Molecular biology approaches to quantify subcellular localisation of mRNA therapeutics

Alfredo Smart

EPSRC CDT

Gene Regulation and RNA Biology

School of Pharmacy

Supervisor: Catherine Jopling

Declaration

Except where acknowledged in the text, I declare that this dissertation is my own work and it is based on the research work that was undertaken by myself, Alfredo Smart, in the Gene Regulation & RNA Biology group, School of Pharmacy, Faculty of Science, University of Nottingham, UK.

Abstract

Messenger RNA (mRNA) therapeutics hold promise for treating diverse diseases but face delivery challenges, as effective cytosolic entry and subsequent translation are limited by endosomal entrapment. Nanoparticles aid mRNA transport across cell membranes, yet only a fraction of mRNA escapes endosomes to reach the cytoplasm, limiting translational efficacy. Current methods to quantify delivered mRNA localisation lack precision, necessitating more quantitative approaches. This thesis addresses this gap by optimising the APEX2-RT-qPCR proximity biotinylation method to enable quantification of mRNA localisation in both the cytosol and endoplasmic reticulum (ER). Through this approach, the study lays groundwork for examining ER-associated localisation and translation, with preliminary investigation into the impact of signal peptides in directing delivered mRNA to the ER.

Stable HEK293T and A549 cell lines expressing APEX2 targeted to either the ER membrane (APEX2-ERM) or cytosol (APEX2-NES) were generated and APEX2 expression was characterised through RT-qPCR and western blotting. Structured illumination microscopy was used to evaluate the subcellular localisation of both the APEX2 epitope tag and APEX2-mediated labelling, with APEX2-ERM in HEK293T cells demonstrating the most effective localisation to its intended subcellular destination.

A refined streptavidin pulldown technique was also developed, using a biotinylated spike-in RNA to optimise the RNA pulldown and to normalise RT-qPCR data from APEX2 RNA labelling experiments. This optimisation significantly advances the application of APEX2 with RT-qPCR, increasing data reproducibility across experiments.

In vitro transcribed (IVT) mRNAs encoding nanoluciferase with and without signal peptides were utilised to assess the extent to which APEX2-RT-qPCR could detect the localisation of delivered IVT mRNA to either the cytosol or the ER. APEX2-RT-qPCR data was most consistent when investigating ER localisation of signal peptide-containing mRNAs in the APEX2-ERM HEK293T cell line. The study then examined how substituting signal peptides

in different coding regions affected both IVT mRNA localisation, as measured by APEX2-RT-qPCR, and translational output, as measured by functional assays. Overall, the PMEL SP led to greater mRNA localisation to the ER compared to the IL6 SP. Other methods were compared with APEX2-RT-qPCR to investigate the localisation of both endogenous and signal peptide-containing delivered mRNA across the cytosol and ER.

These findings highlight the utility of APEX2-RT-qPCR as a powerful tool for quantifying subcellular localisation of delivered mRNA to the ER, providing insights into how SPCS engineering may influence mRNA subcellular trafficking. This method, when combined with reporter assays, offers valuable potential for optimising mRNA therapeutic design by enabling quantitative assessments of localisation and translation output.

Acknowledgements

First and foremost, I owe a tremendous thank you to my supervisor, Catherine, for her invaluable guidance throughout my PhD and for preparing me for my career beyond it, as well as for the numerous opportunities she provided me along this journey. I am also grateful for Snow and Naoto's supervision over the past 4 years.

I also owe a big thank you to Angie and Sasha, the post docs in the GRRB group for their endless knowledge, curiosity and talent. I also thank Athena for teaching wet-lab techniques and general organisation at the beginning of this journey.

I would like to thank the wider GRRB community and lab members for keeping me smiling throughout my PhD. I'd like to acknowledge Ashkan for being a good friend for the past 4 years, Merryn and Rohum who brightened my days towards the end of my project and Hilary Collins for being Hilary Collins.

A big thank you is owed to my partner, Lucy, for putting up with me through the many crises. Your unwavering support and sacrifices over the past 4 years have meant the world to me. Equally I owe a big thank you to my family and my dog, Wallace, for their love and support. I cannot thank you all enough for helping me through this. I am particularly grateful to my father for acting as my confidant and mentor.

A big thank you to Federico and Alex for hosting the Jopling lab with tissue culture facilities at very end of my PhD. I am deeply grateful to Robert Markus, Seema Bagia, and Sal Jones for their assistance with imaging; their expertise and willingness to help have been greatly appreciated.

Finally, I want to express my gratitude to the green parts of the campus, which have provided a serene backdrop for just under eight years of my life.

Contents

List of Tables

List of Figures

List of Abbreviations

A site – Acceptor Site
ac^{4}C – *N4*-acetylcytidine
AES – Amino-Terminal Enhancer of Split
APC – Antigen Presenting Cell
ASO – Antisense Oligonucleotide
ARCA – Anti Reverse Cap Analogue
ATP – Adenosine Triphosphate
bp – Base Pair
BSA – Bovine Serum Albumin
Btn – Biotin
CAR – Chimeric Antigen Receptor
Ccr4 – Catabolite Repressor 4
CBC – Cap Binding Complex
CRISPR – Clustered Regularly Interspaced Short Palindromic Repeats
CDS – Coding Sequence
cDNA – complimentary DNA
circRNA – circular RNA
CPSF – Cleavage/Polyadenylation Specificity Factor
Cre – Cyclic Recombinase
DAPI – Diamidino-2-phenylindole
DNA – deoxyribonucleic acid
DBCO – Dibenzylcyclooctyne
DMEM – Dulbecco's Modified Eagle Medium
DMSO – Dimethyl Sulfoxide
dNTP – Deoxynucleotide Triphosphate
DOPE – dioleoylphosphatidylethanolamine
DTT – dithiothreitol
dsRNA – double-stranded RNA
E site – Exit Site
ECL – Electrochemiluminescence
eIF – eukaryotic Initiation Factor
EJC – Exon Junction Complex
ELISA – Enzyme-Linked Immunosorbent Assay
ER – Endoplasmic Reticulum
GMP – Good Manufacturing Practice
GTP – Guanosine Triphosphate
HBB – Human Haemoglobin β subunit
HCV – Hepatitis C Virus
HRP – Horseradish Peroxidase
IRES – Internal Ribosome Entry Sites
IVT – *In Vitro* Transcription
IVT mRNA – *In Vitro* Transcribed mRNA
LC-MS/MS – Liquid chromatography-tandem mass spectrometry

LB – Luria-Bertani
LNP – Lipid Nanoparticle
LV - Lentiviral
m^1A – *N1*-methyladenosine
$m^1\Psi$ – *N1*-methylpseudouridine
m^5C – 5-methylcytosine
m^5U – 5-methyluridine
m^6A – *N6*-methyladenosine
m^7G – *N7*-methylguanosine
MCP – MS2 Coat Protein
MHC – Major Histocompatibility Complex
miRNA – microRNA
mRNA – messenger RNA
mRNP – messenger Ribonucleoprotein
MS – Mass Spectrometry
NEB – New England Biolabs
NES – Nuclear Export Signal
Not – Negative Regulator Of Transcription
NP – Nanoparticle
NPC – Nuclear Pore Complex
nrRNA – Non-Replicating RNA
nsP – non-structural Proteins
nt – Nucleotide
P site – Peptidyl Site
PABPN1 – Nuclear Poly(A) Binding Protein 1
PAP – Poly(A) Polymerase
PAS – Polyadenylation Signal
PBS – Phosphate Buffered Saline
PCR – Polymerase Chain Reaction
pDNA – plasmid DNA
PDL – Poly-D-lysine
PEG – Poly(ethylene glycol)
PEOZ – Poly(2-ethyl-2-oxazoline)
PI – Protease Inhibitor
UV – Ultraviolet
RT – Room Temperature
RT-qPCR – Reverse Transcription-Quantitative Polymerase Chain Reaction
RIG-I – Retinoic-acid Inducible Gene 1
RIP – RNA Immunoprecipitation
RISC – RNA-Induced Silencing Complex
RNA – Ribonucleic Acid
RNAse – Ribonuclease
rRNA – Ribosomal RNA
S^6G – *S6*-thioguanosine
SA – Signal Anchor

