR reactions were set up with the following reaction components:

Component	Volume (µL) - 25 µL Reaction	Volume (µL) - 50 µL Reaction
Forward Primer (10 µM)	1.25	2.5
Reverse Primer (10 µM)	1.25	2.5
NEB Q5 2X Mastermix	12.5	25
DNA Template (1-10 ng/µL)	1	1
H_2O	9	19

Table 2.1 – PCR reaction components

PCR reactions were run with the following cycle conditions:

Cycle Section	Cycles	Temperature (ºC)	Time (s)
Initial Denaturation	1	98	120
Denaturation	35	98	20
Annealing		50-72	30
Extension		72	20 per kb
Final Extension	1	72	120

Table 2.2 – PCR conditions

PCR products were analysed by agarose gel electrophoresis. Products were purified by Qiagen PCR purification kit or Qiagen gel extraction kit.

2.2.3 Restriction Digestion

PCR products, or plasmid DNA was digested by restriction enzymes (New England Biolabs), in the following reaction. Digestions were incubated for 1 hour at 37 ºC. DNA was purified by Qiagen PCR purification kit.

Component	Volume (µL)
10 X Reaction Buffer	5
Restriction Enzyme A	2
Restriction Enzyme B	2
DNA (Plasmid/PCR Product)	Up to 1 µg
H_2O	To 50 µL

Table 2.3 – Restriction digestion components

2.2.4 DNA Ligation

Digested PCR products and/or vector sequences were covalently joined by ligation with T4 DNA ligase. Reactions contained a 4:1 ratio of insert DNA to vector DNA, to ensure high ligation efficiency. Reactions were set up at room temperature, before subsequent incubation for 30 minutes at room temperature. 2-5 µL of the ligation reaction was transformed in high efficiency DH5α *E. coli*. Reactions contained the following components:

Component	Volume (µL)
10 X T4 Ligase Buffer	2
Digested DNA A - Insert	0.08 pM
Digested DNA B - Vector	0.02 pM
T4 DNA Ligase	1
H_2O	To 20 µL

Table 2.4 – DNA ligation components

2.3 Production of synthetic mRNA

2.3.1 Production of mRNA by *In vitro* transcription

mRNA was synthesised from plasmid templates by run-off *in vitro* transcription reactions using T7 RNA polymerase (New England Biolabs). Plasmids for IVT were linearized with XbaI or BsaXI. Reaction were set up with the following components:

Component	Volume (µL)
10 X T7 RNAP buffer	2
10 mM ATP	2
10 mM CTP	2
10 mM GTP	2
10 mM UTP	2
DNA Template (500 ng/µL)	2
T7 RNAP	2
H_2O	To 20 µL

Table 2.5 – *In vitro* transcription components

Reactions were incubated at 37 ºC for 2 hours, before addition of 2 µL DNase I and further incubation for 20 minutes to remove DNA template. RNA was purified from IVT reactions using the Monarch RNA clean-up kit (New England Biolabs). Product quality was assessed by agarose gel electrophoresis. Concentration was quantified by nanodrop measurement

2.3.2 Production of mRNA in *E. coli*

mRNA encoding plasmids were transformed into BL21 (DE3) or BL21 Star (DE3) (Invitrogen) *E. coli* strains. 5ml starter cultures were inoculated with a single colony, and grown overnight in Luria-Bertani (LB) broth (Thermo Fisher), containing 50 µg/ml kanamycin (Thermo Fisher) at 37 °C, 200 rpm. For small scale expression, 100 µL of

starter culture was used to inoculate 10 ml of LB medium containing 50 µg/ml kanamycin, and cells were grown at 37 °C, 200rpm, until the OD600nm reached 0.4-0.6. For large scale expression, 200 ml of LB was inoculated with 5 ml of overnight culture. IPTG was then added to a final concentration of 1 mM. Cells were harvested by centrifugation at 10000 x*g* for 10 minutes.

2.3.3 RNA extraction from *E. coli*

For small scale culture (cell pellets from <1 ml culture), Total RNA was extracted using the GenElute Total RNA purification kit (Sigma-Aldrich). Cell pellets were resuspended in 100 µL of TE buffer containing 1 mg/ml Lysozyme, and incubated at room temperature for 5 minutes. 300 µL of buffer RL and 200 µL of 96-100% ethanol were added to the lysate before vortexing. Lysate was then applied to the spin column resin, before washing with ethanol solution. RNA was eluted in 50 µL of elution solution. Residual DNA in the RNA sample was then removed through addition of 2 units of RNase free DNase I, and incubation at 37 °C for 30 minutes. RNA was purified from the DNase reaction using the Monarch RNA Cleanup kit (50 µg) (New England Biolabs).

For larger scale extractions, cells were suspended in 5 ml of 5 mg/ml lysozyme solution (Thermo Fisher), and allowed to lyse for 10 minutes at room temperature. 12ml of lysis solution (4% SDS, 0.5 M NaCl (Thermo Fisher)) was added, before incubation for 5 minutes at 65 ºC. 6.8ml of 5M NaCl of was added, and suspensions were placed on ice for 5 minutes to promote precipitation of SDS. Suspensions were centrifuged at 10000 xg for 20 minutes at 4 ºC, and supernatant was transferred to a separate tube.

RNA was precipitated from the supernatant by isopropanol or ethanol, and stored at -80 ºC.

2.3.5 mRNA purification from extracted Total RNA

mRNA was purified from small scale total RNA extractions using the Dynabeads mRNA purification kit (Invitrogen). Approx. 10 µg of Total RNA was adjusted to a volume of 100 µL with nuclease free water. Samples were heated to 65 °C for 2 minutes to disrupt secondary structure, before placing on ice. Magnetic beads were equilibrated in 100 µL of binding buffer (10 mM Tris-HCl, 1M LiCl), before addition of the 100 µL total RNA samples. Beads were incubated with RNA samples for 5 minutes at room temperature with constant rotation, before removal of the supernatant. Beads were washed twice with 200 µL of washing buffer (10 mM Tris-HCl, 1M LiCl). mRNA was eluted from the magnetic beads by addition of 10 mM Tris-HCl, pH 7.5, and its concentration determined by Nanodrop spectrophotometer.

mRNA from large scale extractions was purified by oligo-dT enrichment utilising an ÄKTA pcc chromatography system (Cytiva). 10ml total RNA was bound to a 1 ml volume Oligo-dT(18) column (Sartorius) in binding buffer (50 mM Sodium Phosphate pH 7, 500 mM NaCl). The column was washed in 20 column volumes of wash buffer (50 mM Sodium Phosphate, 150 mM NaCl). Bound mRNA was eluted in 5ml of water, before quantification by Nanodrop spectrophotometer.

2.4 Capillary Gel Electrophoresis

Capillary Gel Electrophoresis (CGE) analysis of RNA integrity and purity was performed with a 5200 Fragment Analyzer System (Agilent, CA, USA), using the DNF-471 RNA Kit (15 nt) (Agilent, CA, USA). The capillary cassette used was FA 12-

Capillary Array Short, 33 cm (Agilent, CA, USA). Samples were diluted to <100 ng/μL in nuclease free water. Before each separation, a pre-run voltage was applied (8 kV for 30 seconds), the capillaries were conditioned with the conditioning solution and the capillaries were dipped twice in the rinse buffer. Following this, the capillaries were filled with RNA separation gel (by pressure) and then the sample was introduced using a voltage injection (5 kV for 4 seconds). The separation was then conducted by applying a voltage of 8 kV for 45 minutes. Detection was carried out using Laser Induced Fluorescence (LIF), by fluorescent dye tagging of the RNA.

2.5 Digital Droplet PCR

Digital droplet PCR was used to determine the concentration of target mRNA molecules in a population of total RNA. Primers were designed to amplify a 150 bp region in the middle of eGFP, universal to every construct under investigation. ddPCR reactions comprised the following components (All Bio-Rad):

Component	Volume (μL)
One-step ddPCR mastermix	5
Reverse Transcriptase	2
300 mM DTT	1
900 nM Forward Primer	1
900 nM Reverse Primer	1
250 nM 5'-FAM Probe	1
Total RNA (1 ng/μL)	1
H_2O	To 20 μL

Table 2.6 – Digital droplet PCR components

The 20 µL reaction mixture and 70 µL of droplet generation oil (Bio–Rad) were loaded into a DG8 Cartridge, and 40 µL of droplets were generated with the Bio-Rad QX200 droplet generator. Droplets were transferred to a 96-well PCR plate (Bio-Rad), sealed with foil, and placed in a Bio-rad C1000 thermal cycler. Cycling conditions were as follows:

Cycle Section	Cycles	Temperature (ºC)	Time (Mins)
Reverse Transcription	1	50	60
Polymerase Activation	1	95	10
Denaturation	40	95	1
Annealing/Extension		60	1
Enzyme Deactivation	1	98	10

Table 2.7 – ddPCR running conditions

A final hold phase at 12 °C for 30 minutes was employed to increase droplet size through condensation. Positive droplets were detected by the QX200 droplet reader (Bio-Rad), using automatically assigned amplitude thresholds determined by QuantaSoft software (Bio-Rad). Samples were only used in analysis if the number of measured droplets exceeded 12000. QuantaSoft software converts positive droplet counts to mRNA copy number by assuming a normal distribution of mRNA copies per positive droplet. The following formula is used to calculate mRNA copy number:

Concentration $= -ln(\frac{N_{Neg}}{N})/V_{droplet}$ Formula 2.1

Whereby; N = Total number of droplets

N_{Neg}= Number of negative droplets

$V_{Droplet}$= Volume of droplets

2.6 Overexpression and Purification of recombinant RNA polymerase

2.6.1 RNAP overexpression

RNAP encoding plasmids were transformed into BL21 (Thermo Fisher) or Shuffle (New England Biolabs) *E. coli* cells. Cells were grown in 5 ml start cultures overnight, which were used to inoculate 500 ml of LB broth. 500 ml cultures were incubated at 37 ºC with 200 rpm shaking until an OD_{600} of 0.4 was reached. Temperature was reduced to 25 ºC, and expression was induced by addition of 1 M IPTG (Merck) for a period of 8 hours. Cells were subsequently centrifuged at 8000 x*g* for 20 minutes.

2.6.2 RNAP Purification

Cells were resuspended in lysis buffer A (50 mM Tris-HCl, 0.5 M NaCl, pH 8.0), and sonicated for 3 x 20 second periods. Cell debris was removed by centrifugation at 70000 x*g* for 15 minutes. Cell free extract was applied to a 5ml HisTrap HP column (Cytivia) at a rate of 5 ml/min. The column was washed with 2 column volumes Buffer A + 40 mM Imidazole (Sigma-Aldrich), before elution of protein in a gradient of imidazole from 0-300 mM over 10 column volumes. 5 ml of eluted protein was applied to a 1.6x60 cm Superdex200 gel filtration column at 1.5 ml/min, with 2 ml fractions collected after void volume. Fractions containing the RNAP of interest were concentrated to a final concentration of 1 mg/ml and exchanged into RNAP storage buffer (50mM Tris-HCl, 100mM NaCl, 50% v/v glycerol, 10mM DTT, 0.1mM EDTA, 0.2% w/v NaN3 (All Sigma-Aldrich)), using a 50 kDA MWCO centrifugal filter (Sigma-Aldrich).

2.7 Cell free coupled transcription-translation assays

Coupled transcription-translation assays to determine RNAP activity were carried out using the TNT SP6 Quick Coupled transcription/translation system (Promega). Reactions were assembled containing 8 µL TnT Quick Master Mix, 1 µL RNAP encoding plasmid (40 ng/µL), 1 µL NanoLuc encoding plasmid (80 ng/µL), and 0.2 µL 1 mM methionine. The assay proceeded at 30 °C for 1 hour. Samples were then diluted 500 fold in nuclease free water, and added at a 1:1 ratio to pre-diluted NanoLuc luciferase assay substrate. Samples were incubated in darkness for 5 minutes, before detection of luminescence by Molecular Devices ID5 plate reader, with an integration time of 10 s.

2.8 Culturing of mammalian cell lines

2.8.1 Routine culturing of CHO cells

CHO cell pools, derived from CHO-S suspension cells (Thermo Fisher), were routinely cultured in CD-CHO medium, supplemented with 6 mM L-Glutamine (Thermo Fisher). 30 ml cultures were maintained in vented 125 ml Erlenmeyer flasks, incubated at 37 C, 5% CO_2, with 140 rpm shaking. Where required, culture volume was increased to 100 ml in 500 ml vented Erlenmeyer flasks. Cell number and viability was determined by trypan blue assay, using a Countess 3 (Thermo Fisher), or Vicell Blue (Beckman-Coulter) automated cell counter. Cells were subcultured to a density of 0.2 x 10^6 - 0.4 x 10^6 viable cells per ml every 3 days.

2.8.2 Routine culturing of HEK cells

A suspension human embryonic kidney (HEK) cell line, Freestyle 293-F (Thermo Fisher), was routinely cultured in Freestyle 293 medium (Thermo Fisher). in CD-CHO

medium, supplemented with 6 mM L-Glutamine (Thermo Fisher). 30 ml cultures were maintained in vented 125 ml Erlenmeyer flasks, incubated at 37 C, 5% CO_2, with 140 rpm shaking. Where required, culture volume was increased to 100 ml in 500 ml vented Erlenmeyer flasks. Cell counting was performed as described in 2.8.1, with subculturing to a density of 0.2 x 10^6 - 0.4 x 10^6 viable cells per ml every 3 days.

2.8.3 Electroporation of CHO cells with mRNA

Suspension CHO cells were electroporated using the Lonza nucleofector system. Cells were subcultured 2 days prior to nucleofection, in order to give a starting VCD of 1-2 x 10^6 cells per ml. 1.846 x 10^6 cells per well transfected were centrifuged at 100 x*g* and media removed. Cells were resuspended in 20 µL total volume, comprising 18 µL nucleofection solution and 2 µL 200 ng/µL mRNA solution. Cells were subsequently transfected in the 4D-nucleofector Core unit, using program FF158. 80 µL of pre-warmed media was added to each nucleofection reaction, before transferring 80 µL of cell suspension to 670 µL pre-warmed media in a 24 well plate. Cells were cultured for 72 hours at 37 ºC, with 5% CO_2 and 230 rpm shaking, before harvesting of cell supernatant for secreted protein quantification.

2.8.4 Quantification of secreted mAb titre from CHO cells by ELISA

mAb titre in cell culture supernatant was quantified by human IgG ELISA (RD Biotech). Supernatant was diluted 1:800 for DTE mAbs and 1:4000 for ETE mAbs in sample diluent buffer. 20 µL of sample was mixed with 100 µL peroxidase conjugated anti-human IgG in the ELISA plate, and incubated at room temperature for 15 minutes. Post-incubation. Sample was then removed, and the plate was subsequently washed 3 times with 300 µL wash solution. 100 µL TMP substrate was added to each well, and the plate was incubated for 10 minutes at room temperature. Reactions were stopped

with stop solution, and absorbance was read at 450 nm and 620 nm using a Molecular Devices iD5 plate reader. Standard curves produced using reference IgG samples were used to convert absorbance values to mAb titre.

2.8.5 Transfection of HEK cells with mRNA by cationic polymer solution

Suspension HEK cells were transfection using the TransIT-mRNA transfection kit (Mirus Bio), comprising a proprietary cationic polymer solution. 24 hours prior to transfection, cells were seeded in 500 µL of Freestyle media in 24 well plates at a density of 2.5 x 10^5 cells per ml. 50 µL of pre-warmed media was mixed with 1 µL mRNA solution (500 ng/µL), 1 µL mRNA boost reagent, and 1 µL TransIT mRNA reagent. The mixture was incubated at room temperature for 5 minutes to allow mRNA complexes to form. Complexes were added dropwise to the wells of the 24 well plate, and gently distributed by rocking. Cells were incubated for 24 hours at 37 ºC, with 5% CO_2 and 230 rpm shaking, before harvesting of cells for protein quantification

2.8.6 Quantification of GFP from mRNA transfected HEK cells

500 µL of cultured HEK cells expressing GFP were centrifuged at 100 x*g*, and the supernatant removed. Cells were resuspended in 200 µL DPBS (Sigma-Aldrich), before determination of GFP signal through reading absorbance at 488 nm and 507 nm using a Molecular devices plate reader

Chapter 3 – Paper I

A platform for the characterisation of novel single subunit RNA polymerases

Edward Curry[1], Sveta Sendelnikova[2], John Rafferty[2], Martin Hulley[3], Adam Brown [1]

[1] Department of Chemical and Biological Engineering, University of Sheffield, Sheffield, UK.

[2] School of Biosciences, University of Sheffield, Sheffield, UK.

[3] Bioprocess Development, AstraZeneca, Cambridge, UK.

Available in Biotechnology Journal

DOI: 10.1002/biot.202400012

Summary

The first paper in this thesis described a new methodology for characterising single subunit RNAPs. A cell free coupled transcription-translation method is used, whereby the expression of the novel RNAP, and the detection of its activity can occur simultaneously. Using this method, 8 new ssRNAP-promoter pairs are detailed, greatly expanding the library of available RNAPs. Further consideration is given to RNAPs which may present optimal targets for future characterisation in the context of enzymes for mRNA manufacture

Contributions

Edward Curry – Conceptualisation, methodology, analysis, investigation, manuscript preparation. **Svetlana Sendelnikova** – Investigation (Protein Purification), analysis. **John Rafferty** – Investigation (Protein Purification), analysis. **Martyn Hulley** – Conceptualisation, analysis. **Adam Brown** – Conceptualisation, methodology, analysis, manuscript preparation, funding acquisition.

Abstract

All mRNA products are currently manufactured in *in vitro* transcription reactions that utilize single-subunit RNA polymerase (RNAP) biocatalysts. Although it is known that discrete polymerases exhibit highly variable bioproduction phenotypes, including different relative processivity rates and impurity generation profiles, only a handful of enzymes are generally available for mRNA biosynthesis. This limited RNAP toolbox restricts strategies to design and troubleshoot new mRNA manufacturing processes, which is particularly undesirable given the continuing diversification of mRNA product lines towards larger and more complex molecules. Herein, we describe development of a high-throughput RNAP screening platform, comprising complementary *in silico* and *in vitro* testing modules, that enables functional characterisation of large enzyme libraries. Utilizing this system, we identified eight novel sequence-diverse RNAPs, with associated active cognate promoters, and subsequently validated their performance as recombinant enzymes in IVT-based mRNA production processes. By increasing the number of available characterized functional RNAPs by > 130% and providing a platform to rapidly identify further potentially useful enzymes, this work significantly expands the RNAP biocatalyst solution space for mRNA manufacture, thereby enhancing capability to achieve application and molecule-specific optimisation of product yield and quality.

3.1 - Introduction

The clinical success of SARS-Cov-2 vaccines established synthetic mRNA as an effective drug format, paving the way for hundreds of new mRNA-based vaccines and gene therapies to enter clinical trials (Qin et al., 2022; Webb et al., 2022). This has resulted in a sharp increase in global demand for mRNA production, and shifted mRNA manufacturing from a relatively niche process to one that underpins current and future strategies to treat monogenic disorders, cancer, and infectious diseases (Al Fayez et al., 2023; Liu et al., 2023; Vavilis et al., 2023b). All such mRNA products are currently produced in standardised *in vitro* transcription (IVT) systems using single-subunit DNA-dependent phage RNA polymerase (RNAP) biocatalysts. While Salmonella phage SP6 and Enterobacteria phage T3 RNAPs can be utilised in certain contexts, the dominant biocatalyst choice for synthetic mRNA manufacture is the *Enterobacteria* phage T7 RNAP. This enzyme has undergone extensive protein engineering to improve bioproduction performance, predominantly via strategies to reduce formation of immunogenic product-related impurities such as short-abortive transcripts (Guilleres et al., 2005; Lyon and Gopalan, 2018) and double-stranded RNA species (Dousis et al., 2023; Wu et al., 2020).

Although T7 RNAP typically generates high product yields and acceptable product quality profiles, it is highly unlikely that a single one-size-fits-all biocatalyst approach will be optimal for all mRNA manufacturing contexts. Indeed, other bioproduction processes rely on biocatalyst toolbox approaches, such as the wide range of evolved and engineered Chinese Hamster Ovary cell factories utilised for recombinant protein manufacture (Fischer et al., 2015). The current unavailability of such an RNAP toolbox for mRNA IVT platforms restricts i) bioprocess design strategies, such as optimising temperature set-point to achieve quality target product

profiles, and ii) molecule/application specific optimisation of product yield and quality. The latter is particularly pertinent given that mRNA product lines are diversifying to include larger and more complex molecular formats that present new biomanufacturing challenges, such as circular RNA (Bai et al., 2023; Qu et al., 2022; Zhu et al., 2022b), long self-amplifying RNA transcripts (Blakney et al., 2021; Pourseif et al., 2022), and linear mRNA incorporating novel cap structures and modified nucleotides (Chen et al., 2022). Indeed, it should be anticipated that mRNA will follow the path of protein therapeutics, where product designers rapidly progressed from relatively simple molecules such as Insulin to highly-engineered formats (e.g. tri-specific antibodies) that require product-specific biocatalyst solutions (Tihanyi and Nyitray, 2020).

The available RNAP toolbox has recently been expanded by studies focussed on identifying and characterising individual enzymes with putative desirable bioproduction phenotypes. KP34 enhances 3' homogeneity of product molecules (Lu et al., 2019), VSW-3 reduces dsRNA impurities (Xia et al., 2022), and Syn5 exhibits increased processivity (Zhu et al., 2013a), as compared to that achieved with T7. While these hypothesis-driven approaches have successfully identified new biocatalysts with novel functionalities, only six characterized RNAPs are currently publicly available for mRNA manufacture (although we note that some additional unpublished enzymes may be utilised in industrial settings). Accordingly, mRNA manufacturing solution spaces are severely limited, and, moreover, currently utilised 'standardised' biocatalysts such as T7 may have relatively poor performance characteristics (E.g. processivity, impurity generation) relative to the hundreds of 'untested' single-subunit phage RNAPs found in nature.

In this study, we address the paucity of biocatalysts available for IVT-based mRNA production. Using a combination of *in silico* and *in vitro* analyses we identify and functionally validate eight new sequence-diverse RNAPs, more than doubling the number of previously described enzymes for mRNA manufacture. In doing so, we describe development of a high-throughput screening system that can be utilised to rapidly select and test future RNAP libraries, facilitating further expansion of the biocatalyst solution space. Provision of a substantially expanded RNAP toolkit enhances capabilities to design and troubleshoot new molecule-specific manufacturing processes, which will be particularly useful for optimising yield and quality of complex next-generation mRNA products.

3.2 - Materials and Methods

3.2.1 - RNAP library creation

A starting library of 351 predicted RNAP sequences was collated from Uniprot, comprising all sequences annotated as predicted phage DNA-directed RNA polymerases. RNAPs were clustered by grouping RNAPs sharing sequence identity >85%, using the Clustal Omega online alignment tool (Madeira et al., 2019). RNAPs were further clustered using the MMSEQ2 online server, with a minimum sequence identity threshold of 85%, and coverage threshold of 70% (Steinegger and Söding, 2017). A representative RNAP from each cluster was chosen by totalling the matrix identity score to determine which polymerase in each cluster was most divergent in sequence to all others. Promoters for remaining RNAPs were predicted by PHIRE (Phage *in silico* regulatory elements) (Lavigne et al., 2004) using parameters of string lengths – 20, window size – 30, and degeneracy – 4. Predicted promoter sequences were verified with the *PhagePromoter* tool (Sampaio et al., 2019), with a probability threshold of 0.5.

3.2.2 - Plasmid Construction

For coupled transcription-translation assay plasmids, RNAP sequences were synthesised and cloned into XhoI and XbaI restriction sites on the pTNT vector (Promega). To create the corresponding transcription templates, the NanoLuc gene (Promega) was cloned into XhoI and XbaI restriction sites in pTNT, before site directed mutagenesis to substitute the SP6 and T7 promoter with the promoter of interest. For RNAP overexpression plasmids, RNAP sequences were inserted between NdeI and XhoI sites on pET-29b (Novagen). Transcription templates were made by site directed

mutagenesis to substitute the T7 promoter with the promoter of interest on pCMV-Cluc2 (New England Biolabs).

3.2.3 - Cell free coupled transcription-translation assays

Coupled transcription-translation assays were carried out using the TNT SP6 Quick Coupled transcription/translation system (Promega). Reactions were assembled containing 8 µL TnT Quick Master Mix, 1 µL RNAP plasmid (40 ng/µL), 1 µL NanoLuc plasmid (80 ng/µL), and 0.2 µL 1 mM methionine. The assay proceeded at 30 °C for 1 hour. Samples were then diluted 500 fold in nuclease free water, and added at a 1:1 ratio to pre-diluted NanoLuc luciferase assay substrate. Samples were incubated in darkness for 5 minutes, before detection of luminescence by Molecular Devices ID5 plate reader, with an integration time of 10 s.

3.2.4 - RNAP expression and purification

pET-29b-RNAP plasmids were transformed into BL21 or NEB Shuffle *E. coli* cells (New England Biolabs), and grown in 5 ml culture overnight at 37 °C. Starter cultures were used to inoculate 500 ml LB, which was incubated at 37 °C until an OD_{600} of 0.4 was reached. At this point incubation temperature was lowered to 25 °C. Protein expression was then induced by addition of 1 M IPTG (Sigma-Aldrich), before harvesting of cells after 8 hours. Cell pellets were re-suspended for purification in buffer A (50 mM Tris-HCl, 0.5 M NaCl, pH 8.0), and lysed by sonication. After removal of cell debris by centrifugation, cell free extract was applied to a 5ml HisTrap HP column (Cytivia) at a rate of 5 ml/min. The column was washed with 2 column volumes Buffer A + 40 mM Imidazole (Sigma-Aldrich), before elution of protein in a gradient of imidazole from 0-300 mM over 10 column volumes. 5 ml of eluted protein was applied to a 1.6x60 cm Superdex200 gel filtration column at 1.5 ml/min, with 2 ml fractions

collected after void volume. Fractions containing the RNAP of interest were concentrated to a final concentration of 1 mg/ml and exchanged into RNAP storage buffer (50mM Tris-HCl, 100mM NaCl, 50% v/v glycerol, 10mM DTT, 0.1mM EDTA, 0.2% w/v NaN3 (All Sigma-Aldrich)), using a 50 kDA MWCO centrifugal filter (Sigma-Aldrich). RNAP preparations were stored at -20 °C. Purity of final preparations was assessed by Tris-Glycine SDS-PAGE.

3.2.5 - *In vitro* transcription

Plasmid templates for IVT were linearised with XbaI, and purified by ethanol precipitation. Transcription reactions using the Hiscribe IVT kit (New England Biolabs), were assembled to a final volume of 20 µL. Reactions contained 2 µL 10X reaction buffer, 2 µL of each NTP, 1 µg of template DNA, and 2 µL of T7 RNAP, or 2 µL of novel RNAP. Transcription reactions were incubated for 2 hours, before addition of 1 µL DNase I, and further incubation for 20 minutes. Transcription reactions were purified using the Monarch RNA cleanup kit (New England Biolabs). mRNA concentration was quantified by nanodrop spectrophotometer, and product integrity assessed by agarose gel electrophoresis.

3.3 - Results and Discussion

3.3.1 - Bioinformatic analysis of the potential RNA polymerase biocatalyst solution space

The biocatalyst solution space for mRNA production is currently limited to a handful of characterised RNAPs. To define the theoretical solution space, we extracted the sequence of all putative single subunit RNAPs from Uniprot. At the time of conducting this analysis, 351 distinct single-stranded RNAPs had been predicted from publicly available genomics data. Accordingly, given that only six of these enzymes had been previously tested, approximately 98% of the potential solution space remained unexplored. We rationalized that determining the function of all 345 previously untested RNAPs would be highly inefficient, and, moreover, unnecessary, given that many of these enzymes will share similar performance characteristics. Indeed, we assumed that variation in bioproduction phenotype (e.g., enzyme processivity, impurity generation profiles) would be underpinned by significant differences in amino acid sequences. Accordingly, we sought to define distinct spots within the potential solution space by identifying RNAP clusters that shared minimal amino acid sequence homology (Fig. 1A).

The Clustal Omega sequence alignment tool (Madeira et al., 2019) was used to define RNAP clusters, whereby enzymes with >85% global sequence identity were grouped into a single distinct family. This analysis identified 93 enzyme clusters, where the smallest and largest groups contained 1 and 32 RNAPs respectively (Fig. 1B). To interrogate local sequence similarities, these families were then analysed using MMSEQ2, grouping RNAPs based on k-mer matching and the Smith-Waterman algorithm. Using sequence identity and coverage thresholds of 85% and 70% respectively (Steinegger and Söding, 2017), the number of discrete RNAP families

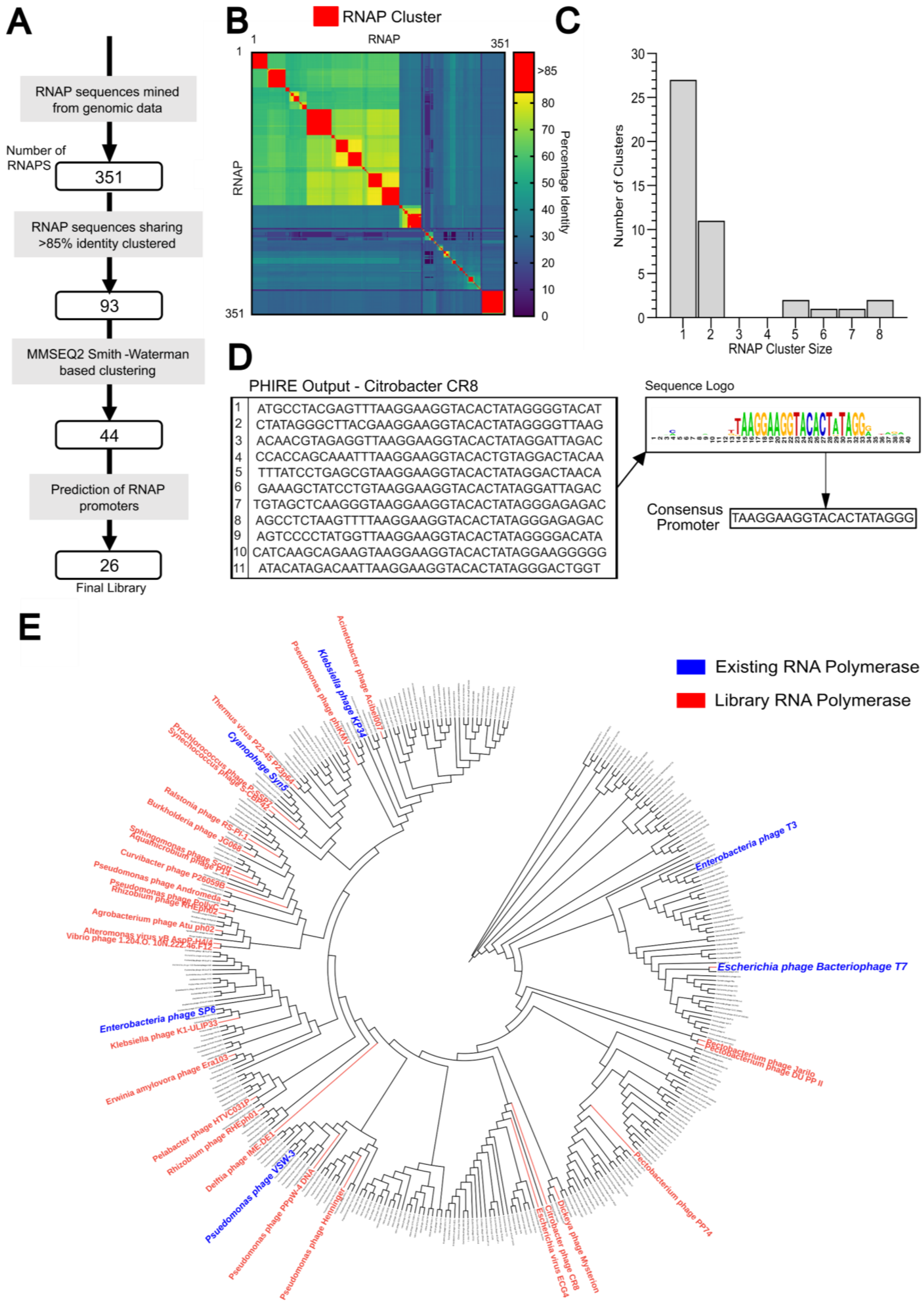
was reduced from 93 to 44, where the majority of clusters (27) contained a single enzyme (Fig. 1C).

Utilisation of novel RNAPs for mRNA manufacture requires concomitant identification of appropriate cognate promoter elements to drive product transcription. This is non-trivial as single subunit RNAPs typically display highly stringent promoter binding activity, where mutation of a single nucleotide can abolish transcriptional output (Rong et al., 1998). Accordingly, identification of novel functional polymerases necessitates highly accurate promoter predictions. However, there are only limited publicly available tools to achieve this, where PHIRE (Lavigne et al., 2004) searches for conserved elements of defined length, and PhagePromoter utilises machine learning models to classify specified phage sequences as ‘promoter’ or ‘non-promoter’ (Sampaio et al., 2019). A representative polymerase was selected from each of the 44 families, and the associated phage genomes were investigated with both of these promoter prediction tools (Fig. 1D). This analysis failed to accurately identify cognate promoters in 18/44 cases. Testing further RNAPs from these 18 families similarly failed to result in identification of useable elements, indicating that the amino acid sequence diversity within these clusters is associated with ‘unusual’ cognate binding motifs that are significantly different to the promoter datasets that were used to train existing prediction tools. Accordingly, ~40% of identified RNAP clusters could not be tested *in vitro* due to limitations in promoter prediction capabilities.

Cognate promoters were successfully predicted for the remaining 26 RNAPs, and optimal reaction temperatures were identified for each enzyme based on the growth temperature of corresponding phage hosts (Table 1). Ten RNAPs had predicted temperature optima ≤ 30 °C, which may be beneficial for mRNA product quality profiles given that IVT reactions performed at reduced temperatures are associated with

decreased levels of product-related impurities (Xia et al., 2022). Phylogenetic analysis of the 26 selected RNAPs confirmed that the final library comprised a panel of evolutionarily diverse enzymes, sharing no significant sequence similarity (≤ 75% global sequence identity) with any the six previously characterised polymerases (Fig. 1E). These polymerase-promoter pairs (Table 1) were taken forward for *in vitro* functional characterisation, facilitating testing of ~60% of the identified RNAP clusters.

Figure 3.1 (Follows below) - Bioinformatics-driven design of an RNAP 'test' library (**A**). Putative RNAPs were clustered using pairwise global sequence identity analysis (**B**), and subsequently grouped into 44 distinct families according to local sequence similarities (**C**). Representative polymerases from families for which accurate cognate promoter prediction was possible (**D**) were phylogenetically analysed to validate evolutionary diversity, represented in the circular cladogram showing all 351 analysed RNAPs (**E**).


A
RNAP sequences mined from genomic data
Number of RNAPS
351
RNAP sequences sharing >85% identity clustered
93
MMSEQ2 Smith -Waterman based clustering
44
Prediction of RNAP promoters
26
Final Library
B
RNAP Cluster
RNAP
1
351
>85
80
70
60
50
40
30
20
10
0
Percentage Identity
C
Number of Clusters
RNAP Cluster Size
D
PHIRE Output - Citrobacter CR8
1 ATGCCTACGAGTTTAAGGAAGGTACACTATAGGGGTACAT
2 CTATAGGGCTTACGAAGGAAGGTACACTATAGGGGTTAAG
3 ACAACGTAGAGGTTAAGGAAGGTACACTATAGGATTAGAC
4 CCACCAGCAAATTTAAGGAAGGTACACTGTAGGACTACAA
5 TTTATCCTGAGCGTAAGGAAGGTACACTATAGGACTAACA
6 GAAAGCTATCCTGTAAGGAAGGTACACTATAGGATTAGAC
7 TGTAGCTCAAGGGTAAGGAAGGTACACTATAGGGAGAGAC
8 AGCCTCTAAGTTTTAAGGAAGGTACACTATAGGGAGAGAC
9 AGTCCCCTATGGTTAAGGAAGGTACACTATAGGGGACATA
10 CATCAAGCAGAAGTAAGGAAGGTACACTATAGGAAGGGGG
11 ATACATAGACAATTAAGGAAGGTACACTATAGGGACTGGT
Sequence Logo
Consensus Promoter
TAAGGAAGGTACACTATAGGG
E
Existing RNA Polymerase
Library RNA Polymerase
Acinetobacter phage Acibel007
Klebsiella phage KP34
Pseudomonas phage phiKMV
Thermus virus P23-45 P23p64
Cyanophage Syn5
Enterobacteria phage T3
Escherichia phage Bacteriophage T7
Enterobacteria phage SP6
Klebsiella phage K1-ULIP33
Erwinia amylovora phage Era103
Pelabacter phage HTVC031P
Rhizobium phage RHEph01
Delftia phage IME-DE1
Psuedomonas phage VSW-3
Pseudomonas phage PPpW-4 DNA
Pseudomonas phage Henninger
Escherichia virus ECG4
Citrobacter phage CR8
Dickeya phage Mysterion
Pectobacterium phage PP74
Pectobacterium phage Jarilo
Pectobacterium phage DU PP II
Agrobacterium phage Atu ph02
Ralstonia phage RS-PI-1
Burkholderia phage JG068

Assigned Number	Phage Name	UniProt Protein ID	Predicted Temperature Optimum (◦C)	Promoter
R1	Prochlorococcus_phage_P-SSP7	Q58N45	14-26	AAAATTCTTCAAGTTTACAA
R2	Synechococcus_phage_S-CBP42	A0A096VKW2	20	CACTTCCACTCAACCAACCG
R3	Erwinia_amylovora_phage_Era103	A2I7X6	28	AATAACCACCCAGTATAGAAGGAA
R4	Agrobacterium_phage_Atu_ph02	A0A223VZI2	28	TTATCCTTCGTATAAGGAATA
R5	Dickeya_phage_Mysterion	A0A385IGY0	28	CTTAAATCATCACTATTAG
R6	Pectobacterium_phage_PP74	A0A1J0MEG1	30	TAATACGACTCACTATTGGGAA
R7	Aquamicrobium_phage_P14	A0A1L5C074	30	TTTCGGTACGCTCTAGCA
R8	Pectobacterium_phage_DU_PP_II	A0A2D2W5U8	30	TTATTAACGACTCACTACTAGGAA
R9	Pectobacterium_phage_Jarilo	A0A2S1GSW7	30	TAATAACGACTCACTATTAGAAG
R10	Sphingomonas_phage_Scott	A0A346FDD2	30	TCGGGTTGTCGATTTCCTTAC
R11	Ralstonia_phage_RS-PI-1	A0A1S6L1D6	35	GTCGAAGTCGTCGAGCAGC
R12	Burkholderia_phage_JG068	U3PFP4	37	TCAGTAGACTATCTAG
R13	Acinetobacter_Acibel007	A0A075DXW8	37	CTGTACTCACAGCTCAATTT
R14	Delftia_phage_IME-DE1	A0A0F7INH1	37	GTTAGCCCACACCATTGAAGACCC
R15	Pseudomonas_phage_Henninger	A0A2K9VHD7	37	TTAAAACCCTCACTATGGCTACA
R16	Pseudomonas_phage_PollyC	A0A2K9VHU7	37	CTCACTCACGACCCAAATTC
R17	Pseudomonas_phage_phiKMV	Q7Y2D9	37	CGACCCTTCCCTACTCCGGCCTTAAAT
R18	Citrobacter_phage_CR8	W6PP41	37	TAAGGAAGGTACACTATAGGG
R19	Thermus_virus_P23-45_P23p64	A7XX94	65	TTATTCCTTTA
R20	Pelagibacter_phage_HTVC031P	A0A4Y1NTX3	16-23	AACTAATGCTCAATTTAGAGATA
R21	Rhizobium_phage_RHEph01	L7TQW5	25-30	ATTACCCCTCCCTTAAGCAAAG
R22	Rhizobium_phage_RHEph02	L7TJC5	25-30	TTAATCCTCACTATTAGGATAA
R23	Curvibacter_phage_P26059B	A0A384UH57	25-30	GCAACATTACAGGTACTGAA
R24	Pseudomonas_phage_PPpW-4	V5YUU1	25-30	TAAAAACCCTCACTGAAACAGGG
R25	Vibrio_phage_10N	A0A2I7RNL7	35-37	ACTTACCTTTCACTATAGCAGCA
R26	Alteromonas_virus_vB_AspP-H44	A0A220YL66	Not Found	TGGTGACTACAGAGCAGCAG

Table 3.1: Bioinformatically-identified single-subunit RNA polymerases selected for *in vitro* functional characterization.

3.3.2 - Identification of novel active RNAP biocatalysts via high-throughput *in vitro* functional characterisation

Previous studies focussed on identifying new RNAP biocatalysts for mRNA manufacture have relied on recombinant production of individual 'test' enzymes in *E. coli* cell-hosts, prior to characterisation in IVT reactions (Lu et al., 2019; Wang et al., 2022; Zhu et al., 2013a). This time-consuming method is undesirable for characterisation of a large RNAP library, particularly given that manufacture of complex proteins at appropriate yield and quality can require significant process optimisation (Bhatwa et al., 2021; Gopal and Kumar, 2013). Moreover, variation in

recombinant protein stability and purity may prevent accurate quantification of relative enzyme activities across the library. Accordingly, to functionally characterise our 26 novel RNAPs in parallel, we developed a high-throughput testing platform that does not require production and purification of each polymerase. This was achieved by adapting a cell-free coupled transcription-translation system that has previously been employed to rapidly assess activity of variant T7 polymerases (Egorova et al., 2021)(Cui et al., 2023). As shown in Figure 2A, this platform utilises a mastermix containing rabbit reticulocyte lysate and recombinant SP6 RNAP to facilitate *in vitro* production of a 'test' RNAP, which then in turn drives expression of a Nano-luciferase reporter-gene under the control of its cognate promoter.

Protein coding sequences and predicted cognate promoter elements for each of the 26 test enzymes were chemically synthesized and inserted into the appropriate RNAP screening platform vectors (Fig. 2A). Resulting plasmid-pairs were individually mixed with the SP6 RNAP-rabbit reticulocyte lysate mastermix, and Luciferase production was measured after incubating the reaction for 1 hr at 30 °C (recommended assay reaction temperature). As shown in Figure 2B, 8/26 enzymes were functionally active, driving luciferase expression levels that ranged from 10% - 161% of that achieved using the control T7 RNAP. Accordingly, approximately 70% of tested enzymes were non-functional, highlighting the difficulty associated with identifying novel RNAP biocatalysts.

There was no significant correlation between predicted enzyme temperature optima and observed activity at 30 °C. However, to further assess the impact of reaction parameters on polymerase performance, we tested enzyme activities at increased (37 °C) and decreased (20 °C) temperatures. While the same eight RNAPs were functional at 20 °C, only three of these enzymes displayed activity at 37 °C.

Moreover, apart from R6 which drove highest Luciferase expression levels under all conditions tested, the relative performance of polymerases varied with temperature. Although our objective was to identify effective polymerase-promoter pairs, rather than to precisely elucidate their relative performance characteristics, these data indicate that enzymes active over a narrow range of temperatures may be incorrectly categorized as non-functional. However, we concluded that this was unlikely when testing across three separate temperature set-points, and that enzyme inactivity in our screening platform was more likely due to either i) inaccurate annotation/sequencing of putative RNAP coding sequences or ii) incorrect promoter prediction.

Although enzyme activity in the cell-free screening system may not be directly predictive of performance in IVT-based mRNA manufacturing processes, it is notable that polymerase R6 drove higher levels of luciferase expression than T7 in all conditions tested (increase ranging between 160% – 620%), including a 220% increase at T7s optimum reaction temperature (37 °C). Five further RNAPs (R3, R5, R9, R14, R15) facilitated luciferase titres greater than or equal to that achieved with T7 in at least one reaction condition. Accordingly, these enzymes may exhibit higher processivity/catalytic activity than T7 and could therefore have potential use in enhancing mRNA production yields. Moreover, their use may permit simplified downstream processing operations via reduced formation of product-related impurities, particularly as many of these RNAPs exhibit relatively high activities at low temperatures (Wang et al., 2022). While further characterisation is required to fully assess their bioindustrial utility, the identification of eight novel functional enzymes more than doubles the number of available RNAPs, expanding the biocatalyst solution space for mRNA manufacture by ~130%.

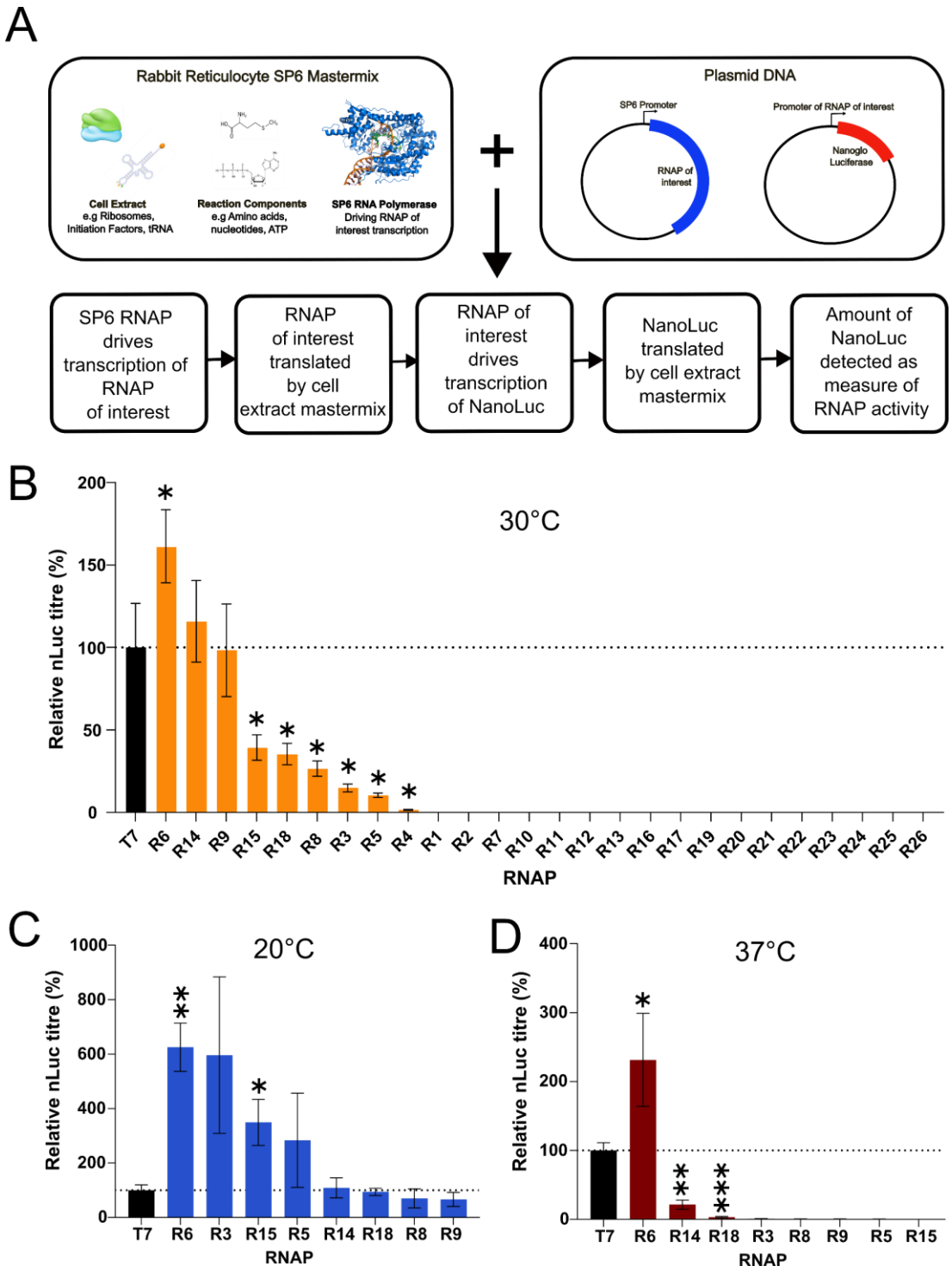


Figure 3.2: RNAPs were functionally characterised in a cell-free coupled transcription-translation system (**A**). Protein coding and cognate promoter sequence pairs were inserted into screening platform vectors and incubated with SP6-rabbit reticulocyte mastermix at 30 °C (**B**), 20 °C (**C**) and 37 °C (**D**). Luciferase expression was quantified 1 hr post-incubation; data are expressed as a percentage of the production achieved using the control T7 RNAP. Values represent the mean + *SD* of three independent experiments ($n = 3$, each performed in triplicate). Statistical significance was determined by two-tailed unpaired t-test compared to T7 nLuc titre, and defined as $p \leq 0.05$ (*=p≤0.05, **=p≤0.01, ***=p≤0.001, ****=p≤0.0001).

3.3.3. Cognate Promoter prediction is the critical limiting factor restricting further expansion of the RNAP biocatalyst solution space.

The finding that ~70% of characterized enzymes were non-functional in *in vitro* tests (Fig. 2) indicates that the vast majority of putative RNAPs cannot be simply extracted from online databases and directly employed in mRNA manufacturing applications. Given that RNAPs are known to display highly stringent promoter recognition requirements (Rong et al., 1998), we hypothesised that enzyme inactivity may have resulted from inaccurate predictions of cognate promoter sequences. To exemplify this, we characterized the ability of R6, the best performing polymerase in *in vitro* screens, to initiate transcription from the promoters of other functional enzymes. As shown in Figure 3A, R6 could not drive quantifiable gene expression from any of these variant elements, where even a single nucleotide change was sufficient to completely abolish transcriptional output. These data highlight that the ability to exploit any given potential RNAP biocatalyst is heavily dependent on highly accurate definition of its cognate promoter sequence.

We reasoned that RNAP promoter prediction tools may be incapable of precisely defining new elements that are significantly divergent from currently known sequences, as evidenced by our inability to derive cognate promoters for ~40% of bioinformatically-determined RNAP clusters (see Section 3.1). Indeed, given the paucity of characterised RNAP promoters, novel ‘test’ enzymes may recognise sequence motifs and architectures that are i) substantially different to those used to train/design current algorithms, and accordingly ii) beyond the predictive capabilities of available tools. Rationalising that divergence in promoter structure/sequence would be underpinned by differences at the amino acid level, we investigated whether enzyme inactivity was associated with DNA binding domain sequences that varied significantly to those of well-studied biocatalysts. As shown in Figure 3B, 8/10

polymerases that share relatively high DNA binding domain sequence identity with T7 (>30%) were found to be active, while all 16 enzymes that share relatively low similarity (<30%) were non-functional. In contrast, cognate promoter sequence similarity with T7 promoter was not a good predictor of RNAP functionality, where 6/8 active and 12/16 inactive elements shared between 40% and 65% sequence identity with T7 (Fig. 2B). These data suggest that we lack the ability to accurately predict divergent promoter elements for new RNAPs when shared DNA binding domain sequence identity with well-studied enzymes falls below a critical threshold.

Our findings indicate that potential RNAPs can be efficiently screened *in silico*, where enzymes that share below ~30% DNA-binding domain sequence identity with T7 are unlikely to be functional *in vitro* owing to incorrect promoter definition. However, as shown in Figure 3C, this cut-off removes approximately 36% of the theoretical biocatalyst solution space for mRNA production. Of the remaining 225 polymerases, 75 share relatively high overall protein sequence identity (>75%) with T7. Such enzymes are considered unlikely to exhibit substantial differences to T7 in key performance criteria such as enzyme processivity and product-related impurity generation. Accordingly, only 150 RNAPs are predicted to be both active *in vitro* and potentially display novel, desirable bioproduction functionalities (including the eight we have identified in this study). This analysis therefore highlights 142 promising additional biocatalyst targets for future investigation, including 88 that do not share high sequence identity (>75%) with either the 6 previously characterized RNAPs or the 8 enzymes identified in this study (listed in Supplementary table 1). However, it also suggests that >120 potentially useful enzymes, are currently difficult to exploit, highlighting promoter prediction capability as the key limiting factor preventing comprehensive exploitation of the theoretical biocatalyst solution space for mRNA

production. Although the cognate promoters of individual polymerases can be elucidated via non-bioinformatic laboratory techniques (Lu et al., 2019), these time-intensive methods are intractable when testing multiple enzymes in parallel. Accordingly, full exploration of the putative RNAP biocatalyst solution space to optimise mRNA production processes will likely require significant advancements in phage promoter prediction tools.

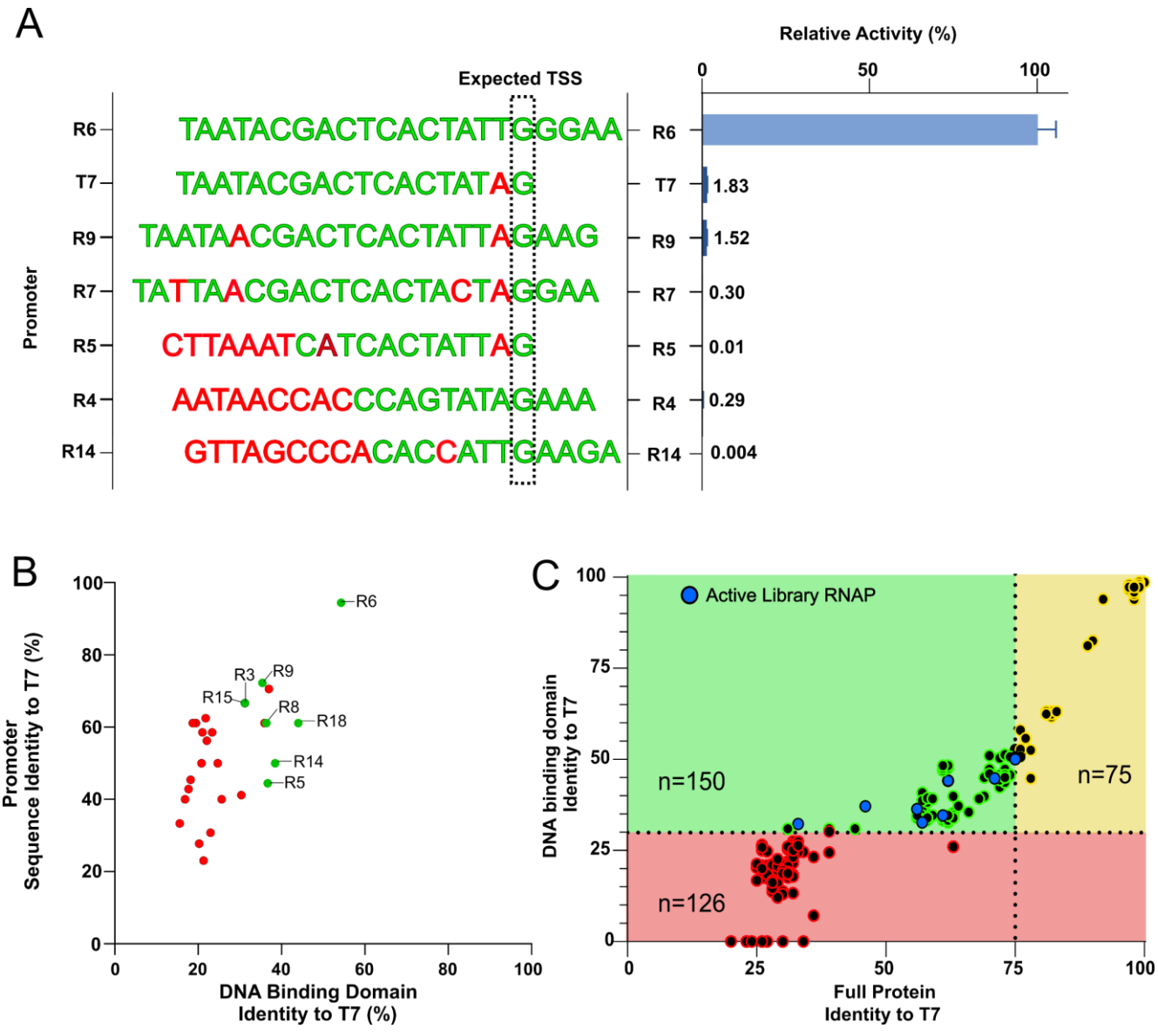


Figure 3: **A**) The ability of RNAP6 (Pectobacterium phage PP74) to drive Luciferase expression from varying non-cognate promoter elements was evaluated in cell-free coupled transcription-translation assays (see Fig. 2). Data are expressed as a percentage of the production achieved using the cognate RNAP6 promoter. Values represent the mean + SD of three independent experiments ($n = 3$, each performed in triplicate). **B**) Test library RNAP-promoter pairs were analysed to determine relative DNA binding domain and promoter sequence identity with T7. Pairs that were found to be active or inactive in functional characterisation tests are shown as green and red dots respectively. **C**) The entire theoretical RNAP biocatalyst solution space was analysed to identify promising future targets for *in vitro* characterisation (green section). Enzymes predicted to have incorrect promoter definitions or similar bioproduction phenotypes to T7 are shown in the red and yellow sections respectively. Functional RNAPs identified in this study are shown as blue dots.

3.3.4 - Novel identified RNAPs enhance the biocatalyst solution space for IVT-based mRNA production.

To validate that novel RNAPs identified via our HT cell-free screening platform have utility in mRNA manufacturing processes, we recombinantly produced polymerases R5 and R6 in *E. coli*. These polymerases were chosen to represent highly- and moderately-active enzymes, where R6 (Pectobacterium phage PP74) was previously shown to be the best performing RNAP in all temperatures tested, and R5 (Dickeya phage Mysterion) drove relatively low-to-medium levels of transcription across varying reaction conditions (Fig 2). Polymerases were manufactured in 0.5 L scale production processes and purified using His-tag affinity and size exclusion chromatographic operations. Purified recombinant RNAPs were then utilised in IVT reactions to manufacture *Cypridina* Luciferase (CLuc) mRNA. As shown in Figure 4, both enzymes drove significant levels of Cluc expression, validating their function as biocatalysts for synthetic mRNA production. To evaluate enzyme robustness, we tested the

performance of each RNAP at a range of pH (predicted optimum ± 1) and temperature (predicted optimum ± 5°C) set-points. Both RNAPs were functional across all conditions tested, where R5 performance was relatively constant, and R6 activity increased with temperature. The latter highlights that host phage growth temperatures are not directly predictive of optimal *in vitro* reaction conditions for recombinant RNAPs.

As shown in Figure 4, utilisation of R6 at 'optimal' reaction parameters (pH 7.9, 35 °C) facilitated mRNA product titers >60% of that achieved when using NEB recombinant T7 at recommended conditions (pH 7.9, 37 °C). Although R6 drove higher levels of gene transcription than T7 in our cell-free system, it is not surprising that T7s relative activity was enhanced in IVT processes given that NEB T7 is a highly-pure engineered enzyme with fully-optimised reaction conditions. Indeed, we anticipate that rational protein engineering/evolution, coupled with improved purification techniques and reaction parameters (E.g., optimised $MgCl_2$ concentration), will significantly increase R6s biocatalytic activity in IVT-based mRNA production. Irrespective of this, by initially facilitating product yields equivalent to ~61% of that achieved by optimised T7, R6 is considered a highly active biocatalyst.