saRNA – Self-Amplifying RNA
SDS – Sodium Dodecyl Sulphate
SDS-PAGE – Sodium Dodecyl Sulphate-Polyacrylamide Gel Electrophoresis
sgRNA – single guide RNA
SIM – Structured Illumination Microscopy
SPCS – Signal Peptide Coding Sequence
SRP – Signal Recognition Particle
ssRNA – single-stranded RNA
taRNA – Trans Amplifying RNA
TBE – Tris/Borate/EDTA
TLR – Toll-Like Receptor
TR – Transreplicon
TREX – Transcription Export
tRNA – transfer RNA
WT – Wildtype
Ψ – Pseudouridine

Chapter 1 - Introduction

1.1 Endogenous mRNA lifecycle

Ribonucleic acid (RNA) is a polymer composed of nucleotides, each featuring a ribose sugar ring with hydroxyl groups at both the 2’ and 3’ positions, a phosphate group attached to the 5’ carbon atom, and a nucleobase linked to the 1’ carbon atom (Figure 1.1). Within the polymer, the 3’ hydroxyl group of one nucleotide forms a phosphodiester bond with the 5' phosphate group of the next nucleotide, creating the RNA strand’s backbone. This structural composition enables RNA to perform a diverse array of functions within cells. The nucleobase allows RNA to carry genetic information, while the whole structure permits the formation of RNA structure-directed functions such as those observed in ribozymes (Wachter, 2014).

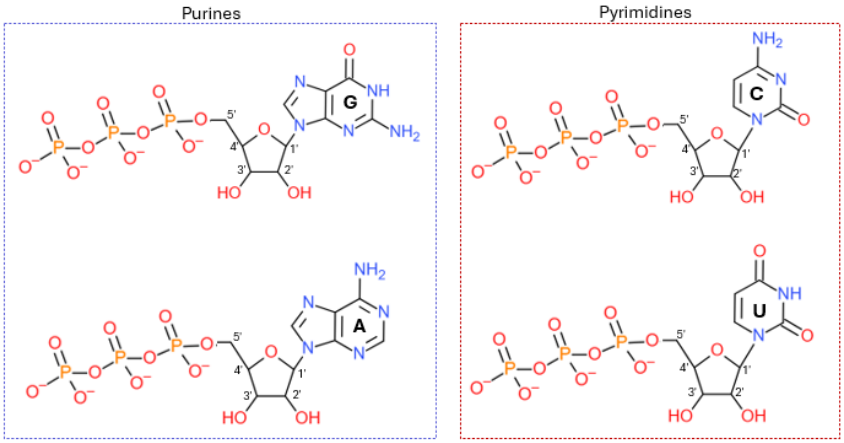


Figure 1.1. The structure of RNA nucleotides.
Each nucleotide consists of a nitrogenous base attached to a ribose sugar, which is further linked to a triphosphate group. The ribose carbons are numerically labelled (1’ to 5’), and the nucleobases are arranged in two lanes: purines (adenine [A] and guanine [G]) on the left and pyrimidines (cytosine [C] and uracil [U]) on the right. Image made using Chemdoodle.com.

Among the different types of RNA, messenger RNA (mRNA) plays a crucial role in gene expression (Alberts et al., 2015). It acts as a temporary carrier of genetic information, transmitting the deoxyribonucleic acid (DNA) code from the nucleus to the ribosomes. At the ribosomes, the sequence of mRNA is read in sets of three nucleotides, known as codons, which determine the sequence of

amino acids in the resulting protein. This process ensures that the genetic instructions encoded in DNA are effectively converted into functional proteins, which are vital for cellular processes and organism development.

Eukaryotic mRNA is produced through transcription and co-transcriptional processing, resulting in a molecule characterised by a 5' cap, a poly(A) tail, and untranslated regions (UTRs) flanking the coding sequence (CDS).

1.1.1 Eukaryotic mRNA biogenesis

Native mRNA within cells undergoes a complex and tightly regulated journey from transcription in the nucleus to translation in the cytoplasm, as reviewed by Buccitelli & Selbach (2020). This process involves numerous interactions with RNA-binding proteins and RNA nucleotide modifications, which collectively contribute to the mRNA's stability, localisation, and translation efficiency. A particular hallmark of the natural pathway is that the processes of mRNA maturation deposit a combination of protein-based markers and RNA modifications that record the transcript's genesis pathway, as reviewed by von der Haar et al. (2023).

1.1.1.1. Transcription

The process of copying a DNA sequence into RNA is known as transcription. DNA is typically stored within highly organised structures called chromatin, where it is wrapped around histone proteins to form nucleosomes, the fundamental units of chromatin (Kornberg, 1974). In mammalian cells, transcription initiation requires chromatin accessibility through histone acetylation, which loosens the chromatin fibre, and through chromatin remodelling complexes that disrupt DNA-histone contacts (Kornberg & Lorch, 1999). Once the chromatin is accessible, RNA Polymerase II (RNA Pol II) binds indirectly to the promoter region of a protein-coding gene through transcription initiation factors (Roeder, 1996). In mammalian cells, many transcription initiation factors are specific to a cell type or developmental stage to ensure tight regulation of transcription according to the needs of cell types or stages of development (Rosenfeld et al., 2006). In contrast, yeast cells exhibit less

complexity in transcription initiation factors, reflecting differences in regulatory needs between unicellular and multicellular organisms. RNA Pol II then moves along the template strand of DNA away from the transcription start site, reading the template DNA strand in a 3' to 5' direction and synthesising a complementary pre-mRNA transcript in the 5' to 3' direction, as illustrated in Figure 1.2A. Transcription is terminated soon after RNA Pol II transcribes a polyadenylation signal (PAS) (Mandel et al., 2008). The PAS is recognised by cleavage/polyadenylation specificity factor (CPSF) and cleavage stimulation factor protein (CstF) complexes, which then recruit additional effector proteins to cleave the pre-mRNA transcript 15–30 nucleotides downstream of the PAS, releasing RNA Pol II from the DNA template (Mandel et al., 2008).

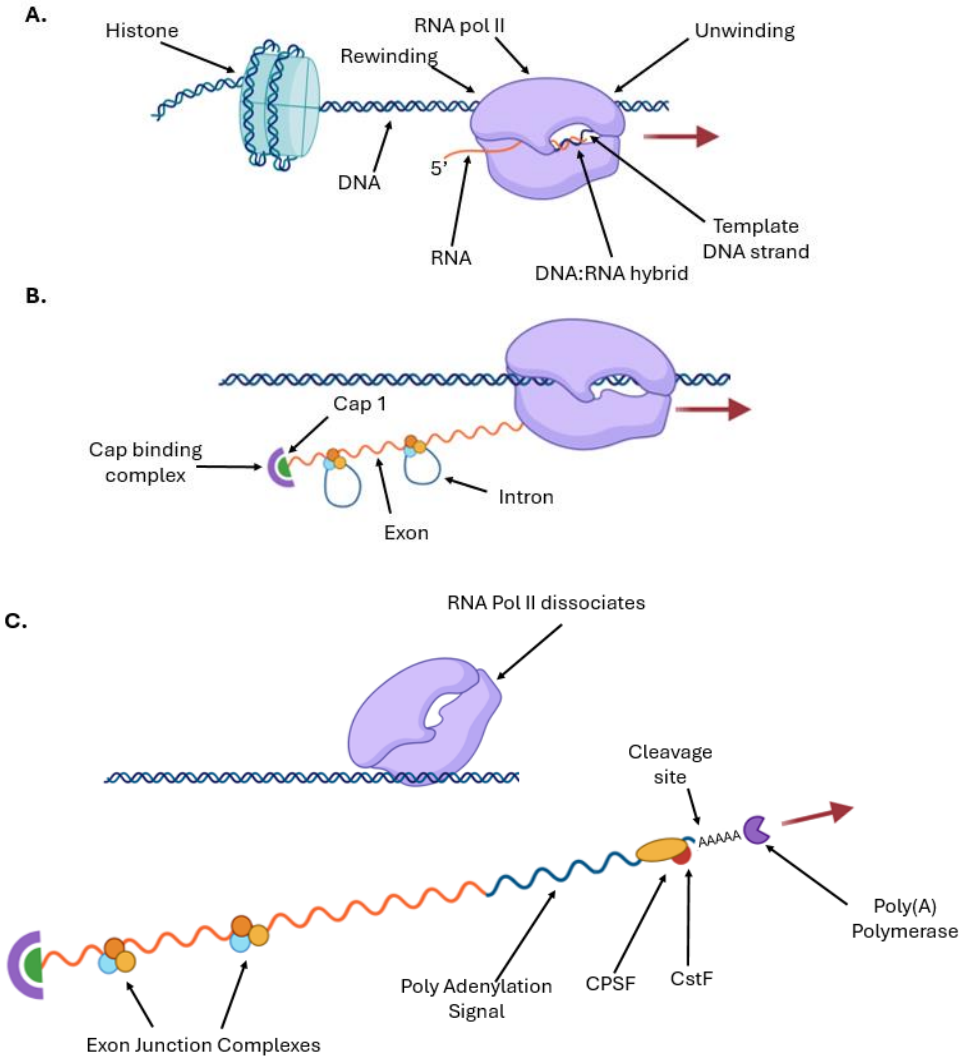


Figure 1.2. Overview of endogenous mRNA synthesis and co-transcriptional processing.
(A) Transcription process from DNA to RNA. (B) Co-transcriptional processing, including the formation of the cap-binding complex and splicing. (C) Poly(A) tail addition following endonucleolytic cleavage of nascent RNA from the DNA template. Image made using Biorender.com.

1.1.1.2. Co-transcriptional events

Although mRNA biogenesis and processing can be reconstituted as distinct reactions *in vitro*, these processes are highly interdependent in cells, with complex interactions that enable precise regulation of gene expression (Bentley, 2014).

In the early stages of transcription by RNA Pol II, an inverted guanosine triphosphate cap is added to the 5' end of the mRNA transcript (Furuichi, 2015;

A. Shatkin, 1976). In mammalian cells, the inverted guanosine cap is then methylated at the *N-7* position by RNA guanine-7 methyltransferase, with additional methylation at the *O*-2 position of the first and second nucleotides catalysed by Cap Methyltransferase 1 and 2, respectively (Galloway & Cowling, 2019). Yeast cells, however, lack *O*-2 methylation, reflecting differences in cap structure across species. The resulting m^7G cap is subsequently recognised and bound by the cap-binding complex (CBC) (Gonatopoulos-Pournatzis & Cowling, 2014). The cap structure is crucial for cap-dependent translation, contributes to mRNA stability by protecting the mRNA molecule from 5'-3' exonucleases and ensures the mRNA is not recognised by the anti-viral innate immune sensor retinoic-acid inducible gene 1 (RIG-I) (as reviewed by Galloway & Cowling, 2019).

Concurrently, nascent mRNA undergoes splicing during transcription to remove introns present in the pre-mRNA (Figure 1.2.2B). This splicing process is carried out by the spliceosome through a complex series of steps, a process well-characterised in mammalian cells. The CBC also aids in recruiting spliceosome-associated proteins to promote splicing of the first intron at the 5' end of the pre-mRNA (Izaurralde et al., 1994; Konarska et al., 1984). In mammalian cells, exon junction protein complexes (EJCs) are deposited at a conserved position 24 nucleotides upstream of each exon boundary of the spliced mRNA (Le Hir et al., 2000).

Following the transcription of the PAS and subsequent endonucleolytic cleavage from the DNA template (Figure 1.2.2C), a stretch of adenosine nucleotides are added by poly(A) polymerase (PAP) (Colgan & Manley, 1997). After 11-14 nucleotides have been added by PAP, the first nuclear poly(A) binding protein (PABPN1) binds to the elongating poly(A) tail, followed by the binding of additional PABPs approximately every 30 nucleotides (Baer & Kornberg, 1983; Kühn et al., 2003; Smith et al., 1997). The poly(A) tail and its associated binding proteins serve multiple functions, including enhancing mRNA stability, promoting cap-dependent translation initiation, and indirectly

facilitating the export of mRNA from the nucleus (Bernstein et al., 1989; Gallie, 1991; Hammell et al., 2002).

1.1.2 Endogenous mRNA modifications

Eukaryotic mRNA undergoes a range of internal chemical base-modifications, which can influence translation efficiency and stability. Among the more extensively studied modifications are N^6-methyladenosine (m^6A), N^1-methyladenosine (m^1A), 5-methylcytosine (m^5C), and N^4-acetylcytidine (ac^4C) (Boo & Kim, 2020). However, the regulatory role of many of these base-modifications remains poorly understood, and their effects on gene expression may vary depending on the developmental stage or cell type.

Among these base-modifications, m^6A is the most thoroughly characterised and plays a particularly dynamic role throughout mRNA biogenesis, as reviewed by Huang et al. (2020). Importantly, m^6A modifications in mRNA are known to enhance the stability of mRNA, as reviewed by Y. Lee et al. (2020). In mammalian cells, the METTL3-METTL14 complex introduces m^6A modifications. Notably, like many methyl modifications, m^6A is reversible, as it can be removed by RNA demethylase enzymes such as fat mass and obesity-associated (FTO) protein (Mauer et al., 2017). Its biological effects are mediated through recognition by specific RNA-binding proteins, and this reversibility adds an extra layer of gene regulation. Evidence suggests that m^6A can be added co-transcriptionally, with early m^6A deposition influencing splicing kinetics (Louloupi et al., 2018).

1.1.3 Nuclear export and trafficking

One of the key functions of the nuclear transport machinery is to regulate the export of mRNA to the cytosol, ensuring that transcripts are retained within the nucleus until they have been fully processed (as reviewed by Katahira (2015)). Central to this process is the nuclear pore complex (NPC), which controls the movement of macromolecules between the nucleus and cytosol.

The Transcription-Export (TREX) complex is essential for coupling transcription and mRNA export, as reviewed by Katahira (2012). TREX complex proteins bind

to mRNA both during and post-transcription and facilitate the formation of export-competent messenger ribonucleoproteins (mRNPs). Components of the TREX complex then facilitate the translocation of the mRNP through the NPC channel. Upon entry into the cytoplasm, the mRNP partially disassembles, removing certain TREX proteins to ensure the unidirectional export of the mRNPs (Xie & Ren, 2019).