Polymerase R5 enabled product titers ~47% of that achieved using R6, suggesting that the comparative performance of novel RNAPs in cell-free testing platforms is broadly predictive of their relative ability to maximise mRNA yields in IVT manufacturing processes. Accordingly, we concluded that the additional 6 novel enzymes identified in this study (Erwinia amylovora phage Era103, Pectobacterium phage DUPP II, Pectobacterium phage Jarilo, Delftia phage IME-DE1, Pseudomonas phage Henninge, Citrobacter phage CR8) are also likely to facilitate moderate-to-high mRNA production yields. While we cannot currently comment on the relative ability of

these new polymerases to enhance product quality, previous work suggests they will generate variable levels of product-related impurities, such as dsRNA and truncated species (Lu et al., 2019; Wang et al., 2022; Xia et al., 2022; Zhu et al., 2013a, 2015). Indeed, this new library is particularly likely to exhibit differential bioproduction phenotypes, given that they were specifically selected based on sharing minimal amino acid sequence similarities. We therefore conclude that addition of these eight novel functional enzymes to the RNAP biocatalyst solution space will significantly enhance IVT-based mRNA manufacturing optimisation strategies.

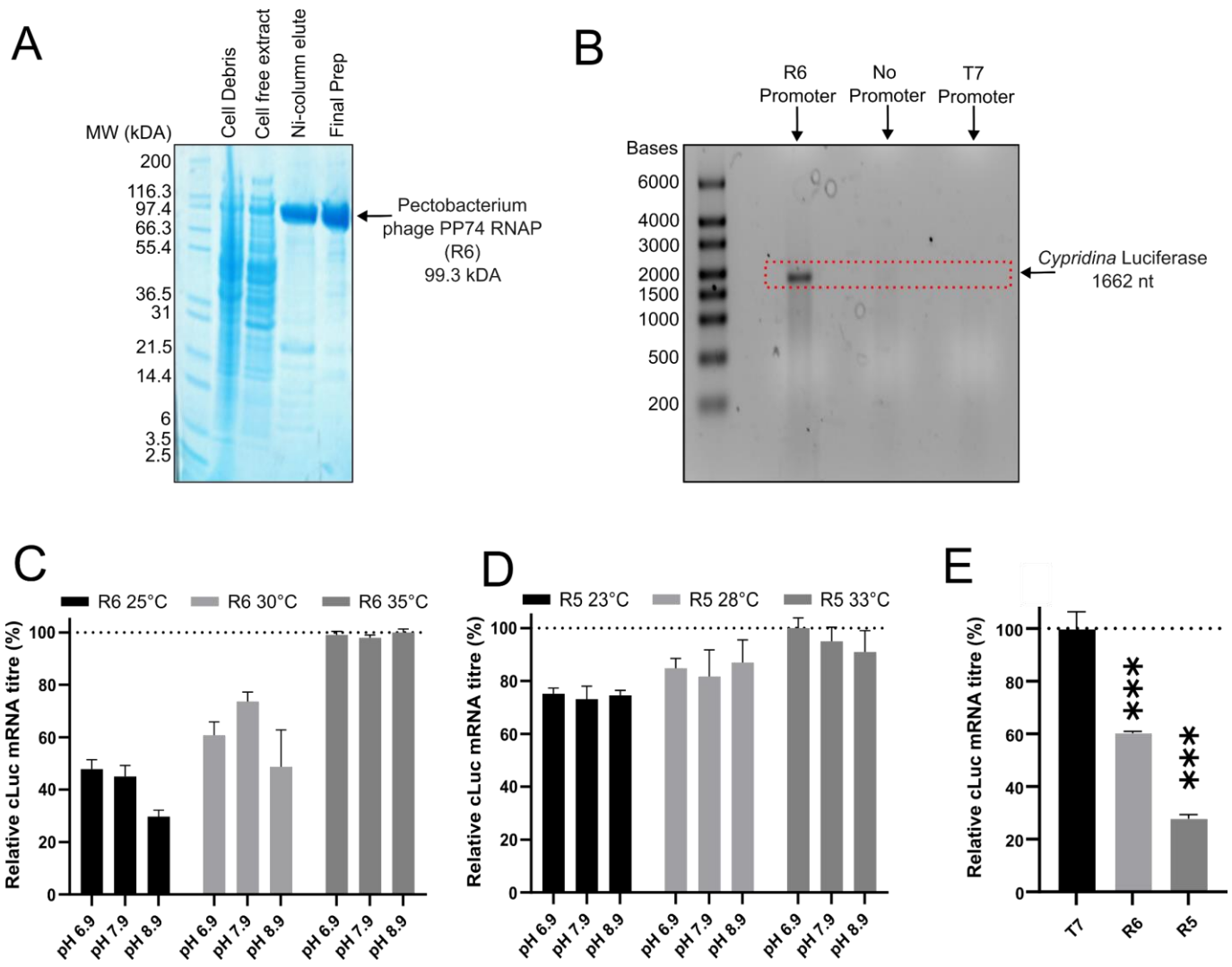


Figure 4: RNAP 5 and 6 were recombinantly produced in *E. coli* and purified using His-tag affinity and size exclusion chromatography (**A**). Purified enzymes were used to manufacture Luciferase mRNA in *in vitro* transcription (IVT) reactions, and production of full-length product was verified by gel electrophoresis analysis (**B**). IVT production processes utilising RNAP6 (**C**) and RNAP5 (**D**) were performed in varying reaction conditions and resulting mRNA titers were quantified using nanodrop spectrophotometry. Using identified optimal reaction parameters for each enzyme, the relative performance of recombinant polymerases was evaluated compared to an NEB T7 control (**E**). Data in E are expressed as a percentage of the production achieved using T7. In C, D and E values represent the mean + SD of three independent experiments ($n = 3$, each performed in triplicate). Statistical significance was determined by two-tailed unpaired t-test compared to T7 RNAP, and defined as $p \leq 0.05$ (*=p≤0.05, **=p≤0.01, ***=p≤0.001, ****=p≤0.0001).

3.4 - Concluding Remarks

The eight novel sequence-diverse functional RNAPs identified in this study substantially increases the number of biocatalysts available for mRNA production. Although further work is required to comprehensively define their relative performance characteristics, particularly their associated impurity generation profiles, this expansion of the biocatalyst solution space significantly enhances design options for molecule-, process-, and application-specific optimisation of mRNA product yield and quality. This improved flexibility will become increasingly useful as mRNA product lines continue to diversify towards large, complex molecules that pose new manufacturing challenges (Bai et al., 2023; Blakney et al., 2021; Chen et al., 2022; Pourseif et al., 2022; Qu et al., 2022). Our combined *in silico* and *in vitro* analysis of the theoretical biocatalyst solution space i) showed that full exploitation of potential RNAPs is restricted by cognate promoter prediction capabilities, but, also ii) identified a panel of enzymes that are particularly promising for future investigation. With respect to the latter, the screening platform we have developed can be utilised to rapidly test additional RNAP libraries, including engineered variants of existing polymerases. In conclusion, by more than doubling the number of available polymerases, and providing associated methods to select and screen further new enzymes, this study has facilitated a significant expansion of the RNAP biocatalyst solution space, enhancing strategies to optimise and troubleshoot IVT-based mRNA production processes.

3.5 Paper I References

Bai, Y., Liu, D., He, Q., Liu, J., Mao, Q., and Liang, Z. (2023). Research progress on circular RNA vaccines. Front. Immunol. 13, 1–12.

Bhatwa, A., Wang, W., Hassan, Y.I., Abraham, N., Li, X.Z., and Zhou, T. (2021). Challenges Associated with the Formation of Recombinant Protein Inclusion Bodies in Escherichia coli and Strategies to Address Them for Industrial Applications. Front. Bioeng. Biotechnol. 9, 630551.

Blakney, A.K., Ip, S., and Geall, A.J. (2021). An update on self-amplifying mRNA vaccine development. Vaccines 9, 1–26.

Chen, T.H., Potapov, V., Dai, N., Ong, J.L., and Roy, B. (2022). N 1-methyl-pseudouridine is incorporated with higher fidelity than pseudouridine in synthetic RNAs. Sci. Rep. 12, 1–13.

Cui, M., Wong, O., Li, Q., and Wang, W. (2023). An Assay Method for Characterizing Bacteriophage T7 RNA Polymerase Activity by Transcription-Translation (TX-TL) System. J. Biochem. 173, 343–352.

Dousis, A., Ravichandran, K., Hobert, E.M., Moore, M.J., and Rabideau, A.E. (2023). An engineered T7 RNA polymerase that produces mRNA free of immunostimulatory byproducts. Nat. Biotechnol. 41, 560–568.

Egorova, T., Shuvalova, E., Mukba, S., Shuvalov, A., Kolosov, P., and Alkalaeva, E. (2021). Method for rapid analysis of mutant rna polymerase activity on templates containing unnatural nucleotides. Int. J. Mol. Sci. 22.

Al Fayez, N., Nassar, M.S., Alshehri, A.A., Alnefaie, M.K., Almughem, F.A., Alshehri, B.Y., Alawad, A.O., and Tawfik, E.A. (2023). Recent Advancement in mRNA Vaccine Development and Applications. Pharmaceutics 15, 1–24.

Fischer, S., Handrick, R., and Otte, K. (2015). The art of CHO cell engineering: A comprehensive retrospect and future perspectives. Biotechnol. Adv. 33, 1878–1896.

Gopal, G.J., and Kumar, A. (2013). Strategies for the production of recombinant protein in escherichia coli. Protein J. 32, 419–425.

Guilleres, J., Lopez, P.J., Proux, F., Launay, H., and Dreyfus, M. (2005). A mutation in T7 RNA polymerase that facilitates promoter clearance. Proc. Natl. Acad. Sci. U. S. A. 102, 5958–5963.

Lavigne, R., Sun, W.D., and Volckaert, G. (2004). PHIRE, a deterministic approach to reveal regulatory elements in bacteriophage genomes. Bioinformatics 20, 629–635.

Liu, C., Shi, Q., Huang, X., Koo, S., Kong, N., and Tao, W. (2023). mRNA-based cancer therapeutics. Nat. Rev. Cancer 23, 526–543.

Lu, X., Wu, H., Xia, H., Huang, F., Yan, Y., Yu, B., Cheng, R., Drulis-Kawa, Z., and Zhu, B. (2019). Klebsiella phage KP34 RNA polymerase and its use in RNA synthesis. Front. Microbiol. 10.

Lyon, S., and Gopalan, V. (2018). A T7 RNA Polymerase Mutant Enhances the Yield of 5′-Thienoguanosine-Initiated RNAs. ChemBioChem 19, 142–146.

Madeira, F., Park, Y.M., Lee, J., Buso, N., Gur, T., Madhusoodanan, N., Basutkar, P., Tivey, A.R.N., Potter, S.C., Finn, R.D., et al. (2019). The EMBL-EBI search and sequence analysis tools APIs in 2019. Nucleic Acids Res. 47, W636–W641.

Pourseif, M.M. (2022). Self-amplifying mRNA vaccines : Mode of action , design , development and optimization. Drug Discov. Today 27, 103341.

Qin, S., Tang, X., Chen, Y., Chen, K., Fan, N., Xiao, W., Zheng, Q., Li, G., Teng, Y., Wu, M., et al. (2022). mRNA-based therapeutics: powerful and versatile tools to combat diseases. Signal Transduct. Target. Ther. 7.

Qu, L., Yi, Z., Shen, Y., Lin, L., Chen, F., Xu, Y., Wu, Z., Tang, H., Zhang, X., Tian, F., et al. (2022). Circular RNA vaccines against SARS-CoV-2 and emerging variants. Cell 185, 1728-1744.e16.

Rong, M., Biao, H., Mcallister, W.T., and Durbin, R.K. (1998). Promoter specificity determinants of T7 RNA polymerase. Proc. Natl. Acad. Sci. U. S. A. 95, 515–519.

Sampaio, M., Rocha, M., Oliveira, H., DIas, O., and Valencia, A. (2019). Predicting promoters in phage genomes using PhagePromoter. Bioinformatics 35, 5301–5302.

Steinegger, M., and Söding, J. (2017). MMseqs2 enables sensitive protein sequence searching for the analysis of massive data sets. Nat. Biotechnol. 35, 1026–1028.

Tihanyi, B., and Nyitray, L. (2020). Recent advances in CHO cell line development for recombinant protein production. Drug Discov. Today Technol. 38, 25–34.

Vavilis, T., Stamoula, E., Ainatzoglou, A., Sachinidis, A., Lamprinou, M., Dardalas, I., and Vizirianakis, I.S. (2023). mRNA in the Context of Protein Replacement Therapy. Pharmaceutics 15.

Wang, G., Cheng, R., Chen, Q., Xu, Y., Yu, B., Zhu, B., Yin, H., and Xia, H. (2022). mRNA produced by VSW-3 RNAP has high-level translation efficiency with low inflammatory stimulation. Cell Insight 1.

Webb, C., Ip, S., Bathula, N. V., Popova, P., Soriano, S.K.V., Ly, H.H., Eryilmaz, B., Nguyen Huu, V.A., Broadhead, R., Rabel, M., et al. (2022). Current Status and Future Perspectives on MRNA Drug Manufacturing. Mol. Pharm. 19, 1047–1058.

Wu, M.Z., Asahara, H., Tzertzinis, G., and Roy, B. (2020). Synthesis of low immunogenicity RNA with high-temperature *in vitro* transcription. RNA 26, 345–360.

Xia, H., Yu, B., Jiang, Y., Cheng, R., Lu, X., Wu, H., and Zhu, B. (2022). Psychrophilic phage VSW-3 RNA polymerase reduces both terminal and full-length dsRNA byproducts in *in vitro* transcription. RNA Biol. 19, 1130–1142.

Zhu, B., Tabor, S., Raytcheva, D.A., Hernandez, A., King, J.A., and Richardson, C.C. (2013). The RNA polymerase of marine cyanophage Syn5. J. Biol. Chem. 288, 3545–3552.

Zhu, B., Hernandez, A., Tan, M., Wollenhaupt, J., Tabor, S., and Richardson, C.C. (2015). Synthesis of 2′-Fluoro RNA by Syn5 RNA polymerase. Nucleic Acids Res. 43, e94–e94.

Zhu, F., Ma, S., Wen, H., Rao, M., Zhang, P., Cui, Y., Yang, H., Tan, C., Chen, J., and Pan, P. (2022). Development of a novel circular mRNA vaccine of six protein combinations against Staphylococcus aureus Development of a novel circular mRNA vaccine of six protein combinations. J. Biomol. Struct. Dyn. 0, 1–21.

Chapter 4 – Paper II

Engineering an *E. coli*-based *in vivo* mRNA manufacturing platform

Edward Curry[1], George Muir[1], Jixin Qu[1], Zoltan Kis[1], Martin Hulley[2], Adam Brown [1]

[1] Department of Chemical and Biological Engineering, University of Sheffield, Sheffield, UK.

[2] Bioprocess Development, AstraZeneca, Cambridge, UK.

Available in Biotechnology and Bioengineering

DOI: 10.1002/bit.28684

Summary

The second paper in this thesis details an alternative platform for the production of synthetic mRNA, using *E. coli* as a cellular factory for the production of the molecule. A coordinated engineering approach focusing on the mRNA sequence, cell chassis, media and DNA sequence allows for the increased accumulation of the molecule within the cell. The described new technology platform may have particular application in contexts where IVT systems are unavailable (e.g., due to reagent shortages), new classes of product formats necessitate process optimisation (e.g., production of very large RNA molecules), or manufacturing costs need to be significantly reduced.

Contributions

Edward Curry – Conceptualisation, methodology, analysis, investigation, manuscript preparation. **George Muir** – Investigation (Fragment analyser), analysis. **Jixin Qu** – Investigation (Large scale purification), analysis. **Zoltan Kis** - Investigation, analysis. **Martyn Hulley** – Conceptualisation, analysis. **Adam Brown** – Conceptualisation, methodology, analysis, manuscript preparation, funding acquisition.

Abstract

Synthetic mRNA is currently produced in standardised *in vitro* transcription systems. However, this one-size-fits-all approach has associated drawbacks in supply chain shortages, high reagent costs, complex product-related impurity profiles and limited design options for molecule-specific optimisation of product yield and quality. Herein, we describe for the first time development of an in vivo mRNA manufacturing platform, utilising an *E. coli* cell chassis. Coordinated mRNA, DNA, cell and media engineering, primarily focussed on disrupting interactions between synthetic mRNA molecules and host cell RNA degradation machinery, increased product yields >40-fold compared to standard unengineered wild-type *E. coli* expression systems. Mechanistic dissection of cell factory performance showed that product mRNA accumulation levels approached theoretical limits, accounting for ~30% of intracellular total RNA mass, and that this was achieved via host-cell's reallocating biosynthetic capacity away from endogenous RNA and cell biomass generation activities. We demonstrate that varying sized functional mRNA molecules can be produced in this system and subsequently purified in large- or small-scale processes. Accordingly, this study introduces a new mRNA production technology, expanding the solution space available for mRNA manufacturing.

4.1 - Introduction

Synthetic mRNA has potential use in a wide range of applications, including cancer immunotherapy, protein replacement therapy, genome editing, pluripotent stem cell generation, and vaccines against infectious diseases (Baden et al., 2021; Breda et al., 2023; Gan et al., 2019; Qin et al., 2022; Vavilis et al., 2023a). In all cases, mRNA molecules are currently produced in standardised *in vitro* transcription (IVT) systems, comprising an RNA Polymerase biocatalyst, DNA template, modified nucleosides, magnesium-containing buffer and a capping enzyme/analog (Ouranidis et al., 2022). These simple, modular, cell-free production platforms embed flexibility and predictability in mRNA manufacture, while substantially reducing process-related impurities (Whitley et al., 2022). However, the requirement for purified input components is associated with relatively high costs, and critical reagent shortages (Kis et al., 2021). Moreover, downstream purification processes are complicated by complex product-related impurity profiles, that include immunostimulatory double-stranded RNA and abortive transcripts (Gholamalipour et al., 2018; Rosa et al., 2021). However, despite these drawbacks, expanding product diversification (particularly with respect to size), highly variable intended applications (with associated variability in required production scale, purity, cost, etc), and the increasing pressure placed on reagent/equipment supplies by growing demand for mRNA synthesis, there are currently no alternative technology platforms available for mRNA manufacture.

Cell-based production systems are the dominant choice for manufacture of other bioproducts, such as AAV vectors, recombinant proteins and recombinant DNA plasmids (Agostinetto et al., 2022; Jiang and Dalby, 2023; McElwain et al., 2022). Although they are associated with relatively complex and costly downstream processing steps to remove host-cell impurities, this is somewhat mitigated by the

availability of well-characterised chromatographic and membrane-based unit operations (Fan et al., 2023; Sripada et al., 2022). As a relatively simple macromolecule, synthetic mRNA could theoretically be produced in virtually any microbial cell factory. *E. coli* is a particularly attractive expression host given that decades of use in recombinant plasmid DNA production has led to development of very low-cost, standardised, easy to scale (up to 100,000L) flexible manufacturing platforms (Pontrelli et al., 2018; Yang et al., 2021). Indeed, these benefits have seen *E. coli* utilised as a biocatalyst for production of RNA aptamers and double stranded RNA (dsRNA) molecules (Delgado-Martín and Velasco, 2021; Ma et al., 2020; Ponchon and Dardel, 2011; Ponchon et al., 2009b, 2013)

The primary limitation of mRNA production in microbial expression hosts is endogenous pathways that encode rapid RNA turnover, where the average mRNA half-life in E.coli is ~5 mins (Esquerré et al., 2015; Mohanty and Kushner, 2022). For dsRNA manufacture, multigram per liter yields have been achieved in *E. coli* bioprocesses by deleting RNase III, a non-essential dsRNA-targeting endonuclease (Pertzev and Nicholson, 2006). However, single stranded mRNA decay is mediated by RNAseE, an essential enzyme required for global RNA metabolism. Although RNAseE has broad substrate specificity, various sequence features have been shown to increase its relative specific activity on individual mRNA species, including unstructured AU rich regions, and, most critically, the presence of a 5'-monophosphate (Bae et al., 2023; Callaghan et al., 2005; Richards and Belasco, 2023). However, other molecule-specific features, such as RNA-binding protein binding sites, codon usage, and secondary structure profiles can reduce RNase E mediated mRNA turnover (Börner et al., 2023; Roux et al., 2022). More generically, global mRNA half-life is affected by both the relative abundance and activity level of RNase E (Mohanty

and Kushner, 2022). Accordingly, the half-life of a specific mRNA molecule within an *E. coli* cell chassis is determined by a complex interplay between the mRNA sequence/structure and the host cell's complement of RNA degradation machinery components.

Herein we report coordinated mRNA, DNA, media and host cell engineering to dramatically increase synthetic mRNA accumulation and maintenance in *E. coli* cell factories. Achieving mRNA yields >40-fold greater than standard 'unengineered' *E. coli* expression systems, we demonstrate rapid production and purification of a range of functional mRNA products. In doing so, we introduce a new technology platform for mRNA manufacturing solution spaces. This may be particularly useful in contexts where IVT systems are unavailable (e.g., due to reagent shortages), product formats necessitate process optimisation (e.g., production of very large RNA molecules), or manufacturing costs need to be significantly reduced.

4.2 - Materials and Methods

4.2.1 Plasmid Design and Construction

The baseline mRNA production construct was designed with 5' and 3' UTR sequences from Xenopus beta-globin, directly flanking a human codon optimised eGFP sequence. All DNA sequences containing stabilisation features of interest flanking the UTRs and coding sequence were synthesized and cloned into pET-24b (Novagen) by Twist Bioscience. Sequences were inserted using the BamHI and XhoI restriction sites, placing the GOI downstream of the T7 promoter and lac operon, and removing the ribosome binding site from the original plasmid. High copy number plasmids were constructed by amplifying the mRNA encoding region from the original pET29b vector by PCR, and inserting it into pRNA128A, a vector with the ColE1 origin of replication. Synthetic triple terminator constructs were assembled using DNA fragments purchased from Genewiz. The fragments consisted of the T3 and *E. coli* endogenous rrnB T1 terminators, and were ligated into vectors using the XhoI and BlpI restriction sites, directly upstream of the T7 terminator. Sequences of DNA parts and plasmids are available in Supplementary table 1 and 2.

4.2.2 Synthetic mRNA Expression

mRNA encoding plasmids were transformed into BL21 (DE3) or BL21 Star (DE3) (Invitrogen) *E. coli* strains, typically used for recombinant protein production. 5ml starter cultures were inoculated with a single colony, and grown overnight in Luria-Bertani (LB) broth (Thermo Fisher), containing 50 µg/ml kanamycin (Thermo Fisher) at 37 °C, 200 rpm. For small scale expression, 100 µL of starter culture was used to inoculate 10 ml of LB medium containing 50 µg/ml kanamycin, and cells were grown at 37 °C, 200rpm, until the OD_{600nm} reached 0.4-0.6. For large scale expression, 200

ml of LB was inoculated with 5 ml of overnight culture. IPTG was then added to a final concentration of 1 mM. Where RNase E inhibitor was used, it was added from a 100x stock at the point of IPTG addition 500 µL of culture was pelleted at 30 minute intervals post induction, and pellets stored at -80 °C. Total cellular capacity (integral of viable cell concentration), cell-specific growth rate and cell-specific productivity were calculated as previously described (Brown et al., 2019).

4.2.3 Small Scale RNA Extraction

Total RNA was extracted using the GenElute Total RNA purification kit (Sigma-Aldrich) according to the manufacturer's instructions. Briefly, cell pellets were resuspended in 100 µL of TE buffer containing 1 mg/ml Lysozyme, and incubated at room temperature for 5 minutes. 300 µL of buffer RL and 200 µL of 96-100% ethanol were added to the lysate before vortexing. Lysate was then applied to the spin column resin, before washing with ethanol solution. RNA was eluted in 50 µL of elution solution. Residual DNA in the RNA sample was then removed through addition of 2 units of RNase free DNase I, and incubation at 37 °C for 30 minutes. RNA was purified from the DNase reaction using the Monarch RNA Cleanup kit (50 µg) (New England Biolabs), following the manufacturer's instructions. Briefly, 2 volumes of RNA binding buffer and 3 volumes of 96-100% ethanol were added to the RNA sample. The RNA was then bound to the spin column resin, before washing with ethanol solution. RNA was eluted from the column in nuclease free water, and stored at -80C. The concentration of samples was determined using a Nanodrop spectrophotometer (Thermo Fisher). Integrity of RNA samples were assessed using denaturing agarose gel electrophoresis. An equal volume of 2X RNA loading dye was added to 200 ng of total RNA, before heating to 65 °C for 5 minutes, and loading on a 1% agarose gel, which was run at 80 V for 40 minutes.

4.2.4 Large Scale RNA Extraction

200 ml of *E. coli* culture was pelleted at 10000 xg for 10 minutes, before extraction by a protocol based upon RNASwift (Nwokeoji et al., 2016). Cells were suspended in 5 ml of 5 mg/ml lysozyme solution (Thermo Fisher), and allowed to lyse for 10 minutes at room temperature. 12ml of lysis solution (4% SDS, 0.5 M NaCl (Thermo Fisher)) was added, before incubation for 5 minutes at 65 ºC. 6.8ml of 5M NaCl of was added, and suspensions were placed on ice for 5 minutes to promote precipitation of SDS. Suspensions were centrifuged at 10000 xg for 20 minutes at 4 ºC, and supernatant was transferred to a separate tube. RNA was precipitated from the supernatant by isopropanol or ethanol, and stored at -20 ºC.

4.2.5 Digital Droplet PCR

mRNA copy number per nanogram of Total RNA was determined by one-step reverse transcription digital droplet PCR (RT-ddPCR). Primers were designed to amplify a 150 bp region in the middle of eGFP, universal to every construct under investigation. ddPCR reaction mixtures of a final volume of 20 µL were comprised of 5 µL one step RT-ddPCR supermix (Bio-Rad), 2 µL reverse transcriptase (Bio-Rad), 1 µL 300 mM DTT (Bio-Rad), 1 µL 900 nM forward/reverse primer mix (Integrated DNA Technologies), 1 µL 250 nM 5'-FAM probe (Integrated DNA Technologies), 9 µL H_20 and 1 µL Total RNA at a concentration of 1 ng/µL. The 20 µL reaction mixture and 70 µL of droplet generation oil (Bio–Rad) were loaded into a DG8 Cartridge, and 40 µL of droplets were generated with the Bio-Rad QX200 droplet generator. Droplets were transferred to a 96-well PCR plate (Bio-Rad), sealed with foil, and placed in a Bio-rad C1000 thermal cycler. Reverse transcription was performed at 50ºC for 1 hour. Polymerase activation was carried out for 10 minutes at 95ºC, before 40 cycles of

denaturation for 1 minute at 95 ºC, before a combined annealing and extension phase for 1 minute at 60 ºC. Enzymes were then deactivated at 98 ºC for 10 minutes, before a final hold phase at 12 ºC for 30 minutes. Positive droplets were detected by the QX200 droplet reader (Bio-Rad), using automatically assigned amplitude thresholds determined by QuantaSoft software (Bio-Rad). Samples were only used in analysis if the number of measured droplets exceeded 12000.

4.2.6 Small Scale mRNA Purification

mRNA was purified from total RNA extraction samples by oligo (dT) enrichment, using the Dynabeads mRNA purification kit (Invitrogen). Approx. 30 µg of Total RNA was adjusted to a volume of 100 µL with nuclease free water. Samples were heated to 65 °C for 2 minutes to disrupt secondary structure, before placing on ice. Magnetic beads were equilibrated in 100 µL of binding buffer (10 mM Tris-HCl, 1M LiCl), before addition of the 100 µL total RNA samples. Beads were incubated with RNA samples for 5 minutes at room temperature with constant rotation, before removal of the supernatant. Beads were washed twice with 200 µL of washing buffer (10 mM Tris-HCl, 1M LiCl). mRNA was eluted from the magnetic beads by addition of 10 mM Tris-HCl, pH 7.5, and its concentration determined by Nanodrop spectrophotometer.

4.2.7 Large Scale mRNA Purification

mRNA from large scale extractions was purified by oligo-dT enrichment utilising an ÄKTA pcc chromatography system (Cytiva). 10ml total RNA was bound to a 1 ml volume Oligo-dt(18) column (Sartorius) in binding buffer (50 mM Sodium Phosphate pH 7, 500 mM NaCl). The column was washed in 20 column volumes of wash buffer (50 mM Sodium Phosphate, 150 mM NaCl). Bound mRNA was eluted in 5ml of water, before quantification by Nanodrop spectrophotometer.

4.2.8 Capillary Gel Electrophoresis

Capillary Gel Electrophoresis (CGE) analysis of RNA integrity and purity was performed with a 5200 Fragment Analyzer System (Agilent, CA, USA), using the DNF-471 RNA Kit (15 nt) (Agilent, CA, USA). The capillary cassette used was FA 12-Capillary Array Short, 33 cm (Agilent, CA, USA). Samples were diluted to <100 ng/µL in nuclease free water. Before each separation, a pre-run voltage was applied (8 kV for 30 seconds), the capillaries were conditioned with the conditioning solution and the capillaries were dipped twice in the rinse buffer. Following this, the capillaries were filled with RNA separation gel (by pressure) and then the sample was introduced using a voltage injection (5 kV for 4 seconds). The separation was then conducted by applying a voltage of 8 kV for 45 minutes. Detection was carried out using Laser Induced Fluorescence (LIF), by fluorescent dye tagging of the RNA.

4.2.9 Transfection of mRNA into HEK293 cells

Suspension adapted HEK293 cells (Thermo Fisher) were routinely cultured in serum-free medium (Thermo Fisher). Cells were maintained in 30 ml volume in 125 ml Erlenmeyer flasks (Corning), at 37 ºC, 85% humidity, and 5% CO_2, with agitation at 140rpm. Cell density and viability was determined by the Countess 3 automated cell counter system (Thermo Fisher). mRNA that required capping prior to transfection was capped using the NEB Vaccinia virus capping system (New England Biolabs), following the manufacturers protocol Cells for transfection were cultured in 24-shallow well plates (Corning), containing 500 µL culture volume, with agitation at 240 rpm. For mRNA transfection, cells were seeded at a density of 0.3 x 10^6 cells per ml in 24-shallow well plates, and incubated for 24 hours. TransIT-mRNA transfection reagent (Mirus) was used to transfect 500 ng of mRNA per well as per manufacturer's

instructions. GFP transfection efficiency was determined by the Countess 3 system using a GFP filter. For fluorescence measurements, cells were by centrifugation at 200 *xg* for 5 minutes, before resuspension in DPBS (Sigma-Aldrich), and determination of fluorescence by plate reader at 488 nm/507 nm.

4.3 Results and Discussion

4.3.1 - mRNA engineering to increase product stability in *E. coli*

The specific activity of RNAse E on a discrete mRNA species directly determines the half-life of that molecule in an *E. coli* cell chassis (Mauger et al., 2019; Mohanty and Kushner, 2022; Viegas et al., 2018). While coding sequence design could theoretically be employed to enhance accumulation of a particular mRNA molecule in *E. coli* (E.g., by optimising codon-usage), the efficacy of such methods would be highly product-specific, dependent on the available design-space for each primary amino acid sequence. Algorithms are available for enhancing mRNA sequence through codon usage, however the design rules applied are unlikely to produce sequences optimal for stability in both an *E. coli* and mammalian context (Leppek et al., 2022; Zhang et al., 2023). Accordingly, to achieve product-agnostic increases in mRNA stability, we focussed on engineering elements that are located outside the protein coding sequence. Firstly, we introduced 'scaffold' tRNA-Lysine motifs at both the 5' and 3' termini (TermtRNA-mRNA, Fig 1A), based on previous findings that i) stable secondary structures significantly reduced the activity of RNase E on mRNA molecules (Richards and Belasco, 2023; Zhang et al., 2021), and ii) incorporation of tRNA motifs increased production of short RNA species in *E. coli* (Nelissen et al., 2012; Ponchon et al., 2009b). Secondly, given that RNase E is preferentially active on RNA species that have a 5' monophosphate (Callaghan et al., 2005), we incorporated ribozyme sequences either side of the untranslated regions (UTRs) to promote self-circularisation of product mRNA (SelfCirc-mRNA, Fig 1B). This design step was aided by recent work describing elements which efficiently catalyse mRNA self-circulation via the 'Permuted Intron Exon' method (Rostain et al., 2020; Wesselhoeft et al., 2018). Finally, we designed a hybrid approach, where sections of the tRNA-Lys motif were

inserted at the 5’ and 3’ termini. Hybridisation of these complementary sequences creates a pseudo-circular molecule, where the gene expression cassette is contained in a single stranded loop attached to the tRNA-Lys structural element (CirctRNA-mRNA, Fig 1C).

TermtRNA-, SelfCirc- and CirctRNA- features were incorporated into a widely used bacterial GFP-expression vector, where transcription of product mRNA was driven by an inducible T7 promoter. These expression plasmids lacked a bacterial ribosome binding site (i.e., to prevent translation of GFP-mRNA in the host-cell) but contained commonly used mammalian 5’ and 3’ untranslated regions, including an encoded polyA tail, to permit translation of purified mRNA in human cells. GFP-vectors were transformed into the *E. coli* protein production strain BL21 (DE3) and small-scale mRNA production processes were carried out. Total RNA was extracted 150 min after induction of expression and GFP mRNA yields were quantified by digital droplet PCR. As shown in Fig 1D, each of the engineered constructs facilitated substantial increases in mRNA product yield, as compared to the standard unengineered control. The best performing construct, SelfCirc-mRNA, enhanced GFP mRNA yield by >11-fold, indicating that these circularised molecules were efficiently shielded from RNAse E-mediated product decay. CirctRNA- (6-fold increase in product yield compared to control) had a significantly reduced stabilising effect compared to TermtRNA- (10-fold increase), which may be due to relatively inefficient formation of the tRNA motif in this molecular context. Unlike the tRNA-based elements, SelfCirc-mRNA is also protected from degradation by 3’-5’ exonucleases, which play a relatively minor role in global mRNA decay in *E. coli* but are particularly active on polyadenylated transcripts (Mohanty and Kushner, 2022). Moreover, circular mRNA molecules can be directly

utilised in downstream applications without requiring the addition of a cap structure or incorporation of modified nucleotides. Indeed, given its low immunogenicity and high molecular stability, both in mammalian cells and during product storage (Deviatkin et al., 2023), circular mRNA is considered a promising molecular format for a variety of applications (Bai et al., 2023; Liu et al., 2022b; Qu et al., 2022). As circularised mRNA facilitated the highest production yields in *E. coli*, while also being associated with simplified downstream processing requirements, we concluded that this emerging product class was particularly well-suited for an *in vivo* biomanufacturing system. Accordingly, we focussed further optimisation of our platform on enhancing production of SelfCirc-mRNA.

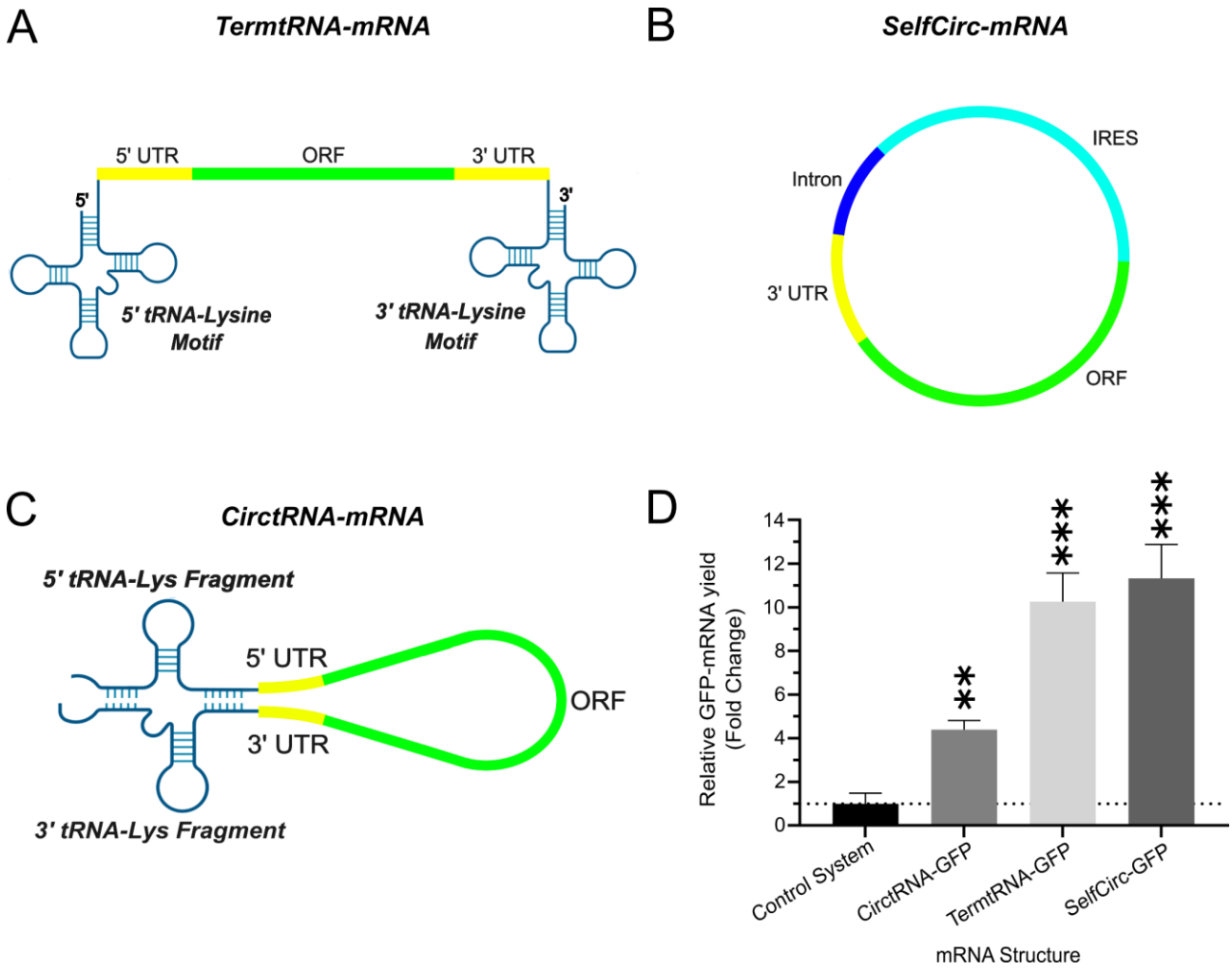


Figure 4.1 - mRNA structures were engineered to increase product stability in *E. coli* host-cells by **A)** including 'scaffold' tRNA-Lysine motifs at both the 5' and 3' termini, **B)** incorporating ribozyme sequences to promote self-circularisation, and **C)** inserting component parts of the tRNA-Lysine motif at the 5' and 3' termini to facilitate formation of pseudo-circular molecules. Designed elements were synthesized, individually inserted into a GFP-expression plasmid and evaluated in 2.5 h production processes **(D)**. Data are expressed as a fold-change of the production exhibited by the unengineered control system. Values represent the mean + *SD* of three independent experiments ($n = 3$, each performed in triplicate). Statistical significance was determined by two-tailed unpaired t-test compared to the control system, and defined as $p \leq 0.05$ (*=p≤0.05, **=p≤0.01, ***=p≤0.001, ****=p≤0.0001). UTR: Untranslated region; IRES: Internal ribosome entry site; ORF: Open reading frame; GFP: Green Fluorescent Protein.

4.3.2 - Cell, DNA and media engineering to maximise mRNA product yield

We previously described a whole pathway engineering approach that maximised production of recombinant proteins in Chinese Hamster Ovary cells (Brown et al., 2019). We hypothesized that a similar strategy could be applied to mRNA manufacturing in *E. Coli* by sequentially improving the host cell chassis, DNA expression vector and cell culture media. Commercially available *E. coli* cell lines have been engineered to reduce RNase E activity to levels that enhance recombinant mRNA stability without impacting global mRNA homeostasis (Heyde and Nørholm, 2021; Miroux and Walker, 1996). Although these strains were originally designed to increase production of 'easy to express' recombinant proteins, they theoretically provide a highly permissive cell background for synthetic mRNA manufacture. To directly test this, we compared SelfCirc-mRNA production in previously utilised standard BL21(DE3) cells and engineered BL21 Star™ (DE3) (*F− ompT hsdSB (rB−, mB−) galdcmrne131* (DE3)) cells. As shown in Fig 2A, cells expressing a mutated RNAse E produced ~1.8-fold more GFP mRNA than the unengineered strain. Although circular mRNA is efficiently protected from RNase E mediated degradation, covalent circularisation requires synthesis of the full-length transcript. A reduction in RNase E activity may therefore enhance synthetic mRNA yields by preventing turnover of nascent product mRNA, increasing the pool of mRNA molecules available for circularisation. Product yields may be further enhanced by cell engineering strategies that increase the host cell's mRNA biosynthesis (E.g., T7 expression level) and/or cell biomass accumulation capacities.

We rationalised that promoter engineering was unlikely to increase product yields, as the expression plasmid already contained a T7 promoter optimised to maximise

recombinant mRNA transcription rates. However, enhancing the number of plasmid copies per cell has previously been shown to enhance manufacture of short dsRNA molecules (Ponchon et al., 2013). Accordingly, we tested the effect of using a pUC origin of replication (Ori), which permits very high plasmid copy numbers per cell (~500-700; (Lee et al., 2006; Lin-Chao et al., 1992)). As shown in Fig 2B, the use of this element did not increase GFP mRNA yields in Star BL21 cells, as compared to the use of the original Rop-ColE1 Ori, despite that construct only encoding maintenance of ~15-20 copies per cell (Bolivar et al., 1977; Lee et al., 2006). This may be caused by the intrinsic metabolic burden associated with replicating and transcribing very high DNA plasmid loads. It is likely that testing a range of synthetic Oris (Joshi et al., 2022; Rouches et al., 2022) will identify a plasmid copy number 'sweet spot' that optimises the quantity of DNA templates available for product biosynthesis without negatively impacting other desirable cellular bioproduction phenotypes.

Beyond the promoter and the Ori, the final DNA plasmid element that can be engineered is the transcriptional terminator. The original expression plasmid utilised a standard class I intrinsic late T7 terminator, TΦ, however this is known to encode a termination efficiency of only ~74% (Carter et al., 1981). Replacing TΦ with a previously described novel triple terminator, comprising a combination of T7 TΦ, T3 and *E. coli* rrnBT1 endogenous terminators, enhanced GFP mRNA yields by ~40% (Fig 2B). This triple terminator has been shown to effectively eliminate read-through transcription by T7 RNA Polymerase (Mairhofer et al., 2015). Accordingly, this terminator facilitates enhanced RNA Polymerase recycling efficiency and increases the total biocatalyst time available for productive synthetic mRNA biosynthesis.

Producing high levels of synthetic mRNA may create product titer-limiting burden/bottlenecks in host cell metabolic pathways. We tested the effect of replacing the commonly utilised protein and plasmid production cell culture media Luria-Bertani broth with other commercially available formulations. Terrific Broth and Bacto CD Supreme Fermentation media were investigated as their use of glycerol, as opposed to oligopeptides, as a carbon source has been reported to increase maximum cell culture densities (Kram and Finkel, 2015). However, both media formulations significantly reduced mRNA product titers (Fig 2C), likely due to the lower cell growth rates achieved (data not shown). We also tested supplementation with L-Glutamine, based on the hypothesis that an additional nitrogen source would enhance mRNA biosynthetic capacity by increasing nucleoside biogenesis, however this did not significantly impact product yields (Fig 2C). Finally, we evaluated the chemical effector design space to identify small molecules that could specifically enhance mRNA production in *E. coli*. The most promising chemicals identified were a range of RNAse E inhibitors that reduce enzyme activity via interactions with the N-terminal domain. However, only one of these inhibitors was commercially available, and accordingly we tested the effect of supplementing LB media with 3-(4-Hydroxy-5-isopropyl-6-oxo-1,6-dihydro-pyrimidin-2-ylsulfanyl)-propionic acid (AS2). It was determined that 2 mM AS2 was the optimal concentration for maximising mRNA maintenance in the cell chassis (Supplementary data, Fig S1), which has previously been shown to reduce RNase E activity in *E. coli* by > 80% (Kime et al., 2015; Mardle et al., 2020) . Utilising AS2 at this concentration increased mRNA yield by ~50% (Fig 2C), where higher concentrations reduced cellular productivity. While a similar increase in titer may be possible via BL21 STAR cell engineering to further attenuate RNAse E activity, AS2

supplementation offers a robust mechanism to precisely optimise the synthetic mRNA-RNAse E interactome in a product-specific manner. Similarly, the use of AS2 in combination with a mutated RNAse E permits use of inhibitor concentrations with reduced off-target effects on the host cell.

The optimal combination of engineered mRNA construct (SelfCirc-mRNA), DNA expression plasmid (Triple terminator), cell host (BL21 STAR) and media formulation (LB + AS2), facilitated a 44x increase in mRNA product yield, compared to the standard control system (Fig 2D). Capillary gel electrophoresis analysis confirmed that product mRNA was full-length and constituted a substantial proportion of total cellular RNA (>20%, compared to <1% for the standard control system; Fig3A). Moreover, high yields of full-length synthetic mRNA were maintained when the relatively small GFP coding sequence (720 nt) was substituted for Cypridina Luciferase (cLuc) (1662 nt) or SARS-COV-2 Spike Protein (3783 nt) (Fig 3B), demonstrating that the engineered *in vivo* biomanufacturing system can produce larger, more complex molecules. Finally, using oligo-dT magnetic beads, we validated that achieved increases in product yield were maintained following small-scale purification processes (Fig 3C). This also demonstrates that mRNA manufactured in an *E. coli* cell-host can be purified using simple low-tech methodologies, facilitated by the absence of abundant endogenous mRNAs with PolyA tails > 5 nucleotides in length (Laalami et al., 2014; Mohanty and Kushner, 2019).

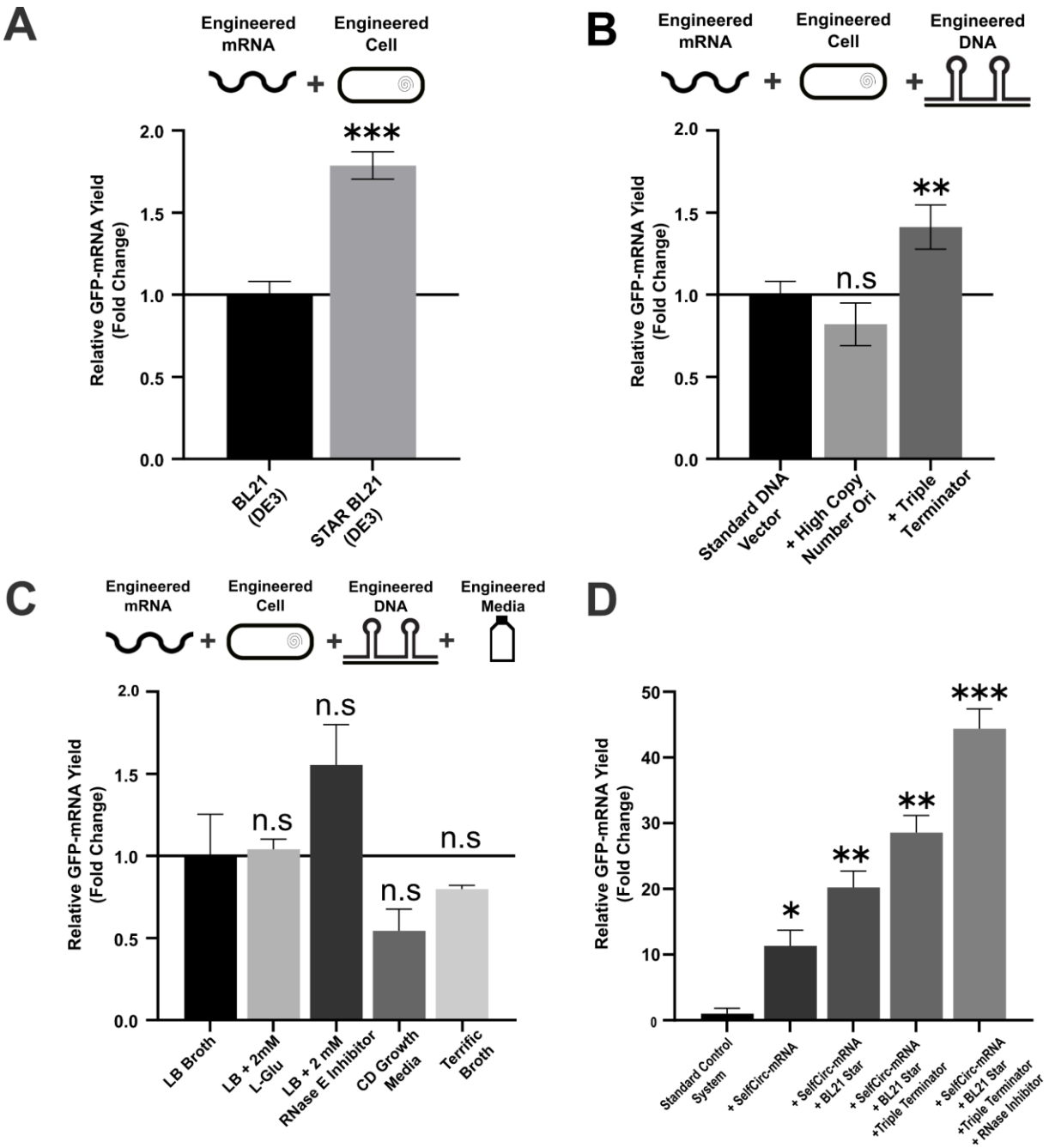


Figure 4.2 - mRNA manufacturing platform components were sequentially optimised by evaluating the function of **A)** engineered host cell chassis, **B)** synthetic DNA expression vectors, and **C)** designed cell culture media formulations. The relative performance of engineered systems was evaluated in 2.5 h production processes. The additive impact of each engineering step on overall product yield is shown in **D**. Data are expressed as a fold-change of the production achieved using standard control components. Values represent the mean + *SD* of three independent experiments ($n = 3$, each performed in triplicate). Statistical significance was determined by two-tailed unpaired t-test compared to the control system, and defined as $p \leq 0.05$ (*=p≤0.05, **=p≤0.01, ***=p≤0.001, ****=p≤0.0001). In all panels, the expressed product is SelfCirc-GFP (see Fig. 1); the effect of whole system engineering on manufacture of TermtRNA-GFP is shown in Supplementary figure 2.

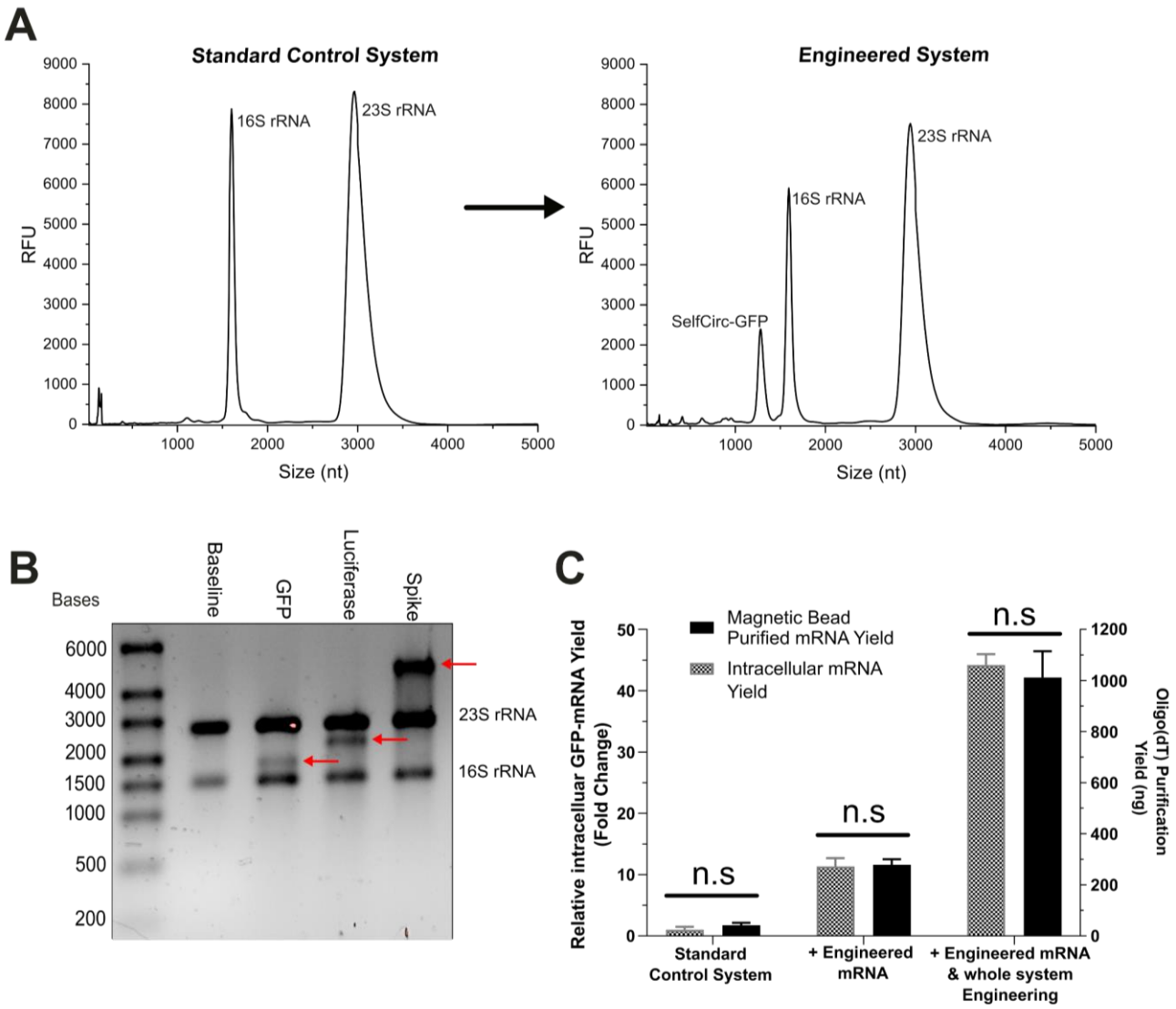


Figure 4.3 - **A)** Capillary electropherograms of total RNA isolated from GFP-mRNA biomanufacturing systems comprising either i) standard control components, or ii) an optimal combination of engineered mRNA construct, DNA expression plasmid, cell host and media formulation (Engineered system, see Fig. 2). **B)** Gel electrophoresis analysis of total RNA isolated from Engineered systems producing Green fluorescent protein (GFP), SARS-COV-2 Spike and Cypridina Luciferase mRNA molecules. Full-length product mRNA molecules are highlighted by red arrows; molecule sizes are enlarged due to the presence of IRES and Intron elements in SelfCirc-mRNA (see Fig. 1). C) Comparison of relative GFP-mRNA yields obtained from Standard control and Engineered systems, quantified before (intracellular) and after purification using oligo-dT magnetic beads. Intracellular yields are expressed as a fold-change of the product yield obtained using the Standard control system. Values represent the mean + *SD* of three independent experiments ($n = 3$, each performed in triplicate). Statistical significance was determined by two-tailed unpaired t-test, and defined as $p \leq 0.05$ (n.s – not significant).

4.3.3 - Mechanistic dissection of synthetic mRNA production in *E. coli* host cell chassis.

In order to understand how *E. coli* host cells utilise available biosynthetic capacity for mRNA production, we profiled cell biomass, total RNA and product mRNA accumulation/maintenance during a 6hr manufacturing time course. We utilised the previously optimised cell-DNA-media composition (see Fig. 2) to manufacture SARS-COV-2 Spike Protein mRNA. To evaluate mechanistic differences between biosynthesis of circular and linear molecules, we separately manufactured SelfCirc- and TermtRNA-mRNA products. With respect to the latter, we confirmed that the optimal BL21 STAR-Triple terminator-LB(AS2) system combination permitted a 36-fold increase in TermtRNA-GFP yields as compared to the standard control system (see Supplementary data, Fig. S2), similar to the 44-fold increase achieved for SelfCirc-GFP.

As shown in Fig 4A, manufacture of circular and linear synthetic mRNA products induced a significant metabolic burden on the host cell. Producer cells reached maximum cell density 2-3 h post induction of mRNA expression, whereas uninduced cells continued to accumulate biomass up to the 6 h harvest timepoint. Indeed, these cells exhibited a 25% reduction in cell specific growth rate during the first 2 h post expression induction (Fig 4B). Moreover, the final maximum cell density achieved was ~50% lower for producer cells, as compared to non-producers, associated with a ~35% reduction in the integral of cell concentration (cumulative cell hours; Fig. 4B). This indicates that producing substantial amounts of synthetic mRNA forces the cell to reallocate biosynthetic capacity away from cell biomass generation activities. Accordingly, approaches to overcome product biosynthesis-associated burden

represent a potentially effective way to enhance total biocatalyst activity and further increase product yields. The simplest method to achieve this may be optimisation of expression induction kinetics, although genetic engineering and/or directed evolution strategies will likely deliver the most significant impact on maximum achievable cell densities (Al'abri et al., 2022; Badran and Liu, 2015; Esvelt et al., 2011). Either way, optimising the biocatalytic capacity available for mRNA product synthesis is critically required to take full-advantage of the ability to scale *E. coli* production processes up to 100,000 L.

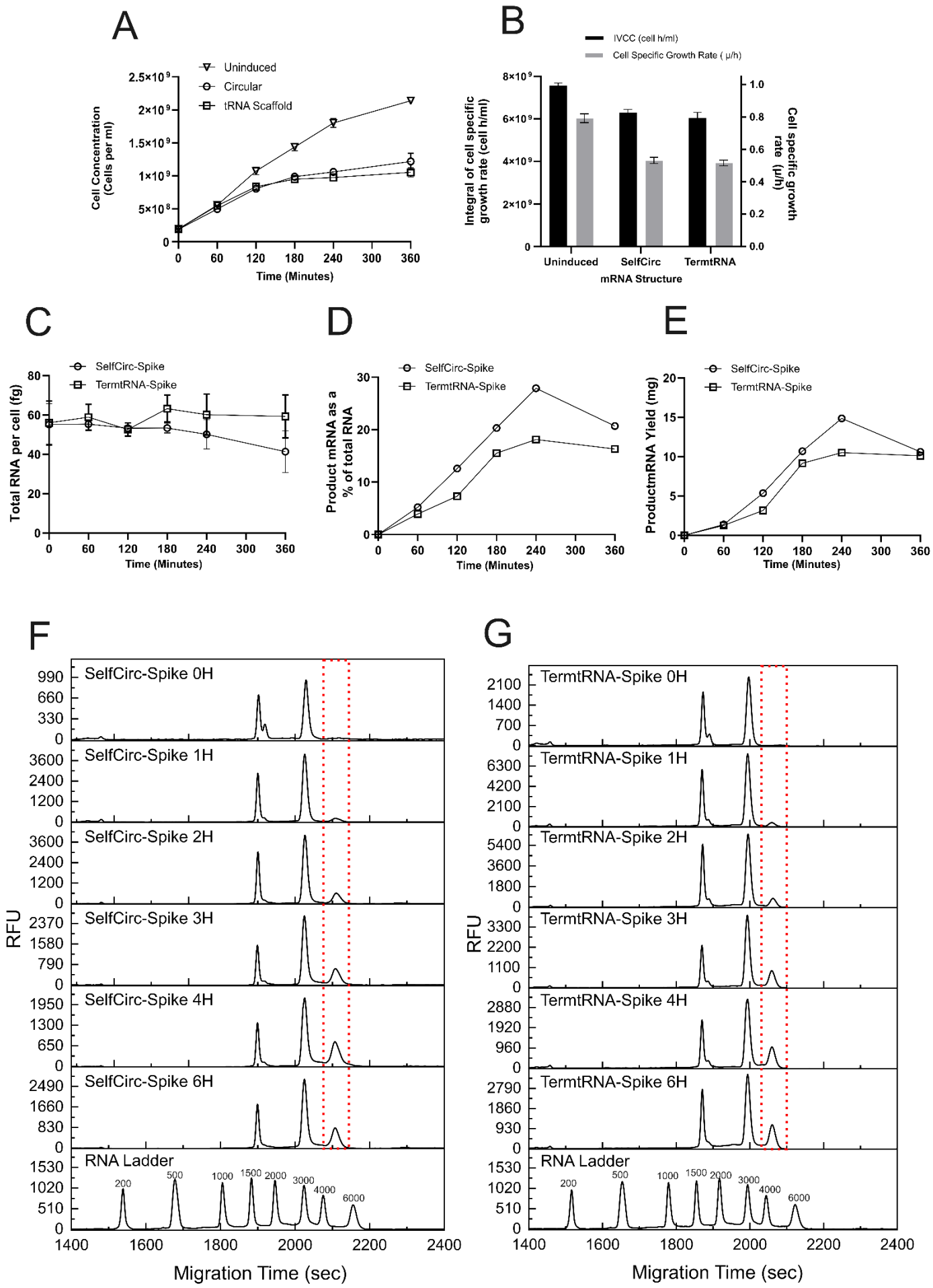
As shown in Figure 4C, total RNA synthesised per cell was stable throughout the production process, at ~60 fg/cell, despite Spike Protein-mRNA accumulating over time (Fig. 4D). This is likely due to feedback mechanisms that act to maintain intracellular concentrations of key macromolecules within relatively narrow concentration ranges (Radoš et al., 2022). Accordingly, accumulation of highly-stable product-mRNA forces the cell to reduce biosynthesis and/or induce degradation of endogenous RNA species, with potential associated off-target effects on desirable bioproduction phenotypes such as cell growth rate. These RNA homeostasis mechanisms place a theoretical limit on the total quantity of product-mRNA that can be maintained per cell, above which concentrations of key endogenous RNA molecules will become critically limiting leading to cell death and/or downregulation of product expression. Indeed, as shown in Fig 4D, intracellular concentrations of Spike Protein-mRNA peaked at 4 h, before decreasing slightly at 6 h. At 4 h, SelfCirc-Spike accounted for ~28% of total RNA mass in the host-cell, which is likely approaching the maximum achievable concentration. Although not a direct comparison, during recombinant protein expression in CHO cells, product molecules typically account for

~30% of intracellular protein mass (in-house data). We concluded that engineering efforts to further enhance intracellular product maintenance are unlikely to be beneficial, and instead should focus on maximising product accumulation rates. For SelfCirc-Spike, cell specific productivity (product-mRNA produced per cell per hour) was relatively constant throughout the first 4 hours of the production process, at ~5 fg $cell^{-1}$ h^{-1}, equating to ~10% of total cellular RNA biosynthetic activity during this time period. DNA vector engineering, for example T7 promoter re-design, may increase synthetic mRNA generation rates, facilitating cells to reach the maximum intracellular product-mRNA concentration level more quickly, permitting shorter production processes with associated benefits in cost and manufacturing flexibility (i.e., ability to rapidly switch between manufacture of different products).

Capillary Gel Electrophoresis analysis clearly shows that product mRNA accumulates intracellularly over time (Fig. 4 F-G). These data exemplify that engineered product molecules are successfully protected from nuclease-mediated decay, facilitating intracellular maintenance over multi-hour time periods, as compared to the typical mRNA half-life in *E. coli* of ~5 min (Bernstein et al., 2002; Mohanty and Kushner, 2022). Moreover, the presence of a single sharp peak at each sampling point indicates that the cell factory is producing full-length Spike Protein-mRNA that is subject to minimal degradation events. Accordingly, i) further system engineering to disrupt the *E. coli* degradasome-synthetic mRNA interactome is not required, and ii) *E. coli* is capable of synthesising homogenous populations of large mRNA molecules, thereby simplifying downstream processing steps. As expected, higher titres were obtained for SelfCirc-Spike than TermtRNA-Spike, where maximum achieved yields were 15 mg/L and 10 mg/L respectively (Fig. 4E). As discussed, we anticipate that significant increases in

product yields will be obtained via further DNA/cell/media engineering to increase maximum cell density, integral of cell concentration, and cell specific productivity. Synthetic mRNA yields > 100 mg/L should be relatively straightforward to obtain, however achieving g/L titres, as is standard for recombinant protein production in *E. coli*, will likely require significant process engineering.

Figure 4.4 (Follows Below) - SelfCirc-Spike and TermtRNA-Spike were produced in biomanufacturing systems comprising an optimal combination of engineered host cell, DNA plasmid and cell culture media (see Fig 2). Growth of non-producer (uninduced) and producer cells was measured at 1 h intervals during 6 h production processes **(A)**, to calculate integral cell concentration (ICC) and cell specific growth rate (u) values **(B)**. Total RNA **(C)** and product mRNA **(E)** yields were also measured at 1 h intervals, and the relative proportion of host cell RNA comprising Spike-mRNA molecules was quantified **(D)**. Total RNA samples from each measured timepoint were analysed by capillary gel electrophoresis; dashed red boxes on capillary electropherograms highlight SelfCirc-Spike (F) and TermtRNA-Spike (G) product accumulating over time. In A-C, values represent the mean ± SD of three independent experiments; D-G show data from a single representative capillary gel electrophoresis timecourse analysis.


A
Uninduced
Circular
tRNA Scaffold
Cell Concentration (Cells per ml)
Time (Minutes)
B
IVCC (cell h/ml)
Cell Specific Growth Rate (μ/h)
Integral of cell specific growth rate (cell h/ml)
Cell specific growth rate (μ/h)
Uninduced
SelfCirc
TermtRNA
mRNA Structure
C
SelfCirc-Spike
TermtRNA-Spike
Total RNA per cell (fg)
Time (Minutes)
D
SelfCirc-Spike
TermtRNA-Spike
Product mRNA as a % of total RNA
Time (Minutes)
E
SelfCirc-Spike
TermtRNA-Spike
ProductmRNA Yield (mg)
Time (Minutes)
F
SelfCirc-Spike 0H
SelfCirc-Spike 1H
SelfCirc-Spike 2H
SelfCirc-Spike 3H
SelfCirc-Spike 4H
SelfCirc-Spike 6H
RNA Ladder
RFU
Migration Time (sec)
G
TermtRNA-Spike 0H
TermtRNA-Spike 1H
TermtRNA-Spike 2H
TermtRNA-Spike 3H
TermtRNA-Spike 4H
TermtRNA-Spike 6H
RNA Ladder
RFU
Migration Time (sec)

4.3.4 - Synthetic mRNA produced in *E. coli* can be purified at large-scale and is functional in human cells

To exemplify the utility of our *E. coli*-based system for large-scale mRNA synthesis we manufactured GFP-mRNA in 1 L production processes. The first purification step required is total RNA extraction from host cell factories. While this can be achieved with commercial kits when production scales are < 1 ml, larger volumes require a scalable cost-efficient procedure. To achieve this, we adapted the RNASwift method previously described by Nwokeoji *et. al* for extraction of dsRNA products from *E. coli* that utilises NaCL and SDS to lyse cells and precipitate macromolecular contaminants (Nwokeoji et al., 2016). To maximise both total RNA yield and RNA quality, we i) introduced a lysozyme digestion step upstream of RNASwift, ii) lowered the lysis incubation temperature from 90 ºC to 65 ºC, and iii) added an ethanol precipitation step downstream of RNASwift. Using this modified RNASwift unit operation we were able to routinely obtain large yields (10 mg per 0.5 g wet cell mass) of high-quality total RNA (RNA Integrity Numbers > 9.5, as determined by capillary gel electrophoresis).

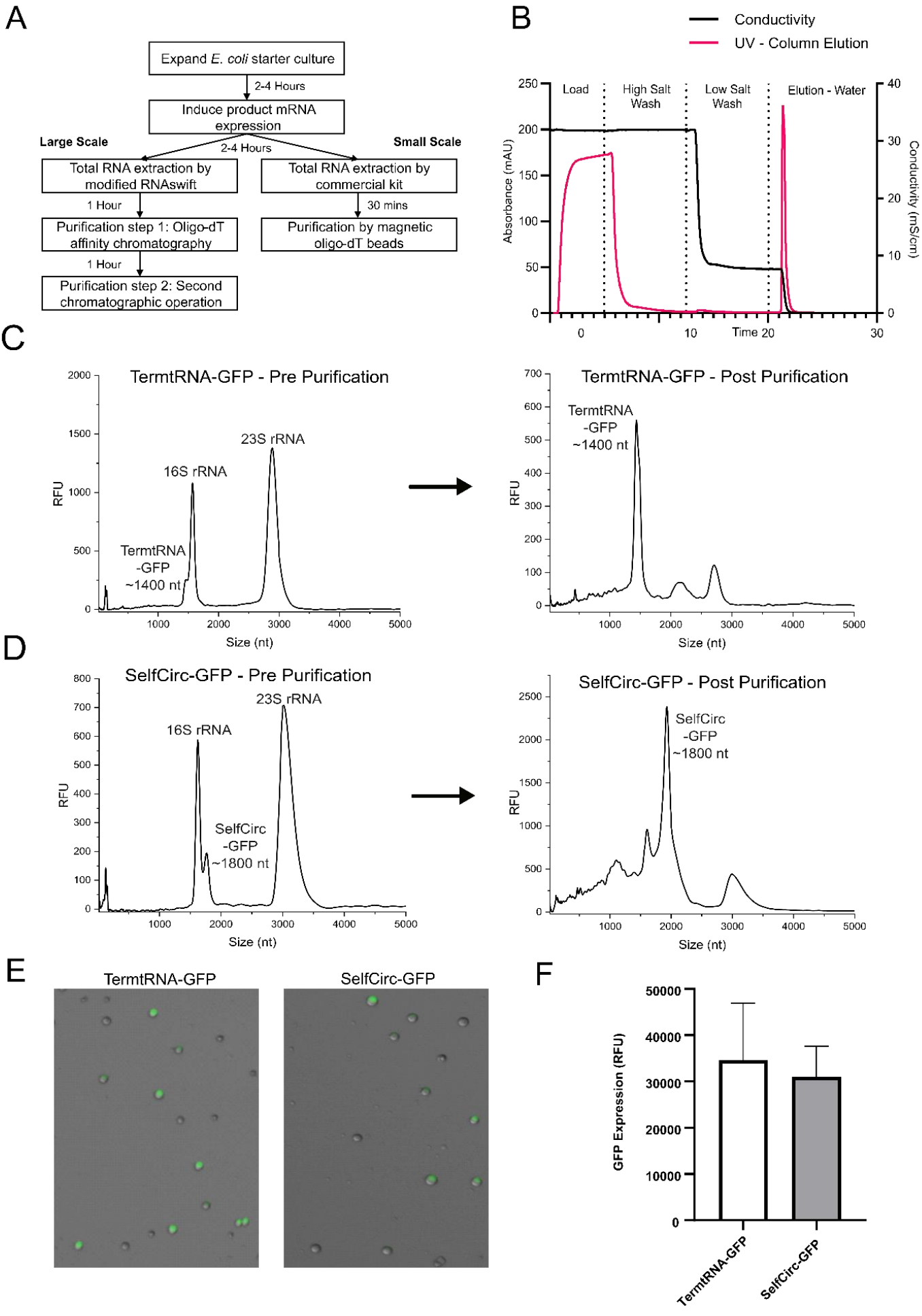
While small amounts of product-mRNA can be purified from total RNA using oligo-dT magnetic beads (see Fig 3C), larger quantities require chromatographic operations. To show that mRNA manufactured in *E. coli* can be purified using a liquid chromatography separation step, we utilised a 1 ml monolithic oligo-dT(18) column in combination with an AKTA PCC system. Figure 5B shows a chromatogram representative of this purification process, indicating conductivity as a measure of salt concentration, and the UV trace of material eluted from the column. Capillary gel electrophoresis analysis of pre- and post-purification samples showed that both SelfCirc-GFP and TermtRNA-GFP molecules could be efficiently purified by an affinity-

capture chromatographic unit operation (Fig. 5C-D). However, TermtRNA-GFP was isolated at much high purity, 71% as compared to 38% for SelfCirc-GFP, where SelfCirc-GFP samples showed a considerable wide peak of impurities representing ~30% of total RNA. This may be due to SelfCirc-GFP molecules having considerably smaller polyA tails than TermtRNA-GFP species, 50 nt and 120 nt respectively, preventing use of elution conditions that deliver both high yield and high purity. Further mRNA/DNA engineering to increase the encoded polyA tail length should permit product isolation with reduced process/product related impurities. Either way, for both molecule-formats, it is clear that for most applications a second chromatographic unit operation would be needed to achieve requisite purity profiles, such as a size-exclusion chromatography step. The use of two chromatographic unit operations is standard for purification of other high-value macromolecules, including recombinant proteins and IVT-derived mRNA (Fan et al., 2023; Rosa et al., 2021; Sripada et al., 2022). A simplified process flow diagram for large-scale and small-scale *in vivo* mRNA production processes is shown in Figure 5A.

Finally, to validate that mRNA products manufactured in *E. coli* were functional in mammalian cells, we transfected purified SelfCirc-GFP and TermtRNA-GFP into Human embryonic kidney cells (HEK). While SelfCirc-GFP contains an internal ribosome binding site (IRES), obviating the need for post-purification processing, TermtRNA-GFP required the enzymatic addition of a Cap-0 structure to enable translation initiation. As shown in Figure 5E-F, both synthetic mRNA molecular formats were translatable in HEK cells, facilitating similar levels of GFP protein expression. Translational efficiency of SelfCirc-GFP molecules would likely be further enhanced via determination and selection of optimal IRES elements (Wesselhoeft et al., 2018).

Indeed, this may provide a route to encode cell-type specificity into mRNA gene therapeutics (Plank et al., 2013).

Figure 4.5 (Follows Overleaf) A) Simplified process flow diagram for large-scale and small-scale *in vivo* mRNA production. **B)** Typical chromatogram from oligo-d(T) affinity chromatography purification of product mRNA manufactured in *E. coli*. **C-D)** Capillary electropherograms showing purification of TermtRNA-GFP (C) and SelfCirc-GFP (D) products using oligo-d(T) affinity chromatography. **E-F)** mRNA products manufactured in *E. coli* were purified and transfected into Human Embryonic Kidney cells. 24 hr posttransfection, GFP protein expression was measured by fluorescent cell imaging (E) and fluorescent plate reader analysis (F). Values in F represent the mean + *SD* of three independent experiments ($n = 3$, each performed in triplicate).


A
Expand E. coli starter culture
2-4 Hours
Induce product mRNA expression
2-4 Hours
Large Scale
Small Scale
Total RNA extraction by modified RNAswift
1 Hour
Purification step 1: Oligo-dT affinity chromatography
1 Hour
Purification step 2: Second chromatographic operation
Total RNA extraction by commercial kit
30 mins
Purification by magnetic oligo-dT beads
B
Conductivity
UV - Column Elution
Load
High Salt Wash
Low Salt Wash
Elution - Water
Absorbance (mAU)
Conductivity (mS/cm)
Time
C
TermtRNA-GFP - Pre Purification
23S rRNA
16S rRNA
TermtRNA -GFP ~1400 nt
RFU
Size (nt)
TermtRNA-GFP - Post Purification
TermtRNA -GFP ~1400 nt
D
SelfCirc-GFP - Pre Purification
23S rRNA
16S rRNA
SelfCirc -GFP ~1800 nt
SelfCirc-GFP - Post Purification
SelfCirc -GFP ~1800 nt
E
TermtRNA-GFP
SelfCirc-GFP
F
GFP Expression (RFU)
TermtRNA-GFP
SelfCirc-GFP

4.4 - Concluding remarks

The *in vivo* mRNA manufacturing system we have presented here can be utilised to produce synthetic mRNA molecules for a wide range of research and commercial applications. Although we have demonstrated for the first time that functional mammalian mRNA can be produced at high-yields in *E. coli* cell factories, widespread adoption of this platform requires further optimisation of both i) downstream processing steps to increase product purities and ii) upstream processes to maximise product titres by enhancing cell specific productivity and cell biomass accumulation/maintenance. While we anticipate that the developed system will be particularly suitable for production of circular mRNA products, its utility for manufacturing linear mRNA molecules would be significantly enhanced by optimising co-expression of a T7-Capping enzyme fusion protein (Qin et al., 2023) to enable synthesis of capped mRNA species in *E. coli*. Similarly, co-expression of nucleotide-modifying enzymes, such as Psuedouridine synthetase (Carlile et al., 2019), would permit production of linear mRNA products with required immunostimulatory properties for therapeutic applications.

The potential utility of microbial cell factories for large-scale mRNA manufacturing is being increasingly recognised. Indeed, earlier this year saw announcements of plans to develop commercial cell-based mRNA production processes using Eukaryotic cell-hosts. The availability of such platforms will become increasingly critical in coming years as product lines begin to diversify (e.g., adoption of more complex molecular formats) and the scale of global mRNA manufacturing continues to increase. By engineering the core components of an *E. coli*-based mRNA production system this study has added a novel technology to the mRNA

manufacturing solution space, providing flexibility to achieve context- and/or application-specific design criteria.

4.5 - Paper II References

Agostinetto, R., Rossi, M., Dawson, J., Lim, A., Simoneau, M.H., Boucher, C., Valldorf, B., Ross-Gillespie, A., Jardine, J.G., Sok, D., et al. (2022). Rapid cGMP manufacturing of COVID-19 monoclonal antibody using stable CHO cell pools. Biotechnol. Bioeng. 119, 663–666.

Al'abri, I.S., Haller, D.J., Li, Z., and Crook, N. (2022). Inducible directed evolution of complex phenotypes in bacteria. Nucleic Acids Res. 50, E58.

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., et al. (2021). Efficacy and Safety of the mRNA-1273 SARS-CoV-2 Vaccine. N. Engl. J. Med. 384, 403–416.

Badran, A.H., and Liu, D.R. (2015). *In vivo* continuous directed evolution. Curr. Opin. Chem. Biol. 24, 1–10.

Bae, D., Hyeon, H., Shin, E., Yeom, J.H., and Lee, K. (2023). Relaxed Cleavage Specificity of Hyperactive Variants of Escherichia coli RNase E on RNA I. J. Microbiol. 61, 211–220.

Bai, Y., Liu, D., He, Q., Liu, J., Mao, Q., and Liang, Z. (2023). Research progress on circular RNA vaccines. Front. Immunol. 13, 1–12.

Bernstein, J.A., Khodursky, A.B., Lin, P.H., Lin-Chao, S., and Cohen, S.N. (2002). Global analysis of mRNA decay and abundance in Escherichia coli at single-gene resolution using two-color fluorescent DNA microarrays. Proc. Natl. Acad. Sci. U. S. A. 99, 9697–9702.

Bolivar, F., Rodriguez, R.L., Greene, P.J., Betlach, M.C., Heyneker, H.L., Boyer, H.W., Crosa, J.H., and Falkow, S. (1977). Construction and characterization of new cloning vehicle. II. A multipurpose cloning system. Gene 2, 95–113.

Börner, J., Friedrich, T., Bartkuhn, M., and Klug, G. (2023). Ribonuclease E strongly impacts bacterial adaptation to different growth conditions. RNA Biol. 20, 120–135.

Breda, L., Papp, T.E., Triebwasser, M.P., Yadegari, A., Fedorky, M.T., Tanaka, N., Abdulmalik, O., Pavani, G., Wang, Y., Grupp, S.A., et al. (2023). *In vivo* hematopoietic stem cell modification by mRNA delivery. Science 381, 436–443.

Brown, A.J., Gibson, S.J., Hatton, D., Arnall, C.L., and James, D.C. (2019). Whole synthetic pathway engineering of recombinant protein production. Biotechnol. Bioeng. 116, 375–387.

Callaghan, A.J., Marcaida, M.J., Stead, J.A., McDowall, K.J., Scott, W.G., and Luisi, B.F. (2005). Structure of Escherichia coli RNase E catalytic domain and implications for RNA turnover. Nature 437, 1187–1191.

Carlile, T.M., Martinez, N.M., Schaening, C., Su, A., Bell, T.A., Zinshteyn, B., and Gilbert, W. V. (2019). mRNA structure determines modification by pseudouridine synthase 1. Nat. Chem. Biol. 15, 966.

Carter, A.D., Morris, C.E., and McAllister, W.T. (1981). Revised transcription map of the late region of bacteriophage T7 DNA. J. Virol. 37, 636–642.

Delgado-Martín, J., and Velasco, L. (2021). An efficient dsRNA constitutive expression system in Escherichia coli. Appl. Microbiol. Biotechnol. 105, 6381–6393.

Deviatkin, A.A., Simonov, R.A., Trutneva, K.A., Maznina, A.A., Soroka, A.B., Kogan, A.A., Feoktistova, S.G., Khavina, E.M., Mityaeva, O.N., and Volchkov, P.Y. (2023). Cap-Independent Circular mRNA Translation Efficiency. Vaccines 11, 1–12.

Esquerré, T., Moisan, A., Chiapello, H., Arike, L., Vilu, R., Gaspin, C., Cocaign-Bousquet, M., and Girbal, L. (2015). Genome-wide investigation of mRNA lifetime determinants in Escherichia coli cells cultured at different growth rates. BMC Genomics 16, 1–13.

Esvelt, K.M., Carlson, J.C., and Liu, D.R. (2011). A system for the continuous directed evolution of biomolecules. Nature 472, 499–503.

Fan, J., Sripada, S.A., Pham, D.N., Linova, M.Y., Woodley, J.M., Menegatti, S., Boi, C., and Carbonell, R.G. (2023). Purification of a monoclonal antibody using a novel high-capacity multimodal cation exchange nonwoven membrane. Sep. Purif. Technol. 317, 123920.

Gan, L.M., Lagerström-Fermér, M., Carlsson, L.G., Arfvidsson, C., Egnell, A.C., Rudvik, A., Kjaer, M., Collén, A., Thompson, J.D., Joyal, J., et al. (2019). Intradermal delivery of modified mRNA encoding VEGF-A in patients with type 2 diabetes. Nat. Commun. 10, 1–9.

Gholamalipour, Y., Karunanayake Mudiyanselage, A., and Martin, C.T. (2018). NAR breakthrough article 3 end additions by T7 RNA polymerase are RNA self-templated, distributive and diverse in character—RNA-Seq analyses. Nucleic Acids Res. 46, 9253–9263.

Heyde, S.A.H., and Nørholm, M.H.H. (2021). Tailoring the evolution of BL21(DE3) uncovers a key role for RNA stability in gene expression toxicity. Nat. Commun. 21, 1–9.

Jiang, Z., and Dalby, P.A. (2023). Challenges in scaling up AAV-based gene therapy manufacturing. Trends Biotechnol. 41, 1268–1281.

Joshi, S.H.N., Yong, C., and Gyorgy, A. (2022). Inducible plasmid copy number control for synthetic biology in commonly used *E. coli* strains. Nat. Commun. 13.

Kime, L., Vincent, H.A., Gendoo, D.M.A., Jourdan, S.S., Fishwick, C.W.G., Callaghan, A.J., and McDowall, K.J. (2015). The first small-molecule inhibitors of members of the ribonuclease E family. Sci. Rep. 5, 8028.

Kis, Z., Kontoravdi, C., Shattock, R., and Shah, N. (2021). Resources, production scales and time required for producing RNA vaccines for the global pandemic demand. Vaccines 9, 1–14.

Kram, K.E., and Finkel, S.E. (2015). Rich medium composition affects Escherichia coli survival, glycation, and mutation frequency during long-term batch culture. Appl. Environ. Microbiol. 81, 4442–4450.

Laalami, S., Zig, L., and Putzer, H. (2014). Initiation of mRNA decay in bacteria. Cell. Mol. Life Sci. 71, 1799–1828.

Lee, C., Kim, J., Shin, S.G., and Hwang, S. (2006). Absolute and relative QPCR quantification of plasmid copy number in Escherichia coli. J. Biotechnol. 123, 273–280.

Leppek, K., Byeon, G.W., Kladwang, W., Wayment-Steele, H.K., Kerr, C.H., Xu, A.F., Kim, D.S., Topkar, V. V., Choe, C., Rothschild, D., et al. (2022). Combinatorial optimization of mRNA structure, stability, and translation for RNA-based therapeutics. Nat. Commun. 13.

Lin-Chao, S., Chen, W. -T, and Wong, T. -T (1992). High copy number of the pUC plasmid results from a Rom/Rop-suppressible point mutation in RNA II. Mol. Microbiol. 6, 3385–3393.

Liu, X., Zhang, Y., Zhou, S., Dain, L., Mei, L., and Zhu, G. (2022). Circular RNA: An emerging frontier in RNA therapeutic targets, RNA therapeutics, and mRNA vaccines. J. Control. Release 348, 84–94.

Ma, Z.Z., Zhou, H., Wei, Y.L., Yan, S., and Shen, J. (2020). A novel plasmid–Escherichia coli system produces large batch dsRNAs for insect gene silencing. Pest Manag. Sci. 76, 2505–2512.

Mairhofer, J., Wittwer, A., Cserjan-Puschmann, M., and Striedner, G. (2015). Preventing T7 RNA polymerase read-through transcription-A synthetic termination signal capable of improving bioprocess stability. ACS Synth. Biol. 4, 265–273.

Mardle, C.E., Goddard, L.R., Spelman, B.C., Atkins, H.S., Butt, L.E., Cox, P.A., Gowers, D.M., Vincent, H.A., and Callaghan, A.J. (2020). Identification and analysis of novel small molecule inhibitors of RNase E: Implications for antibacterial targeting and regulation of RNase E. Biochem. Biophys. Reports 23.

Mauger, D.M., Joseph Cabral, B., Presnyak, V., Su, S. V., Reid, D.W., Goodman, B., Link, K., Khatwani, N., Reynders, J., Moore, M.J., et al. (2019). mRNA structure regulates protein expression through changes in functional half-life. Proc. Natl. Acad. Sci. U. S. A. 116, 24075–24083.

McElwain, L., Phair, K., Kealey, C., and Brady, D. (2022). Current trends in biopharmaceuticals production in Escherichia coli. Biotechnol. Lett. 44, 917–931.

Miroux, B., and Walker, J.E. (1996). Over-production of proteins in Escherichia coli: Mutant hosts that allow synthesis of some membrane proteins and globular proteins at high levels. J. Mol. Biol. 260, 289–298.

Mohanty, B.K., and Kushner, S.R. (2019). New Insights into the Relationship between tRNA Processing and Polyadenylation in Escherichia coli. Trends Genet. 35, 434–445.

Mohanty, B.K., and Kushner, S.R. (2022). Regulation of mRNA decay *in E. coli*. Crit. Rev. Biochem. Mol. Biol. 57, 48–72.

Nelissen, F.H.T., Leunissen, E.H.P., Van De Laar, L., Tessari, M., Heus, H.A., and Wijmenga, S.S. (2012). Fast production of homogeneous recombinant RNA-towards large-scale production of RNA. Nucleic Acids Res. 40.

Nwokeoji, A.O., Kilby, P.M., Portwood, D.E., and Dickman, M.J. (2016). RNASwift: A rapid, versatile RNA extraction method free from phenol and chloroform. Anal. Biochem. 512, 36–46.

Ouranidis, A., Vavilis, T., Mandala, E., Davidopoulou, C., Stamoula, E., Markopoulou, C.K., Karagianni, A., and Kachrimanis, K. (2022). Manufacturing under Pharma 4 . 0 Principles. 1–31.

Pertzev, A. V., and Nicholson, A.W. (2006). Characterization of RNA sequence determinants and antideterminants of processing reactivity for a minimal substrate of Escherichia coli ribonuclease III. Nucleic Acids Res. 34, 3708–3721.

Plank, T.D.M., Whitehurst, J.T., and Kieft, J.S. (2013). Cell type specificity and structural determinants of IRES activity from the 5′ leaders of different HIV-1 transcripts. Nucleic Acids Res. 41, 6698–6714.

Ponchon, L., and Dardel, F. (2011). Large scale expression and purification of recombinant RNA in Escherichia coli. Methods 54, 267–273.

Ponchon, L., Beauvais, G., Nonin-Lecomte, S., and Dardel, F. (2009). A generic protocol for the expression and purification of recombinant RNA in Escherichia coli using a tRNA scaffold. Nat. Protoc. 4, 947–959.

Ponchon, L., Catala, M., Seijo, B., El Khouri, M., Dardel, F., Nonin-Lecomte, S., and Tisné, C. (2013). Co-expression of RNA-protein complexes in Escherichia coli and applications to RNA biology. Nucleic Acids Res. 41, e150.

Pontrelli, S., Chiu, T.Y., Lan, E.I., Chen, F.Y.H., Chang, P., and Liao, J.C. (2018). Escherichia coli as a host for metabolic engineering. Metab. Eng. 50, 16–46.

Qin, C., Xiang, Y., Liu, J., Zhang, R., Liu, Z., Li, T., Sun, Z., Ouyang, X., Zong, Y., Zhang, H.M., et al. (2023). Precise programming of multigene expression stoichiometry in mammalian cells by a modular and programmable transcriptional system. Nat. Commun. 2023 141 14, 1–10.

Qin, S., Tang, X., Chen, Y., Chen, K., Fan, N., Xiao, W., Zheng, Q., Li, G., Teng, Y., Wu, M., et al. (2022). mRNA-based therapeutics: powerful and versatile tools to combat diseases. Signal Transduct. Target. Ther. 7.

Qu, L., Yi, Z., Shen, Y., Lin, L., Chen, F., Xu, Y., Wu, Z., Tang, H., Zhang, X., Tian, F., et al. (2022). Circular RNA vaccines against SARS-CoV-2 and emerging variants. Cell 185, 1728-1744.e16.

Radoš, D., Donati, S., Lempp, M., Rapp, J., and Link, H. (2022). Homeostasis of the biosynthetic *E. coli* metabolome. IScience 25, 104503.

Richards, J., and Belasco, J.G. (2023). Graded impact of obstacle size on scanning by RNase E. Nucleic Acids Res. 51, 1364–1374.

Rosa, S.S., Prazeres, D.M.F., Azevedo, A.M., and Marques, M.P.C. (2021). mRNA vaccines manufacturing: Challenges and bottlenecks. Vaccine 39, 2190–2200.

Rostain, W., Shen, S., Cordero, T., Rodrigo, G., and Jaramillo, A. (2020). Engineering a Circular Riboregulator in Escherichia coli. BioDesign Res. 2020, 1–9.

Rouches, M. V., Xu, Y., Cortes, L.B.G., and Lambert, G. (2022). A plasmid system with tunable copy number. Nat. Commun. 13, 1–12.

Roux, C., Etienne, T.A., Hajnsdorf, E., Ropers, D., Carpousis, A.J., Cocaign-Bousquet, M., and Girbal, L. (2022). The essential role of mRNA degradation in understanding and engineering *E. coli* metabolism. Biotechnol. Adv. 54, 107805.

Sripada, S.A., Chu, W., Williams, T.I., Teten, M.A., Mosley, B.J., Carbonell, R.G., Lenhoff, A.M., Cramer, S.M., Bill, J., Yigzaw, Y., et al. (2022). Towards continuous mAb purification: Clearance of host cell proteins from CHO cell culture harvests via “flow-through affinity chromatography” using peptide-based adsorbents. Biotechnol. Bioeng. 119, 1873–1889.

Vavilis, T., Stamoula, E., Ainatzoglou, A., Sachinidis, A., Lamprinou, M., Dardalas, I., and Vizirianakis, I.S. (2023). mRNA in the Context of Protein Replacement Therapy. Pharmaceutics 15, 1–19.

Viegas, S.C., Apura, P., Martínez-García, E., De Lorenzo, V., and Arraiano, C.M. (2018). Modulating Heterologous Gene Expression with Portable mRNA-Stabilizing 5′-UTR Sequences. ACS Synth. Biol. 7, 2177–2188.

Wesselhoeft, R.A., Kowalski, P.S., and Anderson, D.G. (2018). Engineering circular RNA for potent and stable translation in eukaryotic cells. Nat. Commun. 9, 1–10.

Whitley, J., Zwolinski, C., Denis, C., Maughan, M., Hayles, L., Clarke, D., Snare, M., Liao, H., Chiou, S., Marmura, T., et al. (2022). Development of mRNA manufacturing for vaccines and therapeutics: mRNA platform requirements and development of a scalable production process to support early phase clinical trials. Transl. Res. 242, 38–55.

Yang, D., Pricilia, C., Prabowo, S., Eun, H., Park, S.Y., Cho, I.J., Jiao, S., and Lee, S.Y. (2021). Escherichia coli as a platform microbial host for systems metabolic engineering. Essays Biochem. 65, 225–246.

Zhang, H., Zhang, L., Lin, A., Xu, C., Li, Z., Liu, K., Liu, B., Ma, X., Zhao, F., Jiang, H., et al. (2023). Algorithm for optimized mRNA design improves stability and immunogenicity. Nature 621, 396–403.

Zhang, Q., Ma, D., Wu, F., Standage-Beier, K., Chen, X., Wu, K., Green, A.A., and Wang, X. (2021). Predictable control of RNA lifetime using engineered degradation-tuning RNAs. Nat. Chem. Biol. 17, 828–836.

Chapter 5 – Paper III

Transfection of mRNA control nodes into mAb producing CHO cells reveals rate limiting cellular capacities

THE CONTENTS OF THIS CHAPTER ARE REDACTED DUE TO RESULTS BEING USED FOR COMMERICIALISATION PURPOSES

Edward Curry[1], Melinda Pohle[1], Adam Brown [1]

[1] Department of Chemical and Biological Engineering, University of Sheffield, Sheffield, UK.

The work described in this chapter is currently withheld from publication, due to ongoing commercialisation activities related to data presented herein

Summary

The third paper in this thesis describes the use of synthetic mRNA as a means of revealing critical production capacities in CHO cells producing monoclonal antibodies. 53 mRNA hyperactive 'control nodes', encoding transcription factors and other potent cellular regulators, are introduced to the cell in order to overwhelm various processes. The level of response to each control node gives an indication of the importance of the corresponding process to the production of the mAb. From this data, chemical modulators can be selected which improve mAb titre, by activating or inhibiting components of rate limiting production processes.

Contributions

Edward Curry – Conceptualisation, methodology, analysis, investigation, manuscript preparation. **Melinda Pohle** – Investigation (CHO cell culture), analysis, methodology conceptualisation. **Adam Brown** – Conceptualisation, methodology, analysis, manuscript preparation, funding acquisition.

Chapter 6 – Concluding Remarks & Future Work

6.1 Conclusion

The work presented in this thesis provides two novel advancements in the field of mRNA manufacture, in addition to exemplifying how synthetic mRNA may be utilised in the field of biotechnology, outside of therapeutic applications. Firstly, attention is given to existing IVT manufacturing processes, specifically the RNA polymerase. A new method for identifying active RNA polymerase-promoter pairs is described, which is then utilised to characterise 8 novel RNAPs. These findings allow for speculation of which RNAPs would be suitable engineering targets for future characterisation for mRNA production. Secondly, a non-IVT mRNA manufacture platform is introduced, using *E. coli* as a production host for the first time. Through a sequential engineering approach, mRNA is able to accumulate 40-fold more readily than a baseline control. Finally, by using synthetic mRNA, rate-limiting bottlenecks in the production of monoclonal antibodies were elucidated. This technology allows for the identification of capacities critical to the titre of recombinant protein produced, and provides a basis for improving yield of difficult to express products.

6.2 Paper I – Expanding the RNA polymerase biocatalyst solution space for mRNA manufacture

Current methods for the production of mRNA therapeutics rely upon the activity of T7 RNA polymerase. Whilst yields from such processes are high, extensive multi-stage downstream processing is required as a result of abundant immunostimulatory by-products (Gholamalipour et al., 2018; Karikó et al., 2011). Whilst other single subunit RNA polymerases have been characterised, the toolbox remains small, with only 6 having been described in the literature at the time of writing. Demonstrated in this thesis is a method for the rapid screening of the activity of RNAPs on a putative

promoter sequence. The method allows for an initial characterisation of the RNAP, without the need for time-consuming overexpression and purification strategies. Using this method, presented are a library of 8 novel ssRNAP-promoter pairs, massively expanding the number of RNAPs available for RNA manufacturing applications.

Work by the Zhu group, who described KP34, Syn5 and VSW-3 RNAPs, has focused on extensively characterising a single RNAP that has been overexpressed and purified in *E. coli* (Lu et al., 2019; Xia et al., 2022; Zhu et al., 2013b). A different approach is taken here, whereby a much larger number of RNAPs are initially considered, necessitating a method for quickly screening such enzymes for activity. Producing highly pure preparations of all 27 enzymes would be a time consuming process, and may introduce variability into the library depending on the stability of the isolated enzyme, and the level of purity that could be reached. For this reason, the use of a cell free coupled transcription-translation system, was more appropriate for our purposes. The system allows for the detection of the activity of multiple RNAPs through product luminescence in only 1 hour in a single tube - considerably quicker than the multiple days required for a traditional overexpression and purification process.

The coupled transcriptional-translation system is functional in several RNAP characterisation contexts. Previously described by Egorova and colleagues was the analysis of mutant T7 RNAP species, in particular on DNA templates featuring non-natural nucleotides (Egorova et al., 2021). More recently, Cui showed that such a system may be used to characterised T7 RNAP activity from *E. coli* cell lysate (Cui et al., 2023). The work presented in this theis build upon these previously described methods by showing that the transcription-translation system can be used for the screening of novel RNAP activity. It is also possible to confirm the optimal promoter for a given RNAP. This method could also be applied on engineered RNAPs, featuring

amino acid substitutions, in order to confirm the relative activity and the optimal promoter after altering the protein sequence.

Using the transcription-translation system, we present 8 novel ssRNAPs. As would be expected considering efforts to promote sequence diversity within the final RNAP library, a wide range of relative yields are seen. Of particular interest for mRNA therapeutic manufacturing were R6 (Pectobacterium Phage PP74) and R14 (Delftia Phage IME-DE1), as a result of their higher yields than T7 RNAP in initial assays at 30C. T7 displayed increased relative activity at 37 °C compared to other RNAPs, with the exception of R7, however 6 RNAPs performed better than T7 at 20 °C. The results for R6 were particularly promising, as at all 3 temperatures tested, the RNAP had higher activity than T7. It must be noted, however, that RNAP activity in the transcription-translation assay does not directly correlate to IVT yield, as a result of differences in translatability and enzyme activity in the rabbit reticulocyte lysate used. It would be possible to account for difference in RNAP produced through western blot hybridisation, as described by Egorova, however the utilisation of this method in this work was to quickly screen for active RNAPs to take forward for more extensive characterisation, and therefore the normalisation step was unnecessary.

The results highlight the need for improved tools for the determination of ssRNAP promoters. It is likely that many more RNAPs in our library were only inactive because of an inaccurate promoter prediction. RNAPs that were active exclusively featured sequence similar to the T7 promoter, particularly in the CACTATA region. Several inactive promoters featured ‘T7-like’ elements but may have featured a small number of incorrect nucleotides. Such sequence could not be identified in many of the phage genomes screened, implying a more diverse promoter architecture, or that the RNAP does not transcribed from the phage genome. An example of a non-‘T7 like’ promoter

exists in the form of the KP34 RNAP promoter (TAATGTTACAGGAGTA). The promoter sequence was originally incorrectly identified by bioinformatics, and require more extensive 5'-RACE characterisation, which becomes less feasible with the larger library of RNAPs described here.

Results from the coupled transcription translation assay were verified by overexpressing and purifying a high and low performing RNAP, to see how activity compared to IVT yield. R6 (161% of T7 activity) and R5 (10% of T7 activity) were chosen for this purpose. In IVT reactions, R6 performed better than R5, as would be expected from the transcription-translation assay results. The yield was lower than T7 however, at only 60% of T7, showing the influence that factors other than RNAP activity have on yield indicated by transcription-translation assays. It is important to consider, however, that R6 is being compared to a highly optimised enzyme in T7, which has well defined reaction conditions. Whilst some optimisation was performed for R6, further consideration of factors such as magnesium and salt concentrations could further boost the performance of the RNAP.

In summary, demonstrated here is that coupled transcription-translation reactions are well suited to the initial characterisation of novel ssRNAPs. Using this method, 10 active RNAP-promoter pairs were identified, considerably expanding the number of RNAPs described in the literature. RNAPs active in the transcription-translation assay are also active in IVT reactions, despite relative yields not directly correlating between assays.

6.2.1 Paper I – Future work

One obvious weakness, as previously mentioned, is the current lack of the ability of the RNAP screening system to take into account differences in concentrations of the novel RNAPs produced. Differences in the rates of transcription, translation, and related processes will mean that different relative amounts of each RNAP will be found in each well, meaning the nLuc concentration cannot be directly correlated to enzyme activity. It would therefore be highly advantageous to incorporate a rapid method by which RNAP concentration could be normalised, that required less additional work than the quantitative western blotting method proposed by Egorova (Egorova et al., 2021). It may be possible to fuse the novel RNAP to a fluorescent protein, such as GFP, in the expression vector using a flexible linker. Several examples exist of T7 RNAP fusion proteins, with no discernible impact on RNAP activity (Chen et al., 2020; Komatsu et al., 2023; Moore et al., 2018). Fluorescent signal could be measured immediately after luminescent signal, allowing for efficient normalisation of novel RNAP activity to the RNAP concentration as determined by GFP activity.

The next stage of characterisation for any potentially interesting RNAPs would require an assessment of the quality profile of transcripts produced. Several methods exist for this purpose. Firstly, radioactively labelled NTPs can be used to determine the impurity profile of transcripts on polyacrylamide gels (Dousis et al., 2023). Secondly, transcript homogeneity can be assessed by RNA-seq, or 5’ and 3’-RACE, which whilst costlier and time consuming, present a more detailed profile of transcripts (Lu et al., 2019). Whilst yield is a critical factor to the performance of a RNAP, the homogeneity of transcripts produced must be considered due to the immunogenicity of non-product RNA. Levels of dsRNA production is another key quality attribute that must be assessed for novel RNAPs. dsRNA ELISA methods, or novel dsRNA detection

reagents sold by Promega would allow for quantification of dsRNA with reasonable throughput.

The coupled transcription-translation method can be used as a tool to continue to expand the available toolbox of RNAPs for mRNA manufacture, which will become increasingly important as the diversity of RNA therapeutics increases. With an ever increasing number of putative RNAP sequences becoming available, it is possible that RNAPs more suited to manufacturing than T7 can be found. Two distinct approaches could be taken to find useful RNAPs. Firstly, as the price of recombinant DNA synthesis continues to decrease, it is possible that all deposited RNAP sequences could be screened using the transcription/translation system. The system works well at scale, due to the plate based nature of the reaction, which could be scaled to 384-well plates for simultaneous activity detection of hundreds of RNAPs. This would most likely be facilitated by an industrial research group, where budgets for such work may be higher. Any RNAPs which show high levels of activity would then proceed to a second stage of screening, whereby the quality profiles of transcripts produced is screened.

The second approach, which may be more suitable in an academic setting whereby research budgets are more limited, would be to curate a library of RNAPs for screening, based upon sequence characteristics. The work presented here provides a basis for which RNAPs would be easier engineering targets, based upon the likelihood that their promoter could be predicted successfully. Within these engineering targets, a library of RNAPs could be created based upon the divergence of sequence in different domains of the protein. As an example, if a researcher was particularly interested in finding RNAPs with different levels of processivity, RNAPs could be selected based upon the level of conservation in the DNA binding region.

The screening system would also be particularly useful for screening sequence variants of previously characterised or novel RNAPs. These variants could be produced through rational engineering or random mutagenesis, with the goal of discovering enzymes with enhance functionalities, such as improved yield or processivity. The screening platform would allow for rapid simultaneous detection of variants with advantageous characteristics, without the need to overexpress and purify many enzymes.

Fundamental to the advancement of the field of RNAP characterisation would be improved tools for *de novo* promoter identification in batch. The work presented here is limited by the ability of current tools to accurately detect promoter sequences that are not 'T7-like' in their sequence. Both tools used in this thesis, PHIRE and PhagePromoter were not fully fit for purpose, featuring their own limitations. PHIRE is simplistic, solely looking for conserved sequence within the genome (Lavigne et al., 2004). There is no guarantee that sequence identified is a promoter, and not another type of regulatory sequence. In addition, the program is extremely slow to run, processing one genome in approx. 30 mins. PhagePromoter is more sophisticated in its analysis, using a machine learning approach, however the program is only effective if some previous knowledge of the phage is known, such as phage family, host bacterium species, and phage type (Sampaio et al., 2019). Other tools, such as iProm-phage, do exist, however their sequence input length constraints do not cover full genomes as would be required for this work (Shujaat et al., 2022). A long term goal for the expansion of this work would be novel tools that allow for sophisticated phage promoter prediction from full genome sequences.

It is possible the other cell free coupled transcription-translation systems would be more optimal for use in this screening system. The Promega TXTL SP6 rabbit

reticulocyte system was used due to its previous use in similar contexts, and the fact that the novel RNAPs were large, complex proteins, which may have folded more readily in the mammalian lysate system (Cui et al., 2023; Egorova et al., 2021). Naturally, however, these are bacteriophage enzymes that should be produced efficiently within bacterial lysates. It would therefore be useful to screen alternative cell free lysate system to determine if the choice of cell extract was detrimental to the performance of certain RNAPs. Using a bacterial cell extract would also be more cost effective than the use of a mammalian cell extract, due to the inherent costs of producing the required biomass of both cell types.

6.3 Paper II - Engineering an *E. coli*-based *in vivo* mRNA manufacturing platform

In vitro transcription represents the gold standard for the production of synthetic mRNA in a laboratory setting, and continues to undergo extensive development for industrial scale manufacturing settings (Conrad et al., 2020; Dousis et al., 2023; Woo et al., 2021). Nevertheless, despite the simplicity of the IVT reaction, the requirement for purified recombinant components, which can be costly and have limited availability (Kis et al., 2021; Ohlson, 2020), presents a challenge in a large-scale production context. Further to this, despite optimisation of reaction components, product related impurities represent a key challenge in IVT based manufacturing, with their presence necessitating extensive downstream processing (Karikó et al., 2011; Nelson et al., 2020).

Herein, a proof of concept platform is presented for mRNA manufacture without the requirement for IVT. The fundamental underlying reaction is unchanged; T7 RNA polymerase is used to drive mRNA transcription from a double stranded DNA template encoding the cognate promoter and sequence of interest. This reaction however is

encapsulated in a microbial chassis, which confers the key advantage of being able to synthesise the required components for the transcription reaction. Grown in simple media, *E. coli* is able to produce the required RNA polymerase, nucleotides, and DNA template which normally need to be inputted in the IVT process. *E. coli* was selected as an expression due its simplicity to culture and manipulate, and scalability of existing manufacturing infrastructure.

The use of a microbial host provides several obvious engineering targets for the increased accumulation of synthetic mRNA of various sizes within the cell. Reported is a synergistic approach that looks to find optimal mRNA sequence, DNA template, media and *E. coli* phenotype. Crucial to the approach is the use of two distinct mRNA sequence features that reduce the rate of turnover of the synthetic mRNA molecule by host RNases. First is the introduction of self-splicing introns that lead to the production of a covalently circularised mRNA molecule. Owing to a lack of 5' and 3' ends, the molecule is completely resistant to exonuclease, and is a less optimal substrate for RNase E turnover. Circular mRNA is the focus of extensive research in the field of mRNA vaccines due to their increased stability and reduced innate immunogenicity (Bai et al., 2023; Niu et al., 2023). A crucial benefit of the circular molecule in the *E. coli* manufacturing platform is the lack of necessity for modified nucleotides or a cap structure in the final product, meaning the molecule would need no further modification after extraction from the cell.

The second stabilisation feature uses highly structured tRNA sequences to provide resistance to mRNA turnover. Whilst the molecule is linear, and therefore more in line with current approved mRNA therapies, several further processing steps would be required in order to use the molecule *in vivo*. Endogenous *E. coli* mRNA does not feature a 5' m7G cap, found on mammalian mRNA, and the cell does not encode the

capping enzymes required to produce this structure. This would therefore necessitate an enzymatic capping step post mRNA purification.

One of the key considerations in manufacturing biomolecules in any cellular chassis is the complexity of the downstream processing required. Presented is a scalable option for the extraction of RNA from the *E. coli* host, showing that the molecule has affinity for oligo-d(T) resin. Unlike in mammalian systems, *E. coli* mRNA transcripts are not extensively polyadenylated, with only a small number of transcripts having a tail in the region of 1-5 residues. This allows for efficient separation of the synthetic mRNA from endogenous mRNA.

Synthetic mRNA produced in *E. coli*, encoding eGFP, is shown here to be functional in a human cell line, through transfection of suspension adapted HEK293 cells. As discussed previously, the circular mRNA required no further processing in order to be translated, whilst the tRNA scaffold mRNA was capped before transfection. Transfection efficiency and level of protein production could be improved simply either by modifying the method of transfection, or optimising the DNA sequence for the cell line being used. For the purposes of this study, a cost effective cationic lipid and polymer formulation was used for transfection, which typically yields lower transfection efficiencies than electroporation based methods. Of particular importance for circular mRNA is using an IRES optimal for the cell line of interest, in order to maximise translational rates.

6.3.2 Paper II – Future work

In vivo mRNA production presented in this thesis was carried out in shake-flasks, which whilst cost effective and convenient, do not allow for particularly high densities of *E. coli* cells. In proof of principle experiments, a cell weight of only 2.5 g/L was

typically achieved, as conditions were optimised to ensure cells remained in exponential phase, rather than enhancing cell density. Yields achieved by the *E. coli* manufacturing platform could be dramatically increased through maintaining cells at higher densities in fed batch bioreactor conditions, where the levels of nutrients are maintained throughout. Cells densities of 100 g/L can routinely be achieved in a bioreactor, with reports of densities up to 190 g/L (Beckmann et al., 2017; Shiloach and Fass, 2005). If the achieved yields of approx. 10 mg per 2.5 g of cells are extrapolated to these densities, then product titre may hundreds of milligrams per litre, increasing the viability of platform as an alternative to IVT.

Strategies for capping and modified nucleotide incorporation would need addressing if the platform is to compete with IVT based systems. In terms of capping, one approach would be to express a T7 RNA polymerase & Vaccinia virus capping enzyme fusion protein, previously shown to be functional *in vivo*, producing translatable T7 RNAP transcripts in Chinese Hamster Ovary cells (Qin et al., 2023). The fusion protein would require characterising in *E. coli*, however if functional, this may provide a means of co-transcriptionally adding a cap-0 structure to the transcripts. Approved mRNA vaccines also utilise N1-methyl-psuedouridine in place of uridine in order to ablate immune stimulation (Morais et al., 2021). A potential strategy for incorporation of modified nucleotides into the *E. coli* is the co-expression of nucleotide modifying enzymes such as Pseudouridine sythetase, able to convert uridine into pseudouridine (Carlile et al., 2019). Enzymes able to convert nucleotides to N1-methyl-psuedouridine exist in the Archaeal kingdom, however their activity would again need characterising in an *E. coli* context (Wurm et al., 2012). To summarise, one of the priorities for improving the function of the mRNA manufacturing platform is to incorporate a means of capping and modifying the product mRNA.

The overall purity from the single downstream processing unit operation described here remains relatively low. Two approaches could be taken to improve mRNA purity. Firstly, the oligo-d(T) step itself could be optimised by eluting mRNA in a gentle gradient of buffer to water, and collecting small fractions containing different RNA species. Fractions could then be assessed for their level of purity. Similar approaches are routinely employed in the affinity purification of proteins (Chaga, 2001; Gaberc-Porekar and Menart, 2005). The second approach would be to proceed with further purification operations including size exclusion and ion exchange chromatography. Indeed, existing large scale IVT purification processes are multi-stage, involving several rounds of chromatography and tangential flow filtration, as a consequence of the immunostimulatory properties of dsRNA contaminants. Each sequential operation would increase purity; however, a proportion of product yield may be lost at each stage.

One potential avenue that could be explored further for stabilising the mRNA within the cellular context is the use of virus-like particles (VLPs), as a means of encapsulating the mRNA. This could be achieved through co-expressing MS2 coat proteins with the mRNA of interest, incorporating an MS2 binding loop into the 3' of the mRNA sequence. The coat proteins spontaneously assemble into a VLP structure, which could be purified away from the rest of the cellular material by size exclusion chromatography (Mikel et al., 2017). The VLP provides a barrier to RNases, and proprietary data from several companies indicates that product quality is increased, as dsRNAs are not incorporated into the VLP.

7 - References

Agostinetto, R., Rossi, M., Dawson, J., Lim, A., Simoneau, M.H., Boucher, C., Valldorf, B., Ross-Gillespie, A., Jardine, J.G., Sok, D., et al. (2022). Rapid cGMP manufacturing of COVID-19 monoclonal antibody using stable CHO cell pools. Biotechnol. Bioeng. *119*, 663–666.

Aiuti, A., Cattaneo, F., Galimberti, S., and Roncarolo, M.-G. (2009). Gene therapy for immunodeficiency due to adenosine deaminase deficiency. Pediatrics *124*.

Al'abri, I.S., Haller, D.J., Li, Z., and Crook, N. (2022). Inducible directed evolution of complex phenotypes in bacteria. Nucleic Acids Res. *50*, 1–12.

Alexopoulou, L., Holt, A.C., Medzhitov, R., and Flavell, R.A. (2001). Recognition of double-stranded RNA and activation of NF-κB by Toll-like receptor 3. Nature *413*, 732–738.

Alvarez-Guaita, A., Patel, S., Lim, K., Haider, A., Dong, L., Conway, O.J., Ma, M.K.L., Chiarugi, D., Saudek, V., O'Rahilly, S., et al. (2021). Phenotypic characterization of Adig null mice suggests roles for adipogenin in the regulation of fat mass accrual and leptin secretion. Cell Rep. *34*.

Amaya, L., Grigoryan, L., Li, Z., Lee, A., Wender, P.A., Pulendran, B., and Chang, H.Y. (2023). Circular RNA vaccine induces potent T cell responses. Proc. Natl. Acad. Sci. U. S. A. *120*.

Ambele, M.A., Dhanraj, P., Giles, R., and Pepper, M.S. (2020). Adipogenesis: A complex interplay of multiple molecular determinants and pathways. Int. J. Mol. Sci. *21*, 1–27.

Anderson, B.R., Muramatsu, H., Nallagatla, S.R., Bevilacqua, P.C., Sansing, L.H., Weissman, D., and Karikó, K. (2010). Incorporation of pseudouridine into mRNA enhances translation by diminishing PKR activation. Nucleic Acids Res. *38*, 5884–5892.

Angov, E., Hillier, C.J., Kincaid, R.L., and Lyon, J.A. (2008). Heterologous Protein Expression Is Enhanced by Harmonizing the Codon Usage Frequencies of the Target Gene with those of the Expression Host.

Anttila, V., Saraste, A., Knuuti, J., Hedman, M., Jaakkola, P., Laugwitz, K.L., Krane, M., Jeppsson, A., Sillanmäki, S., Rosenmeier, J., et al. (2023). Direct intramyocardial injection of VEGF mRNA in patients undergoing coronary artery bypass grafting. Mol. Ther. *31*, 866–874.

Asrani, K.H., Farelli, J.D., Stahley, M.R., Miller, R.L., Cheng, C.J., Subramanian, R.R., and Brown, J.M. (2018). Optimization of mRNA untranslated regions for improved expression of therapeutic mRNA. RNA Biol. *15*, 756–762.

Attarwala, H., Lumley, M., Liang, M., Ivaturi, V., and Senn, J. (2023). Translational Pharmacokinetic/Pharmacodynamic Model for mRNA-3927, an Investigational Therapeutic for the Treatment of Propionic Acidemia. Nucleic Acid Ther. *33*, 141–147.

Avello, V., Torres, M., Vergara, M., Berrios, J., Valdez-Cruz, N.A., Acevedo, C., Sampayo, M.M., Dickson, A.J., and Altamirano, C. (2022). Enhanced recombinant

protein production in CHO cell continuous cultures under growth-inhibiting conditions is associated with an arrested cell cycle in G1/G0 phase. PLoS One *17*, e0277620.

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., et al. (2021). Efficacy and Safety of the mRNA-1273 SARS-CoV-2 Vaccine. N. Engl. J. Med. *384*, 403–416.

Badran, A.H., and Liu, D.R. (2015). In vivo continuous directed evolution. Curr. Opin. Chem. Biol. *24*, 1–10.

Bae, D., Hyeon, H., Shin, E., Yeom, J.H., and Lee, K. (2023). Relaxed Cleavage Specificity of Hyperactive Variants of Escherichia coli RNase E on RNA I. J. Microbiol. *61*, 211–220.

Baek, E., Kim, C.L., Kim, M.G., Lee, J.S., and Lee, G.M. (2016). Chemical inhibition of autophagy: Examining its potential to increase the specific productivity of recombinant CHO cell lines. Biotechnol. Bioeng. *113*, 1953–1961.

Bai, Y., Liu, D., He, Q., Liu, J., Mao, Q., and Liang, Z. (2023). Research progress on circular RNA vaccines. Front. Immunol. *13*, 1–12.

Bandyra, K.J., and Luisi, B.F. (2018). RNase E and the High-Fidelity Orchestration of RNA Metabolism. Microbiol. Spectr. *6*.

Baronti, L., Karlsson, H., Marušič, M., and Petzold, K. (2018). A guide to large-scale RNA sample preparation. Anal. Bioanal. Chem. *410*, 3239–3252.

Bause, A.S., and Haigis, M.C. (2013). SIRT3 regulation of mitochondrial oxidative stress. Exp. Gerontol. *48*, 634–639.

Beck, J.D., Reidenbach, D., Salomon, N., Sahin, U., Türeci, Ö., Vormehr, M., and Kranz, L.M. (2021). mRNA therapeutics in cancer immunotherapy. Mol. Cancer *20*, 1–24.

Beckert, B., and Masquida, B. (2011). Synthesis of RNA by In Vitro Transcription.

Beckmann, B., Hohmann, D., Eickmeyer, M., Bolz, S., Brodhagen, C., Derr, P., and Sanders, E.A. (2017). An improved high cell density cultivation—iHCDC—strategy for leucine auxotrophic Escherichia coli K12 ER2507. Eng. Life Sci. *17*, 857–864.

Bengoechea-Alonso, M.T., and Ericsson, J. (2010). The ubiquitin ligase Fbxw7 controls adipocyte differentiation by targeting C/EBPα for degradation. Proc. Natl. Acad. Sci. U. S. A. *107*, 11817–11822.

Bento, F.M.M., Marques, R.N., Campana, F.B., Demétrio, C.G.B., Leandro, R.A., Parra, J.R.P., and Figueira, A. (2020). Gene silencing by RNAi via oral delivery of dsRNA by bacteria in the South American tomato pinworm, Tuta absoluta. Pest Manag. Sci. *76*, 287–295.