1.1.4 Cytoplasmic translation of eukaryotic mRNA

Translation is the process by which the ribosome decodes the genetic material of mRNA to synthesise proteins, as reviewed by Jia et al. (2024). Ribosomes, consisting of a small 40S subunit and a large 60S subunit comprised of ribosomal RNA (rRNA) and proteins, facilitate translation by reading sets of three nucleotides, known as codons, in the 5' to 3' direction (X. Zhou et al., 2015). Each codon corresponds to a specific amino acid or a stop signal. Translation begins at a start codon, typically AUG, which signals the initiation of protein synthesis, and ends when the ribosome encounters a stop codon (e.g. UAA, UAG, or UGA). The coding sequence (CDS), located between the start and stop codons, determines the amino acid composition of the resulting protein, while the regions of RNA flanking the CDS are known as the untranslated regions (UTRs).

The 5' UTR, located upstream of the CDS, plays a key role in regulating translation initiation. A critical feature spanning the 5' UTR and into the CDS is the Kozak sequence, a conserved nucleotide motif that surrounds the start codon (Kozak, 1986). This sequence enhances recognition of the AUG start site by the ribosome, improving translation initiation efficiency.

The process of cap-dependent translation initiation, as reviewed by Brito Querido et al. (2024), begins with the assembly of the eukaryotic initiation factor (eIF) eIF4 complex at the m^7G cap located at the 5' end of the mRNA. This complex includes eIF4E, which binds directly to the 5' cap; eIF4A, an adenosine triphosphate (ATP)-dependent RNA helicase that unwinds the 5' UTR during scanning for the start codon, often with assistance from eIF4B; and eIF4G, a

large scaffold protein that links eIF4E and eIF4A while also interacting with poly(A)-binding proteins associated with the poly(A) tail. Through this eIF4G-mediated interaction, the mRNA forms a closed-loop structure (Figure 1.33.), bringing the 5' and 3' ends into proximity and stabilising eIF4Es binding to the m^7G cap (A. J. Shatkin & Manley, 2000; Wells et al., 1998).

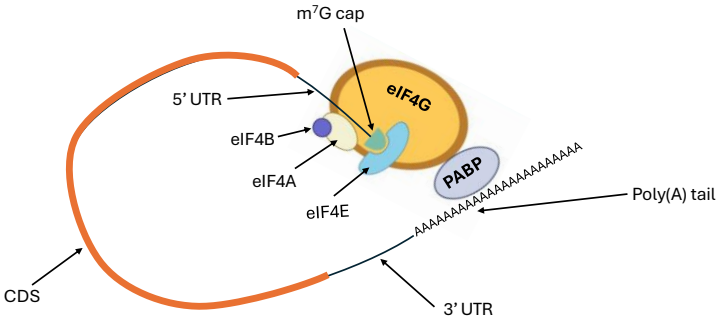


Figure 1.3. Schematic of simplified closed-loop structure of mRNA prior to translation initiation. Image made using Biorender.com.

After the closed-loop structure forms, the 43S pre-initiation complex, comprising the 40S ribosomal subunit with methionyl-transfer RNA (tRNA) positioned in the peptidyl site (P site) and initiation factors eIF1A, eIF2, eIF3, and eIF5, is recruited to the mRNA (as reviewed by Brito Querido et al. (2024)). This recruitment is facilitated by the interaction between eIF3 on the small ribosomal subunit and eIF4G that is indirectly associated with the m^7G cap. The resulting 48S complex then scans the mRNA from the 5' end until the methionyl-tRNA, guided by Kozak's consensus sequence, recognises the AUG start codon. Once the start codon is identified, the guanosine triphosphate (GTP) bound to eIF2 is hydrolysed, triggering the dissociation of most initiation factors. Next, eIF5B is recruited to the small ribosomal subunit, where it hydrolyses GTP, causing eIF1A to dissociate from the A-site. This enables the 60S ribosomal subunit to join, forming the functional 80S ribosome, thereby initiating peptide elongation.

In contrast, cap-independent translation initiation can occur *via* Internal Ribosome Entry Sites (IRES), which enable ribosomes to bind directly to internal regions of the mRNA, bypassing the 5' cap structure (as reviewed by Y. Yang &

Wang (2019)). This provides an alternative route for protein synthesis when cap-dependent mechanisms are compromised.

During translation elongation, as reviewed by Blanchet & Ranjan (2022), aminoacylated tRNAs are delivered to the acceptor site (A site) by elongation factor eEF1A. Within the ribosome, peptide bond formation occurs at the A site, where the growing polypeptide chain is transferred from the tRNA in the P site to the amino acid carried by the tRNA in the A site. As the ribosome translocates along the mRNA, the tRNA in the A site, now attached to the nascent polypeptide chain, shifts to the P site. This cycle continues with the deacylated tRNAs exiting through the exit site (E site) and charged tRNAs entering the A site for the next round of elongation (Bi et al., 2000).

When the ribosome translocates along the mRNA and encounters a stop codon at the A site, release factors eRF1 and eRF3 bind to the ribosome, leading to GTP hydrolysis and the release of the polypeptide chain. The ribosomal subunits are then recycled for new rounds of translation, with ABCE1 facilitating their dissociation by using ATP hydrolysis to separate the 40S and 60S subunits (Borman et al., 2000; Stewart, 2007).

The initial round of translation is marked by the presence of exon junction complexes (EJCs) on the mRNA (Ishigaki et al., 2001). Following the first round of translation, it is thought that EJCs are removed (Ishigaki et al., 2001). It is also speculated that the transition from PABPN to PABPC may occur during the initial round of translation and may be triggered by the ribosome dislodging PABPN (Sato & Maquat, 2009).

Translation is highly responsive to cellular conditions, particularly during stress. One key regulatory mechanism involves the phosphorylation of the α-subunit of eukaryotic initiation factor 2 (eIF2α), which converts eIF2-GDP into a competitive inhibitor of its guanine nucleotide exchange factor, eIF2B. This phosphorylation reduces the assembly of the translation initiation complex and limits global protein synthesis, as reviewed by Boye & Grallert (2020).

The activity of eukaryotic initiation factor 4E (eIF4E) is tightly regulated by its interaction with 4E-Binding Proteins (4E-BPs), which sequester eIF4E and prevent the formation of the eIF4F complex, thereby inhibiting ribosome recruitment to mRNA, as reviewed by Maracci et al. (2022) and M. Yang et al. (2022). In response to stimuli such as nutrient availability and stress, signalling pathways such as the mammalian target of rapamycin (mTOR) phosphorylate 4E-BPs, causing their dissociation from eIF4E and enabling translation initiation.

1.1.5 Eukaryotic translation at the endoplasmic reticulum

mRNAs encoding proteins destined for organelles within the endomembrane system (such as the endoplasmic reticulum (ER), Golgi apparatus, and lysosomes) or for secretion are specifically targeted to the ER for translation. There is also a growing consensus that certain cytosolic proteins are translated at the ER, as reviewed by Reid & Nicchitta (2015).

The mRNA targeting sequence is embedded within the nascent peptide chain, typically as an N-terminal signal peptide or, in the case of transmembrane proteins, as the first transmembrane domain which is known as the signal anchor (Hegde & Keenan, 2022). These sequences within the growing peptide are recognised by the signal recognition particle (SRP) (Figure 1.44A) (Akopian et al., 2013). This interaction temporarily halts translation, and the SRP-ribosome-nascent peptide complex is directed to the ER membrane (Figure 1.44B) (Keenan et al., 2001). Once at the ER, SRP docks with its receptor, releasing SRP and allowing the ribosome to engage with the Sec61 translocon (Keenan et al., 2001). At this point, translation resumes and the nascent peptide chain is co-translationally fed through the Sec61 channel into the lumen of the ER (Keenan et al., 2001).