Bernstein, J.A., Khodursky, A.B., Lin, P.H., Lin-Chao, S., and Cohen, S.N. (2002). Global analysis of mRNA decay and abundance in Escherichia coli at single-gene resolution using two-color fluorescent DNA microarrays. Proc. Natl. Acad. Sci. U. S. A. *99*, 9697–9702.

Bhatwa, A., Wang, W., Hassan, Y.I., Abraham, N., Li, X.Z., and Zhou, T. (2021). Challenges Associated With the Formation of Recombinant Protein Inclusion Bodies

in Escherichia coli and Strategies to Address Them for Industrial Applications. Front. Bioeng. Biotechnol. *9*, 630551.

Binder, J.L., Chander, P., Deretic, V., Weick, J.P., and Bhaskar, K. (2020). Optical induction of autophagy via Transcription factor EB (TFEB) reduces pathological tau in neurons. PLoS One *15*, e0230026.

Björk, C., Subramanian, N., Liu, J., Acosta, J.R., Tavira, B., Eriksson, A.B., Arner, P., and Laurencikiene, J. (2021). An RNAi Screening of Clinically Relevant Transcription Factors Regulating Human Adipogenesis and Adipocyte Metabolism. Endocrinol. (United States) *162*.

Blakney, A.K., Ip, S., and Geall, A.J. (2021). An update on self-amplifying mRNA vaccine development. Vaccines *9*, 1–26.

Bloom, K., van den Berg, F., and Arbuthnot, P. (2021). Self-amplifying RNA vaccines for infectious diseases. Gene Ther. *28*, 117–129.

Boczkowski, D., Nair, S.K., Snyder, D., and Gilboa, E. (1996). Dendritic cells pulsed with RNA are potent antigen-presenting cells in vitro and in vivo. J. Exp. Med. *184*, 465–472.

Bolivar, F., Rodriguez, R.L., Greene, P.J., Betlach, M.C., Heyneker, H.L., Boyer, H.W., Crosa, J.H., and Falkow, S. (1977). Construction and characterization of new cloning vehicle. II. A multipurpose cloning system. Gene *2*, 95–113.

Borkotoky, S., and Murali, A. (2018). The highly efficient T7 RNA polymerase: A wonder macromolecule in biological realm. Int. J. Biol. Macromol. *118*, 49–56.

Börner, J., Friedrich, T., Bartkuhn, M., and Klug, G. (2023). Ribonuclease E strongly impacts bacterial adaptation to different growth conditions. RNA Biol. *20*, 120–135.

Bornstein, C., Winter, D., Barnett-Itzhaki, Z., David, E., Kadri, S., Garber, M., and Amit, I. (2014). A Negative Feedback Loop of Transcription Factors Specifies Alternative Dendritic Cell Chromatin States. Mol. Cell *56*, 749–762.

Boulain, J.C., Dassa, J., Mesta, L., Savatier, A., Costa, N., H.Muller, B., L'hostis, G., A.Stura, E., Troesch, A., and Ducancel, F. (2013). Mutants with higher stability and specific activity from a single thermosensitive variant of T7 RNA polymerase. Protein Eng. Des. Sel. *26*, 725–734.

Breda, L., Papp, T.E., Triebwasser, M.P., Yadegari, A., Fedorky, M.T., Tanaka, N., Abdulmalik, O., Pavani, G., Wang, Y., Grupp, S.A., et al. (2023). In vivo hematopoietic stem cell modification by mRNA delivery. Science *381*, 436–443.

Brenner, S., Jacob, F., and Meselson, M. (1961). An unstable intermediate carrying information from genes to ribosomes for protein synthesis. Nature *190*, 576–581.

Brown, A.J., Gibson, S.J., Hatton, D., Arnall, C.L., and James, D.C. (2019). Whole synthetic pathway engineering of recombinant protein production. Biotechnol. Bioeng. *116*, 375–387.

Budge, J.D., Knight, T.J., Povey, J., Roobol, J., Brown, I.R., Singh, G., Dean, A., Turner, S., Jaques, C.M., Young, R.J., et al. (2020). Engineering of Chinese hamster ovary cell lipid metabolism results in an expanded ER and enhanced recombinant biotherapeutic protein production. Metab. Eng. *57*, 203–216.

Budge, J.D., Roobol, J., Singh, G., Mozzanino, T., Knight, T.J., Povey, J., Dean, A., Turner, S.J., Jaques, C.M., Young, R.J., et al. (2021). A proline metabolism selection system and its application to the engineering of lipid biosynthesis in Chinese hamster ovary cells. Metab. Eng. Commun. *13*, 2214–0301.

Bulcha, J.T., Wang, Y., Ma, H., Tai, P.W.L., and Gao, G. (2021). Viral vector platforms within the gene therapy landscape. Signal Transduct. Target. Ther. *6*.

Bunnell, T.M., Burbach, B.J., Shimizu, Y., and Ervasti, J.M. (2011). β-Actin specifically controls cell growth, migration, and the G-actin pool. Mol. Biol. Cell *22*, 4047–4058.

Bydlinski, N., Coats, M.T., Maresch, D., Strasser, R., and Borth, N. (2020). Transfection of glycoprotein encoding mRNA for swift evaluation of N-glycan engineering strategies. Biotechnol. Prog. *36*.

Cafri, G., Gartner, J.J., Zaks, T., Hopson, K., Levin, N., Paria, B.C., Parkhurst, M.R., Yossef, R., Lowery, F.J., Jafferji, M.S., et al. (2020). mRNA vaccine–induced neoantigen-specific T cell immunity in patients with gastrointestinal cancer. J. Clin. Invest. *130*, 5976–5988.

Callaghan, A.J., Marcaida, M.J., Stead, J.A., McDowall, K.J., Scott, W.G., and Luisi, B.F. (2005). Structure of Escherichia coli RNase E catalytic domain and implications for RNA turnover. Nature *437*, 1187–1191.

Calvopina-Chavez, D.G., Gardner, M.A., and Griffitts, J.S. (2022). Engineering efficient termination of bacteriophage T7 RNA polymerase transcription. G3 Genes, Genomes, Genet. *12*, 3–8.

Cao, J., Cheng, L., and Shi, Y. (2007). Catalytic properties of MGAT3, a putative triacylglycerol synthase. J. Lipid Res. *48*, 583–591.

Cao, J., Choi, M., Guadagnin, E., Soty, M., Silva, M., Verzieux, V., Weisser, E., Markel, A., Zhuo, J., Liang, S., et al. (2021). mRNA therapy restores euglycemia and prevents liver tumors in murine model of glycogen storage disease. Nat. Commun. *12*.

Carlile, T.M., Martinez, N.M., Schaening, C., Su, A., Bell, T.A., Zinshteyn, B., and Gilbert, W. V. (2019). mRNA structure determines modification by pseudouridine synthase 1. Nat. Chem. Biol. *15*, 966.

Carroll, M., and Borden, K.L.B. (2013). The Oncogene eIF4E: Using Biochemical Insights to Target Cancer. J. Interf. Cytokine Res. *33*, 227–238.

Carter, A.D., Morris, C.E., and McAllister, W.T. (1981). Revised transcription map of the late region of bacteriophage T7 DNA. J. Virol. *37*, 636–642.

Cartwright, J.F., Arnall, C.L., Patel, Y.D., Barber, N.O.W., Lovelady, C.S., Rosignoli, G., Harris, C.L., Dunn, S., Field, R.P., Dean, G., et al. (2020). A platform for context-specific genetic engineering of recombinant protein production by CHO cells.

Carver, J., Ng, D., Zhou, M., Ko, P., Zhan, D., Yim, M., Shaw, D., Snedecor, B., Laird, M.W., Lang, S., et al. (2020). Maximizing antibody production in a targeted integration host by optimization of subunit gene dosage and position. Biotechnol. Prog. *36*, e2967.

Ceccaldi, R., Rondinelli, B., and D'Andrea, A.D. (2016). Repair Pathway Choices and Consequences at the Double-Strand Break. Trends Cell Biol. *26*, 52–64.

Cesana, D., Sgualdino, J., Rudilosso, L., Merella, S., Naldini, L., and Montini, E. (2012). Whole transcriptome characterization of aberrant splicing events induced by lentiviral vector integrations. J. Clin. Invest. *122*.

Chaga, G.S. (2001). Twenty-five years of immobilized metal ion affinity chromatography: Past, present and future. J. Biochem. Biophys. Methods *49*, 313–334.

Chan, C.P., Kok, K.H., and Jin, D.Y. (2011). CREB3 subfamily transcription factors are not created equal: Recent insights from global analyses and animal models. Cell Biosci. *1*, 1–7.

Chan, S.H., Molé, C.N., Nye, D., Mitchell, L., Dai, N., Buss, J., Kneller, D.W., Whipple, J.M., and Robb, G.B. (2023). Biochemical characterization of mRNA capping enzyme from Faustovirus.

Chaves-Moreira, D., Mitchell, M.A., Arruza, C., Rawat, P., Sidoli, S., Nameki, R., Reddy, J., Corona, R.I., Afeyan, L.K., Klein, I.A., et al. (2022). The transcription factor PAX8 promotes angiogenesis in ovarian cancer through interaction with SOX17. Sci. Signal. *15*, 2496.

Chen, H., Liu, S., Padula, S., Lesman, D., Griswold, K., Lin, A., Zhao, T., Marshall, J.L., and Chen, F. (2020). Efficient, continuous mutagenesis in human cells using a pseudo-random DNA editor. Nat. Biotechnol. *38*, 165–168.

Chen, T.H., Potapov, V., Dai, N., Ong, J.L., and Roy, B. (2022). N 1-methyl-pseudouridine is incorporated with higher fidelity than pseudouridine in synthetic RNAs. Sci. Rep. *12*, 1–13.

Chen, Y., Su, Y., Run, X., Sun, Z., Wang, T., Sun, S., and Liang, Z. (2013). Pretreatment of PC12 cells with 17β-estradiol prevents aβ-induced down-regulation of CREB phosphorylation and prolongs inhibition of GSK-3β. J. Mol. Neurosci. *50*, 394–401.

Chowdhury, A.M.M.A., Katoh, H., Hatanaka, A., Iwanari, H., Nakamura, N., Hamakubo, T., Natsume, T., Waku, T., and Kobayashi, A. (2017). Multiple regulatory mechanisms of the biological function of NRF3 (NFE2L3) control cancer cell proliferation. Sci. Reports 2017 71 *7*, 1–14.

Christiaens, O., Tardajos, M.G., Reyna, Z.L.M., Dash, M., Dubruel, P., and Smagghe, G. (2018). Increased RNAi efficacy in Spodoptera exigua via the formulation of dsRNA with guanylated polymers. Front. Physiol. *9*.

Chung, J.Y., Thone, M.N., and Kwon, Y.J. (2020). COVID-19 vaccines: The status and perspectives in delivery points of view. Adv. Drug Deliv. Rev.

Chusainow, J., Yang, Y.S., Yeo, J.H.M., Ton, P.C., Asvadi, P., Wong, N.S.C., and Yap, M.G.S. (2009). A study of monoclonal antibody-producing CHO cell lines: what makes a stable high producer? Biotechnol. Bioeng. *102*, 1182–1196.

Coats, M.T., Bydlinski, N., Maresch, D., Diendorfer, A., Klanert, G., and Borth, N. (2020). mRNA Transfection into CHO-Cells Reveals Production Bottlenecks. Biotechnol. J. *15*, 1900198.

Conrad, T., Plumbom, I., Alcobendas, M., Vidal, R., and Sauer, S. (2020). Maximizing transcription of nucleic acids with efficient T7 promoters. Commun. Biol. *3*, 439.

Conzen, S.D., Gottlob, K., Kandel, E.S., Khanduri, P., Wagner, A.J., O'Leary, M., and Hay, N. (2000). Induction of Cell Cycle Progression and Acceleration of Apoptosis Are Two Separable Functions of c-Myc: Transrepression Correlates with Acceleration of Apoptosis. Mol. Cell. Biol. *20*, 6008.

Costello, A., Lao, N.T., Barron, N., and Clynes, M. (2020). Reinventing the Wheel: Synthetic Circular RNAs for Mammalian Cell Engineering. Trends Biotechnol. *38*, 217–230.

Court, D.L., Gan, J., Liang, Y.H., Shaw, G.X., Tropea, J.E., Costantino, N., Waugh, D.S., and Ji, X. (2013). RNase III: Genetics and function; Structure and mechanism. Annu. Rev. Genet. *47*, 405–431.

Crick, F. (1970). Central dogma of molecular biology. Nature *227*, 561–563.

Cui, M., Wong, O., Li, Q., and Wang, W. (2023). An Assay Method for Characterizing Bacteriophage T7 RNA Polymerase Activity by Transcription-Translation (TX-TL) System. J. Biochem. *173*, 343–352.

Dalaisón-Fuentes, L.I., Pascual, A., Gazza, E., Welchen, E., Rivera-Pomar, R., and Catalano, M.I. (2022). Development of efficient RNAi methods in the corn leafhopper Dalbulus maidis, a promising application for pest control. Pest Manag. Sci. *78*, 3108–3116.

Dan, H.C., Sun, M., Kaneko, S., Feldman, R.I., Nicosia, S. V., Wang, H.G., Tsang, B.K., and Cheng, J.Q. (2004). Akt Phosphorylation and Stabilization of X-linked Inhibitor of Apoptosis Protein (XIAP). J. Biol. Chem. *279*, 5405–5412.

Davis, R., Schooley, K., Rasmussen, B., Thomas, J., and Reddy, P. (2000). Effect of PDI overexpression on recombinant protein secretion in CHO cells. Biotechnol. Prog. *16*, 736–743.

Delgado-Martín, J., and Velasco, L. (2021). An efficient dsRNA constitutive expression system in Escherichia coli. Appl. Microbiol. Biotechnol. *105*, 6381–6393.

Denby, C.M., Im, J.H., Yu, R.C., Pesce, C.G., and Brem, R.B. (2012). Negative feedback confers mutational robustness in yeast transcription factor regulation. Proc. Natl. Acad. Sci. U. S. A. *109*, 3874–3878.

Deng, X., Gao, F., Flagg, T., and May, W.S. (2004). Mono- and multisite phosphorylation enhances Bcl2's antiapoptotic function and inhibition of cell cycle entry functions. Proc. Natl. Acad. Sci. U. S. A. *101*, 153–158.

Deviatkin, A.A., Simonov, R.A., Trutneva, K.A., Maznina, A.A., Soroka, A.B., Kogan, A.A., Feoktistova, S.G., Khavina, E.M., Mityaeva, O.N., and Volchkov, P.Y. (2023). Cap-Independent Circular mRNA Translation Efficiency. Vaccines *11*, 1–12.

DiGiusto, D.L., Cannon, P.M., Holmes, M.C., Li, L., Rao, A., Wang, J., Lee, G., Gregory, P.D., Kim, K.A., Hayward, S.B., et al. (2016). Preclinical development and qualification of ZFN-mediated CCR5 disruption in human hematopoietic stem/progenitor cells. Mol. Ther. - Methods Clin. Dev. *3*.

Donaldson, J., Kleinjan, D.J., and Rosser, S. (2022). Synthetic biology approaches

for dynamic CHO cell engineering. Curr. Opin. Biotechnol. *78*.

Dousis, A., Ravichandran, K., Hobert, E.M., Moore, M.J., and Rabideau, A.E. (2023). An engineered T7 RNA polymerase that produces mRNA free of immunostimulatory byproducts. Nat. Biotechnol. *41*, 560–568.

Dumont, J., Euwart, D., Mei, B., Estes, S., and Kshirsagar, R. (2016). Human cell lines for biopharmaceutical manufacturing: history, status, and future perspectives. Crit. Rev. Biotechnol. *36*, 1110–1122.

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

Ebbesen, K.K., Hansen, T.B., and Kjems, J. (2017). Insights into circular RNA biology. RNA Biol. *14*, 1035–1045.

Egorova, T., Shuvalova, E., Mukba, S., Shuvalov, A., Kolosov, P., and Alkalaeva, E. (2021). Method for rapid analysis of mutant rna polymerase activity on templates containing unnatural nucleotides. Int. J. Mol. Sci. *22*, 5186.

Erasmus, J.H., Archer, J., Fuerte-Stone, J., Khandhar, A.P., Voigt, E., Granger, B., Bombardi, R.G., Govero, J., Tan, Q., Durnell, L.A., et al. (2020). Intramuscular Delivery of Replicon RNA Encoding ZIKV-117 Human Monoclonal Antibody Protects against Zika Virus Infection. Mol. Ther. - Methods Clin. Dev. *18*, 402–414.

Esquerré, T., Moisan, A., Chiapello, H., Arike, L., Vilu, R., Gaspin, C., Cocaign-Bousquet, M., and Girbal, L. (2015). Genome-wide investigation of mRNA lifetime determinants in Escherichia coli cells cultured at different growth rates. BMC Genomics *16*, 1–13.

Esvelt, K.M., Carlson, J.C., and Liu, D.R. (2011). A system for the continuous directed evolution of biomolecules. Nature *472*, 499–503.

Evseeva, M.N., Balashova, M.S., Kulebyakin, K.Y., and Rubtsov, Y.P. (2021). Adipocyte Biology from the Perspective of In Vivo Research: Review of Key Transcription Factors. Int. J. Mol. Sci. 2022, Vol. 23, Page 322 *23*, 322.

Fan, J., Sripada, S.A., Pham, D.N., Linova, M.Y., Woodley, J.M., Menegatti, S., Boi, C., and Carbonell, R.G. (2023). Purification of a monoclonal antibody using a novel high-capacity multimodal cation exchange nonwoven membrane. Sep. Purif. Technol. *317*, 123920.

Farley-Barnes, K.I., Deniz, E., Overton, M.M., Khokha, M.K., and Baserga, S.J. (2020). Paired Box 9 (PAX9), the RNA polymerase II transcription factor, regulates human ribosome biogenesis and craniofacial development. PLoS Genet. *16*, 1–36.

Al Fayez, N., Nassar, M.S., Alshehri, A.A., Alnefaie, M.K., Almughem, F.A., Alshehri, B.Y., Alawad, A.O., and Tawfik, E.A. (2023). Recent Advancement in mRNA Vaccine Development and Applications. Pharmaceutics *15*, 1–24.

Feng, Y.X., Jin, D.X., Sokol, E.S., Reinhardt, F., Miller, D.H., and Gupta, P.B. (2017). Cancer-specific PERK signaling drives invasion and metastasis through CREB3L1. Nat. Commun. *8*, 1–10.

Finn, J.D., Smith, A.R., Patel, M.C., Shaw, L., Youniss, M.R., van Heteren, J.,

Dirstine, T., Ciullo, C., Lescarbeau, R., Seitzer, J., et al. (2018). A Single Administration of CRISPR/Cas9 Lipid Nanoparticles Achieves Robust and Persistent In Vivo Genome Editing. Cell Rep. *22*, 2227–2235.

Fiorese, C.J., Schulz, A.M., Lin, Y.F., Rosin, N., Pellegrino, M.W., and Haynes, C.M. (2016). The Transcription Factor ATF5 Mediates a Mammalian Mitochondrial UPR. Curr. Biol. *26*, 2037–2043.

Fischer, S., Handrick, R., and Otte, K. (2015). The art of CHO cell engineering: A comprehensive retrospect and future perspectives. Biotechnol. Adv. *33*, 1878–1896.

Flamme, M., McKenzie, L.K., Sarac, I., and Hollenstein, M. (2019). Chemical methods for the modification of RNA. Methods *161*, 64–82.

Freund, I., Eigenbrod, T., Helm, M., and Dalpke, A.H. (2019). RNA modifications modulate activation of innate toll-like receptors. Genes (Basel). *10*.

Fuchs, A.L., Neu, A., and Sprangers, R. (2016). A general method for rapid and cost-efficient large-scale production of 5' capped RNA. RNA *22*, 1454–1466.

Furuichi, Y. (2015). Discovery of m7G-cap in eukaryotic mRNAs. Proc. Japan Acad. Ser. B Phys. Biol. Sci. *91*, 394–409.

Gaberc-Porekar, V., and Menart, V. (2005). Potential for using histidine tags in purification of proteins at large scale. Chem. Eng. Technol. *28*, 1306–1314.

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

Galbraith, M.D., Donner, A.J., and Espinosa, J.M. (2010). CDK8: A positive regulator of transcription. Transcription *1*, 4–12.

Galdadas, I., Bonis, V., Vgenopoulou, P., Papadourakis, M., Kakoulidis, P., Stergiou, G., Cournia, Z., and Klinakis, A. (2021). The effect of S427F mutation on RXRα activity depends on its dimeric partner. Chem. Sci. *12*, 14700–14710.

Gallotta, I., Mazzarella, N., Donato, A., Esposito, A., Chaplin, J.C., Castro, S., Zampi, G., Battaglia, G.S., Hilliard, M.A., Bazzicalupo, P., et al. (2016). Neuron-specific knock-down of SMN1 causes neuron degeneration and death through an apoptotic mechanism. Hum. Mol. Genet. *25*, 2564–2577.

Gan, L.M., Lagerström-Fermér, M., Carlsson, L.G., Arfvidsson, C., Egnell, A.C., Rudvik, A., Kjaer, M., Collén, A., Thompson, J.D., Joyal, J., et al. (2019). Intradermal delivery of modified mRNA encoding VEGF-A in patients with type 2 diabetes. Nat. Commun. *10*, 1–9.

Ganley, M., Holz, L.E., Minnell, J.J., de Menezes, M.N., Burn, O.K., Poa, K.C.Y., Draper, S.L., English, K., Chan, S.T.S., Anderson, R.J., et al. (2023). mRNA vaccine against malaria tailored for liver-resident memory T cells. Nat. Immunol. *24*, 1487–1498.

Garcia-Llorens, G., Martínez-Sena, T., Pareja, E., Tolosa, L., Castell, J. V., and Bort, R. (2023). A robust reprogramming strategy for generating hepatocyte-like cells usable in pharmaco-toxicological studies. Stem Cell Res. Ther. *14*, 1–19.

García-Niño, W.R., and Zazueta, C. (2021). New insights of Krüppel-like transcription

factors in adipogenesis and the role of their regulatory neighbors. Life Sci. *265*, 118763.

Gardner, L.P., Mookhtiar, K.A., and Coleman, J.E. (1997). Initiation, elongation, and processivity of carboxyl-terminal mutants of T7 RNA polymerase. Biochemistry *36*, 2908–2918.

GC, S., Bellis, S.L., and Hjelmeland, A.B. (2022). ST6Gal1: Oncogenic signaling pathways and targets. Front. Mol. Biosci. *9*, 1–14.

Gholamalipour, Y., Karunanayake Mudiyanselage, A., and Martin, C.T. (2018). NAR breakthrough article 3 end additions by T7 RNA polymerase are RNA self-templated, distributive and diverse in character—RNA-Seq analyses. Nucleic Acids Res. *46*, 9253–9263.

Gholamalipour, Y., Johnson, W.C., and Martin, C.T. (2019). Efficient inhibition of RNA self-primed extension by addition of competing 3'-capture DNA-improved RNA synthesis by T7 RNA polymerase. Nucleic Acids Res.

Ghosh, A., and Lima, C.D. (2010). Enzymology of RNA cap synthesis. Wiley Interdiscip. Rev. RNA *1*, 152–172.

Gibellini, L., Pinti, M., Bartolomeo, R., De Biasi, S., Cormio, A., Musicco, C., Carnevale, G., Pecorini, S., Nasi, M., De Pol, A., et al. (2015). Inhibition of Lon protease by triterpenoids alters mitochondria and is associated to cell death in human cancer cells. Oncotarget *6*, 25466–25483.

Ginn, S.L., Amaya, A.K., Alexander, I.E., Edelstein, M., and Abedi, M.R. (2018). Gene therapy clinical trials worldwide to 2017: An update. J. Gene Med. *20*, e3015.

Gong, H., Wen, J., Luo, R., Feng, Y., Guo, J.J., Fu, H., and Zhou, X. (2023). Integrated mRNA sequence optimization using deep learning. Brief. Bioinform. *24*, 1–11.

Gong, P., Esposito, E.A., and Martin, C.T. (2004). Initial Bubble Collapse Plays a Key Role in the Transition to Elongation in T7 RNA Polymerase*.

Gopal, G.J., and Kumar, A. (2013). Strategies for the production of recombinant protein in escherichia coli. Protein J. *32*, 419–425.

Gopalakrishnan, S., Johnson, W., Valderrama-gomez, M.A., and Icten, E. (2023). Multi-omic characterization of antibody-producing CHO cell lines elucidates.

Gote, V., Bolla, P.K., Kommineni, N., and Butreddy, A. (2023). A Comprehensive Review of mRNA Vaccines. Int. J. Mol. Sci. *2700*.

Gozdecka, M., Lyons, S., Kondo, S., Taylor, J., Li, Y., Walczynski, J., Thiel, G., Breitwieser, W., and Jones, N. (2014). JNK Suppresses Tumor Formation via a Gene-Expression Program Mediated by ATF2. Cell Rep. *9*, 1361–1374.

Grudzien-Nogalska, E., Stepinski, J., Jemielity, J., Zuberek, J., Stolarski, R., Rhoads, R.E., and Darzynkiewicz, E. (2007). Synthesis of Anti-Reverse Cap Analogs (ARCAs) and their Applications in mRNA Translation and Stability. Methods Enzymol. *431*, 203–227.

Gu, L., Zhu, N., Zhang, H., Durden, D.L., Feng, Y., and Zhou, M. (2009). Regulation

of XIAP Translation and Induction by MDM2 following Irradiation. Cancer Cell *15*, 363–375.

Gubelmann, C., Schwalie, P.C., Raghav, S.K., Röder, E., Delessa, T., Kiehlmann, E., Waszak, S.M., Corsinotti, A., Udin, G., Holcombe, W., et al. (2014). Identification of the transcription factor ZEB1 as a central component of the adipogenic gene regulatory network. Elife *3*, 1–30.

Guilleres, J., Lopez, P.J., Proux, F., Launay, H., and Dreyfus, M. (2005). A mutation in T7 RNA polymerase that facilitates promoter clearance. Proc. Natl. Acad. Sci. U. S. A. *102*, 5958–5963.

Ha, T.K., Kim, D., Kim, C.L., Grav, L.M., and Lee, G.M. (2022). Factors affecting the quality of therapeutic proteins in recombinant Chinese hamster ovary cell culture. Biotechnol. Adv. *54*.

Hacein-Bey-Abina, S., Garrigue, A., Wang, G.P., Soulier, J., Lim, A., Morillon, E., Clappier, E., Caccavelli, L., Delabesse, E., Beldjord, K., et al. (2008). Insertional oncogenesis in 4 patients after retrovirus-mediated gene therapy of SCID-X1. J. Clin. Invest. *118*, 3132–3142.

Hadas, Y., Sultana, N., Youssef, E., Sharkar, M.T.K., Kaur, K., Chepurko, E., and Zangi, L. (2019). Optimizing Modified mRNA In Vitro Synthesis Protocol for Heart Gene Therapy. Mol. Ther. - Methods Clin. Dev. *14*, 300–305.

Handlogten, M.W., Lee-O'Brien, A., Roy, G., Levitskaya, S. V., Venkat, R., Singh, S., and Ahuja, S. (2018). Intracellular response to process optimization and impact on productivity and product aggregates for a high-titer CHO cell process. Biotechnol. Bioeng. *115*, 126–138.

Hansen, H.G., Pristovšek, N., Kildegaard, H.F., and Lee, M. (2016). Improving the secretory capacity of Chinese hamster ovary cells by ectopic expression of effector genes: Lessons learned and future directions.

Hansen, H.G., Pristovšek, N., Kildegaard, H.F., and Lee, G.M. (2017). Improving the secretory capacity of Chinese hamster ovary cells by ectopic expression of effector genes: Lessons learned and future directions. Biotechnol. Adv. *35*, 64–76.

Harada, N., Yonemoto, H., Yoshida, M., Yamamoto, H., Yin, Y., Miyamoto, A., Hattori, A., Wu, Q., Nakagawa, T., Nakano, M., et al. (2008). Alternative splicing produces a constitutively active form of human SREBP-1. Biochem. Biophys. Res. Commun. *368*, 820–826.

Harding, H.P., Zhang, Y., Khersonsky, S., Marciniak, S., Scheuner, D., Kaufman, R.J., Javitt, N., Chang, Y.T., and Ron, D. (2005). Bioactive small molecules reveal antagonism between the integrated stress response and sterol-regulated gene expression. Cell Metab. *2*, 361–371.

Hashiro, S., Mitsuhashi, M., and Yasueda, H. (2019). Overexpression system for recombinant RNA in Corynebacterium glutamicum using a strong promoter derived from corynephage BFK20. J. Biosci. Bioeng. *128*, 255–263.

Heil, F., Hemmi, H., Hochrein, H., Ampenberger, F., Kirschning, C., Akira, S., Lipford, G., Wagner, H., and Bauer, S. (2004). Species-Specific Recognition of Single-Stranded RNA via Toll-like Receptor 7 and 8. Science (80-.). *303*, 1526–1529.

Heiser, A., Coleman, D., Dannull, J., Yancey, D., Maurice, M.A., Lallas, C.D., Dahm, P., Niedzwiecki, D., Gilboa, E., and Vieweg, J. (2002). Autologous dendritic cells transfected with prostate-specific antigen RNA stimulate CTL responses against metastatic prostate tumors. J. Clin. Invest. *109*, 409–417.

Henderson, J.M., Ujita, A., Hill, E., Yousif-Rosales, S., Smith, C., Ko, N., McReynolds, T., Cabral, C.R., Escamilla-Powers, J.R., and Houston, M.E. (2021). Cap 1 Messenger RNA Synthesis with Co-transcriptional CleanCap® Analog by In Vitro Transcription. Curr. Protoc. *1*.

Henry, M.N., MacDonald, M.A., Orellana, C.A., Gray, P.P., Gillard, M., Baker, K., Nielsen, L.K., Marcellin, E., Mahler, S., and Martínez, V.S. (2020). Attenuating apoptosis in Chinese hamster ovary cells for improved biopharmaceutical production. Biotechnol. Bioeng. *117*, 1187–1203.

Herzig, S., Hedrick, S., Morantte, I., Koo, S.H., Galimi, F., and Montminy, M. (2003). CREB controls hepatic lipid metabolism through nuclear hormone receptor PPAR-γ. Nature *426*, 190–193.

Heyde, S.A.H., and Nørholm, M.H.H. (2021). Tailoring the evolution of BL21(DE3) uncovers a key role for RNA stability in gene expression toxicity. Nat. Commun. *21*, 1–9.

Hinds, T.D., Stechschulte, L.A., Cash, H.A., Whisler, D., Banerjee, A., Yong, W., Khuder, S.S., Kaw, M.K., Shou, W., Najjar, S.M., et al. (2011). Protein phosphatase 5 mediates lipid metabolism through reciprocal control of glucocorticoid receptor and peroxisome proliferator-activated receptor-γ (PPARγ)*. J. Biol. Chem. *286*, 42911–42922.

Hippach, M.B., Schwartz, I., Pei, J., Huynh, J., Kawai, Y., and Zhu, M.M. (2018). Fluctuations in dissolved oxygen concentration during a CHO cell culture process affects monoclonal antibody productivity and the sulfhydryl-drug conjugation process. Biotechnol. Prog. *34*, 1427–1437.

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

Hong, W.X., Haebe, S., Lee, A.S., Benedikt Westphalen, C., Norton, J.A., Jiang, W., and Levy, R. (2020). Intratumoral immunotherapy for early-stage solid tumors. Clin. Cancer Res. *26*, 3091–3099.

Hou, X., Zaks, T., Langer, R., and Dong, Y. (2021). Lipid nanoparticles for mRNA delivery. Nat. Rev. Mater. *6*, 1078–1094.

Hough, J., Howard, J.D., Brown, S., Portwood, D.E., Kilby, P.M., and Dickman, M.J. (2022). Strategies for the production of dsRNA biocontrols as alternatives to chemical pesticides. Front. Bioeng. Biotechnol. *10*, 1–18.

Hu, Q., Mao, Y., Liu, M., Luo, R., Jiang, R., and Guo, F. (2020). The active nuclear form of SREBP1 amplifies ER stress and autophagy via regulation of PERK. FEBS J. *287*, 2348–2366.

Hu, X., Karthigeyan, K.P., Herbek, S., Valencia, S.M., Jenks, J.A., Webster, H., Miller,

I.G., Connors, M., Pollara, J., Andy, C., et al. (2023). Broad and Functional anti-HCMV IgG responses induced by mRNA-1647. MedRxiv.

Huang, Z., and Yoon, S. (2020). Identifying metabolic features and engineering targets for productivity improvement in CHO cells by integrated transcriptomics and genome-scale metabolic model.

Huang, J., Brieba, L.G., and Sousa, R. (2000). Misincorporation by Wild-Type and Mutant T7 RNA Polymerases: Identification of Interactions That Reduce Misincorporation Rates by Stabilizing the Catalytically Incompetent Open Conformation †.

Hyde, J.L., and Diamond, M.S. (2015). Innate immune restriction and antagonism of viral RNA lacking 2'-O methylation. Virology *479–480*, 66–74.

Ikeda, M., and Takeno, S. (2013). Amino Acid Production by Corynebacterium glutamicum. (Springer, Berlin, Heidelberg), pp. 107–147.

Inoue, K., and Imai, Y. (2015). Fatostatin, an SREBP inhibitor, prevented RANKL-induced bone loss by suppression of osteoclast differentiation. Biochim. Biophys. Acta - Mol. Basis Dis. *1852*, 2432–2441.

Jackson, R.J., T Hellen, C.U., and Pestova, T. V (2010). The mechanism of eukaryotic translation initiation and principles of its regulation.

Jemielity, J., Fowler, T., Zuberek, J., Stepinski, J., Lewdorowicz, M., Niedzwiecka, A., Stolarski, R., Darzynkiewicz, E., and Rhoads, R.E. (2003). Novel "anti-reverse" cap analogs with superior translational properties.

Jiang, Z., and Dalby, P.A. (2023). Challenges in scaling up AAV-based gene therapy manufacturing. Trends Biotechnol. *41*, 1268–1281.

Jiang, L., Berraondo, P., Jericó, D., Guey, L.T., Sampedro, A., Frassetto, A., Benenato, K.E., Burke, K., Santamaría, E., Alegre, M., et al. (2018a). Systemic messenger RNA as an etiological treatment for acute intermittent porphyria. Nat. Med. *24*, 1899–1909.

Jiang, R., Chen, H., and Xu, S. (2018b). pH excursions impact CHO cell culture performance and antibody N-linked glycosylation. Bioprocess Biosyst. Eng. *41*, 1731–1741.

Jiang, S., Koolmeister, C., Misic, J., Siira, S., Kühl, I., Silva Ramos, E., Miranda, M., Jiang, M., Posse, V., Lytovchenko, O., et al. (2019). TEFM regulates both transcription elongation and RNA processing in mitochondria. EMBO Rep. *20*, e48101.

Jin, S., Sun, Y., Liang, X., Gu, X., Ning, J., Xu, Y., Chen, S., and Pan, L. (2022). Emerging new therapeutic antibody derivatives for cancer treatment. Signal Transduct. Target. Ther. *7*, 1–28.

Jin, Y., Li, R., Zhang, Z., Ren, J., Song, X., and Zhang, G. (2020). ZBED1/DREF: A transcription factor that regulates cell proliferation (Review). Oncol. Lett. *20*, 1–8.

Joiner, J., Huang, Z., McHugh, K., Stebbins, M., Aron, K., Borys, M., and Khetan, A. (2022). Process modeling of recombinant adeno-associated virus production in HEK293 cells. Curr. Opin. Chem. Eng. *36*, 100823.

Josephson, S., Lagerholm, E., and Palm, G. (1984). Automatic synthesis of oligodeoxynucleotides and mixed oligodeoxynucleotides using the phosphoamidite method. Acta Chem. Scand. B. *38*, 539–545.

Joshi, S.H.N., Yong, C., and Gyorgy, A. (2022). Inducible plasmid copy number control for synthetic biology in commonly used E. coli strains. Nat. Commun. *13*, 1–16.

Jung, A.E., Fitzsimons, H.L., Bland, R.J., During, M.J., and Young, D. (2008). HSP70 and constitutively active HSF1 mediates protection against CDCrel-1-mediated toxicity. Mol. Ther. *16*, 1048.

Kameda, S., Ohno, H., and Saito, H. (2023). Synthetic circular RNA switches and circuits that control protein expression in mammalian cells. Nucleic Acids Res. *51*, E24.

Kandari, D., and Bhatnagar, R. (2023). Antibody engineering and its therapeutic applications. Int. Rev. Immunol. *42*, 156–183.

Kaneyoshi, K., Kuroda, K., Uchiyama, K., Onitsuka, M., Yamano-Adachi, N., Koga, Y., and Omasa, T. (2019). Secretion analysis of intracellular “difficult-to-express” immunoglobulin G (IgG) in Chinese hamster ovary (CHO) cells. Cytotechnology *71*, 305–316.

Kang, S., Dahl, R., Hsieh, W., Shin, A., Zsebo, K.M., Buettner, C., Hajjar, R.J., and Lebeche, D. (2016). Small molecular allosteric activator of the sarco/endoplasmic reticulum Ca2+-ATPase (SERCA) attenuates diabetes and metabolic disorders. J. Biol. Chem. *291*, 5185–5198.

Karikó, K., Buckstein, M., Ni, H., and Weissman, D. (2005). Suppression of RNA recognition by Toll-like receptors: The impact of nucleoside modification and the evolutionary origin of RNA. Immunity *23*, 165–175.

Karikó, K., Muramatsu, H., Welsh, F.A., Ludwig, J., Kato, H., Akira, S., and Weissman, D. (2008). Incorporation of pseudouridine into mRNA yields superior nonimmunogenic vector with increased translational capacity and biological stability. Mol. Ther. *16*, 1833–1840.

Karikó, K., Muramatsu, H., Ludwig, J., and Weissman, D. (2011). Generating the optimal mRNA for therapy: HPLC purification eliminates immune activation and improves translation of nucleoside-modified, protein-encoding mRNA. Nucleic Acids Res. *39*.

Karottki, K.J. la C., Hefzi, H., Li, S., Pedersen, L.E., Spahn, P.N., Joshi, C., Ruckerbauer, D., Bort, J.A.H., Thomas, A., Lee, J.S., et al. (2021). A metabolic CRISPR-Cas9 screen in Chinese hamster ovary cells identifies glutamine-sensitive genes. Metab. Eng. *66*, 114–122.

Kaur, G., Cheung, H.C., Xu, W., Wong, J.V., Chan, F.F., Li, Y., McReynolds, L., and Huang, L. (2018). Milligram scale production of potent recombinant small interfering RNAs in Escherichia coli. Biotechnol. Bioeng. *115*, 2280–2291.

Kawasaki, S., Fujita, Y., Nagaike, T., Tomita, K., and Saito, H. (2017). Synthetic mRNA devices that detect endogenous proteins and distinguish mammalian cells. Nucleic Acids Res. *45*, e117–e117.

Kersten, S., Dong, D., Lee, W. yi, Reczek, P.R., and Noy, N. (1998). Auto-silencing by the retinoid X receptor. J. Mol. Biol. *284*, 21–32.

Khurana, S., Chakraborty, S., Zhao, X., Liu, Y., Guan, D., Lam, M., Huang, W., Yang, S., and Kao, H.Y. (2012). Identification of a novel LXXLL motif in α-actinin 4-spliced isoform that is critical for its interaction with estrogen receptor α and co-activators. J. Biol. Chem. *287*, 35418–35429.

Kikuchi, Y., and Umekage, S. (2018). Extracellular nucleic acids of the marine bacterium Rhodovulum sulfidophilum and recombinant RNA production technology using bacteria. FEMS Microbiol. Lett. *365*, 268.

Kiliszek, A., B-Laszczyk, L., Kierzek, R., and Rypniewski, W. (2017). Stabilization of RNA hairpins using non-nucleotide linkers and circularization. Nucleic Acids Res. *45*, 92.

Kim, J.Y., Kim, Y.G., and Lee, G.M. (2012). CHO cells in biotechnology for production of recombinant proteins: Current state and further potential. Appl. Microbiol. Biotechnol. *93*, 917–930.

Kim, K.Q., Burgute, B.D., Tzeng, S.C., Jing, C., Jungers, C., Zhang, J., Yan, L.L., Vierstra, R.D., Djuranovic, S., Evans, B.S., et al. (2022a). N1-methylpseudouridine found within COVID-19 mRNA vaccines produces faithful protein products. Cell Rep. *40*, 111300.

Kim, S.C., Sekhon, S.S., Shin, W.R., Ahn, G., Cho, B.K., Ahn, J.Y., and Kim, Y.H. (2022b). Modifications of mRNA vaccine structural elements for improving mRNA stability and translation efficiency. Mol. Cell. Toxicol. *18*, 1–8.

Kimchi-Sarfaty, C., Jung, †, Oh, M., Kim, I.-W., Sauna, Z.E., Calcagno, A.M., Ambudkar, S. V, and Gottesman, M.M. (2004). Supporting Online Material A “Silent” Polymorphism in the MDR1 Gene Changes Substrate Specificity. Nat. Cell Biol *119*, 268.

Kime, L., Vincent, H.A., Gendoo, D.M.A., Jourdan, S.S., Fishwick, C.W.G., Callaghan, A.J., and McDowall, K.J. (2015). The first small-molecule inhibitors of members of the ribonuclease E family. Sci. Rep. *5*, 1–7.

Kis, Z., Kontoravdi, C., Shattock, R., and Shah, N. (2021). Resources, production scales and time required for producing RNA vaccines for the global pandemic demand. Vaccines *9*, 1–14.

Kiyama, T., Chen, C.K., Wang, S.W., Pan, P., Ju, Z., Wang, J., Takada, S., Klein, W.H., and Mao, C.A. (2018). Essential roles of mitochondrial biogenesis regulator Nrf1 in retinal development and homeostasis. Mol. Neurodegener. *13*, 1–23.

Klanert, G., Fernandez, D.J., Weinguny, M., Eisenhut, P., Bühler, E., Melcher, M., Titus, S.A., Diendorfer, A.B., Gludovacz, E., Jadhav, V., et al. (2019). A cross-species whole genome siRNA screen in suspension-cultured Chinese hamster ovary cells identifies novel engineering targets. Sci. Rep. *9*, 1–11.

Klopotek, A., Hirche, F., and Eder, K. (2006). PPARγ ligand troglitazone lowers cholesterol synthesis in HepG2 and Caco-2 cells via a reduced concentration of nuclear SREBP-2. Exp. Biol. Med. *231*, 1365–1372.

Kocmik, I., Piecyk, K., Rudzinska, M., Niedzwiecka, A., Darzynkiewicz, E., Grzela,

R., and Jankowska-Anyszka, M. (2018). Modified ARCA analogs providing enhanced translational properties of capped mRNAs. Cell Cycle *17*.

Koh, J.H., Kim, Y.W., Seo, D.Y., and Sohn, T.S. (2021). Mitochondrial TFAM as a signaling regulator between cellular organelles: A perspective on metabolic diseases. Diabetes Metab. J. *45*, 853–865.

Komatsu, S., Ohno, H., and Saito, H. (2023). Target-dependent RNA polymerase as universal platform for gene expression control in response to intracellular molecules. Nat. Commun. *14*, 1–11.

Konarska, M.M., and Sharp, P.A. (1989). Replication of RNA by the DNA-dependent RNA polymerase of phage T7. Cell *57*, 423–431.

Kong, B., Kim, Y., Kim, E.H., Suk, J.S., and Yang, Y. (2023). mRNA: A promising platform for cancer immunotherapy. Adv. Drug Deliv. Rev. *199*, 114993.

Kram, K.E., and Finkel, S.E. (2015). Rich medium composition affects Escherichia coli survival, glycation, and mutation frequency during long-term batch culture. Appl. Environ. Microbiol. *81*, 4442–4450.

Kramer, S., and Carrington, M. (2014). An AU-rich instability element in the 3′UTR mediates an increase in mRNA stability in response to expression of a dhh1 ATPase mutant. Translation *2*, e28587.

Kulkarni, J.A., Witzigmann, D., Thomson, S.B., Chen, S., Leavitt, B.R., Cullis, P.R., and van der Meel, R. (2021). The current landscape of nucleic acid therapeutics. Nat. Nanotechnol. *16*, 630–643.

Kuo, C.C., Chiang, A.W., Shamie, I., Samoudi, M., Gutierrez, J.M., and Lewis, N.E. (2018). The emerging role of systems biology for engineering protein production in CHO cells. Curr. Opin. Biotechnol. *51*, 64–69.

Kwon, S., Kwon, M., Im, S., Lee, K., and Lee, H. (2022). mRNA vaccines: the most recent clinical applications of synthetic mRNA. Arch. Pharm. Res. *45*, 245–262.

Kyrieleis, O.J.P., Chang, J., De La Peña, M., Shuman, S., and Cusack, S. (2014). Crystal structure of vaccinia virus mRNA capping enzyme provides insights into the mechanism and evolution of the capping apparatus. Structure *22*, 452–465.

Laalami, S., Zig, L., and Putzer, H. (2014). Initiation of mRNA decay in bacteria. Cell. Mol. Life Sci. *71*, 1799–1828.

Lalonde, M.E., and Durocher, Y. (2017). Therapeutic glycoprotein production in mammalian cells. J. Biotechnol. *251*, 128–140.

Lam, C., Carver, J., Ng, D., Zhan, D., Tang, D., Kandamkalam, T., Snedecor, B., Barnard, G., Shen, A., and Misaghi, S. (2023). Combining regulated and constitutive protein expression significantly boosts protein expression by increasing productivity without affecting CHO cell growth. Biotechnol. Prog. *39*, e3337.

Lan, L., Guo, M., Ai, Y., Chen, F., Zhang, Y., Xia, L., Huang, D., Niu, L., Zheng, Y., Suzuki, C.K., et al. (2017). Tetramethylpyrazine blocks TFAM degradation and up-regulates mitochondrial DNA copy number by interacting with TFAM. Biosci. Rep. *37*, 20170319.

Langner, E., Walczak, K., Jeleniewicz, W., Turski, W.A., and Rajtar, G. (2015). Quinaldic acid inhibits proliferation of colon cancer HT-29 cells in vitro: Effects on signaling pathways. Eur. J. Pharmacol. *757*, 21–27.

Lavigne, R., Sun, W.D., and Volckaert, G. (2004). PHIRE, a deterministic approach to reveal regulatory elements in bacteriophage genomes. Bioinformatics *20*, 629–635.

Lee, C., Kim, J., Shin, S.G., and Hwang, S. (2006). Absolute and relative QPCR quantification of plasmid copy number in Escherichia coli. J. Biotechnol. *123*, 273–280.

Lee, I.T., Nachbagauer, R., Ensz, D., Schwartz, H., Carmona, L., Schaefers, K., Avanesov, A., Stadlbauer, D., Henry, C., Chen, R., et al. (2023a). Safety and immunogenicity of a phase 1/2 randomized clinical trial of a quadrivalent, mRNA-based seasonal influenza vaccine (mRNA-1010) in healthy adults: interim analysis. Nat. Commun. *14*, 9–11.

Lee, J.-E., Schmidt, H., Lai, B., and Ge, K. (2019). Transcriptional and Epigenomic Regulation of Adipogenesis. Mol. Cell. Biol. *39*.

Lee, J., Woodruff, M.C., Kim, E.H., and Nam, J.H. (2023b). Knife's edge: Balancing immunogenicity and reactogenicity in mRNA vaccines. Exp. Mol. Med. *55*, 1305–1313.

Lee, S.M., Kim, Y.G., Lee, E.G., and Lee, G.M. (2014). Digital mRNA profiling of N-glycosylation gene expression in recombinant Chinese hamster ovary cells treated with sodium butyrate. J. Biotechnol. *171*, 56–60.

Leppek, K., Das, R., and Barna, M. (2018). Functional 5′ UTR mRNA structures in eukaryotic translation regulation and how to find them. Nat. Rev. Mol. Cell Biol. *19*, 158–174.

Leppek, K., Byeon, G.W., Kladwang, W., Wayment-Steele, H.K., Kerr, C.H., Xu, A.F., Kim, D.S., Topkar, V. V., Choe, C., Rothschild, D., et al. (2022). Combinatorial optimization of mRNA structure, stability, and translation for RNA-based therapeutics. Nat. Commun. *13*, 1–22.

Li, J., Crowley, S.T., Duskey, J., Khargharia, S., Wu, M., and Rice, K.G. (2015). Miniaturization of gene transfection assays in 384- and 1536-well microplates. Anal. Biochem. *470*, 14–21.

Li, L., Howell, K., Sands, M., Banahan, M., and Frohlich, S. (2013). The a and D Isoforms of CREB1 Are Required to Maintain Normal Pulmonary Vascular Resistance. PLoS One *8*, 80637.

Li, L., Liu, S., Li, M., Su, L., Ge, K., Li, L., Li, X., and Liu, X. (2016). Negative feedback regulation of ABA biosynthesis in peanut (Arachis hypogaea): A transcription factor complex inhibits AhNCED1 expression during water stress. Sci. Rep. *6*.

Li, X., Monks, B., Ge, Q., and Birnbaum, M.J. (2007). Akt/PKB regulates hepatic metabolism by directly inhibiting PGC-1α transcription coactivator. Nat. 2007 4477147 *447*, 1012–1016.

Li, X., Molina, H., Huang, H., Zhang, Y.Y., Liu, M., Qian, S.W., Slawson, C., Dias,

W.B., Pandey, A., Hart, G.W., et al. (2009). O-linked N-acetylglucosamine modification on CCAAT enhancer-binding protein β. Role during adipocyte differentiation. J. Biol. Chem. *284*, 19248–19254.

Lin-Chao, S., Chen, W. -T, and Wong, T. -T (1992). High copy number of the pUC plasmid results from a Rom/Rop-suppressible point mutation in RNA II. Mol. Microbiol. *6*, 3385–3393.

Lin, L., Cho, S.F., Xing, L., Wen, K., Li, Y., Yu, T., Hsieh, P.A., Chen, H., Kurtoglu, M., Zhang, Y., et al. (2021a). Preclinical evaluation of CD8+ anti-BCMA mRNA CAR T cells for treatment of multiple myeloma. Leukemia *35*, 752–763.

Lin, M.W., Shen, C.C., Lin, Y.J., Chou, M.Y., Pham, N.N., Chang, Y.H., Chang, C.W., Hwu, J.R., Nguyen, M.T.T., and Hu, Y.C. (2021b). Enhancing the yield and activity of defucosylated antibody produced by CHO-K1 cells using Cas13d-mediated multiplex gene targeting. J. Taiwan Inst. Chem. Eng. *121*, 38–47.

Ling, M.L., Risman, S.S., Klement, J.F., Mcgraw, N., and Mcallister, W.T. (1989). Abortive initiation by bacteriophage T3 and T7 RNA polymerases under conditions of limiting substrate. Nucleic Acids Res. *17*, 4430.

Litke, J.L., and Jaffrey, S.R. (2019). Highly efficient expression of circular RNA aptamers in cells using autocatalytic transcripts. Nat. Biotechnol. *37*, 667–675.

Liu, C., Shi, Q., Huang, X., Koo, S., Kong, N., and Tao, W. (2023). mRNA-based cancer therapeutics. Nat. Rev. Cancer *23*, 526–543.

Liu, C.X., Guo, S.K., Nan, F., Xu, Y.F., Yang, L., and Chen, L.L. (2022a). RNA circles with minimized immunogenicity as potent PKR inhibitors. Mol. Cell *82*, 420-434.e6.

Liu, L., Liang, X. hai, Uliel, S., Unger, R., Ullu, E., and Michaeli, S. (2002). RNA interference of signal peptide-binding protein SRP54 elicits deleterious effects and protein sorting defects in trypanosomes. J. Biol. Chem. *277*, 47348–47357.

Liu, X., Zhang, Y., Zhou, S., Dain, L., Mei, L., and Zhu, G. (2022b). Circular RNA: An emerging frontier in RNA therapeutic targets, RNA therapeutics, and mRNA vaccines. J. Control. Release *348*, 84–94.

Liu, Y., Sun, L., Su, X., and Guo, S. (2016). Inhibition of eukaryotic initiation factor 4E phosphorylation by cercosporamide selectively suppresses angiogenesis, growth and survival of human hepatocellular carcinoma. Biomed. Pharmacother. *84*, 237–243.

Liu, Y., Colby, J.K., Zuo, X., Jaoude, J., Wei, D., and Shureiqi, I. (2018). The role of ppar-δ in metabolism, inflammation, and cancer: Many characters of a critical transcription factor. Int. J. Mol. Sci. *19*.

Loh, W.P., Loo, B., Zhou, L., Zhang, P., Lee, D.Y., Yang, Y., and Lam, K.P. (2014). Overexpression of microRNAs enhances recombinant protein production in Chinese hamster ovary cells. Biotechnol. J. *9*, 1140–1151.

Loh, W.P., Yang, Y., and Lam, K.P. (2017). miR-92a enhances recombinant protein productivity in CHO cells by increasing intracellular cholesterol levels. Biotechnol. J. *12*.

Loo, Y.-M., Fornek, J., Crochet, N., Bajwa, G., Perwitasari, O., Martinez-Sobrido, L.,

Akira, S., Gill, M.A., García-Sastre, A., Katze, M.G., et al. (2008). Distinct RIG-I and MDA5 Signaling by RNA Viruses in Innate Immunity. J. Virol. *82*, 335–345.

Loomis, K.H., Lindsay, K.E., Zurla, C., Bhosle, S.M., Vanover, D.A., Blanchard, E.L., Kirschman, J.L., Bellamkonda, R. V, and Santangelo, P.J. (2018). In Vitro Transcribed mRNA Vaccines with Programmable Stimulation of Innate Immunity. Bioconjug. Chem. *29*, 3072–3083.

Lu, R.M., Hwang, Y.C., Liu, I.J., Lee, C.C., Tsai, H.Z., Li, H.J., and Wu, H.C. (2020). Development of therapeutic antibodies for the treatment of diseases. J. Biomed. Sci. *27*, 1–30.

Lu, X., Wu, H., Xia, H., Huang, F., Yan, Y., Yu, B., Cheng, R., Drulis-Kawa, Z., and Zhu, B. (2019). Klebsiella phage KP34 RNA polymerase and its use in RNA synthesis. Front. Microbiol. *10*.

Luebke, T., Schwarz, L., Beer, Y.Y., Schumann, S., Misterek, M., Sander, F.E., Plaza-Sirvent, C., and Schmitz, I. (2019). c-FLIP and CD95 signaling are essential for survival of renal cell carcinoma. Cell Death Dis. *10*, 1–12.

Lyon, S., and Gopalan, V. (2018). A T7 RNA Polymerase Mutant Enhances the Yield of 5′-Thienoguanosine-Initiated RNAs. ChemBioChem *19*, 142–146.

Ma, Z.Z., Zhou, H., Wei, Y.L., Yan, S., and Shen, J. (2020). A novel plasmid–Escherichia coli system produces large batch dsRNAs for insect gene silencing. Pest Manag. Sci. *76*, 2505–2512.

Machida, K., Shigeta, T., Yamamoto, Y., Ito, T., Svitkin, Y., Sonenberg, N., and Imataka, H. (2018). Dynamic interaction of poly(A)-binding protein with the ribosome. Sci. Rep. *8*.

Madabhushi, S.R., Podtelezhnikov, A.A., Murgolo, N., Xu, S., and Lin, H. (2021). Understanding the effect of increased cell specific productivity on galactosylation of monoclonal antibodies produced using Chinese hamster ovary cells. J. Biotechnol. *329*, 92–103.

Madeira, F., Park, Y.M., Lee, J., Buso, N., Gur, T., Madhusoodanan, N., Basutkar, P., Tivey, A.R.N., Potter, S.C., Finn, R.D., et al. (2019). The EMBL-EBI search and sequence analysis tools APIs in 2019. Nucleic Acids Res. *47*, 636–641.

Maeda, T., Tanaka, Y., Takemoto, N., Hamamoto, N., and Inui, M. (2016). RNase III mediated cleavage of the coding region of mraZ mRNA is required for efficient cell division in Corynebacterium glutamicum. Mol. Microbiol. *99*, 1149–1166.

Mairhofer, J., Wittwer, A., Cserjan-Puschmann, M., and Striedner, G. (2015). Preventing T7 RNA polymerase read-through transcription-A synthetic termination signal capable of improving bioprocess stability. ACS Synth. Biol. *4*, 265–273.

Malone, R.W., Felgner, P.L., and Verma, I.M. (1989). Cationic liposome-mediated RNA transfection. Proc. Natl. Acad. Sci. U. S. A. *86*, 6077–6081.

Mardle, C.E., Goddard, L.R., Spelman, B.C., Atkins, H.S., Butt, L.E., Cox, P.A., Gowers, D.M., Vincent, H.A., and Callaghan, A.J. (2020). Identification and analysis of novel small molecule inhibitors of RNase E: Implications for antibacterial targeting and regulation of RNase E. Biochem. Biophys. Reports *23*, 1–9.

Martin, G., and Keller, W. (1998). Tailing and 3'-end labeling of RNA with yeast poly(A) polymerase and various nucleotides. RNA *4*, 226–230.

Martina, J.A., Diab, H.I., Brady, O.A., and Puertollano, R. (2016). TFEB and TFE3 are novel components of the integrated stress response. EMBO J. *35*, 479.

Maruggi, G., Mallett, C.P., Westerbeck, J.W., Chen, T., Lofano, G., Friedrich, K., Qu, L., Sun, J.T., McAuliffe, J., Kanitkar, A., et al. (2022). A self-amplifying mRNA SARS-CoV-2 vaccine candidate induces safe and robust protective immunity in preclinical models. Mol. Ther. *30*, 1897–1912.

Marx, N., Grünwald-Gruber, C., Bydlinski, N., Dhiman, H., Ngoc Nguyen, L., Klanert, G., and Borth, N. (2018). CRISPR-Based Targeted Epigenetic Editing Enables Gene Expression Modulation of the Silenced Beta-Galactoside Alpha-2,6-Sialyltransferase 1 in CHO Cells. Biotechnol. J. *13*, 1700217.

Mary, C., Scherrer, A., Huck, L., Lakkaraju, A.K.K., Thomas, Y., Johnson, A.E., and Strub, K. (2010). Residues in SRP9/14 essential for elongation arrest activity of the signal recognition particle define a positively charged functional domain on one side of the protein. RNA *16*, 969–979.

Mathias, S., Wippermann, A., Raab, N., Zeh, N., Handrick, R., Gorr, I., Schulz, P., Fischer, S., Gamer, M., and Otte, K. (2020). Unraveling what makes a monoclonal antibody difficult-to-express: From intracellular accumulation to incomplete folding and degradation via ERAD. Biotechnol. Bioeng. *117*, 5–16.

Mauger, D.M., Joseph Cabral, B., Presnyak, V., Su, S. V., Reid, D.W., Goodman, B., Link, K., Khatwani, N., Reynders, J., Moore, M.J., et al. (2019). mRNA structure regulates protein expression through changes in functional half-life. Proc. Natl. Acad. Sci. U. S. A. *116*, 24075–24083.

Mauro, V.P. (2018). Codon Optimization in the Production of Recombinant Biotherapeutics: Potential Risks and Considerations. BioDrugs *32*.

McElwain, L., Phair, K., Kealey, C., and Brady, D. (2022). Current trends in biopharmaceuticals production in Escherichia coli. Biotechnol. Lett. *44*, 917–931.

Medina-Magües, L.G., Gergen, J., Jasny, E., Petsch, B., Lopera-Madrid, J., Medina-Magües, E.S., Salas-Quinchucua, C., and Osorio, J.E. (2021). Mrna vaccine protects against zika virus. Vaccines *9*, 1464.

Meganck, R.M., Liu, J., Hale, A.E., Simon, K.E., Fanous, M.M., Vincent, H.A., Wilusz, J.E., Moorman, N.J., Marzluff, W.F., and Asokan, A. (2021). Engineering highly efficient backsplicing and translation of synthetic circRNAs. Mol. Ther. - Nucleic Acids *23*, 821–834.

Meng, J., Lei, J., Davitt, A., Holt, J.R., Huang, J., Gold, R., Vargo, E.L., Tarone, A.M., and Zhu-Salzman, K. (2020). Suppressing tawny crazy ant (Nylanderia fulva) by RNAi technology. Insect Sci. *27*, 113–121.

de Mey, W., Locy, H., De Ridder, K., De Schrijver, P., Autaers, D., Lakdimi, A., Esprit, A., Franceschini, L., Thielemans, K., Verdonck, M., et al. (2023). An mRNA mix redirects dendritic cells towards an antiviral program, inducing anticancer cytotoxic stem cell and central memory CD8+ T cells. Front. Immunol. *14*, 1–12.

Mikac, S., Rychłowski, M., Dziadosz, A., Szabelska-Beresewicz, A., Fahraeus, R.,

Hupp, T., and Sznarkowska, A. (2021). Identification of a stable, non-canonically regulated nrf2 form in lung cancer cells. Antioxidants *10*.

Mikel, P., Vasickova, P., and Kralik, P. (2017). One-plasmid double-expression His-tag system for rapid production and easy purification of MS2 phage-like particles. Sci. Rep. *7*, 1–12.

Miroux, B., and Walker, J.E. (1996). Over-production of proteins in Escherichia coli: Mutant hosts that allow synthesis of some membrane proteins and globular proteins at high levels. J. Mol. Biol. *260*, 289–298.

Mitchell, R.S., Beitzel, B.F., Schroder, A.R.W., Shinn, P., Chen, H., Berry, C.C., Ecker, J.R., and Bushman, F.D. (2004). Retroviral DNA integration: ASLV, HIV, and MLV show distinct target site preferences. PLoS Biol. *2*.

Mohanty, B.K., and Kushner, S.R. (2019). New Insights into the Relationship between tRNA Processing and Polyadenylation in Escherichia coli. Trends Genet. *35*, 434–445.

Mohanty, B.K., and Kushner, S.R. (2022). Regulation of mRNA decay in E. coli. Crit. Rev. Biochem. Mol. Biol. *57*, 48–72.

Moore, C.L., Papa, L.J., and Shoulders, M.D. (2018). A Processive Protein Chimera Introduces Mutations across Defined DNA Regions in Vivo. J. Am. Chem. Soc. *140*, 11560–11564.

Moorefield, B., Greene, E.A., and Reeder, R.H. (2000). RNA polymerase I transcription factor Rrn3 is functionally conserved between yeast and human. Proc. Natl. Acad. Sci. U. S. A. *97*, 4724–4729.

Morais, P., Adachi, H., and Yu, Y.T. (2021). The Critical Contribution of Pseudouridine to mRNA COVID-19 Vaccines. Front. Cell Dev. Biol. *9*, 1–9.

Morita, T., Mayanagi, T., and Sobue, K. (2007). Reorganization of the actin cytoskeleton via transcriptional regulation of cytoskeletal/focal adhesion genes by myocardin-related transcription factors (MRTFs/MAL/MKLs). Exp. Cell Res. *313*, 3432–3445.

Mu, X., Greenwald, E., Ahmad, S., and Hur, S. (2018). An origin of the immunogenicity of in vitro transcribed RNA. Nucleic Acids Res. *46*, 5239–5249.

Mueller, E. (2014). Understanding the variegation of fat: Novel regulators of adipocyte differentiation and fat tissue biology. Biochim. Biophys. Acta - Mol. Basis Dis. *1842*, 352–357.

Muthukrishnan, S., Both, G.W., Furuichi, Y., and Shatkin, A.J. (1975). 5′-Terminal 7-methylguanosine in eukaryotic mRNA is required for translation. Nature *255*, 33–37.

Nacheva, G.A., and Berzal-Herranz, A. (2003). Preventing nondesired RNA-primed RNA extension catalyzed by T7 RNA polymerase. Eur. J. Biochem. *270*, 1458–1465.

Nagao, N., Suzuki, H., Numano, R., Umekage, S., and Kikuchi, Y. (2015). Short hairpin rnas of designed sequences can be extracellularly produced by the marine bacterium rhodovulum sulfidophilum. J. Gen. Appl. Microbiol. *60*, 222–226.

Nagelreiter, F., Coats, M.T., Klanert, G., Gludovacz, E., Borth, N., Grillari, J., and

Schosserer, M. (2018). OPP Labeling Enables Total Protein Synthesis Quantification in CHO Production Cell Lines at the Single-Cell Level. Biotechnol. J. *13*.

Nam, S.C., and Kang, C. (1988). Transcription initiation site selection and abortive. J. Biol. Chem. *263*, 18123–18127.

Nance, K.D., and Meier, J.L. (2021). Modifications in an Emergency: The Role of N1-Methylpseudouridine in COVID-19 Vaccines. ACS Cent. Sci. *7*, 748–756.

Nelissen, F.H.T., Leunissen, E.H.P., Van De Laar, L., Tessari, M., Heus, H.A., and Wijmenga, S.S. (2012). Fast production of homogeneous recombinant RNA-towards large-scale production of RNA. Nucleic Acids Res. *40*, 1–12.

Nelissen, F.H.T., Heus, H.A., and Wijmenga, S.S. (2015). Production of homogeneous recombinant RNA using a tRNA scaffold and hammerhead ribozymes. Methods Mol. Biol. *1316*, 33–44.

Nelson, J., Sorensen, E.W., Mintri, S., Rabideau, A.E., Zheng, W., Besin, G., Khatwani, N., Su, S. V., Miracco, E.J., Issa, W.J., et al. (2020). Impact of mRNA chemistry and manufacturing process on innate immune activation. Sci. Adv. *6*, 6893.

Niu, D., Wu, Y., and Lian, J. (2023). Circular RNA vaccine in disease prevention and treatment. Signal Transduct. Target. Ther. *8*.

Novoa, E.M., and Ribas de Pouplana, L. (2012). Speeding with control: Codon usage, tRNAs, and ribosomes. Trends Genet. *28*, 574–581.

Nwokeoji, A.O., Kilby, P.M., Portwood, D.E., and Dickman, M.J. (2016). RNASwift: A rapid, versatile RNA extraction method free from phenol and chloroform. Anal. Biochem. *512*, 36–46.

Obi, P., and Chen, Y.G. (2021). The design and synthesis of circular RNAs. Methods.

Obika, S., and Sekine, M. (2018). Synthesis of therapeutic oligonucleotides. Synth. Ther. Oligonucleotides 1–284.

Ohlson, J. (2020). Plasmid manufacture is the bottleneck of the genetic medicine revolution. Drug Discov. Today *25*, 1891.

Okada, T., Haze, K., Nadanaka, S., Yoshida, H., Seidah, N.G., Hirano, Y., Sato, R., Negishi, M., and Mori, K. (2003). A Serine Protease Inhibitor Prevents Endoplasmic Reticulum Stress-induced Cleavage but Not Transport of the Membrane-bound Transcription Factor ATF6. J. Biol. Chem. *278*, 31024–31032.

Onishi, M., Nosaka, T., Misawa, K., Mui, A.L.-F., Gorman, D., McMahon, M., Miyajima, A., and Kitamura, T. (1998). Identification and Characterization of a Constitutively Active STAT5 Mutant That Promotes Cell Proliferation. Mol. Cell. Biol. *18*, 3871–3879.

Opalinska, J.B., and Gewirtz, A.M. (2002). Nucleic-acid therapeutics: Basic principles and recent applications. Nat. Rev. Drug Discov. *1*, 503–514.

Orlandini von Niessen, A.G., Poleganov, M.A., Rechner, C., Plaschke, A., Kranz, L.M., Fesser, S., Diken, M., Löwer, M., Vallazza, B., Beissert, T., et al. (2019). Improving mRNA-Based Therapeutic Gene Delivery by Expression-Augmenting 3′

UTRs Identified by Cellular Library Screening. Mol. Ther. *27*, 824–836.

Ouranidis, A., Vavilis, T., Mandala, E., Davidopoulou, C., Stamoula, E., Markopoulou, C.K., Karagianni, A., and Kachrimanis, K. (2022). Manufacturing under Pharma 4 . 0 Principles. 1–31.

Papamichail, D., Liu, H., MacHado, V., Gould, N., Robert Coleman, J., and Papamichail, G. (2018). Codon Context Optimization in Synthetic Gene Design. IEEE/ACM Trans. Comput. Biol. Bioinforma. *15*, 452–459.

Papić, L., Rivas, J., Toledo, S., and Romero, J. (2018). Double-stranded RNA production and the kinetics of recombinant Escherichia coli HT115 in fed-batch culture. Biotechnol. Reports *20*, e00292.

Pardi, N., Hogan, M.J., Porter, F.W., and Weissman, D. (2018). mRNA vaccines-a new era in vaccinology. Nat. Rev. Drug Discov. *17*, 261–279.

Park, J.H., Noh, S.M., Woo, J.R., Kim, J.W., and Lee, G.M. (2016). Valeric acid induces cell cycle arrest at G1 phase in CHO cell cultures and improves recombinant antibody productivity. Biotechnol. J. *11*, 487–496.

Park, S.M., Kang, T. Il, and So, J.S. (2021). Roles of XBP1s in Transcriptional Regulation of Target Genes. Biomedicines *9*.

Pelletier, J., Thomas, G., and Volarevi, S. (2017). Ribosome biogenesis in cancer: new players and therapeutic avenues. Nat. Rev. Cancer 2017 181 *18*, 51–63.

Perez, C.R., and De Palma, M. (2019). Engineering dendritic cell vaccines to improve cancer immunotherapy. Nat. Commun. *10*, 1–10.

Pertzev, A. V., and Nicholson, A.W. (2006). Characterization of RNA sequence determinants and antideterminants of processing reactivity for a minimal substrate of Escherichia coli ribonuclease III. Nucleic Acids Res. *34*, 3708–3721.

Pesole, G., Mignone, F., Gissi, C., Grillo, G., Licciulli, F., and Liuni, S. (2001). Structural and functional features of eukaryotic mRNA untranslated regions. In Gene, pp. 73–81.

Petkovic, S., and Müller, S. (2015). RNA circularization strategies in vivo and in vitro. Nucleic Acids Res. *43*, 2454–2465.

Piper, K., DePledge, L., Karsy, M., and Cobbs, C. (2021). Glioma Stem Cells as Immunotherapeutic Targets: Advancements and Challenges. Front. Oncol. *11*, 1–13.

Plank, T.D.M., Whitehurst, J.T., and Kieft, J.S. (2013). Cell type specificity and structural determinants of IRES activity from the 5′ leaders of different HIV-1 transcripts. Nucleic Acids Res. *41*, 6698–6714.

Ponchon, L., and Dardel, F. (2011). Large scale expression and purification of recombinant RNA in Escherichia coli. Methods *54*, 267–273.

Ponchon, L., Beauvais, G., Nonin-Lecomte, S., and Dardel, F. (2009a). A generic protocol for the expression and purification of recombinant RNA in Escherichia coli using a tRNA scaffold. Nat. Protoc. *4*, 947–959.

Ponchon, L., Beauvais, G., Nonin-Lecomte, S., and Dardel, F. (2009b). A generic protocol for the expression and purification of recombinant RNA in Escherichia coli

using a tRNA scaffold. Nat. Protoc. *4*, 947–959.

Ponchon, L., Catala, M., Seijo, B., El Khouri, M., Dardel, F., Nonin-Lecomte, S., and Tisné, C. (2013). Co-expression of RNA-protein complexes in Escherichia coli and applications to RNA biology. Nucleic Acids Res. *41*, e150.

Pontrelli, S., Chiu, T.Y., Lan, E.I., Chen, F.Y.H., Chang, P., and Liao, J.C. (2018). Escherichia coli as a host for metabolic engineering. Metab. Eng. *50*, 16–46.

Pourseif, M.M., Masoudi-Sobhanzadeh, Y., Azari, E., Parvizpour, S., Barar, J., Ansari, R., and Omidi, Y. (2022). Self-amplifying mRNA vaccines: Mode of action, design, development and optimization. Drug Discov. Today *27*, 1–19.

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

Qin, C., Xiang, Y., Liu, J., Zhang, R., Liu, Z., Li, T., Sun, Z., Ouyang, X., Zong, Y., Zhang, H.M., et al. (2023). Precise programming of multigene expression stoichiometry in mammalian cells by a modular and programmable transcriptional system. Nat. Commun. 2023 141 *14*, 1–10.

Qin, S., Tang, X., Chen, Y., Chen, K., Fan, N., Xiao, W., Zheng, Q., Li, G., Teng, Y., Wu, M., et al. (2022). mRNA-based therapeutics: powerful and versatile tools to combat diseases. Signal Transduct. Target. Ther. *7*, 1–35.

Qu, L., Yi, Z., Shen, Y., Lin, L., Chen, F., Xu, Y., Wu, Z., Tang, H., Zhang, X., Tian, F., et al. (2022). Circular RNA vaccines against SARS-CoV-2 and emerging variants. Cell *185*, 1728–1744.

Radoš, D., Donati, S., Lempp, M., Rapp, J., and Link, H. (2022). Homeostasis of the biosynthetic E. coli metabolome. IScience *25*, 1–12.

Ramachandran, A., Basu, U., Sultana, S., Nandakumar, D., and Patel, S.S. (2017). Human mitochondrial transcription factors TFAM and TFB2M work synergistically in promoter melting during transcription initiation. Nucleic Acids Res. *45*, 861–874.

Ramamoorth, M., and Narvekar, A. (2015). Non viral vectors in gene therapy - An overview. J. Clin. Diagnostic Res. *9*, GE01–GE06.

Ramírez-Tapia, L.E., and Martin, C.T. (2012). New insights into the mechanism of initial transcription: The T7 RNA polymerase mutant p266l transitions to elongation at longer RNA lengths than wild type. J. Biol. Chem. *287*, 37352–37361.

Ranjan, A., Pang, Y., Butler, M., Merchant, M., Kim, O., Yu, G., Su, Y.T., Gilbert, M.R., Levens, D., and Wu, J. (2021). Targeting CDK9 for the Treatment of Glioblastoma. Cancers 2021, Vol. 13, Page 3039 *13*, 3039.

Rausch, J.W., Heinz, W.F., Payea, M.J., Sherpa, C., Gorospe, M., and Le Grice, S.F.J. (2021). Characterizing and circumventing sequence restrictions for synthesis of circular RNA in vitro. Nucleic Acids Res. *49*, E35–E35.

Reese, C.B. (2005). Oligo- and poly-nucleotides: 50 years of chemical synthesis. Org. Biomol. Chem. *3*, 3851–3868.

Reinhart, D., Sommeregger, W., Debreczeny, M., Gludovacz, E., and Kunert, R.

(2013). Characterization of recombinant IgA producing CHO cell lines by qPCR. BMC Proc. *7*, P114.

Reinhart, D., Sommeregger, W., Debreczeny, M., Gludovacz, E., and Kunert, R. (2014). In search of expression bottlenecks in recombinant CHO cell lines - A case study. Appl. Microbiol. Biotechnol. *98*, 5959–5965.

Reisinger, H., Steinfellner, W., Stern, B., Katinger, H., and Kunert, R. (2008). The absence of effect of gene copy number and mRNA level on the amount of mAb secretion from mammalian cells. Appl. Microbiol. Biotechnol. *81*, 701–710.

Reusch, J.E.B., Colton, L.A., and Klemm, D.J. (2000). CREB Activation Induces Adipogenesis in 3T3-L1 Cells. Mol. Cell. Biol. *20*, 1008–1020.

Rhoads, R.E. (2016). Synthetic mRNA: Production, introduction into cells, and physiological consequences. In Methods in Molecular Biology, pp. 3–27.

Richards, J., and Belasco, J.G. (2023). Graded impact of obstacle size on scanning by RNase E. Nucleic Acids Res. *51*, 1364–1374.

Ritacco, F. V., Wu, Y., and Khetan, A. (2018). Cell culture media for recombinant protein expression in Chinese hamster ovary (CHO) cells: History, key components, and optimization strategies. Biotechnol. Prog. *34*, 1407–1426.

Rong, M., Biao, H., Mcallister, W.T., and Durbin, R.K. (1998). Promoter specificity determinants of T7 RNA polymerase. Proc. Natl. Acad. Sci. U. S. A. *95*, 515–519.

Rosa, S.S., Prazeres, D.M.F., Azevedo, A.M., and Marques, M.P.C. (2021). mRNA vaccines manufacturing: Challenges and bottlenecks. Vaccine *39*, 2190–2200.

Roseman, D.S., Khan, T., Rajas, F., Jun, L.S., Asrani, K.H., Isaacs, C., Farelli, J.D., and Subramanian, R.R. (2018). G6PC mRNA Therapy Positively Regulates Fasting Blood Glucose and Decreases Liver Abnormalities in a Mouse Model of Glycogen Storage Disease 1a. Mol. Ther. *26*, 814–821.

Rosen, E.D., Hsu, C.H., Wang, X., Sakai, S., Freeman, M.W., Gonzalez, F.J., and Spiegelman, B.M. (2002). C/EBPα induces adipogenesis through PPARγ: A unified pathway. Genes Dev. *16*, 22–26.

Ross, J., and Sullivan, T.D. (1985). Half-lives of beta and gamma globin messenger RNAs and of protein synthetic capacity in cultured human reticulocytes. Blood *66*, 1149–1154.

Rostain, W., Shen, S., Cordero, T., Rodrigo, G., and Jaramillo, A. (2020). Engineering a Circular Riboregulator in Escherichia coli. BioDesign Res. *2020*, 1–9.

Rouches, M. V., Xu, Y., Cortes, L.B.G., and Lambert, G. (2022). A plasmid system with tunable copy number. Nat. Commun. *13*, 1–12.

Roux, C., Etienne, T.A., Hajnsdorf, E., Ropers, D., Carpousis, A.J., Cocaign-Bousquet, M., and Girbal, L. (2022). The essential role of mRNA degradation in understanding and engineering E. coli metabolism. Biotechnol. Adv. *54*, 107805.

Rowe, S.M., Zuckerman, J.B., Dorgan, D., Lascano, J., McCoy, K., Jain, M., Schechter, M.S., Lommatzsch, S., Indihar, V., Lechtzin, N., et al. (2023). Inhaled mRNA therapy for treatment of cystic fibrosis: Interim results of a randomized,

double-blind, placebo-controlled phase 1/2 clinical study. J. Cyst. Fibros. *22*, 656–664.

Rupp, J.C., Sokoloski, K.J., Gebhart, N.N., and Hardy, R.W. (2015). Alphavirus RNA synthesis and non-structural protein functions. J. Gen. Virol. *96*, 2483–2500.

Ryczek, M., Pluta, M., Błaszczyk, L., and Kiliszek, A. (2022). Overview of Methods for Large-Scale RNA Synthesis. Appl. Sci. *12*.

Sahin, U., Karikó, K., and Türeci, Ö. (2014). MRNA-based therapeutics-developing a new class of drugs. Nat. Rev. Drug Discov. *13*, 759–780.

Sahin, U., Oehm, P., Derhovanessian, E., Jabulowsky, R.A., Vormehr, M., Gold, M., Maurus, D., Schwarck-Kokarakis, D., Kuhn, A.N., Omokoko, T., et al. (2020). An RNA vaccine drives immunity in checkpoint-inhibitor-treated melanoma. Nature *585*, 107–112.

Salzman, J. (2016). Circular RNA Expression: Its Potential Regulation and Function. Trends Genet. *32*, 309–316.

Sampaio, M., Rocha, M., Oliveira, H., DIas, O., and Valencia, A. (2019). Predicting promoters in phage genomes using PhagePromoter. Bioinformatics *35*, 5301–5302.

Sarcar, S.N., and Miller, D.L. (2018). A specific, promoter-independent activity of T7 RNA polymerase suggests a general model for DNA/RNA editing in single subunit RNA Polymerases. Sci. Rep. *8*, 13885.

Scarrott, J.M., Johari, Y.B., Pohle, T.H., Liu, P., Mayer, A., and James, D.C. (2023). Increased recombinant adeno-associated virus production by HEK293 cells using small molecule chemical additives. Biotechnol. J. *18*, 2200450.

Schmidt, S., Gay, D., Uthe, F.W., Denk, S., Paauwe, M., Matthes, N., Diefenbacher, M.E., Bryson, S., Warrander, F.C., Erhard, F., et al. (2019). A MYC–GCN2–eIF2α negative feedback loop limits protein synthesis to prevent MYC-dependent apoptosis in colorectal cancer. Nat. Cell Biol. *21*, 1413–1424.

Seephetdee, C., Bhukhai, K., Buasri, N., Leelukkanaveera, P., Lerdwattanasombat, P., Manopwisedjaroen, S., Phueakphud, N., Kuhaudomlarp, S., Olmedillas, E., Ollmann Saphire, E., et al. (2022). A circular mRNA vaccine prototype producing VFLIP-X spike confers a broad neutralization of SARS-CoV-2 variants by mouse sera. Antiviral Res. *204*, 105370.

Seibel, R., Maier, S., Schnellbaecher, A., Bohl, S., Wehsling, M., Zeck, A., and Zimmer, A. (2017). Impact of S-sulfocysteine on fragments and trisulfide bond linkages in monoclonal antibodies. MAbs *9*, 889–897.

Senthilvelan, A., Vonderfecht, T., Shanmugasundaram, M., Potter, J., and Kore, A.R. (2023). Click-iT trinucleotide cap analog: Synthesis, mRNA translation, and detection. Bioorg. Med. Chem *77*, 117128.

Shanmugasundaram, M., Senthilvelan, A., and Kore, A.R. (2022). Recent Advances in Modified Cap Analogs: Synthesis, Biochemical Properties, and mRNA Based Vaccines. Chem. Rec. *22*.

Shao, W., and Espenshade, P.J. (2014). Sterol regulatory element-binding protein (SREBP) cleavage regulates golgi-to-endoplasmic reticulum recycling of SREBP

cleavage-activating protein (SCAP). J. Biol. Chem. *289*, 7547–7557.

Sharker, S.M., and Rahman, A. (2020). A Review on the Current Methods of Chinese Hamster Ovary (CHO) Cells Cultivation for the Production of Therapeutic Protein. Curr. Drug Discov. Technol. *18*, 354–364.

Shauchuk, A., Szulc, B., Maszczak-Seneczko, D., Wiertelak, W., Skurska, E., and Olczak, M. (2020). N-glycosylation of the human β1,4-galactosyltransferase 4 is crucial for its activity and Golgi localization. Glycoconj. J. *37*, 577–588.

Shen, C.C., Lin, M.W., Nguyen, B.K.T., Chang, C.W., Shih, J.R., Nguyen, M.T.T., Chang, Y.H., and Hu, Y.C. (2020). CRISPR-Cas13d for Gene Knockdown and Engineering of CHO Cells. ACS Synth. Biol. *9*, 2808–2818.

Shiloach, J., and Fass, R. (2005). Growing E. coli to high cell density - A historical perspective on method development. Biotechnol. Adv. *23*, 345–357.

Shis, D.L., and Bennett, M.R. (2014). Synthetic biology: the many facets of T7 RNA polymerase. Mol. Syst. Biol. *10*, 745.

Shujaat, M., Jin, J.S., Tayara, H., and Chong, K.T. (2022). iProm-phage: A two-layer model to identify phage promoters and their types using a convolutional neural network. Front. Microbiol. *13*.

Solsona-Vilarrasa, E., Fucho, R., Torres, S., Nuñez, S., Nuño-Lámbarri, N., Enrich, C., García-Ruiz, C., and Fernández-Checa, J.C. (2019). Cholesterol enrichment in liver mitochondria impairs oxidative phosphorylation and disrupts the assembly of respiratory supercomplexes. Redox Biol. *24*, 101214.

Sparman, N., Sadeh, S., Wang, J., Whitelegge, J.P., Fried, S.K., Waki, H., Villanueva, C.J., Seldin, M., Sakaguchi, S., Ellmeier, W., et al. (2022). Transcriptional control of adipogenesis by PATZ1. BioRxiv.

Sripada, S.A., Chu, W., Williams, T.I., Teten, M.A., Mosley, B.J., Carbonell, R.G., Lenhoff, A.M., Cramer, S.M., Bill, J., Yigzaw, Y., et al. (2022). Towards continuous mAb purification: Clearance of host cell proteins from CHO cell culture harvests via “flow-through affinity chromatography” using peptide-based adsorbents. Biotechnol. Bioeng. *119*, 1873–1889.

Stadelmann, C., Di Francescantonio, S., Marg, A., Müthel, S., Spuler, S., and Escobar, H. (2022). mRNA-mediated delivery of gene editing tools to human primary muscle stem cells. Mol. Ther. - Nucleic Acids *28*, 47–57.

Steinegger, M., and Söding, J. (2017). MMseqs2 enables sensitive protein sequence searching for the analysis of massive data sets. Nat. Biotechnol. *35*, 1026–1028.

Stepinski, J., Waddell, C., Stolarski, R., Darzynkiewicz, E., and Rhoads, R.E. (2001). Synthesis and properties of mRNAs containing the novel “anti-reverse” cap analogs 7-methyl(39-O-methyl)GpppG and 7-methyl(39-deoxy)GpppG.

Steven, A., Friedrich, M., Jank, P., Heimer, N., Budczies, J., Denkert, C., and Seliger, B. (2020). What turns CREB on? And off? And why does it matter? Cell. Mol. Life Sci. *77*, 4049–4067.

Stolfa, G., Smonskey, M.T., Boniface, R., Hachmann, A.B., Gulde, P., Joshi, A.D., Pierce, A.P., Jacobia, S.J., and Campbell, A. (2018). CHO-Omics Review: The

Impact of Current and Emerging Technologies on Chinese Hamster Ovary Based Bioproduction. Biotechnol. J. *13*, 1700227.

Su, W., Slepenkov, S., Grudzien-Nogalska, E., Kowalska, J., Kulis, M., Zuberek, J., Lukaszewicz, M., Darzynkiewicz, E., Jemielity, J., and Rhoads, R.E. (2011). Translation, stability, and resistance to decapping of mRNAs containing caps substituted in the triphosphate chain with BH3, Se, and NH. RNA *17*, 978–988.

Sun, H., Wang, S., Lu, M., Tinberg, C.E., and Alba, B.M. (2023). Protein production from HEK293 cell linederived stable pools with high protein quality and quantity to support discovery research. PLoS One *18*.

Svitkin, Y. V., Cheng, Y.M., Chakraborty, T., Presnyak, V., John, M., and Sonenberg, N. (2017). N1-methyl-pseudouridine in mRNA enhances translation through eIF2α-dependent and independent mechanisms by increasing ribosome density. Nucleic Acids Res. *45*, 6023–6036.

Swiech, K., Picanço-Castro, V., and Covas, D.T. (2012). Human cells: New platform for recombinant therapeutic protein production. Protein Expr. Purif. *84*, 147–153.

Takahashi, K., and Yamanaka, S. (2006). Induction of Pluripotent Stem Cells from Mouse Embryonic and Adult Fibroblast Cultures by Defined Factors. Cell *126*, 663–676.

Takeda, . K, Kaisho, T., Akira, S., and Ozinsky, ; A (2003). Species-Specific Recognition of Single-Stranded RNA via Toll-like Receptor 7 and 8.

Takiff, H.E., Chen, S.M., and Court, D.L. (1989). Genetic analysis of the rnc operon of Escherichia coli. J. Bacteriol. *171*, 2581–2590.

Tan, E., Chin, C.S.H., Lim, Z.F.S., and Ng, S.K. (2021). HEK293 Cell Line as a Platform to Produce Recombinant Proteins and Viral Vectors. Front. Bioeng. Biotechnol. *9*, 796991.

Tan, K., Hu, Y., Liang, Z., Li, C.Y., Yau, W.L., and Kuang, Y. (2023). Dual Input-Controlled Synthetic mRNA Circuit for Bidirectional Protein Expression Regulation. ACS Synth. Biol. *12*, 2516–2523.

Tang, J.J., Li, J.G., Qi, W., Qiu, W.W., Li, P.S., Li, B.L., and Song, B.L. (2011). Inhibition of SREBP by a Small Molecule, Betulin, Improves Hyperlipidemia and Insulin Resistance and Reduces Atherosclerotic Plaques. Cell Metab. *13*, 44–56.

Tavernier, G., Andries, O., Demeester, J., Sanders, N.N., De Smedt, S.C., and Rejman, J. (2011). MRNA as gene therapeutic: How to control protein expression. J. Control. Release *150*, 238–247.

Taylor, D.R., and Mathews, M.B. (1993). Transcription by SP6 RNA polymerase exhibits an ATP dependence that is influenced by promoter topology.

Tenchov, R., Bird, R., Curtze, A.E., and Zhou, Q. (2021). Lipid Nanoparticles from Liposomes to mRNA Vaccine Delivery, a Landscape of Research Diversity and Advancement. ACS Nano *15*, 16982–17015.

Terpe, K. (2006). Overview of bacterial expression systems for heterologous protein production: From molecular and biochemical fundamentals to commercial systems. Appl. Microbiol. Biotechnol. *72*, 211–222.

Tihanyi, B., and Nyitray, L. (2020). Recent advances in CHO cell line development for recombinant protein production. Drug Discov. Today Technol. *38*, 25–34.

Trepotec, Z., Aneja, M.K., Geiger, J., Hasenpusch, G., Plank, C., and Rudolph, C. (2019a). Maximizing the Translational Yield of mRNA Therapeutics by Minimizing 5′-UTRs. Tissue Eng. - Part A *25*, 69–79.

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

Vavilis, T., Stamoula, E., Ainatzoglou, A., Sachinidis, A., Lamprinou, M., Dardalas, I., and Vizirianakis, I.S. (2023a). mRNA in the Context of Protein Replacement Therapy. Pharmaceutics *15*, 1–19.

Vavilis, T., Stamoula, E., Ainatzoglou, A., Sachinidis, A., Lamprinou, M., Dardalas, I., and Vizirianakis, I.S. (2023b). mRNA in the Context of Protein Replacement Therapy. Pharmaceutics *15*.

Viegas, S.C., Apura, P., Martínez-García, E., De Lorenzo, V., and Arraiano, C.M. (2018). Modulating Heterologous Gene Expression with Portable mRNA-Stabilizing 5′-UTR Sequences. ACS Synth. Biol. *7*, 2177–2188.

Villanueva, C.J., Waki, H., Godio, C., Nielsen, R., Chou, W.L., Vargas, L., Wroblewski, K., Schmedt, C., Chao, L.C., Boyadjian, R., et al. (2011). TLE3 Is a Dual-Function Transcriptional Coregulator of Adipogenesis. Cell Metab. *13*, 413–427.

Vogel, A.B., Lambert, L., Kinnear, E., Busse, D., Erbar, S., Reuter, K.C., Wicke, L., Perkovic, M., Beissert, T., Haas, H., et al. (2018). Self-Amplifying RNA Vaccines Give Equivalent Protection against Influenza to mRNA Vaccines but at Much Lower Doses. Mol. Ther. *26*, 446–455.

Voit, R., and Grummt, I. (2001). Phosphorylation of UBF at serine 388 is required for interaction with RNA polymerase I and activation of rDNA transcription. Proc. Natl. Acad. Sci. U. S. A. *98*, 13631–13636.

Walker, S.C., Avis, J.M., and Conn, G.L. (2003). General plasmids for producing RNA in vitro transcripts with homogeneous ends. Nucleic Acids Res. *31*, 1–6.

Walls, A.C., Park, Y.J., Tortorici, M.A., Wall, A., McGuire, A.T., and Veesler, D. (2020). Structure, Function, and Antigenicity of the SARS-CoV-2 Spike Glycoprotein. Cell *181*, 281-292.e6.

Walsh, G. (2018). Biopharmaceutical benchmarks 2018. Nat. Biotechnol. *36*, 1136–1145.

Wang, G., Cheng, R., Chen, Q., Xu, Y., Yu, B., Zhu, B., Yin, H., and Xia, H. (2022). mRNA produced by VSW-3 RNAP has high-level translation efficiency with low inflammatory stimulation. Cell Insight *1*, 100056.

Wang, W., Li, Y., Wang, Y., Shi, C., Li, C., Li, Q., and Linhardt, R.J. (2018). Bacteriophage T7 transcription system: an enabling tool in synthetic biology. Biotechnol. Adv. *36*, 2129–2137.

Warren, L., Manos, P.D., Ahfeldt, T., Loh, Y.H., Li, H., Lau, F., Ebina, W., Mandal, P.K., Smith, Z.D., Meissner, A., et al. (2010). Highly efficient reprogramming to

pluripotency and directed differentiation of human cells with synthetic modified mRNA. Cell Stem Cell *7*, 618–630.

Weaver, Z.A., and Kane, C.M. (1997). Genomic Characterization of a Testis-Specific TFIIS (TCEA2) Gene. Genomics *46*, 516–519.

Webb, C., Ip, S., Bathula, N. V., Popova, P., Soriano, S.K.V., Ly, H.H., Eryilmaz, B., Nguyen Huu, V.A., Broadhead, R., Rabel, M., et al. (2022). Current Status and Future Perspectives on MRNA Drug Manufacturing. Mol. Pharm. *19*, 1047–1058.

Wei, S., Zhang, M., Zheng, Y., and Yan, P. (2018). ZBTB16 Overexpression Enhances White Adipogenesis and Induces Brown-Like Adipocyte Formation of Bovine White Intramuscular Preadipocytes. Cell. Physiol. Biochem. *48*, 2528–2538.

Wei, Y., Silke, J.R., and Xia, X. (2019). An improved estimation of tRNA expression to better elucidate the coevolution between tRNA abundance and codon usage in bacteria. Sci. Rep. *9*.

Weill, L., Belloc, E., Bava, F.-A., and Méndez, R. (2012). Translational control by changes in poly(A) tail length: recycling mRNAs.

Wesselhoeft, R.A., Kowalski, P.S., and Anderson, D.G. (2018). Engineering circular RNA for potent and stable translation in eukaryotic cells. Nat. Commun. *9*, 1–10.

Wesselhoeft, R.A., Kowalski, P.S., Parker-Hale, F.C., Huang, Y., Bisaria, N., and Anderson, D.G. (2019). RNA Circularization Diminishes Immunogenicity and Can Extend Translation Duration In Vivo. Mol. Cell *74*, 508-520.e4.

Whitley, J., Zwolinski, C., Denis, C., Maughan, M., Hayles, L., Clarke, D., Snare, M., Liao, H., Chiou, S., Marmura, T., et al. (2022). Development of mRNA manufacturing for vaccines and therapeutics: mRNA platform requirements and development of a scalable production process to support early phase clinical trials. Transl. Res. *242*, 38–55.

Williams, D.J., Puhl, H.L., and Ikeda, S.R. (2010). A simple, highly efficient method for heterologous expression in mammalian primary neurons using cationic lipid-mediated mRNA transfection. Front. Neurosci. *4*, 1–20.

Williams, L.M., Sarma, U., Willets, K., Smallie, T., Brennan, F., and Foxwell, B.M.J. (2007). Expression of constitutively active STAT3 can replicate the cytokine-suppressive activity of interleukin-10 in human primary macrophages. J. Biol. Chem. *282*, 6965–6975.

Wolff, J.A., Malone, R.W., Williams, P., Chong, W., Acsadi, G., Jani, A., and Felgner, P.L. (1990). Direct gene transfer into mouse muscle in vivo. Science (80-.). *247*, 1465–1468.

Woo, J., Kim, J.H., Kim, S., and Park, K.S. (2021). Promoter engineering improves transcription efficiency in biomolecular assays †.

Wortel, I.M.N., van der Meer, L.T., Kilberg, M.S., and van Leeuwen, F.N. (2017). Surviving Stress: Modulation of ATF4-Mediated Stress Responses in Normal and Malignant Cells. Trends Endocrinol. Metab. *28*, 794.

Wu, M.Z., Asahara, H., Tzertzinis, G., and Roy, B. (2020). Synthesis of low immunogenicity RNA with high-temperature in vitro transcription. RNA *26*, 345–360.

Wurm, J.P., Griese, M., Bahr, U., Held, M., Heckel, A., Karas, M., Soppa, J., and Wöhnert, J. (2012). Identification of the enzyme responsible for N1-methylation of pseudouridine 54 in archaeal tRNAs. Rna *18*, 412–420.

Xia, X. (2021). Detailed dissection and critical evaluation of the pfizer/biontech and moderna mrna vaccines. Vaccines *9*.

Xia, H., Yu, B., Jiang, Y., Cheng, R., Lu, X., Wu, H., and Zhu, B. (2022). Psychrophilic phage VSW-3 RNA polymerase reduces both terminal and full-length dsRNA byproducts in in vitro transcription. RNA Biol. *19*, 1130–1142.

Xu, W.J., Lin, Y., Mi, C.L., Pang, J.Y., and Wang, T.Y. (2023). Progress in fed-batch culture for recombinant protein production in CHO cells. Appl. Microbiol. Biotechnol. *107*, 1063–1075.

Xue, L., Qi, H., Zhang, H., Ding, L., Huang, Q., Zhao, D., Wu, B.J., and Li, X. (2020). Targeting SREBP-2-Regulated Mevalonate Metabolism for Cancer Therapy. Front. Oncol. *10*, 1–20.

Yamamoto, A., Kormann, M., Rosenecker, J., and Rudolph, C. (2009). Current prospects for mRNA gene delivery. Eur. J. Pharm. Biopharm. *71*, 484–489.

Yan, Q., Bartz, S., Mao, M., Li, L., and Kaelin, W.G. (2007). The Hypoxia-Inducible Factor 2 N-Terminal and C-Terminal Transactivation Domains Cooperate To Promote Renal Tumorigenesis In Vivo †. Mol. Cell. Biol. *27*, 3.

Yang, D., Pricilia, C., Prabowo, S., Eun, H., Park, S.Y., Cho, I.J., Jiao, S., and Lee, S.Y. (2021). Escherichia coli as a platform microbial host for systems metabolic engineering. Essays Biochem. *65*, 225–246.

Yang, Y., Tang, H., Zheng, J., and Yang, K. (2022). The PER1/HIF-1alpha negative feedback loop promotes ferroptosis and inhibits tumor progression in oral squamous cell carcinoma. Transl. Oncol. *18*.

Yarborough, M., and Sharp, R.R. (2009). Public trust and research a decade later: What have we learned since Jesse Gelsinger's death? Mol. Genet. Metab. *97*, 4–5.

Ye, K., Malinina, L., and Patel, D.J. (2003). Recognition of small interfering RNA by a viral suppressor of RNA silencing. Nature *426*, 874–878.

Yoneyama, M., and Fujita, T. (2010). Recognition of viral nucleic acids in innate immunity. Rev. Med. Virol. *20*, 4–22.

Yoshida, H., Matsui, T., Yamamoto, A., Okada, T., and Mori, K. (2001). XBP1 mRNA is induced by ATF6 and spliced by IRE1 in response to ER stress to produce a highly active transcription factor. Cell *107*, 881–891.

Yuan, Q., Harley, C.W., and McLean, J.H. (2003). Mitral cell β1 and 5-HT2A receptor colocalization and camp coregulation: A new model of norepinephrine-induced learning in the olfactory bulb. Learn. Mem. *10*, 5–15.

Zarghampoor, F., Azarpira, N., Khatami, S.R., Behzad-Behbahani, A., and Foroughmand, A.M. (2019). Improved translation efficiency of therapeutic mRNA. Gene *707*, 231–238.

Zeidan, J., Sharma, A.A., Lee, G., Raad, A., Fromentin, R., Fourati, S., Ghneim, K.,

Sanchez, G.P., Benne, C., Canderan, G., et al. (2021). Infusion of CCR5 Gene-Edited T Cells Allows Immune Reconstitution, HIV Reservoir Decay, and Long-Term Virological Control. BioRxiv.

Zhang, F., Parayath, N.N., Ene, C.I., Stephan, S.B., Koehne, A.L., Coon, M.E., Holland, E.C., and Stephan, M.T. (2019a). Genetic programming of macrophages to perform anti-tumor functions using targeted mRNA nanocarriers. Nat. Commun. *10*.

Zhang, H., Zhang, L., Lin, A., Xu, C., Li, Z., Liu, K., Liu, B., Ma, X., Zhao, F., Jiang, H., et al. (2023). Algorithm for optimized mRNA design improves stability and immunogenicity. Nature *621*, 396–403.

Zhang, H.X., Zhang, Y., and Yin, H. (2019b). Genome Editing with mRNA Encoding ZFN, TALEN, and Cas9. Mol. Ther. *27*, 735–746.

Zhang, Q., Ma, D., Wu, F., Standage-Beier, K., Chen, X., Wu, K., Green, A.A., and Wang, X. (2021). Predictable control of RNA lifetime using engineered degradation-tuning RNAs. Nat. Chem. Biol. *17*, 828–836.

Zhang, X., Potty, A.S.R., Jackson, G.W., Stepanov, V., Tang, A., Liu, Y., Kourentzi, K., Strych, U., Fox, G.E., and Willson, R.C. (2009). Engineered 5S ribosomal RNAs displaying aptamers recognizing vascular endothelial growth factor and malachite green. J. Mol. Recognit. *22*, 154–161.

Zhang, X., Han, L., Zong, H., Ding, K., Yuan, Y., Bai, J., Zhou, Y., Zhang, B., and Zhu, J. (2018). Enhanced production of anti-PD1 antibody in CHO cells through transient co-transfection with anti-apoptotic genes Bcl-xL and Mcl-1. Bioprocess Biosyst. Eng. *41*, 633–640.

Zhao, B., Wei, X., Li, W., Udan, R.S., Yang, Q., Kim, J., Xie, J., Ikenoue, T., Yu, J., Li, L., et al. (2007). Inactivation of YAP oncoprotein by the Hippo pathway is involved in cell contact inhibition and tissue growth control. Genes Dev. *21*, 2747–2761.

Zhou, W.Z., Hoon, D.S.B., Huang, S.K.S., Fujii, S., Hashimoto, K., Morishita, R., and Kaneda, Y. (1999). RNA melanoma vaccine: Induction of antitumor immunity by human glycoprotein 100 mRNA immunization. Hum. Gene Ther. *10*, 2719–2724.

Zhou, Y., Wu, J., Yao, S., Xu, Y., Zhao, W., Tong, Y., and Zhou, Z. (2023). DeepCIP: A multimodal deep learning method for the prediction of internal ribosome entry sites of circRNAs. Comput. Biol. Med. *164*, 107288.

Zhu, B., Tabor, S., Raytcheva, D.A., Hernandez, A., King, J.A., and Richardson, C.C. (2013a). The RNA polymerase of marine cyanophage Syn5. J. Biol. Chem. *288*, 3545–3552.

Zhu, B., Tabor, S., Raytcheva, D.A., Hernandez, A., King, J.A., and Richardson, C.C. (2013b). The RNA Polymerase of Marine Cyanophage Syn5. J. Biol. Chem. *288*, 3545–3552.

Zhu, B., Hernandez, A., Tan, M., Wollenhaupt, J., Tabor, S., and Richardson, C.C. (2015). Synthesis of 2′-Fluoro RNA by Syn5 RNA polymerase. Nucleic Acids Res. *43*, e94–e94.

Zhu, F., Ma, S., Wen, H., Rao, M., Zhang, P., Peng, W., Cui, Y., Yang, H., Tan, C., Chen, J., et al. (2022a). Development of a novel circular mRNA vaccine of six protein combinations against Staphylococcus aureus. J. Biomol. Struct. Dyn.

Zhu, F., Ma, S., Wen, H., Rao, M., Zhang, P., Cui, Y., Yang, H., Tan, C., Chen, J., and Pan, P. (2022b). Development of a novel circular mRNA vaccine of six protein combinations against Staphylococcus aureus Development of a novel circular mRNA vaccine of six protein combinations. J. Biomol. Struct. Dyn. *0*, 1–21.

Zou, S., Scarfo, K., Nantz, M.H., and Hecker, J.G. (2010). Lipid-mediated delivery of RNA is more efficient than delivery of DNA in non-dividing cells. Int. J. Pharm. *389*, 232–243.

Zou, S.W., Ai, K.X., Wang, Z.G., Yuan, Z., Yan, J., and Zheng, Q. (2011). The role of Med19 in the proliferation and tumorigenesis of human hepatocellular carcinoma cells. Acta Pharmacol. Sin. *32*, 354–360.

Zucchelli, E., Pema, M., Stornaiuolo, A., Piovan, C., Scavullo, C., Giuliani, E., Bossi, S., Corna, S., Asperti, C., Bordignon, C., et al. (2017). Codon Optimization Leads to Functional Impairment of RD114-TR Envelope Glycoprotein. Mol. Ther. - Methods Clin. Dev. *4*, 102–114.

8 - Appendix

8.1 - Paper I Supplementary Information

Supplementary Table 3.1 – Single Subunit RNA polymerases used in initial library construction

Phage + Polymerase Name	UniProt Protein ID	Genome accession	Identity to T7 (%)
Achromobacter phage vB_AxyP_19-32_Axy10	A0A514CTW9	QDH83928.1	27
Achromobacter phage vB_AxyP_19-32_Axy11	A0A514CU63	QDH84011.1	26
Achromobacter phage vB_AxyP_19-32_Axy13	A0A514CUM4	QDH84174.1	30
Achromobacter phage vB_AxyP_19-32_Axy22	A0A514CVV4	QDH84606.1	26
Acinetobacter phage AB3	L7TIY7	AGC35317.1	26
Acinetobacter phage AbKT21phiIII	A0A3T0IGC3	AZU98485.1	26
Acinetobacter phage Abp1	R4IPS4	AFV51009.1	26
Acinetobacter phage Fri1	A0A0H4THZ7	AKQ06842.1	27
Acinetobacter phage IME200	A0A0P0I313	ALJ97674.1	26
Acinetobacter phage Petty	U5PZM8	AGY47998.1	27
Acinetobacter phage phiAB1	K9J9S8	ACY78114.1	26
Acinetobacter phage phiAB1	E5KJP3	ADQ12732.1	26
Acinetobacter phage phiAB6	A0A162HLG1	ALA12251.1	26
Acinetobacter phage SH-Ab 15519	A0A240EWY5	APD19402.1	26
Acinetobacter phage vB_AbaP_46-62_Aci07	A0A386KLP3	AYD85859.1	27
Acinetobacter phage vB_AbaP_Acibel007	A0A075DXW8	AHY26803.1	26
Acinetobacter phage vB_AbaP_APK14	A0A499SE88	AYR04381.1	26
Acinetobacter phage vB_AbaP_APK2	A0A5H2UKG0	AZU99229.1	26
Acinetobacter phage vB_AbaP_APK2-2	A0A5H2UKJ8	AZU99279.1	26
Acinetobacter phage vB_AbaP_APK32	A0A5H2UHA1	AZU99382.1	27
Acinetobacter phage vB_AbaP_APK37	A0A5H2UHD3	AZU99432.1	27
Acinetobacter phage vB_AbaP_APK92	A0A5H2UKN2	AZU99329.1	26
Acinetobacter phage vB_AbaP_AS11	A0A218KS19	AQN32684.1	27
Acinetobacter phage vB_AbaP_AS12	A0A218KRE5	APW79817.1	26
Acinetobacter phage vB_AbaP_B09_Aci08	A0A386KCC6	AYD82866.1	26
Acinetobacter phage vB_AbaP_B1	A0A221SBH0	ASN73340.1	27
Acinetobacter phage vB_AbaP_B3	A0A221SBM1	ASN73388.1	27
Acinetobacter phage vB_AbaP_B5	A0A221SBR9	ASN73442.1	28
Acinetobacter phage vB_AbaP_D2	A0A2P1MY72	AVP40508.1	26
Acinetobacter phage vB_AbaP_D2M	A0A5J6TX62	QFG15411.1	26
Acinetobacter phage vB_AbaP_PD-6A3	A0A0S1RZT9	ALM01888.1	27
Acinetobacter phage vB_AbaP_PD-AB9	A0A0S1S0Y2	ALM01908.1	26
Acinetobacter phage vB_ApiP_P1	A0A221SBX1	ASN73492.1	26
Acinetobacter phage vB_ApiP_P2	A0A221SC12	ASN73545.1	26
Acinetobacter phage WCHABP5	A0A1X9SFL3	ARQ94903.1	27
Aeromonas phage vB_AsoP_Ca	A0A482MHE4	QBQ72988.1	63
Agrobacterium phage Atu_ph02	A0A223VZI2	ASV44565.1	32

Agrobacterium phage Atu_ph03	A0A223VZW7	ASV44581.1	32
Alteromonas virus vB_AspP-H4/4	A0A220YL66	ASL24412.1	29
Aquamicrobium phage P14	A0A1L5C074	APL99490.1	31
Bordetella phage vB_BbrP_BB8	A0A4Y5TNQ1	QDB70995.1	39
Burkholderia phage AMP1	A0A5C2IFS7	QEP52854.1	32
Burkholderia phage Bp-AMP1	A0A0A1I5M6	CDK30098.1	32
Burkholderia phage Bp-AMP2	A0A0A8KWM2	CDL65184.1	32
Burkholderia phage Bp-AMP3	A0A0A8KWN3	CDL65224.1	32
Burkholderia phage Bp-AMP4	A0A0A8KWP2	CDL65264.1	32
Burkholderia phage JG068	U3PFP4	AGW43609.1	31
Caulobacter phage Cd1	F1ADQ9	ADD21661.1	30
Caulobacter phage Lullwater	A0A291LBB2	ATI16335.1	31
Caulobacter phage Percy	A0A0M4RSN8	ALF01667.1	32
Citrobacter phage CR44b	W6PNY5	CDM21533.1	62
Citrobacter phage CR8	W6PP41	CDM21594.1	62
Citrobacter phage phiCFP-1	A0A0E3JQL5	AKA62124.1	83
Citrobacter phage SH1	A0A172JG22	AMR59436.1	82
Citrobacter phage SH2	A0A172JG53	AMR59464.1	83
Citrobacter phage SH3	A0A172JGA6	AMR59537.1	62
Citrobacter phage SH4	A0A172JGF0	AMR59589.1	62
Citrobacter phage SH5	A0A172JGK4	AMR59637.1	62
Cronobacter phage Dev2	W6PED7	CDM12529.1	61
Cronobacter phage GW1	A0A3S7N8N5	AWY03129.1	63
Curvibacter phage P26059B	A0A384UH57	ASJ79294.1	30
Cyanophage KBS-P-1A	M4T4F9	AGH57729.1	26
Delftia phage IME-DE1	A0A0F7INH1	AKG94487.1	46
Delftia phage RG-2014	A0A097PAJ0	AIU44277.1	30
Dickeya phage BF25/12	A0A219MH23	ALA46497.1	29
Dickeya phage Dagda	A0A2S1GSJ5	AWD92361.1	58
Dickeya phage Dagda_B1	A0A385IFE2	AXY81610.1	58
Dickeya phage Katbat	A0A385IFD3	AXY81723.1	58
Dickeya phage Luksen	A0A385IFU3	AXY81879.1	58
Dickeya phage Mysterion	A0A385IGY0	AXY81987.1	58
Dickeya phage Ninurta	A0A2S1GTB4	AWD92624.1	70
Dickeya phage vB_DsoP_JA10	A0A384ZVU4	AXG66358.1	70
Enterobacter phage E-2	A0A0E3JQ94	AKA61565.1	81
Enterobacter phage E-4	A0A0E3JTE8	AKA61642.1	83
Enterobacter phage phiEap-1	A0A0K2FHD8	ALA45082.1	76
Enterobacteria phage 13a	B3VD33	ACF15888.1	97
Enterobacteria phage 285P	D3W681	ACV32460.1	75
Enterobacteria phage BA14	B3VCM6	ACF15731.1	75
Enterobacteria phage EcoDS1	B3VCT0	ACF15785.1	61
Enterobacteria phage K11 (Bacteriophage K11)	P18147	CAA37330.1	72
Enterobacteria phage K1-5		108861984	32
Enterobacteria phage K1E		83571714	32
Enterobacteria phage K1F		77118167	61
Enterobacteria phage K30		338826807	72

Enterobacteria phage N4 (Bacteriophage N4)	Q859P9	AAO24831.2	32
Enterobacteria phage SP6	Q6UGK9	CAA68288.1	32
Enterobacteria phage T3	P07659	CAA28696.1	81
Enterobacteria phage UAB_Phi78	M1E3S8	ADW95214.1	32
Enterobacteria phage vB_EcoP_ACG-C91	K4FE53	AFH19836.1	31
Enterobacteria phage vB_EcoP_IME390	A0A386KER3	AYD82921.1	98
Erwinia amylovora phage Era103	A2I7X6	ABM63398.1	32
Erwinia phage Ea9-2	W6ASK5	AHI60080.1	20
Erwinia phage FE44	U5NZV0	AGY36902.1	75
Erwinia phage pEp_SNUABM_09	A0A5J6DAB1	QEQ94747.1	73
Erwinia phage phiEa100	E5AGG9	CBX45071.1	32
Erwinia phage phiEa1H	E5AG56	CBX44468.1	32
Erwinia phage vB_EamP_Frozen	A0A191ZCP0	ANJ65156.1	23
Erwinia phage vB_EamP_Gutmeister	A0A191ZD74	ANJ65339.1	23
Erwinia phage vB_EamP_Rexella	A0A191ZCZ2	ANJ65253.1	23
Erwinia phage vB_EamP-L1	G0YQ47	AEJ81490.1	69
Erwinia phage vB_EamP-S2	A0A2K9V4V5	AUV57206.1	32
Escherichia phage 64795_ec1	A0A192Y777	ANM45729.1	98
Escherichia phage C5	A0A386K648	AYD80175.1	99
Escherichia phage CICC 80001	A0A088F7U6	AIM41079.1	97
Escherichia phage DE3	C6ZCU5	ACF32352.1	100
Escherichia phage Ebrios	A0A2P1CL25	AVJ51888.1	90
Escherichia phage ECA2	A0A193GZB1	ANN86245.1	82
Escherichia phage ECBP5	A0A2H4P6B5	AID17660.1	32
Escherichia phage EG1	A0A2K9VJW3	ATW57716.1	98
Escherichia phage HZ2R8	A0A481V9B9	AUV62632.1	98
Escherichia phage HZP2	A0A384WW43	QBI89975.1	98
Escherichia phage IMM-002	A0A1I9SEA0	ATI16970.1	62
Escherichia phage JSS1	Q8SCG8	AOZ65177.1	63
Escherichia phage K1E (Bacteriophage K1E)	Q2WC24	CAJ29407.1	32
Escherichia phage K1E (Bacteriophage K1E)	A0A2I4Q1U4	AQY55033.1	32
Escherichia phage K30	F8R4Q2	AEH41021.1	74
Escherichia phage LL11	A0A385IQB0	AXY85379.1	32
Escherichia phage LL2	A0A385IPR1	AXY85324.1	82
Escherichia phage LM33_P1	A0A1A8YFH6	SBT28095.1	59
Escherichia phage N13	A0A386K7G9	AYD80219.1	99
Escherichia phage N30	A0A386K840	AYD80131.1	99
Escherichia phage P483	A0A0D3QHF3	AJF40472.1	75
Escherichia phage P694	A0A0D3QHD0	AJF40494.1	75
Escherichia phage PE3-1	A0A060D5I6	AIB06962.1	61
Escherichia phage Peacock	A0A5B9NFR5	QEG09673.1	62
Escherichia phage Penshu1	A0A5B9N881	QEG09779.1	63
Escherichia phage PGT2	A0A2D2W380	ATS92446.1	31
Escherichia phage phiKT	H6VUA7	AEZ65095.1	31
Escherichia phage Pisces	A0A5B9NC24	QEG09560.1	62
Escherichia phage SRT7	A0A2Z5H3J4	AXC34604.1	77
Escherichia phage ST31	A0A1Z1LW08	ARW56841.1	63

Escherichia phage T7 (Bacteriophage T7)	C1KTT1	AAA32569.1	100
Escherichia phage vB_Eco_D226	A0A514A6V2	QDH49005.1	76
Escherichia phage vB_EcoP_B	A0A1Q1PUC2	AQN31610.1	32
Escherichia phage vB_EcoP_C	A0A1Q1PUD0	AQN31697.1	32
Escherichia phage vB_EcoP_D	A0A1Q1PUF9	AQN31727.1	32
Escherichia phage vB_EcoP_F	A0A1S6KVE6	AQT25408.1	63
Escherichia phage vB_EcoP_GA2A	A0A1B0TR91	ALP47773.1	63
Escherichia phage vB_EcoP_KAW1A4500	A0A482MT09	QBQ76716.1	34
Escherichia phage vB_EcoP_R	A0A1Q1PVG6	AQN32030.1	32
Escherichia phage vB_EcoP_R4596	A0A482MRF6	QBQ76571.1	32
Escherichia phage vB_EcoP_S523	A0A2Z4Q2U2	AWY04286.1	76
Escherichia phage vB_EcoP-Ro45lw	A0A3S9URU0	AZS12994.1	62
Escherichia phage YZ1	A0A2L0HPX4	AUX83594.1	63
Escherichia phage ZG49	A0A1C9M222	AOQ29248.1	63
Escherichia virus AAPEc6	A0A1P8BK25	ANT40787.1	32
Escherichia virus ECG4	A0A5J6T816	QFG06972.1	66
Escherichia virus K1-5		AAL86891.1	32
Escherichia virus mutPK1A2	A0A2H4MZV8	ATS93311.1	32
Escherichia virus Vec13	A0A345AQE0	AXF38891.1	62
Escherichia virus VEc3	A0A2H4YE34	AUE22234.1	32
Klebsiella phage 2044-307w	A0A249Y210	ASZ78273.1	73
Klebsiella phage IME304	A0A4Y5TVL6	QDB73388.1	74
Klebsiella phage K11	B3VCY2	CAA37330.1	72
Klebsiella phage K1-ULIP33	A0A4P6DBD2	QAU05462.1	33
Klebsiella phage K5	A0A0F7LBY1	AKH49548.1	72
Klebsiella phage K5-2	A0A219YH91	APZ82773.1	74
Klebsiella phage K5-4	A0A219YHD6	APZ82814.1	73
Klebsiella phage KN1-1	A0A3S5IBH2	BBF66827.1	73
Klebsiella phage KN3-1	A0A3Q9WWY9	BBF66848.1	73
Klebsiella phage KN4-1	A0A3Q9WSE6	BBF66872.1	73
Klebsiella phage KOX2	A0A514U6A2	QDK04476.1	78
Klebsiella phage KOX3	A0A5B9NGK0	QEG09953.1	73
Klebsiella phage KOX5	A0A5B9NDA1	QEG10020.1	73
Klebsiella phage KP32	D1L2T7	ACY66666.1	72
Klebsiella phage KPN3	A0A5B9NJV0	QEG11143.1	73
Klebsiella phage kpssk3	A0A3G8F354	AZF88810.1	73
Klebsiella phage Kund-ULIP47	A0A4P6DBN7	QAU05510.1	73
Klebsiella phage Kund-ULIP54	A0A4P6PMA0	QBG78351.1	73
Klebsiella phage Patroon	A0A482MHH9	QBQ72872.1	83
Klebsiella phage Pharr	A0A4D6DY13	QBZ71215.1	73
Klebsiella phage SH-Kp 152410	A0A2K9VGQ6	AUV61470.1	73
Klebsiella phage vB_Kp1	A0A0P0IV82	ALJ98062.1	72
Klebsiella phage vB_KpnP_BIS33	A0A1V0E6J1	ARB12463.1	73
Klebsiella phage vB_KpnP_Emp27	A0A5B9NG02	QEG11865.1	83
Klebsiella phage vB_KpnP_FZ12	A0A4D6T531	QCG76471.1	73
Klebsiella phage vB_KpnP_IME205	A0A0U3DFB5	ALT58461.1	74
Klebsiella phage vB_KpnP_IME321	A0A344UBX8	AXE28399.1	73