The signal peptide is cleaved by a signal peptidase complex after entry into the ER lumen, releasing the fully synthesised protein into the lumen (Liaci et al., 2021). While in the ER, newly synthesised proteins fold into their functional structures, often with the help of molecular chaperones (Tatu & Helenius,

1997). Additionally, many proteins undergo co- and post-translational modifications within the ER, such as N-glycosylation (Barlowe & Miller, 2013).

Proteins that lack specific targeting motifs for retention within the ER, Golgi, or lysosomes and do not possess transmembrane domains are transported from the Golgi to the plasma membrane, where they are secreted from the cell.l (Mellman & Nelson, 2008). This entire pathway ensures that proteins are accurately processed, modified, and directed to their appropriate locations for cellular function (Barlowe & Miller, 2013).

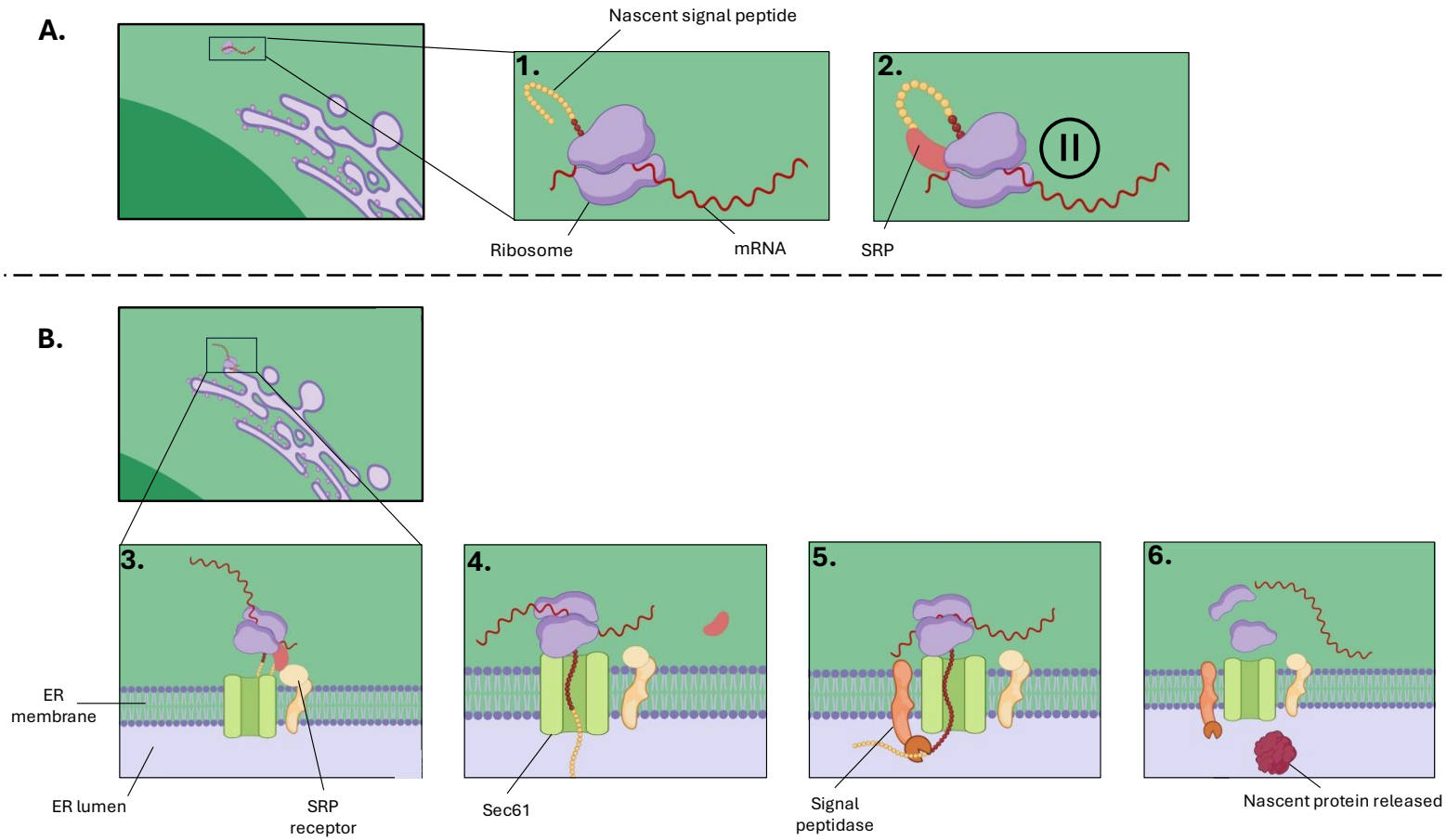


Figure 1.4. Overview of SRP-dependent co-translational trafficking of mRNA to the endoplasmic reticulum. A. Cytosol: (1.) Translation of an mRNA encoding a secreted protein with an N-terminal signal peptide begins on a cytosolic ribosome. (2.) The nascent signal peptide is recognised by the Signal Recognition Particle (SRP), which binds to the ribosome, temporarily halting translation. B. Endoplasmic Reticulum: (3.) The SRP-ribosome-nascent peptide complex is targeted to the ER membrane, where SRP binds to its receptor on the membrane surface. (4.) The ribosome docks with the Sec61 translocon channel, and SRP dissociates, allowing translation to resume. (5.) The growing polypeptide is translocated through the Sec61 channel into the ER lumen and signal peptidase cleaves the signal peptide within the ER lumen. (6.) Translation is terminated, and the nascent protein is released from the ribosome into the ER lumen. Image created using Biorender.com.

1.1.6 mRNA turnover

The turnover of mRNA is essential for regulating gene expression to maintain cellular homeostasis, as well as ensuring mRNA quality control by preventing the production of potentially toxic proteins, as reviewed by Schoenberg & Maquat (2012).

As reviewed by Chen & Shyu (2011) and Passmore & Coller (2022), canonical mRNA decay can initiate during translation, with de-adenylation often occurring while the mRNA is in a closed-loop structure. However, canonical mRNA decay can also occur to mRNAs not undergoing translation, and the extent to which bulk mRNA degradation is co-translational remains unclear.

In mammalian cells, the poly(A)-specific ribonuclease 2 (PAN2)-PAN3 complex trims the distal portion of the poly(A) tail (Figure 1.55A). The catabolite repressor 4 (Ccr4)- negative regulator of transcription (Not) complex then continues this process, removing the poly(A) tail more proximal to the 3' UTR (Figure 1.5B). Once the Ccr4-Not complex reduces the poly(A) tail to a short oligoadenylated form, typically around 10 nucleotides in length, the final PABP is released, and the mRNA's looped structure disassembles.

The oligoadenylated mRNA can then undergo decapping, during which the DCP1-DCP2 complex removes the 5' m^7G cap structure (Figure 1.5C). This step destabilises the mRNA and makes it susceptible to rapid degradation. The mRNA can then be degraded in both the 5'-3' direction by the exoribonuclease 1 (XRN1) and in the 3'-5' direction by the cytoplasmic exosome (Figure 1.5D).