Klebsiella phage vB_KpnP_IME335	A0A5J6CUK8	QEQ50481.1	73
Klebsiella phage vB_KpnP_KpV289	A0A0K2YWK9	CRN12692.1	72
Klebsiella phage vB_KpnP_KpV763	A0A1D8F0C6	AOT28135.1	73
Klebsiella phage vB_KpnP_KpV766	A0A1I9SFA1	AOZ65531.1	73
Klebsiella phage vB_KpnP_KpV767	A0A1I9SF50	AOZ65480.1	73
Klebsiella phage vB_KpnP_NahiliMali	A0A5B9NQY1	QEG13343.1	70
Klebsiella phage vB_KpnP_PRA33	A0A1V0E683	ARB12370.1	73
Klebsiella phage vB_KpnP_Sibilus	A0A5B9NKC7	QEG12916.1	70
Klebsiella virus KP32	A0A2U8USS7	AWN07136.1	73
Klebsiellaphage KP34		ACY66709.1	28
Kluyvera phage Kvp1	B6Z9E6	ACJ14548.1	75
Leclercia phage 10164-302	A0A289YY98	ATA65249.1	82
Leclercia phage 10164RH	A0A289Z6Q5	ATA65295.1	82
Lelliottia phage phD2B	A0A088FT51	AIM51235.1	32
Mesorhizobium phage vB_MloP_Lo5R7ANS		AIK68488.1	32
Morganella phage MmP1	D1FNQ5	ACY74627.1	70
Morganella phage vB_MmoP_MP2	A0A192YBW9	ANM46341.1	72
Pectobacterium phage DU_PP_II	A0A2D2W5U8	ATS93671.1	64
Pectobacterium phage Jarilo	A0A2S1GSW7	AWD92487.1	71
Pectobacterium phage POP72	A0A2R2V0R8	ARB10926.1	31
Pectobacterium phage PP1	I7FXR7	AFP33672.1	31
Pectobacterium phage PP16	A0A1B1PED0	ANT45336.1	30
Pectobacterium phage PP47	A0A1P8L665	APW79749.1	69
Pectobacterium phage PP74	A0A1J0MEG1	APD19621.1	75
Pectobacterium phage PP81	A0A1L7DRY0	APU03032.1	69
Pectobacterium phage PPWS4	A0A286P065	BBA26420.1	68
Pelabacter phage HTVC019P	M1ID95	AGE60589.1	32
Pelabacter phage HTVC021P	A0A4Y1NSE5	AXH68324.1	32
Pelabacter phage HTVC022P	A0A4Y1NU69	AXH71733.1	32
Pelabacter phage HTVC031P	A0A4Y1NTX3	AXH71631.1	34
Pelabacter phage HTVC105P	A0A4Y1NTS2	AXH71416.1	32
Pelabacter phage HTVC109P	A0A4Y1NYD1	AXN54076.1	33
Pelabacter phage HTVC121P	A0A4Y1NTW1	AXH71475.1	32
Pelabacter phage HTVC201P	A0A4Y1NTK5	AXH71527.1	32
Phage MedPE-SWcel-C56	A0A1B1IY27	ANS06226.1	30
Phage NC-A	A0A481S2N0	QBG78787.1	98
Phage Paz		AHB12124.1	31
Podovirus Lau218	A0A060BKZ7	AIA83126.1	30
Prochlorococcus phage P-SSP7		AAX44193.1	26
Proteus phage PM 116	A0A2D0VJZ0	ANU80090.1	32
Proteus phage PM 85	A0A0F6NYB6	AIW03110.1	31
Proteus phage PM 93	A0A0F6NYB9	AIW03132.1	31
Proteus phage vB_PmiP_Pm5460	A0A0G2SSG2	AKA61817.1	31
Pseudomonad phage gh-1	Q859H5	AAO73140.1	58
Pseudomonas phage 22PfluR64PP	A0A3G6V715	AZB48859.1	57
Pseudomonas phage 67PfluR64PP	A0A2S1PGT5	AWH15793.1	57
Pseudomonas phage 71PfluR64PP	A0A2S1PDT9	AWH14736.1	57

Pseudomonas phage Andromeda	A0A1B1SEK9	ANU79101.1	31
Pseudomonas phage Bf7	H2ELW6	AEX65868.1	31
Pseudomonas phage gh-1		AAO73140.1	56
Pseudomonas phage Henninger	A0A2K9VHD7	AUV61717.1	57
Pseudomonas phage inbricus	A0A2H4P7K4	ATW58149.1	26
Pseudomonas phage LKA1	Q0E5X7	CAK25005.1	27
Pseudomonas phage Pf-10	A0A0A0YSI2	AIX12970.1	58
Pseudomonas phage PFP1	A0A2Z4QIP2	AWY10458.1	57
Pseudomonas phage phi15	F0V6X0	CBZ41982.1	57
Pseudomonas phage phiIBB-PF7A	E9KIE1	ADV35666.1	57
Pseudomonas phage phiKMV	Q7Y2D9	CAD44217.1	29
Pseudomonas phage phiPsa17	A0A0G2T693	AKG94352.1	58
Pseudomonas phage phiPSA2	A0A059VA03	AHZ94988.1	56
Pseudomonas phage Phi-S1	M4H3N8	AFO12299.1	58
Pseudomonas phage PollyC	A0A2K9VHU7	AUV61936.1	29
Pseudomonas phage PPPL-1	A0A0S2MVL3	ALO79969.1	57
Pseudomonas phage PPpW-4 DNA	V5YUU1	BAO20676.1	56
Pseudomonas phage PspYZU08	A0A2U7NJN9	ASD52184.1	56
Pseudomonas phage shl2	A0A160SY77	CUR50709.1	58
Pseudomonas phage UNO-SLW1	A0A1B2ANA8	ANY29164.1	59
Pseudomonas phage UNO-SLW2	A0A1B2AN40	ANY29117.1	59
Pseudomonas phage UNO-SLW3	A0A1B2AMZ7	ANY29070.1	59
Pseudomonas phage UNO-SLW4	A0A1B2AMX6	ANY29024.1	59
Pseudomonas phage YMC11/06/C171_PPU_BP	A0A127KP47	AMO43666.1	28
Pseudomonas virus KNP	A0A1W6JS11	ARM69622.1	58
Pseudomonas virus Pf1 ERZ-2017	A0A2H4YGR2	AUE23184.1	58
Pseudomonas virus WRT	A0A1W6JRT9	ARM69573.1	58
Ralstonia phage DU_RP_I	A0A2D2W578	ATS93364.1	39
Ralstonia phage RSB1	A0A5P8D3T3	BAG70384.1	31
Ralstonia phage RSB2 DNA	A0A5P8D447	BAJ51806.1	44
Ralstonia phage RSB3	U3TFM9	BAN92342.1	28
Ralstonia phage RSJ2	A0A068Q5R3	BAP15831.1	32
Ralstonia phage RSJ5	A0A077KYK0	BAP34919.1	31
Ralstonia phage RsoP1EGY	A0A2R2ZGE5	AUO78209.1	39
Ralstonia phage RS-PI-1	A0A1S6L1D6	AQT27784.1	32
Ralstonia phage RS-PII-1	A0A1L7DQC4	APU00310.1	32
Ralstonia virus phiAp1	A0A1L7DS98	APU03175.1	28
Rhizobium phage RHEph01	L7TQW5	AGC35521.1	36
Rhizobium phage RHEph02	L7TJC5	AGC35603.1	30
Rhizobium phage RHEph03	L7TJH2	AGC35663.1	30
Rhizobium phage RHEph08	L7TS16	AGC35961.1	30
Rhizobium phage RHEph09	L7TLU8	AGC36015.1	30
Salmonella phage 3A_8767	A0A2Z5HC42	AXC37072.1	99
Salmonella phage BP12A	A0A140XFL3	AIT13633.1	76
Salmonella phage BP12B	A0A140XFR2	AIT13682.1	33
Salmonella phage BSP161	A0A3G1L319	ATW58438.1	76

Salmonella phage LPST144	A0A5C2ID52	QEP53480.1	76
Salmonella phage phiSG-JL2	B3FYH6	ACD75668.1	81
Salmonella phage Vi06		CBV65202.1	92
Serratia phage 2050H2	A0A249Y2R9	ASZ78985.1	83
Serratia phage Pila	A0A5J6TE08	QFG06779.1	99
Serratia phage SM9-3Y	A0A1I9S3U3	AOZ61237.1	83
Shigella phage VB_Ship_A7	A0A4D6DRR8	QBZ69017.1	76
Sinorhizobium phage ort11	A0A5C2H2X6	QEP29836.1	24
SPBeta prophage YonO		CAB14022.1	29
Sphingomonas phage Scott	A0A346FDD2	AXN53746.1	31
Stenotrophomonas phage IME15		AFV51444.1	89
Synechococcus phage S-CBP3	I3ULU9	AFK66464.1	25
Synechococcus phage S-CBP42	A0A096VKW2	AGK86693.1	25
Synechococcus phage S-RIP1	M4NMG5	AGG91275.1	26
Synechococcus phage S-RIP2	M4NNU7	AGG91331.1	26
Synechococcus T7-like virus S-TIP37	A0A345AY93	AXF42073.1	25
Synechococcus virus P60	L0CNW4	AGA17876.1	26
Thermus virus P23-45 P23p64		ABU96897.1	36
Vibrio phage 1.204.O._10N.222.46.F12	A0A2I7RNL7	AUR95242.1	33
Vibrio phage ICP3	F1D002	ADX87446.1	62
Vibrio phage ICP3_2007_A	F1D0J5	ADX87639.1	62
Vibrio phage ICP3_2008_A	F1D0E7	ADX87591.1	61
Vibrio phage ICP3_2009_A	F1D0A0	ADX87544.1	62
Vibrio phage ICP3_2009_B	F1D053	ADX87497.1	62
Vibrio phage JSF11	A0A2D0Z112	ASV42767.1	62
Vibrio phage JSF18	A0A2D0YMX9	ASV43222.1	62
Vibrio phage JSF20	A0A2D0YL99	ASV43270.1	62
Vibrio phage JSF24	A0A2D0Z841	ASV43318.1	62
Vibrio phage JSF25	A0A2D0XR33	ASU01137.1	62
Vibrio phage JSF30	A0A2D0YV87	ASV42904.1	62
Vibrio phage JSF31	A0A2D0Z1P3	ASV42947.1	61
Vibrio phage JSF32	A0A2D0Z2L0	ASV43021.1	61
Vibrio phage JSF34	A0A2D0YLK2	ASV43366.1	62
Vibrio phage JSF35	A0A2D0YKN7	ASV43086.1	62
Vibrio phage JSF36	A0A2D0Z259	ASV43147.1	61
Vibrio phage N4	D0Q187	ACR16468.1	61
Vibrio phage phi-A318		AFB82771.1	33
Vibrio phage Rostov-1	A0A2P0ZKC2	AVH85446.1	62
Vibrio phage VP3	H9YAF4	AFH14406.1	62
Vibriophage VP4	Q4TVY1	AAY46276.1	61
Xanthomonas phage phiL7		ACE75775.1	31
Yersinia pestis phage phiA1122	Q858N4	AAP20500.1	97
Yersinia phage Berlin	A0ZXJ9	CAJ70654.1	74
Yersinia phage fPS-10	A0A2H1X8U9	SOR54291.1	73
Yersinia phage fPS16	A0A2H1X8Z7	SOR54342.1	73
Yersinia phage fPS-19	A0A2D0PDM8	SOO46384.1	73
Yersinia phage fPS-21	A0A2D0PE32	SOO46739.1	73

Yersinia phage fPS-26	A0A2D0PD60	SOO46435.1	73
Yersinia phage fPS-50	A0A2D0PDY6	SOO46689.1	73
Yersinia phage fPS-52	A0A2D0PDI2	SOO46334.1	73
Yersinia phage fPS-53	A0A2H1UJD0	SOO56471.1	73
Yersinia phage fPS-54-ocr	A0A2H1UJE6	SOP75972.1	73
Yersinia phage fPS-59	A0A2D0PE84	SOO46791.1	73
Yersinia phage fPS-64	A0A2D0PEF0	SOO46837.1	73
Yersinia phage fPS-7	A0A2D0PDD4	SOO46486.1	73
Yersinia phage fPS-85	A0A2H1UKL9	SOO56422.1	73
Yersinia phage fPS-86	A0A2D0PEP1	SOO46638.1	73
Yersinia phage fPS-89	A0A2D0PDP8	SOO46589.1	73
Yersinia phage fPS-9	A0A2C9D120	SOL37498.1	73
Yersinia phage phiYe-F10	A0A0U2DV42	AKQ06767.1	83
Yersinia phage phiYeO3-12	Q9T145	CAB63592.1	81
Yersinia phage PYPS50	A0A3G8F3B0	AZF87572.1	76
Yersinia phage R	L0HSI1	AGB07325.1	99
Yersinia phage vB_YenP_AP10	A0A0P0M6B3	ALK86936.1	78
Yersinia phage vB_YenP_AP5	A0A088F657	AIM40350.1	81
Yersinia phage Y	M9PKG8	AGC35464.1	99
Yersinia phage Yepe2	B3VCH9	ACF15684.1	75
Yersinia phage Yep-phi	E5L7B9	ADQ83157.1	74
Yersinia phage YpP-G	I6Q9Q7	AFK13447.1	75
Yersinia phage YpP-R	I6Q993	AFK13398.1	99
Yersinia phage YpsP-G	I6Q9B2	AFK13493.1	98

Supplementary Table 3.2 – Final library of single subunit RNA polymerases used in initial transcription-translation experiments. The predicted optimum temperature is determined through literature reports of the optimal temperature at which the bacteriophage grows at, from which the RNAP originates

Assigned Number	Phage Name	Genome accession	UniProt Protein ID	Protein Length	Predicted Temperature Optimum (Literature Reports)	Promoter
R1	Prochlorococcus_phage_P-SSP7	AAX44193.1	Q58N45	778	14-26	AAAATTCTTCAAGTTTACAA
R2	Synechococcus_phage_S-CBP42	AGK86693.1	A0A096VKW2	777	20	CACTTCCACTCAACCAACCG
R3	Erwinia_amylovora_phage_Era103	ABM63398.1	A2I7X6	885	28	AATAACCACCCAGTATAGAAGGAA
R4	Agrobacterium_phage_Atu_ph02	ASV44565.1	A0A223VZI2	815	28	TTATCCTTCGTATAAGGAATA
R5	Dickeya_phage_Mysterion	AXY81987.1	A0A385IGY0	876	28	CTTAAATCATCACTATTAG
R6	Pectobacterium_phage_PP74	APD19621.1	A0A1J0MEG1	883	30	TAATACGACTCACTATTGGGAA
R7	Aquamicrobium_phage_P14	APL99490.1	A0A1L5C074	804	30	TTTCGGTACGCTCTAGCA
R8	Pectobacterium_phage_DU_PP_II	ATS93671.1	A0A2D2W5U8	910	30	TTATTAACGACTCACTACTAGGAA
R9	Pectobacterium_phage_Jarilo	AWD92487.1	A0A2S1GSW7	900	30	TAATAACGACTCACTATTAGAAG
R10	Sphingomonas_phage_Scott	AXN53746.1	A0A346FDD2	862	30	TCGGGTTGTCGATTTCCTTAC
R11	Ralstonia_phage_RS-PI-1	AQT27784.1	A0A1S6L1D6	830	35	GTCGAAGTCGTCGAGCAGC
R12	Burkholderia_phage_JG068	AGW43609.1	U3PFP4	828	37	TCAGTAGACTATCTAG
R13	Acinetobacter_Acibel007	AHY26803.1	A0A075DXW8	809	37	CTGTACTCACAGCTCAATTT
R14	Delftia_phage_IME-DE1	AKG94487.1	A0A0F7INH1	859	37	GTTAGCCCACACCATTGAAGACCC
R15	Pseudomonas_phage_Henninger	AUV61717.1	A0A2K9VHD7	884	37	TTAAAACCCTCACTATGGCTACA
R16	Pseudomonas_phage_PollyC	AUV61936.1	A0A2K9VHU7	807	37	CTCACTCACGACCCAAATTC
R17	Pseudomonas_phage_phiKMV	CAD44217.1	Q7Y2D9	816	37	CGACCCTTCCCTACTCCGGCCTTAAAT
R18	Citrobacter_phage_CR8	CDM21594.1	W6PP41	890	37	TAAGGAAGGTACACTATAGGG
R19	Thermus_virus_P23-45_P23p64	ABU96897.1	A7XX94	616	65	TTATTCCTTTA
R20	Pelagibacter_phage_HTVC031P	AXH71631.1	A0A4Y1NTX3	808	16-23	AACTAATGCTCAATTTAGAGATA
R21	Rhizobium_phage_RHEph01	AGC35521.1	L7TQW5	858	25-30	ATTACCCCTCCCTTAAGCAAAG
R22	Rhizobium_phage_RHEph02	AGC35603.1	L7TJC5	812	25-30	TTAATCCTCACTATTAGGATAA
R23	Curvibacter_phage_P26059B	ASJ79294.1	A0A384UH57	816	25-30	GCAACATTACAGGTACTGAA
R24	Pseudomonas_phage_PPpW-4	BAO20676.1	V5YUU1	890	25-30	TAAAAACCCTCACTGAAACAGGG
R25	Vibrio_phage_10N	AUR95242.1	A0A2I7RNL7	808	35-37	ACTTACCTTTCACTATAGCAGCA
R26	Alteromonas_virus_vB_AspP-H44	ASL24412.1	A0A220YL66	796	Not Found	TGGTGACTACAGAGCAGCAG

Supplementary Table 3.3 – The 88 RNAPs that have <75% overall identity with any previously characterised RNAP and >30% DNA binding domain identity with T7 RNAP. Identity to T7 and DNA binding domain identity are calculated through amino acid sequence alignment using Clustal Omega. The DNA binding domain is defined as amino acids 740-769 of T7 RNAP.

RNAP Name	Identity to T7 (%)	DNA binding domain identity (%)	Uniprot ID
Dickeya phage Ninurta	70	47.33	A0A2S1GTB4
Dickeya phage vB_DsoP_JA10	70	46	A0A384ZVU4
Enterobacteria phage K11	72	42.38	P18147
Erwinia phage pEp_SNUABM_09	73	47.33	A0A5J6DAB1
Erwinia phage vB_EamP-L1	69	45.03	G0YQ47
Escherichia phage K30	74	45.7	F8R4Q2
Escherichia phage Pisces	62	32.68	A0A5B9NC24
Klebsiella phage 2044-307w	73	45.03	A0A249Y210
Klebsiella phage IME304	74	45.7	A0A4Y5TVL6
Klebsiella phage K11	72	45.03	B3VCY2
Klebsiella phage K5	72	45.03	A0A0F7LBY1
Klebsiella phage K5-4	73	45.7	A0A219YHD6
Klebsiella phage KN1-1	73	45.03	A0A3S5IBH2
Klebsiella phage KN3-1	73	45.03	A0A3Q9WWY9
Klebsiella phage KN4-1	73	45.03	A0A3Q9WSE6
Klebsiella phage KOX3	73	45.7	A0A5B9NGK0
Klebsiella phage KOX5	73	45.7	A0A5B9NDA1
Klebsiella phage KP32	72	45.7	D1L2T7
Klebsiella phage kpssk3	73	45.03	A0A3G8F354
Klebsiella phage Kund-ULIP47	73	45.7	A0A4P6DBN7
Klebsiella phage Kund-ULIP54	73	45.7	A0A4P6PMA0
Klebsiella phage Pharr	73	45.03	A0A4D6DY13
Klebsiella phage SH-Kp 152410	73	44.37	A0A2K9VGQ6
Klebsiella phage vB_Kp1	72	45.03	A0A0P0IV82
Klebsiella phage vB_KpnP_BIS33	73	43.71	A0A1V0E6J1
Klebsiella phage vB_KpnP_IME205	74	45.7	A0A0U3DFB5
Klebsiella phage vB_KpnP_IME321	73	45.7	A0A344UBX8
Klebsiella phage vB_KpnP_IME335	73	45.03	A0A5J6CUK8
Klebsiella phage vB_KpnP_KpV289	72	44.37	A0A0K2YWK9
Klebsiella phage vB_KpnP_KpV763	73	45.03	A0A1D8F0C6
Klebsiella phage vB_KpnP_KpV766	73	45.03	A0A1I9SFA1
Klebsiella phage vB_KpnP_KpV767	73	45.7	A0A1I9SF50
Klebsiella phage vB_KpnP_NahiliMali	70	47.33	A0A5B9NQY1
Klebsiella phage vB_KpnP_PRA33	73	45.7	A0A1V0E683
Klebsiella phage vB_KpnP_Sibilus	70	46	A0A5B9NKC7
Morganella phage MmP1	70	51.02	D1FNQ5
Morganella phage vB_MmoP_MP2	72	50.34	A0A192YBW9
Pectobacterium phage DU_PP_II	64	34.57	A0A2D2W5U8
Pseudomonas phage 22PfluR64PP	57	35.71	A0A3G6V715

Pseudomonas phage 67PfluR64PP	57	35.71	A0A2S1PGT5
Pseudomonas phage 71PfluR64PP	57	35.71	A0A2S1PDT9
Pseudomonas phage Pf-10	58	38.31	A0A0A0YSI2
Pseudomonas phage PFP1	57	36.36	A0A2Z4QIP2
Pseudomonas phage phi15	57	40.91	F0V6X0
Pseudomonas phage Phi-S1	58	38.31	M4H3N8
Pseudomonas phage PPPL-1	57	38.71	A0A0S2MVL3
Pseudomonas phage PspYZU08	56	34.64	A0A2U7NJN9
Pseudomonas phage shl2	58	39.35	A0A160SY77
Ralstonia phage DU_RP_I	39	30.94	A0A2D2W578
Ralstonia phage RSB1	31	30.94	A0A5P8D3T3
Ralstonia phage RSB2 DNA	44	30.94	A0A5P8D447
Ralstonia phage RsoP1EGY	39	30.22	A0A2R2ZGE5
Vibrio phage ICP3	62	48.32	F1D002
Vibrio phage ICP3_2007_A	62	48.32	F1D0J5
Vibrio phage ICP3_2008_A	61	48.32	F1D0E7
Vibrio phage ICP3_2009_A	62	48.32	F1D0A0
Vibrio phage ICP3_2009_B	62	47.65	F1D053
Vibrio phage JSF11	62	48.32	A0A2D0Z112
Vibrio phage JSF18	62	48.32	A0A2D0YMX9
Vibrio phage JSF20	62	48.32	A0A2D0YL99
Vibrio phage JSF24	62	48.32	A0A2D0Z841
Vibrio phage JSF25	62	48.32	A0A2D0XR33
Vibrio phage JSF30	62	48.32	A0A2D0YV87
Vibrio phage JSF31	61	46.98	A0A2D0Z1P3
Vibrio phage JSF32	61	46.98	A0A2D0Z2L0
Vibrio phage JSF34	62	48.32	A0A2D0YLK2
Vibrio phage JSF35	62	47.65	A0A2D0YKN7
Vibrio phage JSF36	61	47.65	A0A2D0Z259
Vibrio phage N4	61	48.32	D0Q187
Vibrio phage Rostov-1	62	48.32	A0A2P0ZKC2
Vibrio phage VP3	62	48.32	H9YAF4
Vibriophage VP4	61	48.32	Q4TVY1
Yersinia phage fPS-10	73	51.33	A0A2H1X8U9
Yersinia phage fPS16	73	51.33	A0A2H1X8Z7
Yersinia phage fPS-19	73	51.33	A0A2D0PDM8
Yersinia phage fPS-21	73	51.33	A0A2D0PE32
Yersinia phage fPS-26	73	51.33	A0A2D0PD60
Yersinia phage fPS-50	73	51.33	A0A2D0PDY6
Yersinia phage fPS-52	73	51.33	A0A2D0PDI2
Yersinia phage fPS-53	73	51.33	A0A2H1UJD0
Yersinia phage fPS-54-ocr	73	51.33	A0A2H1UJE6
Yersinia phage fPS-59	73	51.33	A0A2D0PE84
Yersinia phage fPS-64	73	51.33	A0A2D0PEF0
Yersinia phage fPS-7	73	51.33	A0A2D0PDD4
Yersinia phage fPS-85	73	51.33	A0A2H1UKL9
Yersinia phage fPS-86	73	51.33	A0A2D0PEP1
Yersinia phage fPS-89	73	51.33	A0A2D0PDP8

Yersinia phage fPS-9	73	51.33	A0A2C9D120

8.2 Paper II Supplementary Information

Supplementary Table 4.1 – DNA element sequences used in this study

eGFP

ATGGTGAGCAAGGGCGAGGAGCTGTTCACCGGGGTGGTGCCCATCCTGGTCGAGCTGGACGGCGACGTAAACGGCCAC
AAGTTCAGCGTGTCTGGCGAGGGCGAGGGCGATGCCACCTACGGCAAGCTGACCCTGAAGTTCATCTGCACCACCGGCA
AGCTGCCCGTGCCCTGGCCCACCCTCGTGACCACCCTGACCTACGGCGTGCAGTGCTTCAGCCGCTACCCCGACCACAT
GAAGCAGCACGACTTCTTCAAGTCCGCCATGCCCGAAGGCTACGTCCAGGAGCGCACCATCTTCTTCAAGGACGACGGCA
ACTACAAGACCCGCGCCGAGGTGAAGTTCGAGGGCGACACCCTGGTGAACCGCATCGAGCTGAAGGGCATCGACTTCAA
GGAGGACGGCAACATCCTGGGGCACAAGCTGGAGTACAACTACAACAGCCACAACGTCTATATCATGGCCGACAAGCAGA
AGAACGGCATCAAGGCGAACTTCAAGATCCGCCACAACATCGAGGACGGCAGCGTGCAGCTCGCCGACCACTACCAGCA
GAACACCCCCATCGGCGACGGCCCCGTGCTGCTGCCCGACAACCACTACCTGAGCACCCAGTCCGCCCTGAGCAAAGAC
CCCAACGAGAAGCGCGATCACATGGTCCTGCTGGAGTTCGTGACCGCCGCCGGGATCACTCTCGGCATGGACGAGCTGT
ACAAGTAA

Cypridina Luciferase

ATGAAGACCTTAATTCTTGCCGTTGCATTAGTCTACTGCGCCACTGTTCATTGCCAGGACTGTCCTTACGAACCTGATCCAC
CAAACACAGTTCCAACTTCCTGTGAAGCTAAAGAAGGAGAATGTATTGATAGCAGCTGTGGCACCTGCACGAGAGACATAC
TATCAGATGGACTGTGTGAAAATAAACCAGGAAAAACATGTTGCCGAATGTGTCAGTATGTAATTGAATGCAGAGTAGAGGC
CGCAGGATGGTTTAGAACATTCTATGGAAAGAGATTCCAGTTCCAGGAACCTGGTACATACGTGTTGGGTCAAGGAACCAA
GGGCGGCGACTGGAAGGTGTCCATCACCCTGGAGAACCTGGATGGAACCAAGGGGGCTGTGCTGACCAAGACAAGACT
GGAAGTGGCTGGAGACATCATTGACATCGCTCAAGCTACTGAGAATCCCATCACTGTAAACGGTGGAGCTGACCCTATCAT
CGCCAACCCGTACACCATCGGCGAGGTCACCATCGCTGTTGTTGAGATGCCAGGCTTCAACATCACCGTCATTGAGTTCTT
CAAACTGATCGTGATCGACATCCTCGGAGGAAGATCTGTAAGAATCGCCCCAGACACAGCAAACAAAGGAATGATCTCTGG
CCTCTGTGGAGATCTTAAAATGATGGAAGATACAGACTTCACTTCAGATCCAGAACAACTCGCTATTCAGCCTAAGATCAACC
AGGAGTTTGACGGTTGTCCACTCTATGGAAATCCTGATGACGTTGCATACTGCAAAGGTCTTCTGGAGCCGTACAAGGACA
GCTGCCGCAACCCCATCAACTTCTACTACTACACCATCTCCTGCGCCTTCGCCCGCTGTATGGGTGGAGACGAGCGAGCC
TCACACGTGCTGCTTGACTACAGGGAGACGTGCGCTGCTCCCGAAACTAGAGGAACCTGCGTTTTGTCTGGACATACTTTC
TACGATACATTTGACAAAGCAAGATACCAATTCCAGGGTCCCTGCAAGGAGATTCTTATGGCCGCCGACTGTTTCTGGAACA
CTTGGGATGTGAAGGTTTCACACAGGAATGTTGACTCTTACACTGAAGTAGAGAAAGTACGAATCAGGAAACAATCGACTGT
AGTAGAACTCATTGTTGATGGAAAACAGATTCTGGTTGGAGGAGAAGCCGTGTCCGTCCCGTACAGCTCTCAGAACACTTC
CATCTACTGGCAAGATGGTGACATACTGACTACAGCCATCCTACCTGAAGCTCTGGTGGTCAAGTTCAACTTCAAGCAACTG
CTCGTCGTACATATTAGAGATCCATTCGATGGTAAGACTTGCGGTATTTGCGGTAACTACAACCAGGATTTCAGTGATGATTC
TTTTGATGCTGAAGGAGCCTGTGATCTGACCCCCAACCCACCGGGATGCACCGAAGAACAGAAACCTGAAGCTGAACGAC
TCTGCAATAGTCTCTTCGCCGGTCAAAGTGATCTTGATCAGAAATGTAACGTGTGCCACAAGCCTGACCGTGTCGAACGAT
GCATGTACGAGTATTGCCTGAGGGGACAACAGGGTTTCTGTGACCACGCATGGGAGTTCAAGAAAGAATGCTACATAAAGC
ATGGAGACACCCTAGAAGTACCAGATGAATGCAAATAG

Sars-nCov2 spike protein

ATGGTGTTCGTGTTTCTGGTGCTGCTGCCTCTGGTGTCCAGCCAGTGTGTGAACCTGACCACAAGAACCCAGCTGCCTCC
AGCCTACACCAACAGCTTTACCAGAGGCGTGTACTACCCCGACAAGGTGTTCAGATCCAGCGTGCTGCACTCTACCCAGG
ACCTGTTCCTGCCTTTCTTCAGCAACGTGACCTGGTTCCACGCCATCCACGTGTCCGGCACCAATGGCACCAAGAGATTC
GACAACCCCGTGCTGCCCTTCAACGACGGGGTGTACTTTGCCAGCACCGAGAAGTCCAACATCATCAGAGGCTGGATCTT
CGGCACCACACTGGACAGCAAGACCCAGAGCCTGCTGATCGTGAACAACGCCACCAACGTGGTCATCAAAGTGTGCGAG
TTCCAGTTCTGCAACGACCCCTTCCTGGGCGTCTACTATCACAAGAACAACAAGAGCTGGATGGAAAGCGAGTTCCGGGT
GTACAGCAGCGCCAACAACTGCACCTTCGAGTACGTGTCCCAGCCTTTCCTGATGGACCTGGAAGGCAAGCAGGGCAACT
TCAAGAACCTGCGCGAGTTCGTGTTCAAGAACATCGACGGCTACTTCAAGATCTACAGCAAGCACACCCCTATCAACCTCG
TGCGGGATCTGCCTCAGGGCTTCTCTGCTCTGGAACCCCTGGTGGATCTGCCCATCGGCATCAACATCACCCGGTTTCAG
ACACTGCTGGCCCTGCACAGAAGCTACCTGACACCTGGCGATAGCAGCAGCGGATGGACAGCTGGTGCCGCCGCTTACTA
TGTGGGCTACCTGCAGCCTAGAACCTTTCTGCTGAAGTACAACGAGAACGGCACCATCACCGACGCCGTGGATTGTGCTC
TGGATCCTCTGAGCGAGACAAAGTGCACCCTGAAGTCCTTCACCGTGGAAAAGGGCATCTACCAGACCAGCAACTTCCGG
GTGCAGCCCACCGAATCCATCGTGCGGTTCCCCAATATCACCAATCTGTGCCCCTTCGGCGAGGTGTTCAATGCCACCAGA
TTCGCCTCTGTGTACGCCTGGAACCGGAAGCGGATCAGCAATTGCGTGGCCGACTACTCCGTGCTGTACAACTCCGCCAG
CTTCAGCACCTTCAAGTGCTACGGCGTGTCCCCTACCAAGCTGAACGACCTGTGCTTCACAAACGTGTACGCCGACAGCT
TCGTGATCCGGGGAGATGAAGTGCGGCAGATTGCCCCTGGACAGACAGGCAAGATCGCCGACTACAACTACAAGCTGCCC
GACGACTTCACCGGCTGTGTGATTGCCTGGAACAGCAACAACCTGGACTCCAAAGTCGGCGGCAACTACAATTACCTGTA
CCGGCTGTTCCGGAAGTCCAATCTGAAGCCCTTCGAGCGGGACATCTCCACCGAGATCTATCAGGCCGGCAGCACCCCTT
GTAACGGCGTGGAAGGCTTCAACTGCTACTTCCCACTGCAGTCCTACGGCTTTCAGCCCACAAATGGCGTGGGCTATCAG
CCCTACAGAGTGGTGGTGCTGAGCTTCGAACTGCTGCATGCCCCTGCCACAGTGTGCGGCCCTAAGAAAAGCACCAATCT
CGTGAAGAACAAATGCGTGAACTTCAACTTCAACGGCCTGACCGGCACCGGCGTGCTGACAGAGAGCAACAAGAAGTTCC
TGCCATTCCAGCAGTTTGGCCGGGATATCGCCGATACCACAGACGCCGTTAGAGATCCCCAGACACTGGAAATCCTGGACA
TCACCCCTTGCAGCTTCGGCGGAGTGTCTGTGATCACCCCTGGCACCAACACCAGCAATCAGGTGGCAGTGCTGTACCAG
GACGTGAACTGTACCGAAGTGCCCGTGGCCATTCACGCCGATCAGCTGACACCTACATGGCGGGTGTACTCCACCGGCAG
CAATGTGTTTCAGACCAGAGCCGGCTGTCTGATCGGAGCCGAGCACGTGAACAATAGCTACGAGTGCGACATCCCCATCG
GCGCTGGCATCTGTGCCAGCTACCAGACACAGACAAACAGCCCCGCCTCTGTGGCCAGCCAGAGCATCATTGCCTACACA
ATGTCTCTGGGCGCCGAGAACAGCGTGGCCTACTCCAACAACTCTATCGCTATCCCCACCAACTTCACCATCAGCGTGACC
ACAGAGATCCTGCCTGTGTCCATGACCAAGACCAGCGTGGACTGCACCATGTACATCTGCGGCGATTCCACCGAGTGCTC
CAACCTGCTGCTGCAGTACGGCAGCTTCTGCACCCAGCTGAATAGAGCCCTGACAGGGATCGCCGTGGAACAGGACAAG
AACACCCAAGAGGTGTTCGCCCAAGTGAAGCAGATCTACAAGACCCCTCCTATCAAGGACTTCGGCGGCTTCAATTTCAGC
CAGATTCTGCCCGATCCTAGCAAGCCCAGCAAGCGGAGCTTCATCGAGGACCTGCTGTTCAACAAAGTGACACTGGCCGA
CGCCGGCTTCATCAAGCAGTATGGCGATTGTCTGGGCGACATTGCCGCCAGGGATCTGATTTGCGCCCAGAAGTTTAACG
GACTGACAGTGCTGCCTCCTCTGCTGACCGATGAGATGATCGCCCAGTACACATCTGCCCTGCTGGCCGGCACAATCACA
AGCGGCTGGACATTTGGAGCTGGCGCCGCTCTGCAGATCCCCTTTGCTATGCAGATGGCCTACCGGTTCAACGGCATCGG

	AGTGACCCAGAATGTGCTGTACGAGAACCAGAAGCTGATCGCCAACCAGTTCAACAGCGCCATCGGCAAGATCCAGGACA GCCTGAGCAGCACAGCAAGCGCCCTGGGAAAGCTGCAGGACGTGGTCAACCAGAATGCCCAGGCACTGAACACCCTGGT CAAGCAGCTGTCCTCCAACTTCGGCGCCATCAGCTCTGTGCTGAACGATATCCTGAGCAGACTGGACcctcctGAAGCCGAG GTGCAGATCGACAGACTGATCACCGGAAGGCTGCAGTCCCTGCAGACCTACGTTACCCAGCAGCTGATCAGAGCCGCCG AGATTAGAGCCTCTGCCAATCTGGCCGCCACCAAGATGTCTGAGTGTGTGCTGGGCCAGAGCAAGAGAGTGGACTTTTGC GGCAAGGGCTACCACCTGATGAGCTTCCCTCAGTCTGCCCCTCACGGCGTGGTGTTTCTGCACGTGACATACGTGCCCGC TCAAGAGAAGAATTTCACCACCGCTCCAGCCATCTGCCACGACGGCAAAGCCCACTTTCCTAGAGAAGGCGTGTTCGTGT CCAACGGCACCCATTGGTTCGTGACCCAGCGGAACTTCTACGAGCCCCAGATCATCACCACCGACAACACCTTCGTGTCT GGCAACTGCGACGTCGTGATCGGCATTGTGAACAATACCGTGTACGACCCTCTGCAGCCCGAGCTGGACAGCTTCAAAGA GGAACTGGATAAGTACTTTAAGAACCACACAAGCCCCGACGTGGACCTGGGCGATATCAGCGGAATCAATGCCAGCGTCGT GAACATCCAGAAAGAGATCGACCGGCTGAACGAGGTGGCCAAGAATCTGAACGAGAGCCTGATCGACCTGCAAGAACTG GGGAAGTACGAGCAGTACATCAAGTGGCCCAGCGGCCGCTTGGTCCCACGTGGCTCACCCGGATCTGGATACATCCCGG AGGCCCCTAGGGACGGTCAAGCTTACGTGAGAAAGGACGGCGAATGGGTTCTGCTGTCGACCTTCTTGGGACATCATCAT CATCATCACCGAATTCAGTAA
Xenopus globin 5’ UTR	CTTGTTCTTTTTGCAGAAGCTCAGAATAAACGCTCAACTTTGG
Xenopus globin 3’ UTR	AACCAGCCTCAAGAACACCCGAATGGAGTCTCTAAGCTACATAATACCAACTTACACTTTACAAAATGTTGTCCCCCAAAATG TAGCCATTCGTATCTGCTCCTAATAAAAAGAAAGTTTCTTCAC
tRNA-Lysine-Full	GCCCGGATAGCTCAGGCGGGAGAGCAGCGGCCGAGTAATTTACGTCGACGGTGACGTCGATGGTTGCGGCCGCGGGTC CAGGGTTCAAGTCCCTGTTCGGGCGCCA
3'-tRNA-Lysine	CTCCCGGGTGGTTGCGGATATCCGGGTCCAGGGTTCAAGTCCCTGTTCGGGCGCCA
5'-tRNA-Lysine	GCCCGGATAGCTCAGACGGAAGAGCAGGATATCCGAGTAATTTACGCCCGGGAG
EMCV IRES	CCCCTCTCCCTCCCCCCCTAACGTTACTGGCCGAAGCCGCTTGGAATAAGGCCGGTGTGCGTTTGTCTATATGTTATTTTCC ACCATATTGCCGTCTTTTGGCAATGTGAGGGCCCGGAAACCTGGCCCTGTCTTCTTGACGAGCATTCCTAGGGGTCTTTCC CCTCTCGCCAAAGGAATGCAAGGTCTGTTGAATGTCGTGAAGGAAGCAGTTCCTCTGGAAGCTTCTTGAAGACAAACAAC GTCTGTAGCGACCCTTTGCAGGCAGCGGAACCCCCCACCTGGCGACAGGTGCCTCTGCGGCCAAAAGCCACGTGTATAA GATACACCTGCAAAGGCGGCACAACCCCAGTGCCACGTTGTGAGTTGGATAGTTGTGGAAAGAGTCAAATGGCTCTCCTC AAGCGTATTCAACAAGGGGCTGAAGGATGCCCAGAAGGTACCCCATTGTATGGGATCTGATCTGGGGCCTCGGTGCACAT GCTTTACATGTGTTTAGTCGAGGTTAAAAAACGTCTAGGCCCCCCGAACCACGGGGACGTGGTTTTCCTTTGAAAAACACG ATGATAATATGGCCACAACC
SelfCirc-5’	GGGAGACCCTCGACCGTCGATTGTCCACTGGTCAACAATAGATGACTTACAACTAATCGGAAGGTGCAGAGACTCGACGG GAGCTACCCTAACGTCAAGACGAGGGTAAAGAGAGAGTCCAATTCTCAAAGCCAATAGGCAGTAGCGAAAGCTGCAAGAG AATGAAAATCCGTTGACCTTAAACGGTCGTGTGGGTTCAAGTCCCTCCACCCCCACGCCGGAAACGCAATAGCCGAAAAAC AAAAAACAAAAAAACC
SelfCirc-3’	AAAAAACAAAAAACAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAACGGCTATTATGCGTTACCGGCGAGACGCTACGGA CTTAAATAATTGAGCCTTAAAGAAGAAATTCTTTAAGTGGATGCTCTCAAACTCAGGGAAACCTAAATCTAGTTATAGACAAGG CAATCCTGAGCCAAGCCGAAGTAGTAATTAGTAAGACCAGTGGACAATCGACGGATAACAGCATATCTAG
Triple Terminator	CTGAGCAATAACCTAGCATAAACCCCTTGGGTTCCCTCTTTAGGAGTCTGAGGGGTTTTTTGCTGAAAGATCAAATAAAACG AAAGGCTCAGTCGAAAGACTGGGCCTTTCGTTTTATCTGTTGTTTGTCGCTGAGCAATAACTAGCATAACCCCTTGGGGCCT CTAAACGGGTCTTGAGGGGTTTTTTG

Supplementary Table 4.2 – Plasmid Sequences used in this study

Control GFP Sequence	CTTGTTCTTTTTGCAGAAGCTCAGAATAAACGCTCAACTTTGGGCCGCCACCATGGTGAGCAAGGGCGAGGAGCTGTTCA CCGGGGTGGTGCCCATCCTGGTCGAGCTGGACGGCGACGTAAACGGCCACAAGTTCAGCGTGTCTGGCGAGGGCGAGG GCGATGCCACCTACGGCAAGCTGACCCTGAAGTTCATCTGCACCACCGGCAAGCTGCCCGTGCCCTGGCCCACCCTCGT GACCACCCTGACCTACGGCGTGCAGTGCTTCAGCCGCTACCCCGACCACATGAAGCAGCACGACTTCTTCAAGTCCGCCA TGCCCGAAGGCTACGTCCAGGAGCGCACCATCTTCTTCAAGGACGACGGCAACTACAAGACCCGCGCCGAGGTGAAGTT CGAGGGCGACACCCTGGTGAACCGCATCGAGCTGAAGGGCATCGACTTCAAGGAGGACGGCAACATCCTGGGGCACAAG CTGGAGTACAACTACAACAGCCACAACGTCTATATCATGGCCGACAAGCAGAAGAACGGCATCAAGGCGAACTTCAAGATC CGCCACAACATCGAGGACGGCAGCGTGCAGCTCGCCGACCACTACCAGCAGAACACCCCCATCGGCGACGGCCCCGTG CTGCTGCCCGACAACCACTACCTGAGCACCCAGTCCGCCCTGAGCAAAGACCCCAACGAGAAGCGCGATCACATGGTCC TGCTGGAGTTCGTGACCGCCGCCGGGATCACTCTCGGCATGGACGAGCTGTACAAGTAAAACCAGCCTCAAGAACACCCG AATGGAGTCTCTAAGCTACATAATACCAACTTACACTTTACAAAATGTTGTCCCCCAAAATGTAGCCATTCGTATCTGCTCCTA ATAAAAAGAAAGTTTCTTCACCTCGAGCACCACCACCACCACCACTGAGATCCGGCTGCTAACAAAGCCCGAAAGGAAGCT GAGTTGGCTGCTGCCACCGCTGAGCAATAACTAGCATAACCCCTTGGGGCCTCTAAACGGGTCTTGAGGGGTTTTTTGCTG

AAAGGAGGAACTATATCCGGATTGGCGAATGGGACGCGCCCTGTAGCGGCGCATTAAGCGCGGCGGGTGTGGTGGTTAC
GCGCAGCGTGACCGCTACACTTGCCAGCGCCCTAGCGCCCGCTCCTTTCGCTTTCTTCCCTTCCTTTCTCGCCACGTTCG
CCGGCTTTCCCCGTCAAGCTCTAAATCGGGGGCTCCCTTTAGGGTTCCGATTTAGTGCTTTACGGCACCTCGACCCCAAAA
AACTTGATTAGGGTGATGGTTCACGTAGTGGGCCATCGCCCTGATAGACGGTTTTTCGCCCTTTGACGTTGGAGTCCACGT
TCTTTAATAGTGGACTCTTGTTCCAAACTGGAACAACACTCAACCCTATCTCGGTCTATTCTTTTGATTTATAAGGGATTTTGC
CGATTTCGGCCTATTGGTTAAAAAATGAGCTGATTTAACAAAAATTTAACGCGAATTTTAACAAAATATTAACGCTTACAATTTA
GGTGGCACTTTTCGGGGAAATGTGCGCGGAACCCCTATTTGTTTATTTTTCTAAATACATTCAAATATGTATCCGCTCATGAAT
TAATTCTTAGAAAAACTCATCGAGCATCAAATGAAACTGCAATTTATTCATATCAGGATTATCAATACCATATTTTTGAAAAAGC
CGTTTCTGTAATGAAGGAGAAAACTCACCGAGGCAGTTCCATAGGATGGCAAGATCCTGGTATCGGTCTGCGATTCCGACT
CGTCCAACATCAATACAACCTATTAATTTCCCCTCGTCAAAAATAAGGTTATCAAGTGAGAAATCACCATGAGTGACGACTGA
ATCCGGTGAGAATGGCAAAAGTTTATGCATTTCTTTCCAGACTTGTTCAACAGGCCAGCCATTACGCTCGTCATCAAAATCA
CTCGCATCAACCAAACCGTTATTCATTCGTGATTGCGCCTGAGCGAGACGAAATACGCGATCGCTGTTAAAAGGACAATTAC
AAACAGGAATCGAATGCAACCGGCGCAGGAACACTGCCAGCGCATCAACAATATTTTCACCTGAATCAGGATATTCTTCTAA
TACCTGGAATGCTGTTTTCCCGGGGATCGCAGTGGTGAGTAACCATGCATCATCAGGAGTACGGATAAAATGCTTGATGGTC
GGAAGAGGCATAAATTCCGTCAGCCAGTTTAGTCTGACCATCTCATCTGTAACATCATTGGCAACGCTACCTTTGCCATGTTT
CAGAAACAACTCTGGCGCATCGGGCTTCCCATACAATCGATAGATTGTCGCACCTGATTGCCCGACATTATCGCGAGCCCAT
TTATACCCATATAAATCAGCATCCATGTTGGAATTTAATCGCGGCCTAGAGCAAGACGTTTCCCGTTGAATATGGCTCATAACA
CCCCTTGTATTACTGTTTATGTAAGCAGACAGTTTTATTGTTCATGACCAAAATCCCTTAACGTGAGTTTTCGTTCCACTGAGC
GTCAGACCCCGTAGAAAAGATCAAAGGATCTTCTTGAGATCCTTTTTTTCTGCGCGTAATCTGCTGCTTGCAAACAAAAAAA
CCACCGCTACCAGCGGTGGTTTGTTTGCCGGATCAAGAGCTACCAACTCTTTTTCCGAAGGTAACTGGCTTCAGCAGAGC
GCAGATACCAAATACTGTCCTTCTAGTGTAGCCGTAGTTAGGCCACCACTTCAAGAACTCTGTAGCACCGCCTACATACCTC
GCTCTGCTAATCCTGTTACCAGTGGCTGCTGCCAGTGGCGATAAGTCGTGTCTTACCGGGTTGGACTCAAGACGATAGTTA
CCGGATAAGGCGCAGCGGTCGGGCTGAACGGGGGGTTCGTGCACACAGCCCAGCTTGGAGCGAACGACCTACACCGAA
CTGAGATACCTACAGCGTGAGCTATGAGAAAGCGCCACGCTTCCCGAAGGGAGAAAGGCGGACAGGTATCCGGTAAGCG
GCAGGGTCGGAACAGGAGAGCGCACGAGGGAGCTTCCAGGGGGAAACGCCTGGTATCTTTATAGTCCTGTCGGGTTTCG
CCACCTCTGACTTGAGCGTCGATTTTTGTGATGCTCGTCAGGGGGGCGGAGCCTATGGAAAAACGCCAGCAACGCGGCCT
TTTTACGGTTCCTGGCCTTTTGCTGGCCTTTTGCTCACATGTTCTTTCCTGCGTTATCCCCTGATTCTGTGGATAACCGTATT
ACCGCCTTTGAGTGAGCTGATACCGCTCGCCGCAGCCGAACGACCGAGCGCAGCGAGTCAGTGAGCGAGGAAGCGGAA
GAGCGCCTGATGCGGTATTTTCTCCTTACGCATCTGTGCGGTATTTCACACCGCAATGGTGCACTCTCAGTACAATCTGCTC
TGATGCCGCATAGTTAAGCCAGTATACACTCCGCTATCGCTACGTGACTGGGTCATGGCTGCGCCCCGACACCCGCCAACA
CCCGCTGACGCGCCCTGACGGGCTTGTCTGCTCCCGGCATCCGCTTACAGACAAGCTGTGACCGTCTCCGGGAGCTGCA
TGTGTCAGAGGTTTTCACCGTCATCACCGAAACGCGCGAGGCAGCTGCGGTAAAGCTCATCAGCGTGGTCGTGAAGCGAT
TCACAGATGTCTGCCTGTTCATCCGCGTCCAGCTCGTTGAGTTTCTCCAGAAGCGTTAATGTCTGGCTTCTGATAAAGCGG
GCCATGTTAAGGGCGGTTTTTTCCTGTTTGGTCACTGATGCCTCCGTGTAAGGGGGATTTCTGTTCATGGGGGTAATGATAC
CGATGAAACGAGAGAGGATGCTCACGATACGGGTTACTGATGATGAACATGCCCGGTTACTGGAACGTTGTGAGGGTAAAC
AACTGGCGGTATGGATGCGGCGGGACCAGAGAAAAATCACTCAGGGTCAATGCCAGCGCTTCGTTAATACAGATGTAGGTG
TTCCACAGGGTAGCCAGCAGCATCCTGCGATGCAGATCCGGAACATAATGGTGCAGGGCGCTGACTTCCGCGTTTCCAGA
CTTTACGAAACACGGAAACCGAAGACCATTCATGTTGTTGCTCAGGTCGCAGACGTTTTGCAGCAGCAGTCGCTTCACGTT
CGCTCGCGTATCGGTGATTCATTCTGCTAACCAGTAAGGCAACCCCGCCAGCCTAGCCGGGTCCTCAACGACAGGAGCAC
GATCATGCGCACCCGTGGGGCCGCCATGCCGGCGATAATGGCCTGCTTCTCGCCGAAACGTTTGGTGGCGGGACCAGTG
ACGAAGGCTTGAGCGAGGGCGTGCAAGATTCCGAATACCGCAAGCGACAGGCCGATCATCGTCGCGCTCCAGCGAAAGC
GGTCCTCGCCGAAAATGACCCAGAGCGCTGCCGGCACCTGTCCTACGAGTTGCATGATAAAGAAGACAGTCATAAGTGCG
GCGACGATAGTCATGCCCCGCGCCCACCGGAAGGAGCTGACTGGGTTGAAGGCTCTCAAGGGCATCGGTCGAGATCCCG
GTGCCTAATGAGTGAGCTAACTTACATTAATTGCGTTGCGCTCACTGCCCGCTTTCCAGTCGGGAAACCTGTCGTGCCAGC
TGCATTAATGAATCGGCCAACGCGCGGGGAGAGGCGGTTTGCGTATTGGGCGCCAGGGTGGTTTTTCTTTTCACCAGTGA
GACGGGCAACAGCTGATTGCCCTTCACCGCCTGGCCCTGAGAGAGTTGCAGCAAGCGGTCCACGCTGGTTTGCCCCAGC
AGGCGAAAATCCTGTTTGATGGTGGTTAACGGCGGGATATAACATGAGCTGTCTTCGGTATCGTCGTATCCCACTACCGAGA
TATCCGCACCAACGCGCAGCCCGGACTCGGTAATGGCGCGCATTGCGCCCAGCGCCATCTGATCGTTGGCAACCAGCATC
GCAGTGGGAACGATGCCCTCATTCAGCATTTGCATGGTTTGTTGAAAACCGGACATGGCACTCCAGTCGCCTTCCCGTTCC
GCTATCGGCTGAATTTGATTGCGAGTGAGATATTTATGCCAGCCAGCCAGACGCAGACGCGCCGAGACAGAACTTAATGGG
CCCGCTAACAGCGCGATTTGCTGGTGACCCAATGCGACCAGATGCTCCACGCCCAGTCGCGTACCGTCTTCATGGGAGAA
AATAATACTGTTGATGGGTGTCTGGTCAGAGACATCAAGAAATAACGCCGGAACATTAGTGCAGGCAGCTTCCACAGCAATG
GCATCCTGGTCATCCAGCGGATAGTTAATGATCAGCCCACTGACGCGTTGCGCGAGAAGATTGTGCACCGCCGCTTTACAG
GCTTCGACGCCGCTTCGTTCTACCATCGACACCACCACGCTGGCACCCAGTTGATCGGCGCGAGATTTAATCGCCGCGAC
AATTTGCGACGGCGCGTGCAGGGCCAGACTGGAGGTGGCAACGCCAATCAGCAACGACTGTTTGCCCGCCAGTTGTTGT
GCCACGCGGTTGGGAATGTAATTCAGCTCCGCCATCGCCGCTTCCACTTTTTCCCGCGTTTTCGCAGAAACGTGGCTGGC
CTGGTTCACCACGCGGGAAACGGTCTGATAAGAGACACCGGCATACTCTGCGACATCGTATAACGTTACTGGTTTCACATTC
ACCACCCTGAATTGACTCTCTTCCGGGCGCTATCATGCCATACCGCGAAAGGTTTTGCGCCATTCGATGGTGTCCGGGATC
TCGACGCTCTCCCTTATGCGACTCCTGCATTAGGAAGCAGCCCAGTAGTAGGTTGAGGCCGTTGAGCACCGCCGCCGCAA
GGAATGGTGCATGCAAGGAGATGGCGCCCAACAGTCCCCCGGCCACGGGGCCTGCCACCATACCCACGCCGAAACAAGC
GCTCATGAGCCCGAAGTGGCGAGCCCGATCTTCCCCATCGGTGATGTCGGCGATATAGGCGCCAGCAACCGCACCTGTG
GCGCCGGTGATGCCGGCCACGATGCGTCCGGCGTAGAGGATCGAGATCTCGATCCCGCGAAATTAATACGACTCACTATAG
GGGAATTGTGAGCGGATAACAATTCCCCTCTAGGATCC

SelfCirc-GFP

GGGAGACCCTCGACCGTCGATTGTCCACTGGTCAACAATAGATGACTTACAACTAATCGGAAGGTGCAGAGACTCGACGG
GAGCTACCCTAACGTCAAGACGAGGGTAAAGAGAGAGTCCAATTCTCAAAGCCAATAGGCAGTAGCGAAAGCTGCAAGAG
AATGAAAATCCGTTGACCTTAAACGGTCGTGTGGGTTCAAGTCCCTCCACCCCCACGCCGGAAACGCAATAGCCGAAAAAC
AAAAAACAAAAAAACCCCCCTCTCCCTCCCCCCCTAACGTTACTGGCCGAAGCCGCTTGGAATAAGGCCGGTGTGCGTTT
GTCTATATGTTATTTTCCACCATATTGCCGTCTTTTGGCAATGTGAGGGCCCGGAAACCTGGCCCTGTCTTCTTGACGAGCA
TTCCTAGGGGTCTTTCCCCTCTCGCCAAAGGAATGCAAGGTCTGTTGAATGTCGTGAAGGAAGCAGTTCCTCTGGAAGCTT
CTTGAAGACAAACAACGTCTGTAGCGACCCTTTGCAGGCAGCGGAACCCCCCACCTGGCGACAGGTGCCTCTGCGGCCA
AAAGCCACGTGTATAAGATACACCTGCAAAGGCGGCACAACCCCAGTGCCACGTTGTGAGTTGGATAGTTGTGGAAAGAGT
CAAATGGCTCTCCTCAAGCGTATTCAACAAGGGGCTGAAGGATGCCCAGAAGGTACCCCATTGTATGGGATCTGATCTGGG
GCCTCGGTGCACATGCTTTACATGTGTTTAGTCGAGGTTAAAAAACGTCTAGGCCCCCCGAACCACGGGGACGTGGTTTTC
CTTTGAAAAACACGATGATAATATGGCCACAACCATGGTGAGCAAGGGCGAGGAGCTGTTCACCGGGGTGGTGCCCATCC
TGGTCGAGCTGGACGGCGACGTAAACGGCCACAAGTTCAGCGTGTCTGGCGAGGGCGAGGGCGATGCCACCTACGGCA
AGCTGACCCTGAAGTTCATCTGCACCACCGGCAAGCTGCCCGTGCCCTGGCCCACCCTCGTGACCACCCTGACCTACGG
CGTGCAGTGCTTCAGCCGCTACCCCGACCACATGAAGCAGCACGACTTCTTCAAGTCCGCCATGCCCGAAGGCTACGTCC
AGGAGCGCACCATCTTCTTCAAGGACGACGGCAACTACAAGACCCGCGCCGAGGTGAAGTTCGAGGGCGACACCCTGGT
GAACCGCATCGAGCTGAAGGGCATCGACTTCAAGGAGGACGGCAACATCCTGGGGCACAAGCTGGAGTACAACTACAACA