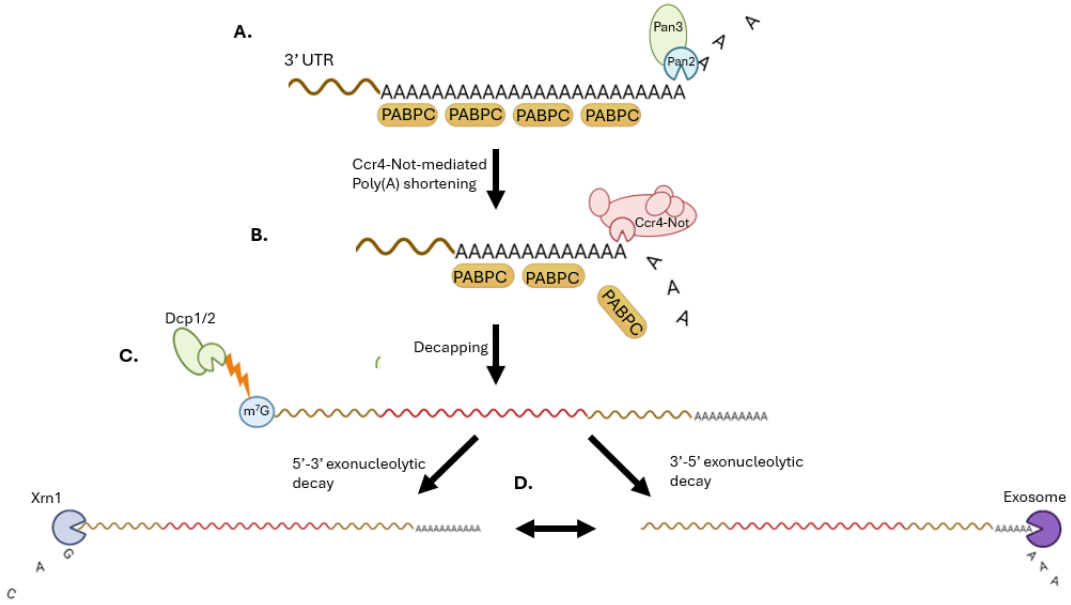


Figure 1.5. Overview of canonical de-adenylation dependent mRNA decay. (A.) Shows Pan2-Pan3 mediated Poly(A) shortening of the distal poly(A) tail. (B.) Shows Ccr4-Not-mediated Poly(A) shortening of the proximal poly(A) tail and removal of cytosolic Poly(A) binding proteins (PABPC). (C.) Shows m^7G decapping by Dcp1-Dcp2 complex. (D.) Shows 5'-3' and 3'-5' exonucleotlytic decay by Xrn1 and the cytosolic exonuclease respectively. Image made using Biorender.com.

The 3' UTR plays a crucial role in determining mRNA stability, localisation, and translational efficiency (Mayr, 2019). Regulatory elements within this region bind to RNA-binding proteins, which act as mediators by recruiting effector proteins that can either stabilise or destabilise the mRNA or regulate its translation (Mayr, 2019). Additionally, the 3' UTR contains microRNA (miRNA) target sites, where miRNAs guide the RNA-Induced Silencing Complex (RISC) to the mRNA, leading to translation repression or mRNA degradation (Bartel, 2004).

Several other mechanisms regulate mRNA decay to maintain cellular function. For example, pathways such as Nonsense-Mediated Decay and Ribosome Quality Control are essential for removing faulty or aberrant mRNAs, ensuring the quality and fidelity of gene expression (as reviewed by Joazeiro (2019) and Nickless et al. (2017)).

1.2 Overview of RNA therapeutics

RNA therapeutics encompass a diverse range of strategies for modulating gene expression or correcting genetic defects, including the use of short interfering RNAs (siRNAs), antisense oligonucleotides (ASOs), and clustered regularly interspaced short palindromic repeats (CRISPR)-Cas9 gene editing, each with distinct mechanisms and therapeutic applications.

siRNAs are 21-nucleotide double-stranded RNA (dsRNA) molecules, where one strand is incorporated into the RNA-Induced Silencing Complex (RISC), as reviewed by B. Hu et al. (2020). The siRNA then guides RISC to complementary mRNA sequences, leading to mRNA cleavage and subsequent degradation. Unlike miRNAs, which can bind to multiple mRNA targets with varying degrees of affinity and regulate gene expression through various mechanisms, therapeutic siRNAs are designed to target unique sequences of 7-8 nucleotides that are as selective as possible for the mRNA encoding the target protein.

Antisense oligonucleotides (ASOs) are short, single-stranded nucleic acid molecules, typically 15-30 nucleotides in length, composed of both ribonucleotides and deoxyribonucleotides that are designed to bind to mRNA (as reviewed by Roberts et al. (2020)). Depending on the types of chemical modifications present in their structure, ASOs can modulate gene expression by binding to either pre-mRNA or mature mRNA. When ASOs bind to pre-mRNA, they can influence splicing patterns through a process known as splice-switching, effectively altering the composition of the resulting mRNA. In contrast, when targeting mature mRNA, ASOs promote degradation *via* ribonuclease (RNAse) H-mediated mechanisms, leading to reduced protein expression. This versatility provides substantial therapeutic potential for genetic disorders and conditions requiring precise control of protein levels.

CRISPR-Cas9 gene editing relies on a single guide RNA (sgRNA) to direct the Cas9 enzyme to a specific DNA sequence, where it induces precise genome modifications, as reviewed by Popovitz et al. (2023). This technology has shown promise in treating genetic diseases, such as β-thalassemia, by directly

correcting the underlying DNA mutations. Cas9 can be delivered either as an mRNA encoding the protein alongside delivery of the sgRNA or as a ribonucleoprotein (RNP) complex for immediate action.

While siRNA, ASO, and CRISPR modalities provide valuable tools for reducing or editing specific genes, mRNA can be used therapeutically to directly increase protein of interest levels, as reviewed by Deyhimfar et al. (2024) and Sahin et al. (2014). Unlike DNA-based therapies that share the same therapeutic aim, mRNA does not pose a risk of integrating into the genome, which could lead to harmful mutations. Additionally, mRNA harnesses the cell's own post-translational machinery, allowing for complex protein modifications that are difficult to replicate in recombinant protein-based therapies. However, despite these benefits, progress in advancing mRNA technology beyond vaccines to replacement therapies remains limited (Dolgin, 2021).

1.2.1 Development of mRNA into a therapeutic modality

The development of mRNA therapeutics began with early experiments that highlighted the potential of exogenous mRNA to induce protein expression in recipient cells (Dimitriadis, 1978; Ostro et al., 1978).

A key breakthrough came when Wolff et al. (1990) demonstrated that direct injection of naked *in vitro* transcribed (IVT) mRNA, as well as plasmid DNA, into the skeletal muscle of mice led to the expression of the encoded protein. This finding challenged previous concerns about mRNA stability *in vivo*, showing that protein expression could be achieved without delivery vehicles. However, the required dose was substantial (100 µg), exceeding the enzymatic manufacturing capabilities of that time.

Further progress was achieved when IVT mRNA was used for protein replacement therapy, successfully treating a metabolic disorder in rats (Jirikowski et al., 1992), and when naked IVT mRNA delivered to mice was shown to induce antigen-specific antibody responses against influenza virus (Martinon et al., 1993).

Despite enthusiasm in the early 1990s, major pharmaceutical companies remained sceptical about mRNA technology and progress halted (Dolgin, 2021). However, following innovations in incorporating modified nucleosides into IVT mRNA to reduce the innate immune response to delivered RNA (Karikó et al., 2008), companies such as CureVac, Moderna, and BioNTech emerged as leaders in mRNA therapeutics, aiming to explore the commercial viability of this technology. mRNA vaccines have since achieved remarkable success in the fight against COVID-19, and hold promise for a range of other therapeutic indications.

1.2.2. General features of mRNA therapeutics

Therapeutic mRNAs share many of the same structural features as endogenous mRNAs, including a 5' untranslated region (UTR), a Kozak sequence for efficient translation initiation, a coding sequence possibly with a signal sequence at the 5’ end, a 3' UTR, and a poly(A) tail. Optimising these components is crucial for enhancing protein output from therapeutic mRNAs, as overall protein yield is primarily determined by translational efficiency and stability of the mRNA molecule (de Sousa Abreu et al., 2009).