GCCACAACGTCTATATCATGGCCGACAAGCAGAAGAACGGCATCAAGGCGAACTTCAAGATCCGCCACAACATCGAGGAC
GGCAGCGTGCAGCTCGCCGACCACTACCAGCAGAACACCCCCATCGGCGACGGCCCCGTGCTGCTGCCCGACAACCAC
TACCTGAGCACCCAGTCCGCCCTGAGCAAAGACCCCAACGAGAAGCGCGATCACATGGTCCTGCTGGAGTTCGTGACCG
CCGCCGGGATCACTCTCGGCATGGACGAGCTGTACAAGTAAAAAAAACAAAAAACAAAACGGCTATTATGCGTTACCGGCG
AGACGCTACGGACTTAAATAATTGAGCCTTAAAGAAGAAATTCTTTAAGTGGATGCTCTCAAACTCAGGGAAACCTAAATCTA
GTTATAGACAAGGCAATCCTGAGCCAAGCCGAAGTAGTAATTAGTAAGACCAGTGGACAATCGACGGATAACAGCATATCTA
GCTCGAGCACCACCACCACCACCACCACTGAGATCCGGCTGCTAACAAAGCCCGAAAGGAAGCTGAGTTGGCTGCTGCCACC
GCTGAGCAATAACTAGCATAACCCCTTGGGGCCTCTAAACGGGTCTTGAGGGGTTTTTTGCTGAAAGGAGGAACTATATCC
GGATTGGCGAATGGGACGCGCCCTGTAGCGGCGCATTAAGCGCGGCGGGTGTGGTGGTTACGCGCAGCGTGACCGCTAC
ACTTGCCAGCGCCCTAGCGCCCGCTCCTTTCGCTTTCTTCCCTTCCTTTCTCGCCACGTTCGCCGGCTTTCCCCGTCAAG
CTCTAAATCGGGGGCTCCCTTTAGGGTTCCGATTTAGTGCTTTACGGCACCTCGACCCCAAAAAACTTGATTAGGGTGATGG
TTCACGTAGTGGGCCATCGCCCTGATAGACGGTTTTTCGCCCTTTGACGTTGGAGTCCACGTTCTTTAATAGTGGACTCTTG
TTCCAAACTGGAACAACACTCAACCCTATCTCGGTCTATTCTTTTGATTTATAAGGGATTTTGCCGATTTCGGCCTATTGGTTA
AAAAATGAGCTGATTTAACAAAAATTTAACGCGAATTTTAACAAAATATTAACGCTTACAATTTAGGTGGCACTTTTCGGGGAA
ATGTGCGCGGAACCCCTATTTGTTTATTTTTCTAAATACATTCAAATATGTATCCGCTCATGAATTAATTCTTAGAAAAACTCATC
GAGCATCAAATGAAACTGCAATTTATTCATATCAGGATTATCAATACCATATTTTTGAAAAAGCCGTTTCTGTAATGAAGGAGAA
AACTCACCGAGGCAGTTCCATAGGATGGCAAGATCCTGGTATCGGTCTGCGATTCCGACTCGTCCAACATCAATACAACCTA
TTAATTTCCCCTCGTCAAAAATAAGGTTATCAAGTGAGAAATCACCATGAGTGACGACTGAATCCGGTGAGAATGGCAAAAG
TTTATGCATTTCTTTCCAGACTTGTTCAACAGGCCAGCCATTACGCTCGTCATCAAAATCACTCGCATCAACCAAACCGTTAT
TCATTCGTGATTGCGCCTGAGCGAGACGAAATACGCGATCGCTGTTAAAAGGACAATTACAAACAGGAATCGAATGCAACC
GGCGCAGGAACACTGCCAGCGCATCAACAATATTTTCACCTGAATCAGGATATTCTTCTAATACCTGGAATGCTGTTTTCCCG
GGGATCGCAGTGGTGAGTAACCATGCATCATCAGGAGTACGGATAAAATGCTTGATGGTCGGAAGAGGCATAAATTCCGTC
AGCCAGTTTAGTCTGACCATCTCATCTGTAACATCATTGGCAACGCTACCTTTGCCATGTTTCAGAAACAACTCTGGCGCAT
CGGGCTTCCCATACAATCGATAGATTGTCGCACCTGATTGCCCGACATTATCGCGAGCCCATTTATACCCATATAAATCAGCA
TCCATGTTGGAATTTAATCGCGGCCTAGAGCAAGACGTTTCCCGTTGAATATGGCTCATAACACCCCTTGTATTACTGTTTAT
GTAAGCAGACAGTTTTATTGTTCATGACCAAAATCCCTTAACGTGAGTTTTCGTTCCACTGAGCGTCAGACCCCGTAGAAAA
GATCAAAGGATCTTCTTGAGATCCTTTTTTTCTGCGCGTAATCTGCTGCTTGCAAACAAAAAAACCACCGCTACCAGCGGTG
GTTTGTTTGCCGGATCAAGAGCTACCAACTCTTTTTCCGAAGGTAACTGGCTTCAGCAGAGCGCAGATACCAAATACTGTCC
TTCTAGTGTAGCCGTAGTTAGGCCACCACTTCAAGAACTCTGTAGCACCGCCTACATACCTCGCTCTGCTAATCCTGTTACC
AGTGGCTGCTGCCAGTGGCGATAAGTCGTGTCTTACCGGGTTGGACTCAAGACGATAGTTACCGGATAAGGCGCAGCGGT
CGGGCTGAACGGGGGGTTCGTGCACACAGCCCAGCTTGGAGCGAACGACCTACACCGAACTGAGATACCTACAGCGTGA
GCTATGAGAAAGCGCCACGCTTCCCGAAGGGAGAAAGGCGGACAGGTATCCGGTAAGCGGCAGGGTCGGAACAGGAGA
GCGCACGAGGGAGCTTCCAGGGGGAAACGCCTGGTATCTTTATAGTCCTGTCGGGTTTCGCCACCTCTGACTTGAGCGTC
GATTTTTGTGATGCTCGTCAGGGGGGCGGAGCCTATGGAAAAACGCCAGCAACGCGGCCTTTTTACGGTTCCTGGCCTTT
TGCTGGCCTTTTGCTCACATGTTCTTTCCTGCGTTATCCCCTGATTCTGTGGATAACCGTATTACCGCCTTTGAGTGAGCTGA
TACCGCTCGCCGCAGCCGAACGACCGAGCGCAGCGAGTCAGTGAGCGAGGAAGCGGAAGAGCGCCTGATGCGGTATTTT
CTCCTTACGCATCTGTGCGGTATTTCACACCGCAATGGTGCACTCTCAGTACAATCTGCTCTGATGCCGCATAGTTAAGCCA
GTATACACTCCGCTATCGCTACGTGACTGGGTCATGGCTGCGCCCCGACACCCGCCAACACCCGCTGACGCGCCCTGACG
GGCTTGTCTGCTCCCGGCATCCGCTTACAGACAAGCTGTGACCGTCTCCGGGAGCTGCATGTGTCAGAGGTTTTCACCGT
CATCACCGAAACGCGCGAGGCAGCTGCGGTAAAGCTCATCAGCGTGGTCGTGAAGCGATTCACAGATGTCTGCCTGTTCA
TCCGCGTCCAGCTCGTTGAGTTTCTCCAGAAGCGTTAATGTCTGGCTTCTGATAAAGCGGGCCATGTTAAGGGCGGTTTTT
TCCTGTTTGGTCACTGATGCCTCCGTGTAAGGGGGATTTCTGTTCATGGGGGTAATGATACCGATGAAACGAGAGAGGATG
CTCACGATACGGGTTACTGATGATGAACATGCCCGGTTACTGGAACGTTGTGAGGGTAAACAACTGGCGGTATGGATGCGG
CGGGACCAGAGAAAAATCACTCAGGGTCAATGCCAGCGCTTCGTTAATACAGATGTAGGTGTTCCACAGGGTAGCCAGCA
GCATCCTGCGATGCAGATCCGGAACATAATGGTGCAGGGCGCTGACTTCCGCGTTTCCAGACTTTACGAAACACGGAAAC
CGAAGACCATTCATGTTGTTGCTCAGGTCGCAGACGTTTTGCAGCAGCAGTCGCTTCACGTTCGCTCGCGTATCGGTGATT
CATTCTGCTAACCAGTAAGGCAACCCCGCCAGCCTAGCCGGGTCCTCAACGACAGGAGCACGATCATGCGCACCCGTGG
GGCCGCCATGCCGGCGATAATGGCCTGCTTCTCGCCGAAACGTTTGGTGGCGGGACCAGTGACGAAGGCTTGAGCGAGG
GCGTGCAAGATTCCGAATACCGCAAGCGACAGGCCGATCATCGTCGCGCTCCAGCGAAAGCGGTCCTCGCCGAAAATGAC
CCAGAGCGCTGCCGGCACCTGTCCTACGAGTTGCATGATAAAGAAGACAGTCATAAGTGCGGCGACGATAGTCATGCCCC
GCGCCCACCGGAAGGAGCTGACTGGGTTGAAGGCTCTCAAGGGCATCGGTCGAGATCCCGGTGCCTAATGAGTGAGCTA
ACTTACATTAATTGCGTTGCGCTCACTGCCCGCTTTCCAGTCGGGAAACCTGTCGTGCCAGCTGCATTAATGAATCGGCCA
ACGCGCGGGGAGAGGCGGTTTGCGTATTGGGCGCCAGGGTGGTTTTTCTTTTCACCAGTGAGACGGGCAACAGCTGATT
GCCCTTCACCGCCTGGCCCTGAGAGAGTTGCAGCAAGCGGTCCACGCTGGTTTGCCCCAGCAGGCGAAAATCCTGTTTG
ATGGTGGTTAACGGCGGGATATAACATGAGCTGTCTTCGGTATCGTCGTATCCCACTACCGAGATATCCGCACCAACGCGCA
GCCCGGACTCGGTAATGGCGCGCATTGCGCCCAGCGCCATCTGATCGTTGGCAACCAGCATCGCAGTGGGAACGATGCC
CTCATTCAGCATTTGCATGGTTTGTTGAAAACCGGACATGGCACTCCAGTCGCCTTCCCGTTCCGCTATCGGCTGAATTTGA
TTGCGAGTGAGATATTTATGCCAGCCAGCCAGACGCAGACGCGCCGAGACAGAACTTAATGGGCCCGCTAACAGCGCGAT
TTGCTGGTGACCCAATGCGACCAGATGCTCCACGCCCAGTCGCGTACCGTCTTCATGGGAGAAAATAATACTGTTGATGGG
TGTCTGGTCAGAGACATCAAGAAATAACGCCGGAACATTAGTGCAGGCAGCTTCCACAGCAATGGCATCCTGGTCATCCAG
CGGATAGTTAATGATCAGCCCACTGACGCGTTGCGCGAGAAGATTGTGCACCGCCGCTTTACAGGCTTCGACGCCGCTTC
GTTCTACCATCGACACCACCACGCTGGCACCCAGTTGATCGGCGCGAGATTTAATCGCCGCGACAATTTGCGACGGCGCG
TGCAGGGCCAGACTGGAGGTGGCAACGCCAATCAGCAACGACTGTTTGCCCGCCAGTTGTTGTGCCACGCGGTTGGGAA
TGTAATTCAGCTCCGCCATCGCCGCTTCCACTTTTTCCCGCGTTTTCGCAGAAACGTGGCTGGCCTGGTTCACCACGCGG
GAAACGGTCTGATAAGAGACACCGGCATACTCTGCGACATCGTATAACGTTACTGGTTTCACATTCACCACCCTGAATTGAC
TCTCTTCCGGGGCGCTATCATGCCATACCGCGAAAGGTTTTGCGCCATTCGATGGTGTCCGGGATCTCGACGCTCTCCCTTA
TGCGACTCCTGCATTAGGAAGCAGCCCAGTAGTAGGTTGAGGCCGTTGAGCACCGCCGCCGCAAGGAATGGTGCATGCAA
GGAGATGGCGCCCAACAGTCCCCCGGCCACGGGGCCTGCCACCATACCCACGCCGAAACAAGCGCTCATGAGCCCGAA
GTGGCGAGCCCGATCTTCCCCATCGGTGATGTCGGCGATATAGGCGCCAGCAACCGCACCTGTGGCGCCGGTGATGCCG
GCCACGATGCGTCCGGCGTAGAGGATCGAGATCTCGATCCCGCGAAATTAATACGACTCACTATAGGGGAATTGTGAGCGG
ATAACAATTCCCCTCTAGGATCC

SelfCirc-Luc

CTCGAGCACCACCACCACCACCACTGAGATCCGGCTGCTAACAAAGCCCGAAAGGAAGCTGAGTTGGCTGCTGCCACCG
CTGAGCAATAACTAGCATAACCCCTTGGGGCCTCTAAACGGGTCTTGAGGGGTTTTTTGCTGAAAGGAGGAACTATATCCG
GATTGGCGAATGGGACGCGCCCTGTAGCGGCGCATTAAGCGCGGCGGGTGTGGTGGTTACGCGCAGCGTGACCGCTACA
CTTGCCAGCGCCCTAGCGCCCGCTCCTTTCGCTTTCTTCCCTTCCTTTCTCGCCACGTTCGCCGGCTTTCCCCGTCAAGCT
CTAAATCGGGGGCTCCCTTTAGGGTTCCGATTTAGTGCTTTACGGCACCTCGACCCCAAAAAACTTGATTAGGGTGATGGTT
CACGTAGTGGGCCATCGCCCTGATAGACGGTTTTTCGCCCTTTGACGTTGGAGTCCACGTTCTTTAATAGTGGACTCTTGTT
CCAAACTGGAACAACACTCAACCCTATCTCGGTCTATTCTTTTGATTTATAAGGGATTTTGCCGATTTCGGCCTATTGGTTAAA
AAATGAGCTGATTTAACAAAAATTTAACGCGAATTTTAACAAAATATTAACGCTTACAATTTAGGTGGCACTTTTCGGGGAAAT

GTGCGCGGAACCCCTATTTGTTTATTTTTCTAAATACATTCAAATATGTATCCGCTCATGAATTAATTCTTAGAAAAACTCATCG
AGCATCAAATGAAACTGCAATTTATTCATATCAGGATTATCAATACCATATTTTTGAAAAAGCCGTTTCTGTAATGAAGGAGAAA
ACTCACCGAGGCAGTTCCATAGGATGGCAAGATCCTGGTATCGGTCTGCGATTCCGACTCGTCCAACATCAATACAACCTAT
TAATTTCCCCTCGTCAAAAATAAGGTTATCAAGTGAGAAATCACCATGAGTGACGACTGAATCCGGTGAGAATGGCAAAAGT
TTATGCATTTCTTTCCAGACTTGTTCAACAGGCCAGCCATTACGCTCGTCATCAAAATCACTCGCATCAACCAAACCGTTATT
CATTCGTGATTGCGCCTGAGCGAGACGAAATACGCGATCGCTGTTAAAAGGACAATTACAAACAGGAATCGAATGCAACCG
GCGCAGGAACACTGCCAGCGCATCAACAATATTTTCACCTGAATCAGGATATTCTTCTAATACCTGGAATGCTGTTTTCCCGG
GGATCGCAGTGGTGAGTAACCATGCATCATCAGGAGTACGGATAAAATGCTTGATGGTCGGAAGAGGCATAAATTCCGTCA
GCCAGTTTAGTCTGACCATCTCATCTGTAACATCATTGGCAACGCTACCTTTGCCATGTTTCAGAAACAACTCTGGCGCATC
GGGCTTCCCATACAATCGATAGATTGTCGCACCTGATTGCCCGACATTATCGCGAGCCCATTTATACCCATATAAATCAGCAT
CCATGTTGGAATTTAATCGCGGCCTAGAGCAAGACGTTTCCCGTTGAATATGGCTCATAACACCCCTTGTATTACTGTTTATG
TAAGCAGACAGTTTTATTGTTCATGACCAAAATCCCTTAACGTGAGTTTTCGTTCCACTGAGCGTCAGACCCCGTAGAAAAG
ATCAAAGGATCTTCTTGAGATCCTTTTTTTCTGCGCGTAATCTGCTGCTTGCAAACAAAAAAACCACCGCTACCAGCGGTGG
TTTGTTTGCCGGATCAAGAGCTACCAACTCTTTTTCCGAAGGTAACTGGCTTCAGCAGAGCGCAGATACCAAATACTGTCCT
TCTAGTGTAGCCGTAGTTAGGCCACCACTTCAAGAACTCTGTAGCACCGCCTACATACCTCGCTCTGCTAATCCTGTTACCA
GTGGCTGCTGCCAGTGGCGATAAGTCGTGTCTTACCGGGTTGGACTCAAGACGATAGTTACCGGATAAGGCGCAGCGGTC
GGGCTGAACGGGGGGTTCGTGCACACAGCCCAGCTTGGAGCGAACGACCTACACCGAACTGAGATACCTACAGCGTGAG
CTATGAGAAAGCGCCACGCTTCCCGAAGGGAGAAAGGCGGACAGGTATCCGGTAAGCGGCAGGGTCGGAACAGGAGAG
CGCACGAGGGAGCTTCCAGGGGGAAACGCCTGGTATCTTTATAGTCCTGTCGGGTTTCGCCACCTCTGACTTGAGCGTCG
ATTTTTGTGATGCTCGTCAGGGGGGCGGAGCCTATGGAAAAACGCCAGCAACGCGGCCTTTTTACGGTTCCTGGCCTTTT
GCTGGCCTTTTGCTCACATGTTCTTTCCTGCGTTATCCCCTGATTCTGTGGATAACCGTATTACCGCCTTTGAGTGAGCTGAT
ACCGCTCGCCGCAGCCGAACGACCGAGCGCAGCGAGTCAGTGAGCGAGGAAGCGGAAGAGCGCCTGATGCGGTATTTT
CTCCTTACGCATCTGTGCGGTATTTCACACCGCAATGGTGCACTCTCAGTACAATCTGCTCTGATGCCGCATAGTTAAGCCA
GTATACACTCCGCTATCGCTACGTGACTGGGTCATGGCTGCGCCCCGACACCCGCCAACACCCGCTGACGCGCCCTGACG
GGCTTGTCTGCTCCCGGCATCCGCTTACAGACAAGCTGTGACCGTCTCCGGGAGCTGCATGTGTCAGAGGTTTTCACCGT
CATCACCGAAACGCGCGAGGCAGCTGCGGTAAAGCTCATCAGCGTGGTCGTGAAGCGATTCACAGATGTCTGCCTGTTCA
TCCGCGTCCAGCTCGTTGAGTTTCTCCAGAAGCGTTAATGTCTGGCTTCTGATAAAGCGGGCCATGTTAAGGGCGGTTTTT
TCCTGTTTGGTCACTGATGCCTCCGTGTAAGGGGGATTTCTGTTCATGGGGGTAATGATACCGATGAAACGAGAGAGGATG
CTCACGATACGGGTTACTGATGATGAACATGCCCGGTTACTGGAACGTTGTGAGGGTAAACAACTGGCGGTATGGATGCGG
CGGGACCAGAGAAAAATCACTCAGGGTCAATGCCAGCGCTTCGTTAATACAGATGTAGGTGTTCCACAGGGTAGCCAGCA
GCATCCTGCGATGCAGATCCGGAACATAATGGTGCAGGGCGCTGACTTCCGCGTTTCCAGACTTTACGAAACACGGAAAC
CGAAGACCATTCATGTTGTTGCTCAGGTCGCAGACGTTTTGCAGCAGCAGTCGCTTCACGTTCGCTCGCGTATCGGTGATT
CATTCTGCTAACCAGTAAGGCAACCCCGCCAGCCTAGCCGGGTCCTCAACGACAGGAGCACGATCATGCGCACCCGTGG
GGCCGCCATGCCGGCGATAATGGCCTGCTTCTCGCCGAAACGTTTGGTGGCGGGACCAGTGACGAAGGCTTGAGCGAGG
GCGTGCAAGATTCCGAATACCGCAAGCGACAGGCCGATCATCGTCGCGCTCCAGCGAAAGCGGTCCTCGCCGAAAATGAC
CCAGAGCGCTGCCGGCACCTGTCCTACGAGTTGCATGATAAAGAAGACAGTCATAAGTGCGGCGACGATAGTCATGCCCC
GCGCCCACCGGAAGGAGCTGACTGGGTTGAAGGCTCTCAAGGGCATCGGTCGAGATCCCGGTGCCTAATGAGTGAGCTA
ACTTACATTAATTGCGTTGCGCTCACTGCCCGCTTTCCAGTCGGGAAACCTGTCGTGCCAGCTGCATTAATGAATCGGCCA
ACGCGCGGGGAGAGGCGGTTTGCGTATTGGGCGCCAGGGTGGTTTTTCTTTTCACCAGTGAGACGGGCAACAGCTGATT
GCCCTTCACCGCCTGGCCCTGAGAGAGTTGCAGCAAGCGGTCCACGCTGGTTTGCCCCAGCAGGCGAAAATCCTGTTTG
ATGGTGGTTAACGGCGGGATATAACATGAGCTGTCTTCGGTATCGTCGTATCCCACTACCGAGATATCCGCACCAACGCGCA
GCCCGGACTCGGTAATGGCGCGCATTGCGCCCAGCGCCATCTGATCGTTGGCAACCAGCATCGCAGTGGGAACGATGCC
CTCATTCAGCATTTGCATGGTTTGTTGAAAACCGGACATGGCACTCCAGTCGCCTTCCCGTTCCGCTATCGGCTGAATTTGA
TTGCGAGTGAGATATTTATGCCAGCCAGCCAGACGCAGACGCGCCGAGACAGAACTTAATGGGCCCGCTAACAGCGCGAT
TTGCTGGTGACCCAATGCGACCAGATGCTCCACGCCCAGTCGCGTACCGTCTTCATGGGAGAAAATAATACTGTTGATGGG
TGTCTGGTCAGAGACATCAAGAAATAACGCCGGAACATTAGTGCAGGCAGCTTCCACAGCAATGGCATCCTGGTCATCCAG
CGGATAGTTAATGATCAGCCCACTGACGCGTTGCGCGAGAAGATTGTGCACCGCCGCTTTACAGGCTTCGACGCCGCTTC
GTTCTACCATCGACACCACCACGCTGGCACCCAGTTGATCGGCGCGAGATTTAATCGCCGCGACAATTTGCGACGGCGCG
TGCAGGGCCAGACTGGAGGTGGCAACGCCAATCAGCAACGACTGTTTGCCCGCCAGTTGTTGTGCCACGCGGTTGGGAA
TGTAATTCAGCTCCGCCATCGCCGCTTCCACTTTTTCCCGCGTTTTCGCAGAAACGTGGCTGGCCTGGTTCACCACGCGG
GAAACGGTCTGATAAGAGACACCGGCATACTCTGCGACATCGTATAACGTTACTGGTTTCACATTCACCACCCTGAATTGAC
TCTCTTCCGGGCGCTATCATGCCATACCGCGAAAGGTTTTGCGCCATTCGATGGTGTCCGGGATCTCGACGCTCTCCCTTA
TGCGACTCCTGCATTAGGAAGCAGCCCAGTAGTAGGTTGAGGCCGTTGAGCACCGCCGCCGCAAGGAATGGTGCATGCAA
GGAGATGGCGCCCAACAGTCCCCCGGCCACGGGGCCTGCCACCATACCCACGCCGAAACAAGCGCTCATGAGCCCGAA
GTGGCGAGCCCGATCTTCCCCATCGGTGATGTCGGCGATATAGGCGCCAGCAACCGCACCTGTGGCGCCGGTGATGCCG
GCCACGATGCGTCCGGCGTAGAGGATCGAGATCTCGATCCCGCGAAATTAATACGACTCACTATAGGGGAATTGTGAGCGG
ATAACAATTCCCCTCTAGGATCCGGGAGACCCTCGACCGTCGATTGTCCACTGGTCAACAATAGATGACTTACAACTAATCG
GAAGGTGCAGAGACTCGACGGGAGCTACCCTAACGTCAAGACGAGGGTAAAGAGAGAGTCCAATTCTCAAAGCCAATAGG
CAGTAGCGAAAGCTGCAAGAGAATGAAAATCCGTTGACCTTAAACGGTCGTGTGGGTTCAAGTCCCTCCACCCCCACGCC
GGAAACGCAATAGCCGAAAAACAAAAAACAAAAAAACCCCCCCTCTCCCTCCCCCCCTAACGTTACTGGCCGAAGCCGCTTG
GAATAAGGCCGGTGTGCGTTTGTCTATATGTTATTTTCCACCATATTGCCGTCTTTTGGCAATGTGAGGGCCCGGAAACCTG
GCCCTGTCTTCTTGACGAGCATTCCTAGGGGTCTTTCCCCTCTCGCCAAAGGAATGCAAGGTCTGTTGAATGTCGTGAAGG
AAGCAGTTCCTCTGGAAGCTTCTTGAAGACAAACAACGTCTGTAGCGACCCTTTGCAGGCAGCGGAACCCCCCACCTGGC
GACAGGTGCCTCTGCGGCCAAAAGCCACGTGTATAAGATACACCTGCAAAGGCGGCACAACCCCAGTGCCACGTTGTGAG
TTGGATAGTTGTGGAAAGAGTCAAATGGCTCTCCTCAAGCGTATTCAACAAGGGGCTGAAGGATGCCCAGAAGGTACCCA
TTGTATGGGATCTGATCTGGGGCCTCGGTGCACATGCTTTACATGTGTTTAGTCGAGGTTAAAAAACGTCTAGGCCCCCCGA
ACCACGGGGACGTGGTTTTCCTTTGAAAAACACGATGATAATATGGCCACAACCATGAAGACCTTAATTCTTGCCGTTGCAT
TAGTCTACTGCGCCACTGTTCATTGCCAGGACTGTCCTTACGAACCTGATCCACCAAACACAGTTCCAACTTCCTGTGAAG
CTAAAGAAGGAGAATGTATTGATAGCAGCTGTGGCACCTGCACGAGAGACATACTATCAGATGGACTGTGTGAAAATAAACC
AGGAAAAACATGTTGCCGAATGTGTCAGTATGTAATTGAATGCAGAGTAGAGGCCGCAGGATGGTTTAGAACATTCTATGGA
AAGAGATTCCAGTTCCAGGAACCTGGTACATACGTGTTGGGTCAAGGAACCAAGGGCGGCGACTGGAAGGTGTCCATCAC
CCTGGAGAACCTGGATGGAACCAAGGGGGCTGTGCTGACCAAGACAAGACTGGAAGTGGCTGGAGACATCATTGACATC
GCTCAAGCTACTGAGAATCCCATCACTGTAAACGGTGGAGCTGACCCTATCATCGCCAACCCGTACACCATCGGCGAGGTC
ACCATCGCTGTTGTTGAGATGCCAGGCTTCAACATCACCGTCATTGAGTTCTTCAAACTGATCGTGATCGACATCCTCGGAG
GAAGATCTGTAAGAATCGCCCCAGACACAGCAAACAAAGGAATGATCTCTGGCCTCTGTGGAGATCTTAAAATGATGGAAG
ATACAGACTTCACTTCAGATCCAGAACAACTCGCTATTCAGCCTAAGATCAACCAGGAGTTTGACGGTTGTCCACTCTATGG
AAATCCTGATGACGTTGCATACTGCAAAGGTCTTCTGGAGCCGTACAAGGACAGCTGCCGCAACCCCATCAACTTCTACTA
CTACACCATCTCCTGCGCCTTCGCCCGCTGTATGGGTGGAGACGAGCGAGCCTCACACGTGCTGCTTGACTACAGGGAGA
CGTGCGCTGCTCCCGAAACTAGAGGAACCTGCGTTTTGTCTGGACATACTTTCTACGATACATTTGACAAAGCAAGATACCA
ATTCCAGGGTCCCTGCAAGGAGATTCTTATGGCCGCCGACTGTTTCTGGAACACTTGGGATGTGAAGGTTTCACACAGGAA

TGTTGACTCTTACACTGAAGTAGAGAAAGTACGAATCAGGAAACAATCGACTGTAGTAGAACTCATTGTTGATGGAAAACAG
ATTCTGGTTGGAGGAGAAGCCGTGTCCGTCCCGTACAGCTCTCAGAACACTTCCATCTACTGGCAAGATGGTGACATACTG
ACTACAGCCATCCTACCTGAAGCTCTGGTGGTCAAGTTCAACTTCAAGCAACTGCTCGTCGTACATATTAGAGATCCATTCG
ATGGTAAGACTTGCGGTATTTGCGGTAACTACAACCAGGATTTCAGTGATGATTCTTTTGATGCTGAAGGAGCCTGTGATCT
GACCCCCAACCCACCGGGATGCACCGAAGAACAGAAACCTGAAGCTGAACGACTCTGCAATAGTCTCTTCGCCGGTCAAA
GTGATCTTGATCAGAAATGTAACGTGTGCCACAAGCCTGACCGTGTCGAACGATGCATGTACGAGTATTGCCTGAGGGGAC
AACAGGGTTTCTGTGACCACGCATGGGAGTTCAAGAAAGAATGCTACATAAAGCATGGAGACACCCTAGAAGTACCAGATG
AATGCAAATAGAAAAAACAAAAAACAAAACAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAGGCTATTATGCGTTACCGGCGAG
ACGCTACGGACTTAAATAATTGAGCCTTAAAGAAGAAATTCTTTAAGTGGATGCTCTCAAACTCAGGGAAACCTAAATCTAGT
TATAGACAAGGCAATCCTGAGCCAAGCCGAAGTAGTAATTAGTAAGACCAGTGGACAATCGACGGATAACAGCATATCTAG

SelfCirc-Spike

CTCGAGCACCACCACCACCACCACTGAGATCCGGCTGCTAACAAAGCCCGAAAGGAAGCTGAGTTGGCTGCTGCCACCG
CTGAGCAATAACTAGCATAACCCCTTGGGGCCTCTAAACGGGTCTTGAGGGGTTTTTTGCTGAAAGGAGGAACTATATCCG
GATTGGCGAATGGGACGCGCCCTGTAGCGGCGCATTAAGCGCGGCGGGTGTGGTGGTTACGCGCAGCGTGACCGCTACA
CTTGCCAGCGCCCTAGCGCCCGCTCCTTTCGCTTTCTTCCCTTCCTTTCTCGCCACGTTCGCCGGCTTTCCCCGTCAAGCT
CTAAATCGGGGGCTCCCTTTAGGGTTCCGATTTAGTGCTTTACGGCACCTCGACCCCAAAAAACTTGATTAGGGTGATGGTT
CACGTAGTGGGCCATCGCCCTGATAGACGGTTTTTCGCCCTTTGACGTTGGAGTCCACGTTCTTTAATAGTGGACTCTTGTT
CCAAACTGGAACAACACTCAACCCTATCTCGGTCTATTCTTTTGATTTATAAGGGATTTTGCCGATTTCGGCCTATTGGTTAAA
AAATGAGCTGATTTAACAAAAATTTAACGCGAATTTTAACAAAATATTAACGCTTACAATTTAGGTGGCACTTTTCGGGGAAAT
GTGCGCGGAACCCCTATTTGTTTATTTTTCTAAATACATTCAAATATGTATCCGCTCATGAATTAATTCTTAGAAAAACTCATCG
AGCATCAAATGAAACTGCAATTTATTCATATCAGGATTATCAATACCATATTTTTGAAAAAGCCGTTTCTGTAATGAAGGAGAAA
ACTCACCGAGGCAGTTCCATAGGATGGCAAGATCCTGGTATCGGTCTGCGATTCCGACTCGTCCAACATCAATACAACCTAT
TAATTTCCCCTCGTCAAAAATAAGGTTATCAAGTGAGAAATCACCATGAGTGACGACTGAATCCGGTGAGAATGGCAAAAGT
TTATGCATTTCTTTCCAGACTTGTTCAACAGGCCAGCCATTACGCTCGTCATCAAAATCACTCGCATCAACCAAACCGTTATT
CATTCGTGATTGCGCCTGAGCGAGACGAAATACGCGATCGCTGTTAAAAGGACAATTACAAACAGGAATCGAATGCAACCG
GCGCAGGAACACTGCCAGCGCATCAACAATATTTTCACCTGAATCAGGATATTCTTCTAATACCTGGAATGCTGTTTTCCCGG
GGATCGCAGTGGTGAGTAACCATGCATCATCAGGAGTACGGATAAAATGCTTGATGGTCGGAAGAGGCATAAATTCCGTCA
GCCAGTTTAGTCTGACCATCTCATCTGTAACATCATTGGCAACGCTACCTTTGCCATGTTTCAGAAACAACTCTGGCGCATC
GGGCTTCCCATACAATCGATAGATTGTCGCACCTGATTGCCCGACATTATCGCGAGCCCATTTATACCCATATAAATCAGCAT
CCATGTTGGAATTTAATCGCGGCCTAGAGCAAGACGTTTCCCGTTGAATATGGCTCATAACACCCCTTGTATTACTGTTTATG
TAAGCAGACAGTTTTATTGTTCATGACCAAAATCCCTTAACGTGAGTTTTCGTTCCACTGAGCGTCAGACCCCGTAGAAAAG
ATCAAAGGATCTTCTTGAGATCCTTTTTTTCTGCGCGTAATCTGCTGCTTGCAAACAAAAAAACCACCGCTACCAGCGGTGG
TTTGTTTGCCGGATCAAGAGCTACCAACTCTTTTTCCGAAGGTAACTGGCTTCAGCAGAGCGCAGATACCAAATACTGTCCT
TCTAGTGTAGCCGTAGTTAGGCCACCACTTCAAGAACTCTGTAGCACCGCCTACATACCTCGCTCTGCTAATCCTGTTACCA
GTGGCTGCTGCCAGTGGCGATAAGTCGTGTCTTACCGGGTTGGACTCAAGACGATAGTTACCGGATAAGGCGCAGCGGTC
GGGCTGAACGGGGGGTTCGTGCACACAGCCCAGCTTGGAGCGAACGACCTACACCGAACTGAGATACCTACAGCGTGAG
CTATGAGAAAGCGCCACGCTTCCCGAAGGGAGAAAGGCGGACAGGTATCCGGTAAGCGGCAGGGTCGGAACAGGAGAG
CGCACGAGGGAGCTTCCAGGGGGAAACGCCTGGTATCTTTATAGTCCTGTCGGGTTTCGCCACCTCTGACTTGAGCGTCG
ATTTTTGTGATGCTCGTCAGGGGGGCGGAGCCTATGGAAAAACGCCAGCAACGCGGCCTTTTTACGGTTCCTGGCCTTTT
GCTGGCCTTTTGCTCACATGTTCTTTCCTGCGTTATCCCCTGATTCTGTGGATAACCGTATTACCGCCTTTGAGTGAGCTGAT
ACCGCTCGCCGCAGCCGAACGACCGAGCGCAGCGAGTCAGTGAGCGAGGAAGCGGAAGAGCGCCTGATGCGGTATTTT
CTCCTTACGCATCTGTGCGGTATTTCACACCGCAATGGTGCACTCTCAGTACAATCTGCTCTGATGCCGCATAGTTAAGCCA
GTATACACTCCGCTATCGCTACGTGACTGGGTCATGGCTGCGCCCCGACACCCGCCAACACCCGCTGACGCGCCCTGACG
GGCTTGTCTGCTCCCGGCATCCGCTTACAGACAAGCTGTGACCGTCTCCGGGAGCTGCATGTGTCAGAGGTTTTCACCGT
CATCACCGAAACGCGCGAGGCAGCTGCGGTAAAGCTCATCAGCGTGGTCGTGAAGCGATTCACAGATGTCTGCCTGTTCA
TCCGCGTCCAGCTCGTTGAGTTTCTCCAGAAGCGTTAATGTCTGGCTTCTGATAAAGCGGGCCATGTTAAGGGCGGTTTTT
TCCTGTTTGGTCACTGATGCCTCCGTGTAAGGGGGATTTCTGTTCATGGGGGTAATGATACCGATGAAACGAGAGAGGATG
CTCACGATACGGGTTACTGATGATGAACATGCCCGGTTACTGGAACGTTGTGAGGGTAAACAACTGGCGGTATGGATGCGG
CGGGACCAGAGAAAAATCACTCAGGGTCAATGCCAGCGCTTCGTTAATACAGATGTAGGTGTTCCACAGGGTAGCCAGCA
GCATCCTGCGATGCAGATCCGGAACATAATGGTGCAGGGCGCTGACTTCCGCGTTTCCAGACTTTACGAAACACGGAAAC
CGAAGACCATTCATGTTGTTGCTCAGGTCGCAGACGTTTTGCAGCAGCAGTCGCTTCACGTTCGCTCGCGTATCGGTGATT
CATTCTGCTAACCAGTAAGGCAACCCCGCCAGCCTAGCCGGGTCCTCAACGACAGGAGCACGATCATGCGCACCCGTGG
GGCCGCCATGCCGGCGATAATGGCCTGCTTCTCGCCGAAACGTTTGGTGGCGGGACCAGTGACGAAGGCTTGAGCGAGG
GCGTGCAAGATTCCGAATACCGCAAGCGACAGGCCGATCATCGTCGCGCTCCAGCGAAAGCGGTCCTCGCCGAAAATGAC
CCAGAGCGCTGCCGGCACCTGTCCTACGAGTTGCATGATAAAGAAGACAGTCATAAGTGCGGCGACGATAGTCATGCCCC
GCGCCCACCGGAAGGAGCTGACTGGGTTGAAGGCTCTCAAGGGCATCGGTCGAGATCCCGGTGCCTAATGAGTGAGCTA
ACTTACATTAATTGCGTTGCGCTCACTGCCCGCTTTCCAGTCGGGAAACCTGTCGTGCCAGCTGCATTAATGAATCGGCCA
ACGCGCGGGGAGAGGCGGTTTGCGTATTGGGCGCCAGGGTGGTTTTTCTTTTCACCAGTGAGACGGGCAACAGCTGATT
GCCCTTCACCGCCTGGCCCTGAGAGAGTTGCAGCAAGCGGTCCACGCTGGTTTGCCCCAGCAGGCGAAAATCCTGTTTG
ATGGTGGTTAACGGCGGGATATAACATGAGCTGTCTTCGGTATCGTCGTATCCCACTACCGAGATATCCGCACCAACGCGCA
GCCCGGACTCGGTAATGGCGCGCATTGCGCCCAGCGCCATCTGATCGTTGGCAACCAGCATCGCAGTGGGAACGATGCC
CTCATTCAGCATTTGCATGGTTTGTTGAAAACCGGACATGGCACTCCAGTCGCCTTCCCGTTCCGCTATCGGCTGAATTTGA
TTGCGAGTGAGATATTTATGCCAGCCAGCCAGACGCAGACGCGCCGAGACAGAACTTAATGGGCCCGCTAACAGCGCGAT
TTGCTGGTGACCCAATGCGACCAGATGCTCCACGCCCAGTCGCGTACCGTCTTCATGGGAGAAAATAATACTGTTGATGGG
TGTCTGGTCAGAGACATCAAGAAATAACGCCGGAACATTAGTGCAGGCAGCTTCCACAGCAATGGCATCCTGGTCATCCAG
CGGATAGTTAATGATCAGCCCACTGACGCGTTGCGCGAGAAGATTGTGCACCGCCGCTTTACAGGCTTCGACGCCGCTTC
GTTCTACCATCGACACCACCACGCTGGCACCCAGTTGATCGGCGCGAGATTTAATCGCCGCGACAATTTGCGACGGCGCG
TGCAGGGCCAGACTGGAGGTGGCAACGCCAATCAGCAACGACTGTTTGCCCGCCAGTTGTTGTGCCACGCGGTTGGGAA
TGTAATTCAGCTCCGCCATCGCCGCTTCCACTTTTTCCCGCGTTTTCGCAGAAACGTGGCTGGCCTGGTTCACCACGCGG
GAAACGGTCTGATAAGAGACACCGGCATACTCTGCGACATCGTATAACGTTACTGGTTTCACATTCACCACCCTGAATTGAC
TCTCTTCCGGGGCGCTATCATGCCATACCGCGAAAGGTTTTGCGCCATTCGATGGTGTCCGGGATCTCGACGCTCTCCCTTA
TGCGACTCCTGCATTAGGAAGCAGCCCAGTAGTAGGTTGAGGCCGTTGAGCACCGCCGCCGCAAGGAATGGTGCATGCAA
GGAGATGGCGCCCAACAGTCCCCCGGCCACGGGGCCTGCCACCATACCCACGCCGAAACAAGCGCTCATGAGCCCGAA
GTGGCGAGCCCGATCTTCCCCATCGGTGATGTCGGCGATATAGGCGCCAGCAACCGCACCTGTGGCGCCGGTGATGCCG
GCCACGATGCGTCCGGCGTAGAGGATCGAGATCTCGATCCCGCGAAATTAATACGACTCACTATAGGGGAATTGTGAGCGG
ATAACAATTCCCCTCTAGGATCCGGGAGACCCTCGACCGTCGATTGTCCACTGGTCAACAATAGATGACTTACAACTAATCG
GAAGGTGCAGAGACTCGACGGGAGCTACCCTAACGTCAAGACGAGGGTAAAGAGAGAGTCCAATTCTCAAAGCCAATAGG
CAGTAGCGAAAGCTGCAAGAGAATGAAAATCCGTTGACCTTAAACGGTCGTGTGGGTTCAAGTCCCTCCACCCCCACGCC
GGAAACGCAATAGCCGAAAAACAAAAAACAAAAAAACCCCCCTCTCCCTCCCCCCCTAACGTTACTGGCCGAAGCCGCTTG
GAATAAGGCCGGTGTGCGTTTGTCTATATGTTATTTTCCACCATATTGCCGTCTTTTGGCAATGTGAGGGCCCGGAAACCTG

GCCCTGTCTTCTTGACGAGCATTCCTAGGGGTCTTTCCCCTCTCGCCAAAGGAATGCAAGGTCTGTTGAATGTCGTGAAGG
AAGCAGTTCCTCTGGAAGCTTCTTGAAGACAAACAACGTCTGTAGCGACCCTTTGCAGGCAGCGGAACCCCCCACCTGGC
GACAGGTGCCTCTGCGGCCAAAAGCCACGTGTATAAGATACACCTGCAAAGGCGGCACAACCCCAGTGCCACGTTGTGAG
TTGGATAGTTGTGGAAAGAGTCAAATGGCTCTCCTCAAGCGTATTCAACAAGGGGCTGAAGGATGCCCAGAAGGTACCCCA
TTGTATGGGATCTGATCTGGGGCCTCGGTGCACATGCTTTACATGTGTTTAGTCGAGGTTAAAAAACGTCTAGGCCCCCCGA
ACCACGGGGACGTGGTTTTCCTTTGAAAAACACGATGATAATATGGCCACAACCATGGTGTTCGTGTTTCTGGTGCTGCTG
CCTCTGGTGTCCAGCCAGTGTGTGAACCTGACCACAAGAACCCAGCTGCCTCCAGCCTACACCAACAGCTTTACCAGAGG
CGTGTACTACCCCGACAAGGTGTTCAGATCCAGCGTGCTGCACTCTACCCAGGACCTGTTCCTGCCTTTCTTCAGCAACGT
GACCTGGTTCCACGCCATCCACGTGTCCGGCACCAATGGCACCAAGAGATTCGACAACCCCGTGCTGCCCTTCAACGACG
GGGTGTACTTTGCCAGCACCGAGAAGTCCAACATCATCAGAGGCTGGATCTTCGGCACCACACTGGACAGCAAGACCCAG
AGCCTGCTGATCGTGAACAACGCCACCAACGTGGTCATCAAAGTGTGCGAGTTCCAGTTCTGCAACGACCCCTTCCTGGG
CGTCTACTATCACAAGAACAACAAGAGCTGGATGGAAAGCGAGTTCCGGGTGTACAGCAGCGCCAACAACTGCACCTTCG
AGTACGTGTCCCAGCCTTTCCTGATGGACCTGGAAGGCAAGCAGGGCAACTTCAAGAACCTGCGCGAGTTCGTGTTCAAG
AACATCGACGGCTACTTCAAGATCTACAGCAAGCACACCCCTATCAACCTCGTGCGGGATCTGCCTCAGGGCTTCTCTGCT
CTGGAACCCCTGGTGGATCTGCCCATCGGCATCAACATCACCCGGTTTCAGACACTGCTGGCCCTGCACAGAAGCTACCT
GACACCTGGCGATAGCAGCAGCGGATGGACAGCTGGTGCCGCCGCTTACTATGTGGGCTACCTGCAGCCTAGAACCTTTC
TGCTGAAGTACAACGAGAACGGCACCATCACCGACGCCGTGGATTGTGCTCTGGATCCTCTGAGCGAGACAAAGTGCACC
CTGAAGTCCTTCACCGTGGAAAAGGGCATCTACCAGACCAGCAACTTCCGGGTGCAGCCCACCGAATCCATCGTGCGGTT
CCCCAATATCACCAATCTGTGCCCCTTCGGCGAGGTGTTCAATGCCACCAGATTCGCCTCTGTGTACGCCTGGAACCGGAA
GCGGATCAGCAATTGCGTGGCCGACTACTCCGTGCTGTACAACTCCGCCAGCTTCAGCACCTTCAAGTGCTACGGCGTGT
CCCCTACCAAGCTGAACGACCTGTGCTTCACAAACGTGTACGCCGACAGCTTCGTGATCCGGGGAGATGAAGTGCGGCAG
ATTGCCCCTGGACAGACAGGCAAGATCGCCGACTACAACTACAAGCTGCCCGACGACTTCACCGGCTGTGTGATTGCCTG
GAACAGCAACAACCTGGACTCCAAAGTCGGCGGCAACTACAATTACCTGTACCGGCTGTTCCGGAAGTCCAATCTGAAGC
CCTTCGAGCGGGACATCTCCACCGAGATCTATCAGGCCGGCAGCACCCCTTGTAACGGCGTGGAAGGCTTCAACTGCTAC
TTCCCACTGCAGTCCTACGGCTTTCAGCCCACAAATGGCGTGGGCTATCAGCCCTACAGAGTGGTGGTGCTGAGCTTCGA
ACTGCTGCATGCCCCTGCCACAGTGTGCGGCCCTAAGAAAAGCACCAATCTCGTGAAGAACAAATGCGTGAACTTCAACTT
CAACGGCCTGACCGGCACCGGCGTGCTGACAGAGAGCAACAAGAAGTTCCTGCCATTCCAGCAGTTTGGCCGGGATATC
GCCGATACCACAGACGCCGTTAGAGATCCCCAGACACTGGAAATCCTGGACATCACCCCTTGCAGCTTCGGCGGAGTGTC
TGTGATCACCCCTGGCACCAACACCAGCAATCAGGTGGCAGTGCTGTACCAGGACGTGAACTGTACCGAAGTGCCCGTGG
CCATTCACGCCGATCAGCTGACACCTACATGGCGGGTGTACTCCACCGGCAGCAATGTGTTTCAGACCAGAGCCGGCTGT
CTGATCGGAGCCGAGCACGTGAACAATAGCTACGAGTGCGACATCCCCATCGGCGCTGGCATCTGTGCCAGCTACCAGAC
ACAGACAAACAGCCCCGCCTCTGTGGCCAGCCAGAGCATCATTGCCTACACAATGTCTCTGGGCGCCGAGAACAGCGTG
GCCTACTCCAACAACTCTATCGCTATCCCCACCAACTTCACCATCAGCGTGACCACAGAGATCCTGCCTGTGTCCATGACC
AAGACCAGCGTGGACTGCACCATGTACATCTGCGGCGATTCCACCGAGTGCTCCAACCTGCTGCTGCAGTACGGCAGCTT
CTGCACCCAGCTGAATAGAGCCCTGACAGGGATCGCCGTGGAACAGGACAAGAACACCCAAGAGGTGTTCGCCCAAGTG
AAGCAGATCTACAAGACCCCTCCTATCAAGGACTTCGGCGGCTTCAATTTCAGCCAGATTCTGCCCGATCCTAGCAAGCCC
AGCAAGCGGAGCTTCATCGAGGACCTGCTGTTCAACAAAGTGACACTGGCCGACGCCGGCTTCATCAAGCAGTATGGCGA
TTGTCTGGGCGACATTGCCGCCAGGGATCTGATTTGCGCCCAGAAGTTTAACGGACTGACAGTGCTGCCTCCTCTGCTGA
CCGATGAGATGATCGCCCAGTACACATCTGCCCTGCTGGCCGGCACAATCACAAGCGGCTGGACATTTGGAGCTGGCGCC
GCTCTGCAGATCCCCTTTGCTATGCAGATGGCCTACCGGTTCAACGGCATCGGAGTGACCCAGAATGTGCTGTACGAGAAC
CAGAAGCTGATCGCCAACCAGTTCAACAGCGCCATCGGCAAGATCCAGGACAGCCTGAGCAGCACAGCAAGCGCCCTGG
GAAAGCTGCAGGACGTGGTCAACCAGAATGCCCAGGCACTGAACACCCTGGTCAAGCAGCTGTCCTCCAACTTCGGCGC
CATCAGCTCTGTGCTGAACGATATCCTGAGCAGACTGGACcctcctGAAGCCGAGGTGCAGATCGACAGACTGATCACCGGA
AGGCTGCAGTCCCTGCAGACCTACGTTACCCAGCAGCTGATCAGAGCCGCCGAGATTAGAGCCTCTGCCAATCTGGCCGC
CACCAAGATGTCTGAGTGTGTGCTGGGCCAGAGCAAGAGAGTGGACTTTTGCGGCAAGGGCTACCACCTGATGAGCTTCC
CTCAGTCTGCCCCTCACGGCGTGGTGTTTCTGCACGTGACATACGTGCCCGCTCAAGAGAAGAATTTCACCACCGCTCCA
GCCATCTGCCACGACGGCAAAGCCCACTTTCCTAGAGAAGGCGTGTTCGTGTCCAACGGCACCCATTGGTTCGTGACCCA
GCGGAACTTCTACGAGCCCCAGATCATCACCACCGACAACACCTTCGTGTCTGGCAACTGCGACGTCGTGATCGGCATTG
TGAACAATACCGTGTACGACCCTCTGCAGCCCGAGCTGGACAGCTTCAAAGAGGAACTGGATAAGTACTTTAAGAACCACA
CAAGCCCCGACGTGGACCTGGGCGATATCAGCGGAATCAATGCCAGCGTCGTGAACATCCAGAAAGAGATCGACCGGCTG
AACGAGGTGGCCAAGAATCTGAACGAGAGCCTGATCGACCTGCAAGAACTGGGGAAGTACGAGCAGTACATCAAGTGGCC
CAGCGGCCGCTTGGTCCCACGTGGCTCACCCGGATCTGGATACATCCCGGAGGCCCCTAGGGACGGTCAAGCTTACGTG
AGAAAGGACGGCGAATGGGTTCTGCTGTCGACCTTCTTGGGACATCATCATCATCACCGAATTCAGTAAAAAAAACAAA
AAACAAAACAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAGGCTATTATGCGTTACCGGCGAGACGCTACGGACTTAAATAATT
GAGCCTTAAAGAAGAAATTCTTTAAGTGGATGCTCTCAAACTCAGGGAAACCTAAATCTAGTTATAGACAAGGCAATCCTGAG
CCAAGCCGAAGTAGTAATTAGTAAGACCAGTGGACAATCGACGGATAACAGCATATCTAG

TermtRNA-GFP

GCCCGGATAGCTCAGGCGGGAGAGCAGCGGCCGAGTAATTTACGTCGACGGTGACGTCGATGGTTGCGGCCGCGGGTC
CAGGGTTCAAGTCCCTGTTCGGGCGCCACTTGTTCTTTTTGCAGAAGCTCAGAATAAACGCTCAACTTTGGGCCGCCACCA
TGGTGAGCAAGGGCGAGGAGCTGTTCACCGGGGTGGTGCCCATCCTGGTCGAGCTGGACGGCGACGTAAACGGCCACA
AGTTCAGCGTGTCTGGCGAGGGCGAGGGCGATGCCACCTACGGCAAGCTGACCCTGAAGTTCATCTGCACCACCGGCAA
GCTGCCCGTGCCCTGGCCCACCCTCGTGACCACCCTGACCTACGGCGTGCAGTGCTTCAGCCGCTACCCCGACCACATG
AAGCAGCACGACTTCTTCAAGTCCGCCATGCCCGAAGGCTACGTCCAGGAGCGCACCATCTTCTTCAAGGACGACGGCAA
CTACAAGACCCGCGCCGAGGTGAAGTTCGAGGGCGACACCCTGGTGAACCGCATCGAGCTGAAGGGCATCGACTTCAAG
GAGGACGGCAACATCCTGGGGCACAAGCTGGAGTACAACTACAACAGCCACAACGTCTATATCATGGCCGACAAGCAGAA
GAACGGCATCAAGGCGAACTTCAAGATCCGCCACAACATCGAGGACGGCAGCGTGCAGCTCGCCGACCACTACCAGCAG
AACACCCCCATCGGCGACGGCCCCGTGCTGCTGCCCGACAACCACTACCTGAGCACCCAGTCCGCCCTGAGCAAAGACC
CCAACGAGAAGCGCGATCACATGGTCCTGCTGGAGTTCGTGACCGCCGCCGGGATCACTCTCGGCATGGACGAGCTGTA
CAAGTAAAACCAGCCTCAAGAACACCCGAATGGAGTCTCTAAGCTACATAATACCAACTTACACTTTACAAAATGTTGTCCCC
CAAAATGTAGCCATTCGTATCTGCTCCTAATAAAAAGAAAGTTTCTTCACACCGCGGGCTTGTCCCTGAACTTGGGACCTGG
GCGCCGGCGTTGGTAGCTGCAGTGGCAGCTGCATTTAATGAGCCGGCGACGAGAGGGCGGACTCGATAGGCCCGCTCGA
GCACCACCACCACCACCACTGAGATCCGGCTGCTAACAAAGCCCGAAAGGAAGCTGAGTTGGCTGCTGCCACCGCTGAG
CAATAACTAGCATAACCCCTTGGGGCCTCTAAACGGGTCTTGAGGGGTTTTTTGCTGAAAGGAGGAACTATATCCGGATTGG
CGAATGGGACGCGCCCTGTAGCGGCGCATTAAGCGCGGCGGGTGTGGTGGTTACGCGCAGCGTGACCGCTACACTTGCC
AGCGCCCTAGCGCCCGCTCCTTTCGCTTTCTTCCCTTCCTTTCTCGCCACGTTCGCCGGCTTTCCCCGTCAAGCTCTAAAT
CGGGGGCTCCCTTTAGGGTTCCGATTTAGTGCTTTACGGCACCTCGACCCCAAAAAACTTGATTAGGGTGATGGTTCACGT
AGTGGGCCATCGCCCTGATAGACGGTTTTTCGCCCTTTGACGTTGGAGTCCACGTTCTTTAATAGTGGACTCTTGTTCCAAA
CTGGAACAACACTCAACCCTATCTCGGTCTATTCTTTTGATTTATAAGGGATTTTGCCGATTTCGGCCTATTGGTTAAAAAATG
AGCTGATTTAACAAAAATTTAACGCGAATTTTAACAAAATATTAACGCTTACAATTTAGGTGGCACTTTTCGGGGAAATGTGCG
CGGAACCCCTATTTGTTTATTTTTCTAAATACATTCAAATATGTATCCGCTCATGAATTAATTCTTAGAAAAACTCATCGAGCATC