1.2.2.1. 5’ Cap design and introduction

IVT mRNAs require a 5' cap structure to mimic endogenous mRNA and ensure both stability and efficient translation (Sahin et al., 2014). The addition of a 5' cap to IVT mRNA can be achieved post-transcription *in vitro* using viral capping enzymes and 2'-O-methyltransferases (Fuchs et al., 2016). However, co-transcriptional capping is generally preferred for maximum cost-efficiency. A challenge in the IVT process is that the m^7GpppG cap analogue used during capping can be incorporated in the reverse orientation, reducing translational efficiency. To address this issue, anti-reverse cap analogues (ARCA) which feature a 3'-O-methylated m^7G structure (Figure 1.6A) can be utilised co-transcriptionally to prevent incorrect incorporation of the cap analogue (Mockey et al., 2006; Stepinski et al., 2001).

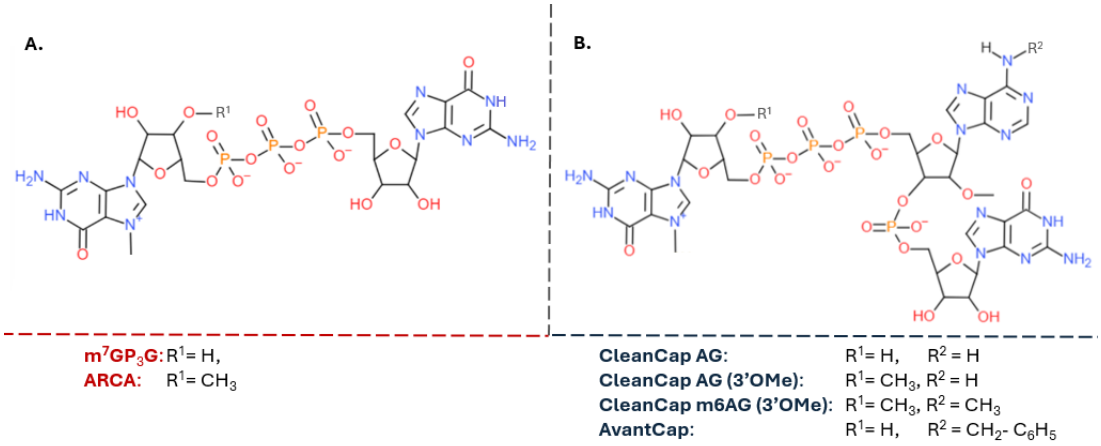


Figure 1.6. Schematic illustration of cap structures. (A.) m^7GP_3G cap and $m_2^{7,3'\text{-}O}GP_3G$ Anti-Reverse Cap Analogue (ARCA). (B.) Cap 1 structures, including CleanCap AG, CleanCap AG 3'OMe, CleanCap m6AG and AvantCap. Image made using chemdoodle.com.

Co-transcriptional capping using ARCA necessitates careful control of GTP concentrations to ensure proper cap incorporation, which can reduce IVT yield and has been reported to lead to truncated mRNAs (Henderson et al., 2021). CleanCap AG ($m^7Gppp^{2'OMe}ApG$) (Figure 1.6B) improves co-transcriptional capping efficiency compared to ARCA (Henderson et al., 2021). The use of CleanCap structures requires mutation of the initiation sequence downstream of the T7 promoter from GG to AG, which is thought to minimise the incorporation of mononucleotides at the 5' end of the mRNA instead of the cap structure (Henderson et al., 2021). Additionally, incorporation of Cap 1 structure in therapeutic mRNA may enhance mRNA stability and increase translation (Henderson et al., 2021). A similar CleanCap analogue, with 3' O-methylation of the m^7G, known as CleanCap AG 3' OMe, was used in the Pfizer-BioNTech COVID-19 mRNA vaccine (Vogel et al., 2021). More recently, TriLink introduced CleanCap m6AG, where the second nucleotide in the mRNA (A) is modified to m^6A to enhance *in vivo* performance, although comprehensive studies have yet to be published. Warminski et al. (2024) further advanced this technology by substituting the *N6*-methyl group in m^6A of CleanCap m6AG with a methyl-benzyl group. This modified cap structure, known as AvantCap, has shown improved protein output in certain cell lines compared to CleanCap AG and

CleanCap m6AG, as well as enhanced protein production *in vivo* relative to CleanCap AG.

1.2.2.2. 5’ UTR Design and optimisation

The 5' UTR plays a major role in regulating translation initiation efficiency, and its structure has a significant impact on ribosome scanning. Therapeutic mRNAs often use 5' UTRs derived from stable endogenous mRNAs, such as those from human α and β-globin mRNAs, which are well-characterised for their stability and simple secondary structures (Kozak, 1994; Reshetnikov et al., 2024). A minimal and non-complex secondary structure in the 5' UTR is thought to be essential for efficient ribosomal binding and scanning, ensuring that translation initiation occurs smoothly (S. Castillo-Hair et al., 2024). Secondary structures in the 5' UTR, such as hairpins or stem-loops, can inhibit ribosome recruitment and scanning for cap-dependent translation (Leppek et al., 2018). A strong Kozak sequence is typically incorporated into the 5’ UTRs of therapeutic mRNAs to enhance the likelihood of successful translation initiation (Xia, 2021).

Significant effort has also been devoted to maximising UTR design to enhance total protein output. Much of the research on optimising 5' UTRs has focused on increasing translation initiation efficiency, with traditional metrics often centred on mean ribosomal load as a major determinant of translation efficiency (S. M. Castillo-Hair & Seelig, 2022; Reshetnikov et al., 2024; Sample et al., 2019).

Leppek et al. (2022) demonstrated that while enhanced translation efficiency is typically desirable, excessive polysome loading can paradoxically destabilise mRNA. Further research by Bicknell et al. (2024) found that delivered mRNAs with moderate translation initiation rates (approximately one ribosome per minute) and ribosome loads (around one ribosome per 170 nucleotides) produced the highest protein output. This evidence underscores the need for a more nuanced approach to 5' UTR optimisation that balances the promotion of effective translation initiation with the preservation of mRNA stability.

1.2.2.3. Coding sequence design and optimisation

The nucleotide sequence chosen for both the CDS and UTRs can significantly influence both the magnitude and duration of protein production. Within the CDS, the choice of codons is a key factor that influences protein production by modulating the speed of translation elongation, as reviewed by Hanson & Coller (2018). Codon usage bias, the preferential use of certain codons over others, plays a central role in this process. This bias is often quantified using the tRNA adaptation index (tAI), which measures the match between codons in the mRNA and the availability of their corresponding tRNAs in the cell (Bahiri-Elitzur & Tuller, 2021). Codons with abundant matching tRNAs are considered "optimal" and contribute to faster translation elongation and improved decoding accuracy, thereby enhancing overall translation efficiency (as reviewed by Y. Liu et al. (2021)). Importantly, the optimality of specific codons can vary across cell types due to differences in tRNA availability, making translation efficiency cell-type-specific and context-dependent.

Although much is understood about codon usage, its role in translation remains nuanced. For example, studies suggest that codon optimality may not be uniform across entire transcripts (Tuller et al., 2010) and replacing non-optimal codons in engineered mRNA sequences carries risk. Some non-optimal codons are evolutionarily conserved to promote translational pausing, which can help ensure proper co-translational protein folding (J.-R. Yang et al., 2014). When non-optimal codons are replaced with optimal ones, translation speed may increase, but this can disrupt the protein folding process, leading to a reduction in overall functional protein yield (Komar et al., 1999; Yu et al., 2015).