AAATGAAACTGCAATTTATTCATATCAGGATTATCAATACCATATTTTTGAAAAAGCCGTTTCTGTAATGAAGGAGAAAACTCAC
CGAGGCAGTTCCATAGGATGGCAAGATCCTGGTATCGGTCTGCGATTCCGACTCGTCCAACATCAATACAACCTATTAATTT
CCCCTCGTCAAAAATAAGGTTATCAAGTGAGAAATCACCATGAGTGACGACTGAATCCGGTGAGAATGGCAAAAGTTTATGC
ATTTCTTTCCAGACTTGTTCAACAGGCCAGCCATTACGCTCGTCATCAAAATCACTCGCATCAACCAAACCGTTATTCATTCG
TGATTGCGCCTGAGCGAGACGAAATACGCGATCGCTGTTAAAAGGACAATTACAAACAGGAATCGAATGCAACCGGCGCAG
GAACACTGCCAGCGCATCAACAATATTTTCACCTGAATCAGGATATTCTTCTAATACCTGGAATGCTGTTTTCCCGGGGATCG
CAGTGGTGAGTAACCATGCATCATCAGGAGTACGGATAAAATGCTTGATGGTCGGAAGAGGCATAAATTCCGTCAGCCAGTT
TAGTCTGACCATCTCATCTGTAACATCATTGGCAACGCTACCTTTGCCATGTTTCAGAAACAACTCTGGCGCATCGGGCTTC
CCATACAATCGATAGATTGTCGCACCTGATTGCCCGACATTATCGCGAGCCCATTTATACCCATATAAATCAGCATCCATGTTG
GAATTTAATCGCGGCCTAGAGCAAGACGTTTCCCGTTGAATATGGCTCATAACACCCCTTGTATTACTGTTTATGTAAGCAGA
CAGTTTTATTGTTCATGACCAAAATCCCTTAACGTGAGTTTTCGTTCCACTGAGCGTCAGACCCCGTAGAAAAGATCAAAGG
ATCTTCTTGAGATCCTTTTTTTCTGCGCGTAATCTGCTGCTTGCAAACAAAAAAACCACCGCTACCAGCGGTGGTTTGTTTG
CCGGATCAAGAGCTACCAACTCTTTTTCCGAAGGTAACTGGCTTCAGCAGAGCGCAGATACCAAATACTGTCCTTCTAGTGT
AGCCGTAGTTAGGCCACCACTTCAAGAACTCTGTAGCACCGCCTACATACCTCGCTCTGCTAATCCTGTTACCAGTGGCTG
CTGCCAGTGGCGATAAGTCGTGTCTTACCGGGTTGGACTCAAGACGATAGTTACCGGATAAGGCGCAGCGGTCGGGCTGA
ACGGGGGGTTCGTGCACACAGCCCAGCTTGGAGCGAACGACCTACACCGAACTGAGATACCTACAGCGTGAGCTATGAGA
AAGCGCCACGCTTCCCGAAGGGAGAAAGGCGGACAGGTATCCGGTAAGCGGCAGGGTCGGAACAGGAGAGCGCACGAG
GGAGCTTCCAGGGGGAAACGCCTGGTATCTTTATAGTCCTGTCGGGTTTCGCCACCTCTGACTTGAGCGTCGATTTTTGTG
ATGCTCGTCAGGGGGGCGGAGCCTATGGAAAAACGCCAGCAACGCGGCCTTTTTACGGTTCCTGGCCTTTTGCTGGCCTT
TTGCTCACATGTTCTTTCCTGCGTTATCCCCTGATTCTGTGGATAACCGTATTACCGCCTTTGAGTGAGCTGATACCGCTCGC
CGCAGCCGAACGACCGAGCGCAGCGAGTCAGTGAGCGAGGAAGCGGAAGAGCGCCTGATGCGGTATTTTCTCCTTACGC
ATCTGTGCGGTATTTCACACCGCAATGGTGCACTCTCAGTACAATCTGCTCTGATGCCGCATAGTTAAGCCAGTATACACTC
CGCTATCGCTACGTGACTGGGTCATGGCTGCGCCCCGACACCCGCCAACACCCGCTGACGCGCCCTGACGGGCTTGTCT
GCTCCCGGCATCCGCTTACAGACAAGCTGTGACCGTCTCCGGGAGCTGCATGTGTCAGAGGTTTTCACCGTCATCACCGA
AACGCGCGAGGCAGCTGCGGTAAAGCTCATCAGCGTGGTCGTGAAGCGATTCACAGATGTCTGCCTGTTCATCCGCGTCC
AGCTCGTTGAGTTTCTCCAGAAGCGTTAATGTCTGGCTTCTGATAAAGCGGGCCATGTTAAGGGCGGTTTTTTCCTGTTTGG
TCACTGATGCCTCCGTGTAAGGGGGATTTCTGTTCATGGGGGTAATGATACCGATGAAACGAGAGAGGATGCTCACGATAC
GGGTTACTGATGATGAACATGCCCGGTTACTGGAACGTTGTGAGGGTAAACAACTGGCGGTATGGATGCGGCGGGACCAG
AGAAAAATCACTCAGGGTCAATGCCAGCGCTTCGTTAATACAGATGTAGGTGTTCCACAGGGTAGCCAGCAGCATCCTGCG
ATGCAGATCCGGAACATAATGGTGCAGGGCGCTGACTTCCGCGTTTCCAGACTTTACGAAACACGGAAACCGAAGACCATT
CATGTTGTTGCTCAGGTCGCAGACGTTTTGCAGCAGCAGTCGCTTCACGTTCGCTCGCGTATCGGTGATTCATTCTGCTAA
CCAGTAAGGCAACCCCGCCAGCCTAGCCGGGTCCTCAACGACAGGAGCACGATCATGCGCACCCGTGGGGCCGCCATGC
CGGCGATAATGGCCTGCTTCTCGCCGAAACGTTTGGTGGCGGGACCAGTGACGAAGGCTTGAGCGAGGGCGTGCAAGAT
TCCGAATACCGCAAGCGACAGGCCGATCATCGTCGCGCTCCAGCGAAAGCGGTCCTCGCCGAAAATGACCCAGAGCGCT
GCCGGCACCTGTCCTACGAGTTGCATGATAAAGAAGACAGTCATAAGTGCGGCGACGATAGTCATGCCCCGCGCCCACCG
GAAGGAGCTGACTGGGTTGAAGGCTCTCAAGGGCATCGGTCGAGATCCCGGTGCCTAATGAGTGAGCTAACTTACATTAAT
TGCGTTGCGCTCACTGCCCGCTTTCCAGTCGGGAAACCTGTCGTGCCAGCTGCATTAATGAATCGGCCAACGCGCGGGGA
GAGGCGGTTTGCGTATTGGGCGCCAGGGTGGTTTTTCTTTTCACCAGTGAGACGGGCAACAGCTGATTGCCCTTCACCGC
CTGGCCCTGAGAGAGTTGCAGCAAGCGGTCCACGCTGGTTTGCCCCAGCAGGCGAAAATCCTGTTTGATGGTGGTTAACG
GCGGGATATAACATGAGCTGTCTTCGGTATCGTCGTATCCCACTACCGAGATATCCGCACCAACGCGCAGCCCGGACTCGG
TAATGGCGCGCATTGCGCCCAGCGCCATCTGATCGTTGGCAACCAGCATCGCAGTGGGAACGATGCCCTCATTCAGCATTT
GCATGGTTTGTTGAAAACCGGACATGGCACTCCAGTCGCCTTCCCGTTCCGCTATCGGCTGAATTTGATTGCGAGTGAGAT
ATTTATGCCAGCCAGCCAGACGCAGACGCGCCGAGACAGAACTTAATGGGCCCGCTAACAGCGCGATTTGCTGGTGACCC
AATGCGACCAGATGCTCCACGCCCAGTCGCGTACCGTCTTCATGGGAGAAAATAATACTGTTGATGGGTGTCTGGTCAGAG
ACATCAAGAAATAACGCCGGAACATTAGTGCAGGCAGCTTCCACAGCAATGGCATCCTGGTCATCCAGCGGATAGTTAATGA
TCAGCCCACTGACGCGTTGCGCGAGAAGATTGTGCACCGCCGCTTTACAGGCTTCGACGCCGCTTCGTTCTACCATCGAC
ACCACCACGCTGGCACCCAGTTGATCGGCGCGAGATTTAATCGCCGCGACAATTTGCGACGGCGCGTGCAGGGCCAGAC
TGGAGGTGGCAACGCCAATCAGCAACGACTGTTTGCCCGCCAGTTGTTGTGCCACGCGGTTGGGAATGTAATTCAGCTCC
GCCATCGCCGCTTCCACTTTTTCCCGCGTTTTCGCAGAAACGTGGCTGGCCTGGTTCACCACGCGGGAAACGGTCTGATA
AGAGACACCGGCATACTCTGCGACATCGTATAACGTTACTGGTTTCACATTCACCACCCTGAATTGACTCTCTTCCGGGCGC
TATCATGCCATACCGCGAAAGGTTTTGCGCCATTCGATGGTGTCCGGGATCTCGACGCTCTCCCTTATGCGACTCCTGCATT
AGGAAGCAGCCCAGTAGTAGGTTGAGGCCGTTGAGCACCGCCGCCGCAAGGAATGGTGCATGCAAGGAGATGGCGCCCA
ACAGTCCCCCGGCCACGGGGCCTGCCACCATACCCACGCCGAAACAAGCGCTCATGAGCCCGAAGTGGCGAGCCCGATC
TTCCCCATCGGTGATGTCGGCGATATAGGCGCCAGCAACCGCACCTGTGGCGCCGGTGATGCCGGCCACGATGCGTCCG
GCGTAGAGGATCGAGATCTCGATCCCGCGAAATTAATACGACTCACTATAGGGGAATTGTGAGCGGATAACAATTCCCCTCT
AGGATCC

TermtRNA-Luc

CTCGAGCACCACCACCACCACCACTGAGATCCGGCTGCTAACAAAGCCCGAAAGGAAGCTGAGTTGGCTGCTGCCACCG
CTGAGCAATAACTAGCATAACCCCTTGGGGCCTCTAAACGGGTCTTGAGGGGTTTTTTGCTGAAAGGAGGAACTATATCCG
GATTGGCGAATGGGACGCGCCCTGTAGCGGCGCATTAAGCGCGGCGGGTGTGGTGGTTACGCGCAGCGTGACCGCTACA
CTTGCCAGCGCCCTAGCGCCCGCTCCTTTCGCTTTCTTCCCTTCCTTTCTCGCCACGTTCGCCGGCTTTCCCCGTCAAGCT
CTAAATCGGGGGCTCCCTTTAGGGTTCCGATTTAGTGCTTTACGGCACCTCGACCCCAAAAAACTTGATTAGGGTGATGGTT
CACGTAGTGGGCCATCGCCCTGATAGACGGTTTTTCGCCCTTTGACGTTGGAGTCCACGTTCTTTAATAGTGGACTCTTGTT
CCAAACTGGAACAACACTCAACCCTATCTCGGTCTATTCTTTTGATTTATAAGGGATTTTGCCGATTTCGGCCTATTGGTTAAA
AAATGAGCTGATTTAACAAAAATTTAACGCGAATTTTAACAAAATATTAACGCTTACAATTTAGGTGGCACTTTTCGGGGAAAT
GTGCGCGGAACCCCTATTTGTTTATTTTTCTAAATACATTCAAATATGTATCCGCTCATGAATTAATTCTTAGAAAAACTCATCG
AGCATCAAATGAAACTGCAATTTATTCATATCAGGATTATCAATACCATATTTTTGAAAAAGCCGTTTCTGTAATGAAGGAGAAA
ACTCACCGAGGCAGTTCCATAGGATGGCAAGATCCTGGTATCGGTCTGCGATTCCGACTCGTCCAACATCAATACAACCTAT
TAATTTCCCCTCGTCAAAAATAAGGTTATCAAGTGAGAAATCACCATGAGTGACGACTGAATCCGGTGAGAATGGCAAAAGT
TTATGCATTTCTTTCCAGACTTGTTCAACAGGCCAGCCATTACGCTCGTCATCAAAATCACTCGCATCAACCAAACCGTTATT
CATTCGTGATTGCGCCTGAGCGAGACGAAATACGCGATCGCTGTTAAAAGGACAATTACAAACAGGAATCGAATGCAACCG
GCGCAGGAACACTGCCAGCGCATCAACAATATTTTCACCTGAATCAGGATATTCTTCTAATACCTGGAATGCTGTTTTCCCGG
GGATCGCAGTGGTGAGTAACCATGCATCATCAGGAGTACGGATAAAATGCTTGATGGTCGGAAGAGGCATAAATTCCGTCA
GCCAGTTTAGTCTGACCATCTCATCTGTAACATCATTGGCAACGCTACCTTTGCCATGTTTCAGAAACAACTCTGGCGCATC
GGGCTTCCCATACAATCGATAGATTGTCGCACCTGATTGCCCGACATTATCGCGAGCCCATTTATACCCATATAAATCAGCAT
CCATGTTGGAATTTAATCGCGGCCTAGAGCAAGACGTTTCCCGTTGAATATGGCTCATAACACCCCTTGTATTACTGTTTATG
TAAGCAGACAGTTTTATTGTTCATGACCAAAATCCCTTAACGTGAGTTTTCGTTCCACTGAGCGTCAGACCCCGTAGAAAAG
ATCAAAGGATCTTCTTGAGATCCTTTTTTTCTGCGCGTAATCTGCTGCTTGCAAACAAAAAAACCACCGCTACCAGCGGTGG
TTTGTTTGCCGGATCAAGAGCTACCAACTCTTTTTCCGAAGGTAACTGGCTTCAGCAGAGCGCAGATACCAAATACTGTCCT
TCTAGTGTAGCCGTAGTTAGGCCACCACTTCAAGAACTCTGTAGCACCGCCTACATACCTCGCTCTGCTAATCCTGTTACCA

GTGGCTGCTGCCAGTGGCGATAAGTCGTGTCTTACCGGGTTGGACTCAAGACGATAGTTACCGGATAAGGCGCAGCGGTC
GGGCTGAACGGGGGGTTCGTGCACACAGCCCAGCTTGGAGCGAACGACCTACACCGAACTGAGATACCTACAGCGTGAG
CTATGAGAAAGCGCCACGCTTCCCGAAGGGAGAAAGGCGGACAGGTATCCGGTAAGCGGCAGGGTCGGAACAGGAGAG
CGCACGAGGGAGCTTCCAGGGGGAAACGCCTGGTATCTTTATAGTCCTGTCGGGTTTCGCCACCTCTGACTTGAGCGTCG
ATTTTTGTGATGCTCGTCAGGGGGGCGGAGCCTATGGAAAAACGCCAGCAACGCGGCCTTTTTACGGTTCCTGGCCTTTT
GCTGGCCTTTTGCTCACATGTTCTTTCCTGCGTTATCCCCTGATTCTGTGGATAACCGTATTACCGCCTTTGAGTGAGCTGAT
ACCGCTCGCCGCAGCCGAACGACCGAGCGCAGCGAGTCAGTGAGCGAGGAAGCGGAAGAGCGCCTGATGCGGTATTTT
CTCCTTACGCATCTGTGCGGTATTTCACACCGCAATGGTGCACTCTCAGTACAATCTGCTCTGATGCCGCATAGTTAAGCCA
GTATACACTCCGCTATCGCTACGTGACTGGGTCATGGCTGCGCCCCGACACCCGCCAACACCCGCTGACGCGCCCTGACG
GGCTTGTCTGCTCCCGGCATCCGCTTACAGACAAGCTGTGACCGTCTCCGGGAGCTGCATGTGTCAGAGGTTTTCACCGT
CATCACCGAAACGCGCGAGGCAGCTGCGGTAAAGCTCATCAGCGTGGTCGTGAAGCGATTCACAGATGTCTGCCTGTTCA
TCCGCGTCCAGCTCGTTGAGTTTCTCCAGAAGCGTTAATGTCTGGCTTCTGATAAAGCGGGCCATGTTAAGGGCGGTTTTT
TCCTGTTTGGTCACTGATGCCTCCGTGTAAGGGGGATTTCTGTTCATGGGGGTAATGATACCGATGAAACGAGAGAGGATG
CTCACGATACGGGTTACTGATGATGAACATGCCCGGTTACTGGAACGTTGTGAGGGTAAACAACTGGCGGTATGGATGCGG
CGGGACCAGAGAAAAATCACTCAGGGTCAATGCCAGCGCTTCGTTAATACAGATGTAGGTGTTCCACAGGGTAGCCAGCA
GCATCCTGCGATGCAGATCCGGAACATAATGGTGCAGGGCGCTGACTTCCGCGTTTCCAGACTTTACGAAACACGGAAAC
CGAAGACCATTCATGTTGTTGCTCAGGTCGCAGACGTTTTGCAGCAGCAGTCGCTTCACGTTCGCTCGCGTATCGGTGATT
CATTCTGCTAACCAGTAAGGCAACCCCGCCAGCCTAGCCGGGTCCTCAACGACAGGAGCACGATCATGCGCACCCGTGG
GGCCGCCATGCCGGCGATAATGGCCTGCTTCTCGCCGAAACGTTTGGTGGCGGGACCAGTGACGAAGGCTTGAGCGAGG
GCGTGCAAGATTCCGAATACCGCAAGCGACAGGCCGATCATCGTCGCGCTCCAGCGAAAGCGGTCCTCGCCGAAAATGAC
CCAGAGCGCTGCCGGCACCTGTCCTACGAGTTGCATGATAAAGAAGACAGTCATAAGTGCGGCGACGATAGTCATGCCCC
GCGCCCACCGGAAGGAGCTGACTGGGTTGAAGGCTCTCAAGGGCATCGGTCGAGATCCCGGTGCCTAATGAGTGAGCTA
ACTTACATTAATTGCGTTGCGCTCACTGCCCGCTTTCCAGTCGGGAAACCTGTCGTGCCAGCTGCATTAATGAATCGGCCA
ACGCGCGGGGAGAGGCGGTTTGCGTATTGGGCGCCAGGGTGGTTTTTCTTTTCACCAGTGAGACGGGCAACAGCTGATT
GCCCTTCACCGCCTGGCCCTGAGAGAGTTGCAGCAAGCGGTCCACGCTGGTTTGCCCCAGCAGGCGAAAATCCTGTTTG
ATGGTGGTTAACGGCGGGATATAACATGAGCTGTCTTCGGTATCGTCGTATCCCACTACCGAGATATCCGCACCAACGCGCA
GCCCGGACTCGGTAATGGCGCGCATTGCGCCCAGCGCCATCTGATCGTTGGCAACCAGCATCGCAGTGGGAACGATGCC
CTCATTCAGCATTTGCATGGTTTGTTGAAAACCGGACATGGCACTCCAGTCGCCTTCCCGTTCCGCTATCGGCTGAATTTGA
TTGCGAGTGAGATATTTATGCCAGCCAGCCAGACGCAGACGCGCCGAGACAGAACTTAATGGGCCCGCTAACAGCGCGAT
TTGCTGGTGACCCAATGCGACCAGATGCTCCACGCCCAGTCGCGTACCGTCTTCATGGGAGAAAATAATACTGTTGATGGG
TGTCTGGTCAGAGACATCAAGAAATAACGCCGGAACATTAGTGCAGGCAGCTTCCACAGCAATGGCATCCTGGTCATCCAG
CGGATAGTTAATGATCAGCCCACTGACGCGTTGCGCGAGAAGATTGTGCACCGCCGCTTTACAGGCTTCGACGCCGCTTC
GTTCTACCATCGACACCACCACGCTGGCACCCAGTTGATCGGCGCGAGATTTAATCGCCGCGACAATTTGCGACGGCGCG
TGCAGGGCCAGACTGGAGGTGGCAACGCCAATCAGCAACGACTGTTTGCCCGCCAGTTGTTGTGCCACGCGGTTGGGAA
TGTAATTCAGCTCCGCCATCGCCGCTTCCACTTTTTCCCGCGTTTTCGCAGAAACGTGGCTGGCCTGGTTCACCACGCGG
GAAACGGTCTGATAAGAGACACCGGCATACTCTGCGACATCGTATAACGTTACTGGTTTCACATTCACCACCCTGAATTGAC
TCTCTTCCGGGCGCTATCATGCCATACCGCGAAAGGTTTTGCGCCATTCGATGGTGTCCGGGATCTCGACGCTCTCCCTTA
TGCGACTCCTGCATTAGGAAGCAGCCCAGTAGTAGGTTGAGGCCGTTGAGCACCGCCGCCGCAAGGAATGGTGCATGCAA
GGAGATGGCGCCCAACAGTCCCCCGGCCACGGGGCCTGCCACCATACCCACGCCGAAACAAGCGCTCATGAGCCCGAA
GTGGCGAGCCCGATCTTCCCCATCGGTGATGTCGGCGATATAGGCGCCAGCAACCGCACCTGTGGCGCCGGTGATGCCG
GCCACGATGCGTCCGGCGTAGAGGATCGAGATCTCGATCCCGCGAAATTAATACGACTCACTATAGGGGAATTGTGAGCGG
ATAACAATTCCCCTCTAGGATCCGCCCGGATAGCTCAGGCGGGAGAGCAGCGGCCGAGTAATTTACGTCGACGGTGACGT
CGATGGTTGCGGCCGCGGGTCCAGGGTTCAAGTCCCTGTTCGGGCGCCACTTGTTCTTTTTGCAGAAGCTCAGAATAAAC
GCTCAACTTTGGGCCGCCACCATGAAGACCTTAATTCTTGCCGTTGCATTAGTCTACTGCGCCACTGTTCATTGCCAGGACT
GTCCTTACGAACCTGATCCACCAAACACAGTTCCAACTTCCTGTGAAGCTAAAGAAGGAGAATGTATTGATAGCAGCTGTGG
CACCTGCACGAGAGACATACTATCAGATGGACTGTGTGAAAATAAACCAGGAAAAACATGTTGCCGAATGTGTCAGTATGTA
ATTGAATGCAGAGTAGAGGCCGCAGGATGGTTTAGAACATTCTATGGAAAGAGATTCCAGTTCCAGGAACCTGGTACATACG
TGTTGGGTCAAGGAACCAAGGGCGGCGACTGGAAGGTGTCCATCACCCTGGAGAACCTGGATGGAACCAAGGGGGCTGT
GCTGACCAAGACAAGACTGGAAGTGGCTGGAGACATCATTGACATCGCTCAAGCTACTGAGAATCCCATCACTGTAAACGG
TGGAGCTGACCCTATCATCGCCAACCCGTACACCATCGGCGAGGTCACCATCGCTGTTGTTGAGATGCCAGGCTTCAACAT
CACCGTCATTGAGTTCTTCAAACTGATCGTGATCGACATCCTCGGAGGAAGATCTGTAAGAATCGCCCCAGACACAGCAAA
CAAAGGAATGATCTCTGGCCTCTGTGGAGATCTTAAAATGATGGAAGATACAGACTTCACTTCAGATCCAGAACAACTCGCT
ATTCAGCCTAAGATCAACCAGGAGTTTGACGGTTGTCCACTCTATGGAAATCCTGATGACGTTGCATACTGCAAAGGTCTTC
TGGAGCCGTACAAGGACAGCTGCCGCAACCCCATCAACTTCTACTACTACACCATCTCCTGCGCCTTCGCCCGCTGTATGG
GTGGAGACGAGCGAGCCTCACACGTGCTGCTTGACTACAGGGAGACGTGCGCTGCTCCCGAAACTAGAGGAACCTGCGT
TTTGTCTGGACATACTTTCTACGATACATTTGACAAAGCAAGATACCAATTCCAGGGTCCCTGCAAGGAGATTCTTATGGCCG
CCGACTGTTTCTGGAACACTTGGGATGTGAAGGTTTCACACAGGAATGTTGACTCTTACACTGAAGTAGAGAAAGTACGAAT
CAGGAAACAATCGACTGTAGTAGAACTCATTGTTGATGGAAAACAGATTCTGGTTGGAGGAGAAGCCGTGTCCGTCCCGTA
CAGCTCTCAGAACACTTCCATCTACTGGCAAGATGGTGACATACTGACTACAGCCATCCTACCTGAAGCTCTGGTGGTCAA
GTTCAACTTCAAGCAACTGCTCGTCGTACATATTAGAGATCCATTCGATGGTAAGACTTGCGGTATTTGCGGTAACTACAACC
AGGATTTCAGTGATGATTCTTTTGATGCTGAAGGAGCCTGTGATCTGACCCCCAACCCACCGGGATGCACCGAAGAACAGA
AACCTGAAGCTGAACGACTCTGCAATAGTCTCTTCGCCGGTCAAAGTGATCTTGATCAGAAATGTAACGTGTGCCACAAGC
CTGACCGTGTCGAACGATGCATGTACGAGTATTGCCTGAGGGGACAACAGGGTTTCTGTGACCACGCATGGGAGTTCAAG
AAAGAATGCTACATAAAGCATGGAGACACCCTAGAAGTACCAGATGAATGCAAATAGAACCAGCCTCAAGAACACCCGAATG
GAGTCTCTAAGCTACATAATACCAACTTACACTTTACAAAATGTTGTCCCCCAAAATGTAGCCATTCGTATCTGCTCCTAATAA
AAAGAAAGTTTCTTCACACCGCGGGCTTGTCCCTGAACTTGGGACCTGGGCGCCGGCGTTGGTAGCTGCAGTGGCAGCT
GCATTTAATGAGCCGGCGACGAGAGGGCGGACTCGATAGGCCCG

TermtRNA-Spike

CTCGAGCACCACCACCACCACCACTGAGATCCGGCTGCTAACAAAGCCCGAAAGGAAGCTGAGTTGGCTGCTGCCACCG
CTGAGCAATAACTAGCATAACCCCTTGGGGCCTCTAAACGGGTCTTGAGGGGTTTTTTGCTGAAAGGAGGAACTATATCCG
GATTGGCGAATGGGACGCGCCCTGTAGCGGCGCATTAAGCGCGGCGGGTGTGGTGGTTACGCGCAGCGTGACCGCTACA
CTTGCCAGCGCCCTAGCGCCCGCTCCTTTCGCTTTCTTCCCTTCCTTTCTCGCCACGTTCGCCGGCTTTCCCCGTCAAGCT
CTAAATCGGGGGCTCCCTTTAGGGTTCCGATTTAGTGCTTTACGGCACCTCGACCCCAAAAAACTTGATTAGGGTGATGGTT
CACGTAGTGGGCCATCGCCCTGATAGACGGTTTTTCGCCCTTTGACGTTGGAGTCCACGTTCTTTAATAGTGGACTCTTGTT
CCAAACTGGAACAACACTCAACCCTATCTCGGTCTATTCTTTTGATTTATAAGGGATTTTGCCGATTTCGGCCTATTGGTTAAA
AAATGAGCTGATTTAACAAAAATTTAACGCGAATTTTAACAAAATATTAACGCTTACAATTTAGGTGGCACTTTTCGGGGAAAT
GTGCGCGGAACCCCTATTTGTTTATTTTTCTAAATACATTCAAATATGTATCCGCTCATGAATTAATTCTTAGAAAAACTCATCG
AGCATCAAATGAAACTGCAATTTATTCATATCAGGATTATCAATACCATATTTTTGAAAAAGCCGTTTCTGTAATGAAGGAGAAA
ACTCACCGAGGCAGTTCCATAGGATGGCAAGATCCTGGTATCGGTCTGCGATTCCGACTCGTCCAACATCAATACAACCTAT
TAATTTCCCCTCGTCAAAAATAAGGTTATCAAGTGAGAAATCACCATGAGTGACGACTGAATCCGGTGAGAATGGCAAAAGT

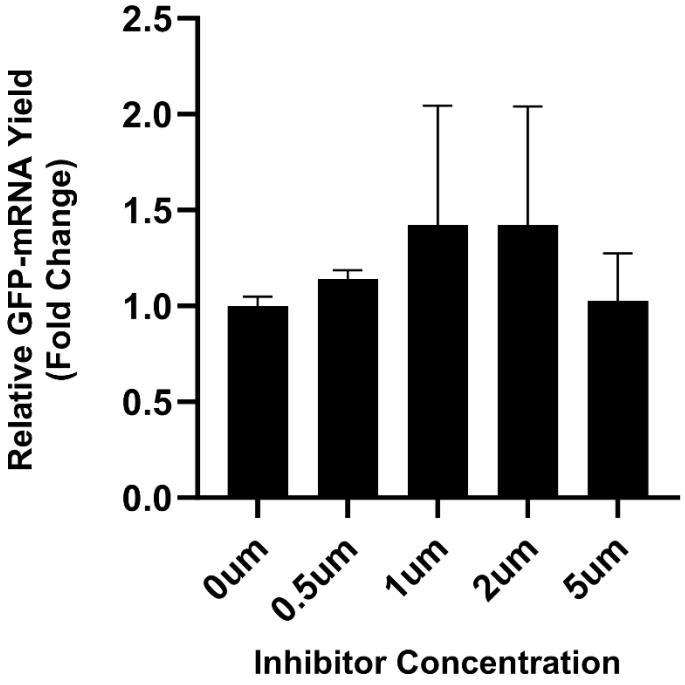
TTATGCATTTCTTTCCAGACTTGTTCAACAGGCCAGCCATTACGCTCGTCATCAAAATCACTCGCATCAACCAAACCGTTATT
CATTCGTGATTGCGCCTGAGCGAGACGAAATACGCGATCGCTGTTAAAAGGACAATTACAAACAGGAATCGAATGCAACCG
GCGCAGGAACACTGCCAGCGCATCAACAATATTTTCACCTGAATCAGGATATTCTTCTAATACCTGGAATGCTGTTTTCCCGG
GGATCGCAGTGGTGAGTAACCATGCATCATCAGGAGTACGGATAAAATGCTTGATGGTCGGAAGAGGCATAAATTCCGTCA
GCCAGTTTAGTCTGACCATCTCATCTGTAACATCATTGGCAACGCTACCTTTGCCATGTTTCAGAAACAACTCTGGCGCATC
GGGCTTCCCATACAATCGATAGATTGTCGCACCTGATTGCCCGACATTATCGCGAGCCCATTTATACCCATATAAATCAGCAT
CCATGTTGGAATTTAATCGCGGCCTAGAGCAAGACGTTTCCCGTTGAATATGGCTCATAACACCCCTTGTATTACTGTTTATG
TAAGCAGACAGTTTTATTGTTCATGACCAAAATCCCTTAACGTGAGTTTTCGTTCCACTGAGCGTCAGACCCCGTAGAAAAG
ATCAAAGGATCTTCTTGAGATCCTTTTTTTCTGCGCGTAATCTGCTGCTTGCAAACAAAAAAACCACCGCTACCAGCGGTGG
TTTGTTTGCCGGATCAAGAGCTACCAACTCTTTTTCCGAAGGTAACTGGCTTCAGCAGAGCGCAGATACCAAATACTGTCCT
TCTAGTGTAGCCGTAGTTAGGCCACCACTTCAAGAACTCTGTAGCACCGCCTACATACCTCGCTCTGCTAATCCTGTTACCA
GTGGCTGCTGCCAGTGGCGATAAGTCGTGTCTTACCGGGTTGGACTCAAGACGATAGTTACCGGATAAGGCGCAGCGGTC
GGGCTGAACGGGGGGTTCGTGCACACAGCCCAGCTTGGAGCGAACGACCTACACCGAACTGAGATACCTACAGCGTGAG
CTATGAGAAAGCGCCACGCTTCCCGAAGGGAGAAAGGCGGACAGGTATCCGGTAAGCGGCAGGGTCGGAACAGGAGAG
CGCACGAGGGAGCTTCCAGGGGGAAACGCCTGGTATCTTTATAGTCCTGTCGGGTTTCGCCACCTCTGACTTGAGCGTCG
ATTTTTGTGATGCTCGTCAGGGGGGCGGAGCCTATGGAAAAACGCCAGCAACGCGGCCTTTTTACGGTTCCTGGCCTTTT
GCTGGCCTTTTGCTCACATGTTCTTTCCTGCGTTATCCCCTGATTCTGTGGATAACCGTATTACCGCCTTTGAGTGAGCTGAT
ACCGCTCGCCGCAGCCGAACGACCGAGCGCAGCGAGTCAGTGAGCGAGGAAGCGGAAGAGCGCCTGATGCGGTATTTT
CTCCTTACGCATCTGTGCGGTATTTCACACCGCAATGGTGCACTCTCAGTACAATCTGCTCTGATGCCGCATAGTTAAGCCA
GTATACACTCCGCTATCGCTACGTGACTGGGTCATGGCTGCGCCCCGACACCCGCCAACACCCGCTGACGCGCCCTGACG
GGCTTGTCTGCTCCCGGCATCCGCTTACAGACAAGCTGTGACCGTCTCCGGGAGCTGCATGTGTCAGAGGTTTTCACCGT
CATCACCGAAACGCGCGAGGCAGCTGCGGTAAAGCTCATCAGCGTGGTCGTGAAGCGATTCACAGATGTCTGCCTGTTCA
TCCGCGTCCAGCTCGTTGAGTTTCTCCAGAAGCGTTAATGTCTGGCTTCTGATAAAGCGGGCCATGTTAAGGGCGGTTTTT
TCCTGTTTGGTCACTGATGCCTCCGTGTAAGGGGGATTTCTGTTCATGGGGGTAATGATACCGATGAAACGAGAGAGGATG
CTCACGATACGGGTTACTGATGATGAACATGCCCGGTTACTGGAACGTTGTGAGGGTAAACAACTGGCGGTATGGATGCGG
CGGGACCAGAGAAAAATCACTCAGGGTCAATGCCAGCGCTTCGTTAATACAGATGTAGGTGTTCCACAGGGTAGCCAGCA
GCATCCTGCGATGCAGATCCGGAACATAATGGTGCAGGGCGCTGACTTCCGCGTTTCCAGACTTTACGAAACACGGAAAC
CGAAGACCATTCATGTTGTTGCTCAGGTCGCAGACGTTTTGCAGCAGCAGTCGCTTCACGTTCGCTCGCGTATCGGTGATT
CATTCTGCTAACCAGTAAGGCAACCCCGCCAGCCTAGCCGGGTCCTCAACGACAGGAGCACGATCATGCGCACCCGTGG
GGCCGCCATGCCGGCGATAATGGCCTGCTTCTCGCCGAAACGTTTGGTGGCGGGACCAGTGACGAAGGCTTGAGCGAGG
GCGTGCAAGATTCCGAATACCGCAAGCGACAGGCCGATCATCGTCGCGCTCCAGCGAAAGCGGTCCTCGCCGAAAATGAC
CCAGAGCGCTGCCGGCACCTGTCCTACGAGTTGCATGATAAAGAAGACAGTCATAAGTGCGGCGACGATAGTCATGCCCC
GCGCCCACCGGAAGGAGCTGACTGGGTTGAAGGCTCTCAAGGGCATCGGTCGAGATCCCGGTGCCTAATGAGTGAGCTA
ACTTACATTAATTGCGTTGCGCTCACTGCCCGCTTTCCAGTCGGGAAACCTGTCGTGCCAGCTGCATTAATGAATCGGCCA
ACGCGCGGGGAGAGGCGGTTTGCGTATTGGGCGCCAGGGTGGTTTTTCTTTTCACCAGTGAGACGGGCAACAGCTGATT
GCCCTTCACCGCCTGGCCCTGAGAGAGTTGCAGCAAGCGGTCCACGCTGGTTTGCCCCAGCAGGCGAAAATCCTGTTTG
ATGGTGGTTAACGGCGGGATATAACATGAGCTGTCTTCGGTATCGTCGTATCCCACTACCGAGATATCCGCACCAACGCGCA
GCCCGGACTCGGTAATGGCGCGCATTGCGCCCAGCGCCATCTGATCGTTGGCAACCAGCATCGCAGTGGGAACGATGCC
CTCATTCAGCATTTGCATGGTTTGTTGAAAACCGGACATGGCACTCCAGTCGCCTTCCCGTTCCGCTATCGGCTGAATTTGA
TTGCGAGTGAGATATTTATGCCAGCCAGCCAGACGCAGACGCGCCGAGACAGAACTTAATGGGCCCGCTAACAGCGCGAT
TTGCTGGTGACCCAATGCGACCAGATGCTCCACGCCCAGTCGCGTACCGTCTTCATGGGAGAAAATAATACTGTTGATGGG
TGTCTGGTCAGAGACATCAAGAAATAACGCCGGAACATTAGTGCAGGCAGCTTCCACAGCAATGGCATCCTGGTCATCCAG
CGGATAGTTAATGATCAGCCCACTGACGCGTTGCGCGAGAAGATTGTGCACCGCCGCTTTACAGGCTTCGACGCCGCTTC
GTTCTACCATCGACACCACCACGCTGGCACCCAGTTGATCGGCGCGAGATTTAATCGCCGCGACAATTTGCGACGGCGCG
TGCAGGGCCAGACTGGAGGTGGCAACGCCAATCAGCAACGACTGTTTGCCCGCCAGTTGTTGTGCCACGCGGTTGGGAA
TGTAATTCAGCTCCGCCATCGCCGCTTCCACTTTTTCCCGCGTTTTCGCAGAAACGTGGCTGGCCTGGTTCACCACGCGG
GAAACGGTCTGATAAGAGACACCGGCATACTCTGCGACATCGTATAACGTTACTGGTTTCACATTCACCACCCTGAATTGAC
TCTCTTCCGGGCGCTATCATGCCATACCGCGAAAGGTTTTGCGCCATTCGATGGTGTCCGGGATCTCGACGCTCTCCCTTA
TGCGACTCCTGCATTAGGAAGCAGCCCAGTAGTAGGTTGAGGCCGTTGAGCACCGCCGCCGCAAGGAATGGTGCATGCAA
GGAGATGGCGCCCAACAGTCCCCCGGCCACGGGGCCTGCCACCATACCCACGCCGAAACAAGCGCTCATGAGCCCGAA
GTGGCGAGCCCGATCTTCCCCATCGGTGATGTCGGCGATATAGGCGCCAGCAACCGCACCTGTGGCGCCGGTGATGCCG
GCCACGATGCGTCCGGCGTAGAGGATCGAGATCTCGATCCCGCGAAATTAATACGACTCACTATAGGGGAATTGTGAGCGG
ATAACAATTCCCCTCTAGGATCCGCCCGGATAGCTCAGGCGGGAGAGCAGCGGCCGAGTAATTTACGTCGACGGTGACGT
CGATGGTTGCGGCCGCGGGTCCAGGGTTCAAGTCCCTGTTCGGGCGCCACTTGTTCTTTTTGCAGAAGCTCAGAATAAAC
GCTCAACTTTGGGCCGCCACCATGGTGTTCGTGTTTCTGGTGCTGCTGCCTCTGGTGTCCAGCCAGTGTGTGAACCTGAC
CACAAGAACCCAGCTGCCTCCAGCCTACACCAACAGCTTTACCAGAGGCGTGTACTACCCCGACAAGGTGTTCAGATCCA
GCGTGCTGCACTCTACCCAGGACCTGTTCCTGCCTTTCTTCAGCAACGTGACCTGGTTCCACGCCATCCACGTGTCCGGC
ACCAATGGCACCAAGAGATTCGACAACCCCGTGCTGCCCTTCAACGACGGGGTGTACTTTGCCAGCACCGAGAAGTCCAA
CATCATCAGAGGCTGGATCTTCGGCACCACACTGGACAGCAAGACCCAGAGCCTGCTGATCGTGAACAACGCCACCAACG
TGGTCATCAAAGTGTGCGAGTTCCAGTTCTGCAACGACCCCTTCCTGGGCGTCTACTATCACAAGAACAACAAGAGCTGGA
TGGAAAGCGAGTTCCGGGTGTACAGCAGCGCCAACAACTGCACCTTCGAGTACGTGTCCCAGCCTTTCCTGATGGACCTG
GAAGGCAAGCAGGGCAACTTCAAGAACCTGCGCGAGTTCGTGTTCAAGAACATCGACGGCTACTTCAAGATCTACAGCAA
GCACACCCCTATCAACCTCGTGCGGGATCTGCCTCAGGGCTTCTCTGCTCTGGAACCCCTGGTGGATCTGCCCATCGGCA
TCAACATCACCCGGTTTCAGACACTGCTGGCCCTGCACAGAAGCTACCTGACACCTGGCGATAGCAGCAGCGGATGGACA
GCTGGTGCCGCCGCTTACTATGTGGGCTACCTGCAGCCTAGAACCTTTCTGCTGAAGTACAACGAGAACGGCACCATCAC
CGACGCCGTGGATTGTGCTCTGGATCCTCTGAGCGAGACAAAGTGCACCCTGAAGTCCTTCACCGTGGAAAAGGGCATCT
ACCAGACCAGCAACTTCCGGGTGCAGCCCACCGAATCCATCGTGCGGTTCCCCAATATCACCAATCTGTGCCCCTTCGGC
GAGGTGTTCAATGCCACCAGATTCGCCTCTGTGTACGCCTGGAACCGGAAGCGGATCAGCAATTGCGTGGCCGACTACTC
CGTGCTGTACAACTCCGCCAGCTTCAGCACCTTCAAGTGCTACGGCGTGTCCCCTACCAAGCTGAACGACCTGTGCTTCA
CAAACGTGTACGCCGACAGCTTCGTGATCCGGGGAGATGAAGTGCGGCAGATTGCCCCTGGACAGACAGGCAAGATCGC
CGACTACAACTACAAGCTGCCCGACGACTTCACCGGCTGTGTGATTGCCTGGAACAGCAACAACCTGGACTCCAAAGTCG
GCGGCAACTACAATTACCTGTACCGGCTGTTCCGGAAGTCCAATCTGAAGCCCTTCGAGCGGGACATCTCCACCGAGATCT
ATCAGGCCGGCAGCACCCCTTGTAACGGCGTGGAAGGCTTCAACTGCTACTTCCCACTGCAGTCCTACGGCTTTCAGCCC
ACAAATGGCGTGGGCTATCAGCCCTACAGAGTGGTGGTGCTGAGCTTCGAACTGCTGCATGCCCCTGCCACAGTGTGCGG
CCCTAAGAAAAGCACCAATCTCGTGAAGAACAAATGCGTGAACTTCAACTTCAACGGCCTGACCGGCACCGGCGTGCTGA
CAGAGAGCAACAAGAAGTTCCTGCCATTCCAGCAGTTTGGCCGGGATATCGCCGATACCACAGACGCCGTTAGAGATCCC
CAGACACTGGAAATCCTGGACATCACCCCTTGCAGCTTCGGCGGAGTGTCTGTGATCACCCCTGGCACCAACACCAGCAA
TCAGGTGGCAGTGCTGTACCAGGACGTGAACTGTACCGAAGTGCCCGTGGCCATTCACGCCGATCAGCTGACACCTACAT
GGCGGGTGTACTCCACCGGCAGCAATGTGTTTCAGACCAGAGCCGGCTGTCTGATCGGAGCCGAGCACGTGAACAATAG
CTACGAGTGCGACATCCCCATCGGCGCTGGCATCTGTGCCAGCTACCAGACACAGACAAACAGCCCCGCCTCTGTGGCCA

GCCAGAGCATCATTGCCTACACAATGTCTCTGGGCGCCGAGAACAGCGTGGCCTACTCCAACAACTCTATCGCTATCCCCA
CCAACTTCACCATCAGCGTGACCACAGAGATCCTGCCTGTGTCCATGACCAAGACCAGCGTGGACTGCACCATGTACATCT
GCGGCGATTCCACCGAGTGCTCCAACCTGCTGCTGCAGTACGGCAGCTTCTGCACCCAGCTGAATAGAGCCCTGACAGG
GATCGCCGTGGAACAGGACAAGAACACCCAAGAGGTGTTCGCCCAAGTGAAGCAGATCTACAAGACCCCTCCTATCAAGG
ACTTCGGCGGCTTCAATTTCAGCCAGATTCTGCCCGATCCTAGCAAGCCCAGCAAGCGGAGCTTCATCGAGGACCTGCTG
TTCAACAAAGTGACACTGGCCGACGCCGGCTTCATCAAGCAGTATGGCGATTGTCTGGGCGACATTGCCGCCAGGGATCT
GATTTGCGCCCAGAAGTTTAACGGACTGACAGTGCTGCCTCCTCTGCTGACCGATGAGATGATCGCCCAGTACACATCTGC
CCTGCTGGCCGGCACAATCACAAGCGGCTGGACATTTGGAGCTGGCGCCGCTCTGCAGATCCCCTTTGCTATGCAGATGG
CCTACCGGTTCAACGGCATCGGAGTGACCCAGAATGTGCTGTACGAGAACCAGAAGCTGATCGCCAACCAGTTCAACAGC
GCCATCGGCAAGATCCAGGACAGCCTGAGCAGCACAGCAAGCGCCCTGGGAAAGCTGCAGGACGTGGTCAACCAGAATG
CCCAGGCACTGAACACCCTGGTCAAGCAGCTGTCCTCCAACTTCGGCGCCATCAGCTCTGTGCTGAACGATATCCTGAGC
AGACTGGACcctcctGAAGCCGAGGTGCAGATCGACAGACTGATCACCGGAAGGCTGCAGTCCCTGCAGACCTACGTTACCC
AGCAGCTGATCAGAGCCGCCGAGATTAGAGCCTCTGCCAATCTGGCCGCCACCAAGATGTCTGAGTGTGTGCTGGGCCA
GAGCAAGAGAGTGGACTTTTGCGGCAAGGGCTACCACCTGATGAGCTTCCCTCAGTCTGCCCCTCACGGCGTGGTGTTTC
TGCACGTGACATACGTGCCCGCTCAAGAGAAGAATTTCACCACCGCTCCAGCCATCTGCCACGACGGCAAAGCCCACTTT
CCTAGAGAAGGCGTGTTCGTGTCCAACGGCACCCATTGGTTCGTGACCCAGCGGAACTTCTACGAGCCCCAGATCATCAC
CACCGACAACACCTTCGTGTCTGGCAACTGCGACGTCGTGATCGGCATTGTGAACAATACCGTGTACGACCCTCTGCAGC
CCGAGCTGGACAGCTTCAAAGAGGAACTGGATAAGTACTTTAAGAACCACACAAGCCCCGACGTGGACCTGGGCGATATC
AGCGGAATCAATGCCAGCGTCGTGAACATCCAGAAAGAGATCGACCGGCTGAACGAGGTGGCCAAGAATCTGAACGAGA
GCCTGATCGACCTGCAAGAACTGGGGAAGTACGAGCAGTACATCAAGTGGCCCAGCGGCCGCTTGGTCCCACGTGGCTC
ACCCGGATCTGGATACATCCCGGAGGCCCCTAGGGACGGTCAAGCTTACGTGAGAAAGGACGGCGAATGGGTTCTGCTGT
CGACCTTCTTGGGACATCATCATCATCATCACCGAATTCAGTAAAACCAGCCTCAAGAACACCCGAATGGAGTCTCTAAGCT
ACATAATACCAACTTACACTTTACAAAATGTTGTCCCCCAAAATGTAGCCATTCGTATCTGCTCCTAATAAAAAGAAAGTTTCTT
CACACCGCGGGCTTGTCCCTGAACTTGGGACCTGGGCGCCGGCGTTGGTAGCTGCAGTGGCAGCTGCATTTAATGAGCC
GGCGACGAGAGGGCGGACTCGATAGGCCCG

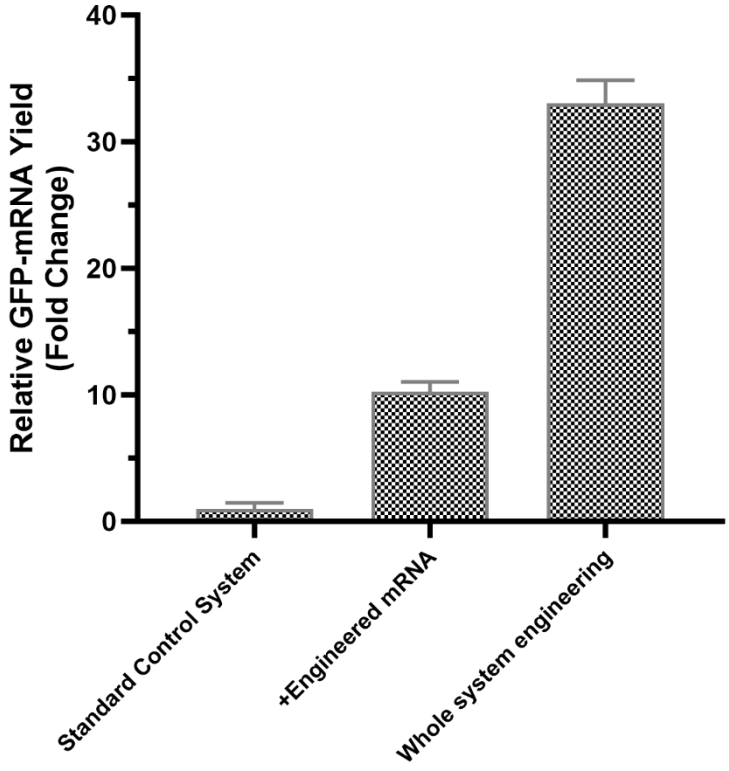
CirctRNA-GFP

CGCCCGGATAGCTCAGACGGAAGAGCAGGATATCCGAGTAATTTACGCCCGGGAGTGCTCACCATCTGATGAGGCCGAAA
GGCCGAAACCCGTGAAAACGGGTCCTTGTTCTTTTTGCAGAAGCTCAGAATAAACGCTCAACTTTGGGCCGCCACCATGGT
GAGCAAGGGCGAGGAGCTGTTCACCGGGGTGGTGCCCATCCTGGTCGAGCTGGACGGCGACGTAAACGGCCACAAGTT
CAGCGTGTCTGGCGAGGGCGAGGGCGATGCCACCTACGGCAAGCTGACCCTGAAGTTCATCTGCACCACCGGCAAGCTG
CCCGTGCCCTGGCCCACCCTCGTGACCACCCTGACCTACGGCGTGCAGTGCTTCAGCCGCTACCCCGACCACATGAAGC
AGCACGACTTCTTCAAGTCCGCCATGCCCGAAGGCTACGTCCAGGAGCGCACCATCTTCTTCAAGGACGACGGCAACTAC
AAGACCCGCGCCGAGGTGAAGTTCGAGGGCGACACCCTGGTGAACCGCATCGAGCTGAAGGGCATCGACTTCAAGGAG
GACGGCAACATCCTGGGGCACAAGCTGGAGTACAACTACAACAGCCACAACGTCTATATCATGGCCGACAAGCAGAAGAA
CGGCATCAAGGCGAACTTCAAGATCCGCCACAACATCGAGGACGGCAGCGTGCAGCTCGCCGACCACTACCAGCAGAAC
ACCCCCATCGGCGACGGCCCCGTGCTGCTGCCCGACAACCACTACCTGAGCACCCAGTCCGCCCTGAGCAAAGACCCCA
ACGAGAAGCGCGATCACATGGTCCTGCTGGAGTTCGTGACCGCCGCCGGGATCACTCTCGGCATGGACGAGCTGTACAA
GTAAAACCAGCCTCAAGAACACCCGAATGGAGTCTCTAAGCTACATAATACCAACTTACACTTTACAAAATGTTGTCCCCCAA
AATGTAGCCATTCGTATCTGCTCCTAATAAAAAGAAAGTTTCTTCACGACGGGTCTACCCGTCCTGATGAGTCCACGCGTGT
GGACGAATTACTTGTACCTCCCGGGTGGTTGCGGATATCCGGGTCCAGGGTTCAAGTCCCTGTTCGGGCGCCACTCGAGC
ACCACCACCACCACCACTGAGATCCGGCTGCTAACAAAGCCCGAAAGGAAGCTGAGTTGGCTGCTGCCACCGCTGAGCA
ATAACTAGCATAACCCCTTGGGGCCTCTAAACGGGTCTTGAGGGGTTTTTTGCTGAAAGGAGGAACTATATCCGGATTGGC
GAATGGGACGCGCCCTGTAGCGGCGCATTAAGCGCGGCGGGTGTGGTGGTTACGCGCAGCGTGACCGCTACACTTGCCA
GCGCCCTAGCGCCCGCTCCTTTCGCTTTCTTCCCTTCCTTTCTCGCCACGTTCGCCGGCTTTCCCCGTCAAGCTCTAAATC
GGGGGCTCCCTTTAGGGTTCCGATTTAGTGCTTTACGGCACCTCGACCCCAAAAAACTTGATTAGGGTGATGGTTCACGTA
GTGGGCCATCGCCCTGATAGACGGTTTTTCGCCCTTTGACGTTGGAGTCCACGTTCTTTAATAGTGGACTCTTGTTCCAAAC
TGGAACAACACTCAACCCTATCTCGGTCTATTCTTTTGATTTATAAGGGATTTTGCCGATTTCGGCCTATTGGTTAAAAAATGA
GCTGATTTAACAAAAATTTAACGCGAATTTTAACAAAATATTAACGCTTACAATTTAGGTGGCACTTTTCGGGGAAATGTGCGC
GGAACCCCTATTTGTTTATTTTTCTAAATACATTCAAATATGTATCCGCTCATGAATTAATTCTTAGAAAAACTCATCGAGCATCA
AATGAAACTGCAATTTATTCATATCAGGATTATCAATACCATATTTTTGAAAAAGCCGTTTCTGTAATGAAGGAGAAAACTCACC
GAGGCAGTTCCATAGGATGGCAAGATCCTGGTATCGGTCTGCGATTCCGACTCGTCCAACATCAATACAACCTATTAATTTC
CCCTCGTCAAAAATAAGGTTATCAAGTGAGAAATCACCATGAGTGACGACTGAATCCGGTGAGAATGGCAAAAGTTTATGCA
TTTCTTTCCAGACTTGTTCAACAGGCCAGCCATTACGCTCGTCATCAAAATCACTCGCATCAACCAAACCGTTATTCATTCGT
GATTGCGCCTGAGCGAGACGAAATACGCGATCGCTGTTAAAAGGACAATTACAAACAGGAATCGAATGCAACCGGCGCAG
GAACACTGCCAGCGCATCAACAATATTTTCACCTGAATCAGGATATTCTTCTAATACCTGGAATGCTGTTTTCCCGGGGATCG
CAGTGGTGAGTAACCATGCATCATCAGGAGTACGGATAAAATGCTTGATGGTCGGAAGAGGCATAAATTCCGTCAGCCAGTT
TAGTCTGACCATCTCATCTGTAACATCATTGGCAACGCTACCTTTGCCATGTTTCAGAAACAACTCTGGCGCATCGGGCTTC
CCATACAATCGATAGATTGTCGCACCTGATTGCCCGACATTATCGCGAGCCCATTTATACCCATATAAATCAGCATCCATGTTG
GAATTTAATCGCGGCCTAGAGCAAGACGTTTCCCGTTGAATATGGCTCATAACACCCCTTGTATTACTGTTTATGTAAGCAGA
CAGTTTTATTGTTCATGACCAAAATCCCTTAACGTGAGTTTTCGTTCCACTGAGCGTCAGACCCCGTAGAAAAGATCAAAGG
ATCTTCTTGAGATCCTTTTTTTCTGCGCGTAATCTGCTGCTTGCAAACAAAAAAACCACCGCTACCAGCGGTGGTTTGTTTG
CCGGATCAAGAGCTACCAACTCTTTTTCCGAAGGTAACTGGCTTCAGCAGAGCGCAGATACCAAATACTGTCCTTCTAGTGT
AGCCGTAGTTAGGCCACCACTTCAAGAACTCTGTAGCACCGCCTACATACCTCGCTCTGCTAATCCTGTTACCAGTGGCTG
CTGCCAGTGGCGATAAGTCGTGTCTTACCGGGTTGGACTCAAGACGATAGTTACCGGATAAGGCGCAGCGGTCGGGCTGA
ACGGGGGGTTCGTGCACACAGCCCAGCTTGGAGCGAACGACCTACACCGAACTGAGATACCTACAGCGTGAGCTATGAGA
AAGCGCCACGCTTCCCGAAGGGAGAAAGGCGGACAGGTATCCGGTAAGCGGCAGGGTCGGAACAGGAGAGCGCACGAG
GGAGCTTCCAGGGGGAAACGCCTGGTATCTTTATAGTCCTGTCGGGTTTCGCCACCTCTGACTTGAGCGTCGATTTTTGTG
ATGCTCGTCAGGGGGGCGGAGCCTATGGAAAAACGCCAGCAACGCGGCCTTTTTACGGTTCCTGGCCTTTTGCTGGCCTT
TTGCTCACATGTTCTTTCCTGCGTTATCCCCTGATTCTGTGGATAACCGTATTACCGCCTTTGAGTGAGCTGATACCGCTCGC
CGCAGCCGAACGACCGAGCGCAGCGAGTCAGTGAGCGAGGAAGCGGAAGAGCGCCTGATGCGGTATTTTCTCCTTACGC
ATCTGTGCGGTATTTCACACCGCAATGGTGCACTCTCAGTACAATCTGCTCTGATGCCGCATAGTTAAGCCAGTATACACTC
CGCTATCGCTACGTGACTGGGTCATGGCTGCGCCCCGACACCCGCCAACACCCGCTGACGCGCCCTGACGGGCTTGTCT
GCTCCCGGCATCCGCTTACAGACAAGCTGTGACCGTCTCCGGGAGCTGCATGTGTCAGAGGTTTTCACCGTCATCACCGA
AACGCGCGAGGCAGCTGCGGTAAAGCTCATCAGCGTGGTCGTGAAGCGATTCACAGATGTCTGCCTGTTCATCCGCGTCC
AGCTCGTTGAGTTTCTCCAGAAGCGTTAATGTCTGGCTTCTGATAAAGCGGGCCATGTTAAGGGCGGTTTTTTCCTGTTTGG
TCACTGATGCCTCCGTGTAAGGGGGATTTCTGTTCATGGGGGTAATGATACCGATGAAACGAGAGAGGATGCTCACGATAC
GGGTTACTGATGATGAACATGCCCGGTTACTGGAACGTTGTGAGGGTAAACAACTGGCGGTATGGATGCGGCGGGACCAG
AGAAAAATCACTCAGGGTCAATGCCAGCGCTTCGTTAATACAGATGTAGGTGTTCCACAGGGTAGCCAGCAGCATCCTGCG
ATGCAGATCCGGAACATAATGGTGCAGGGCGCTGACTTCCGCGTTTCCAGACTTTACGAAACACGGAAACCGAAGACCATT
CATGTTGTTGCTCAGGTCGCAGACGTTTTGCAGCAGCAGTCGCTTCACGTTCGCTCGCGTATCGGTGATTCATTCTGCTAA

CCAGTAAGGCAACCCCGCCAGCCTAGCCGGGTCCTCAACGACAGGAGCACGATCATGCGCACCCGTGGGGCCGCCATGC
CGGCGATAATGGCCTGCTTCTCGCCGAAACGTTTGGTGGCGGGACCAGTGACGAAGGCTTGAGCGAGGGCGTGCAAGAT
TCCGAATACCGCAAGCGACAGGCCGATCATCGTCGCGCTCCAGCGAAAGCGGTCCTCGCCGAAAATGACCCAGAGCGCT
GCCGGCACCTGTCCTACGAGTTGCATGATAAAGAAGACAGTCATAAGTGCGGCGACGATAGTCATGCCCCGCGCCCACCG
GAAGGAGCTGACTGGGTTGAAGGCTCTCAAGGGCATCGGTCGAGATCCCGGTGCCTAATGAGTGAGCTAACTTACATTAAT
TGCGTTGCGCTCACTGCCCGCTTTCCAGTCGGGAAACCTGTCGTGCCAGCTGCATTAATGAATCGGCCAACGCGCGGGGA
GAGGCGGTTTGCGTATTGGGCGCCAGGGTGGTTTTTCTTTTCACCAGTGAGACGGGCAACAGCTGATTGCCCTTCACCGC
CTGGCCCTGAGAGAGTTGCAGCAAGCGGTCCACGCTGGTTTGCCCCAGCAGGCGAAAATCCTGTTTGATGGTGGTTAACG
GCGGGATATAACATGAGCTGTCTTCGGTATCGTCGTATCCCACTACCGAGATATCCGCACCAACGCGCAGCCCGGACTCGG
TAATGGCGCGCATTGCGCCCAGCGCCATCTGATCGTTGGCAACCAGCATCGCAGTGGGAACGATGCCCTCATTCAGCATTT
GCATGGTTTGTTGAAAACCGGACATGGCACTCCAGTCGCCTTCCCGTTCCGCTATCGGCTGAATTTGATTGCGAGTGAGAT
ATTTATGCCAGCCAGCCAGACGCAGACGCGCCGAGACAGAACTTAATGGGCCCGCTAACAGCGCGATTTGCTGGTGACCC
AATGCGACCAGATGCTCCACGCCCAGTCGCGTACCGTCTTCATGGGAGAAAATAATACTGTTGATGGGTGTCTGGTCAGAG
ACATCAAGAAATAACGCCGGAACATTAGTGCAGGCAGCTTCCACAGCAATGGCATCCTGGTCATCCAGCGGATAGTTAATGA
TCAGCCCACTGACGCGTTGCGCGAGAAGATTGTGCACCGCCGCTTTACAGGCTTCGACGCCGCTTCGTTCTACCATCGAC
ACCACCACGCTGGCACCCAGTTGATCGGCGCGAGATTTAATCGCCGCGACAATTTGCGACGGCGCGTGCAGGGCCAGAC
TGGAGGTGGCAACGCCAATCAGCAACGACTGTTTGCCCGCCAGTTGTTGTGCCACGCGGTTGGGAATGTAATTCAGCTCC
GCCATCGCCGCTTCCACTTTTTCCCGCGTTTTCGCAGAAACGTGGCTGGCCTGGTTCACCACGCGGGAAACGGTCTGATA
AGAGACACCGGCATACTCTGCGACATCGTATAACGTTACTGGTTTCACATTCACCACCCTGAATTGACTCTCTTCCGGGCGC
TATCATGCCATACCGCGAAAGGTTTTGCGCCATTCGATGGTGTCCGGGATCTCGACGCTCTCCCTTATGCGACTCCTGCATT
AGGAAGCAGCCCAGTAGTAGGTTGAGGCCGTTGAGCACCGCCGCCGCAAGGAATGGTGCATGCAAGGAGATGGCGCCCA
ACAGTCCCCCGGCCACGGGGCCTGCCACCATACCCACGCCGAAACAAGCGCTCATGAGCCCGAAGTGGCGAGCCCGATC
TTCCCCATCGGTGATGTCGGCGATATAGGCGCCAGCAACCGCACCTGTGGCGCCGGTGATGCCGGCCACGATGCGTCCG
GCGTAGAGGATCGAGATCTCGATCCCGCGAAATTAATACGACTCACTATAGGGGAATTGTGAGCGGATAACAATTCCCCTCT
AGGATCC



Supplementary Figure 4.1 - SelfCirc-GFP was produced in BL21 (STAR) cells using a DNA expression vector containing a synthetic triple terminator element. The relative effect of supplementing LB Broth cell culture media with varying concentrations of the RNAse E inhibitor AS2 was evaluated in 2.5 h production processes. Data are expressed as a fold-change of the production achieved using unsupplemented media. Values represent the mean + SD of three independent experiments (n = 3, each performed in triplicate).



Supplementary Figure 4.2 – TermtRNA-GFP was produced in systems comprising either i) standard control components or ii) an optimised combination of host cell chassis, DNA expression vector and cell culture media (see Fig. 2). The relative performance of each system was evaluated in 2.5 h production processes. Data are expressed as a fold-change of the GFP-mRNA production achieved using standard control components (including the unengineered mRNA construct). Values represent the mean + SD of three independent experiments (n = 3, each performed in triplicate)