The structural properties of the CDS also play a critical role with increasing evidence suggesting that a structured CDS is associated with increased protein expression (Mauger et al., 2019). Codon optimisation is also intricately linked to RNA secondary structure, as optimal codons tend to be GC-rich and thus more structured (Bicknell et al., 2024). This interplay between codon usage and RNA structure complicates efforts to disentangle their respective effects on translation dynamics and stability, though a few sequence optimisation

algorithms have been designed with both in mind (Mauger et al., 2019; H. Zhang et al., 2023).

Signal peptide coding sequences (SPCS) within the CDS are included in therapeutic mRNAs, encoding secreted or membrane-bound proteins. In mRNA vaccines, SPCS play a crucial role in antigen presentation by directing the translated protein to major histocompatibility complex (MHC) class II compartments in antigen-presenting cells (APCs) (Sahin et al., 2014). This facilitates the activation of helper T cells, leading to a stronger and more durable immune response. Moreover, exogenous antigens taken up by dendritic cells can be cross-presented on MHC class I molecules (Bevan, 2006). While research into mRNA encoding secreted proteins continues, specific developments are often difficult to access due to recent patent protections held by Moderna for “modified polynucleotides for the production of secreted proteins” (Dolgin, 2021).

1.2.2.4. 3’ UTR design and optimisation

On the 3' UTR side, therapeutic mRNAs are designed to incorporate stabilising 3’ UTR elements that improve mRNA stability and translation persistence. The most common approach is to utilise α and β-globin 3’ UTRs as they are known to confer high mRNA stability (Sahin et al., 2014). However, more tailored approaches are increasingly being explored to optimise stability for specific target tissues or cell types, given that different cells express varying levels of miRNAs and RNA-binding proteins that affect mRNA stability (S. Castillo-Hair et al., 2024; Sahin et al., 2014).

One innovative approach developed by Orlandini von Niessen et al. (2019) involves screening for the most stable endogenous 3' UTR structures within a given cell type. This process is followed by reporter assays to assess how effectively each 3' UTR enhances the stability of an intentionally destabilised luciferase reporter mRNA, helping to reduce microRNA binding sites and increase overall stability (Orlandini von Niessen et al., 2019; Xia, 2021).

It was recently shown that integrating viral RNA elements into the 3' UTR of IVT mRNA can substantially increase mRNA stability (Seo et al., 2023). Additionally, other innovations, such as using two human β-globin 3' UTRs arranged head-to-tail, have been found to further enhance the stability of reporter mRNAs (Holtkamp et al., 2006).

1.2.2.5. Poly(A) tail introduction

Poly(A) tails of IVT mRNAs are generally designed to consist of around 100-150 adenosines to enhance stability (Trepotec et al., 2019; J. Zhao et al., 1999). These tails can be introduced either at the DNA template stage or post-transcriptionally using template independent recombinant poly(A) polymerases. However, the alkaline conditions required for these polymerases are also conducive to degradation of RNA, meaning treatment times must be limited (Trepotec et al., 2019). Moreover, enzymatic addition of poly(A) tails results in heterogeneity in tail length, posing challenges for good manufacturing practice (GMP) quality assurance (Weissman, 2015).

Adding poly(A) tails directly to DNA templates is the preferred approach but has presented technical difficulties, as long stretches of homo-nucleotide sequences can lead to plasmid instability and recombination (Grier et al., 2016). Once a plasmid is generated and confirmed to be free of recombination issues, it can be amplified by polymerase chain reaction for small scale IVT mRNA production to mitigate future concerns (Perenkov et al., 2023).

One effective solution to minimise the risk of plasmid recombination is the use of spacers within the poly(A) tail, which allows for scalable plasmid production in E. coli without the risk of recombination. Studies have indicated that these spacers do not adversely affect the translation or stability of IVT mRNA (Trepotec et al., 2019), and this technique was adopted in the Pfizer-BioNTech COVID-19 vaccine (Xia, 2021).

The use of linear plasmids, which are better suited for stable cloning of poly(A) tails up to 500 bp, has also been proposed (Grier et al., 2016); however, these vectors are very low-copy and exceed 12 kb in size. More recently, companies

such as 4basebio and TouchLight Genetics have developed synthetic DNA platform technologies that enable rapid enzymatic amplification of double-stranded DNA, as reviewed by Lenk et al. (2024). This innovation enhances the feasibility of scaling up DNA production while effectively managing homo-polymeric regions and may replace the use of poly(A) tail spacers.

Despite these technical hurdles in tail addition, recent research has prompted a re-evaluation of the necessity of long poly(A) tails in IVT mRNA. Poly(A) tails of newly transcribed endogenous mRNA vary greatly, ranging from approximately 30 nucleotides to over 250 nucleotides (as reviewed by Passmore & Coller, 2022). Notably, stable and highly translated mRNA often features poly(A) tails as short as 30 adenosines, which is sufficient for maintaining mRNA stability and translational efficiency (Lima et al., 2017).

1.2.2.6. Common nucleoside modifications in mRNA therapeutics

Among the various strategies to enhance the effectiveness of IVT mRNA therapeutics, the incorporation of specific chemical modifications has emerged as a pivotal approach for managing immunogenicity. Much of the foundational work on the immunogenicity of exogenous RNA was pioneered by Katalin Karikó and Drew Weissman, who were awarded the Nobel Prize in 2023 for their contributions to the development of the COVID-19 mRNA vaccines from Moderna and Pfizer/BioNTech.

Karikó et al. (2005) demonstrated that uridines in unmodified mRNA activate Toll-like receptor 7 (TLR7) in endosomes. Building on this, they then showed that incorporating pseudouridine (Ψ) into IVT mRNA significantly reduces immunogenicity and enhances translation efficiency both *in vitro* and *in vivo* (Karikó et al., 2008). Subsequent studies revealed that Ψ-modified RNA dampens the innate immune response by inhibiting several interferon-stimulated pathways, including protein kinase R (PKR) activation, which limits eIF2α phosphorylation (B. R. Anderson et al., 2010).

The discovery that *N1*-methylpseudouridine ($m^1\Psi$), a naturally occurring modification in 18S rRNA, enhances IVT mRNA translation even more than Ψ was another key advancement. Andries et al. (2015) showed that $m^1\Psi$ reduces TLR3 activation and increases translation efficiency across multiple cell lines. Later work, such as that by Svitkin et al. (2017), confirmed that $m^1\Psi$ reduces PKR activation, thus preventing eIF2α phosphorylation and further enhancing protein synthesis relative to IVT mRNA generated with Ψ. This enhancement is thought to be due to $m^1\Psi$-containing mRNA being less structured, making it less likely to activate dsRNA sensors such as TLR3 and PKR.

$m^1\Psi$ is commonly used in mRNA therapeutics, such as the Pfizer and Moderna COVID-19 mRNA vaccines. CureVac pursued uridine depletion as a strategy to reduce immunogenicity and enhance mRNA stability of their COVID-19 mRNA vaccine, as reviewed by Verbeke et al. (2022). However, this approach proved less effective, resulting in higher reactogenicity, and CureVac have since begun to explore $m^1\Psi$ in their current pre-clinical mRNA vaccines (Bernard et al., 2023). More recent studies have shown that $m^1\Psi$ induces +1 frameshifting in the Pfizer COVID-19 mRNA vaccine, although the extent to which this poses a problem *in vivo* remains unclear (Mulroney et al., 2024).

1.2.3. Alternatives to mRNA for therapeutic protein expression

1.2.3.1. Circular RNA

Circular RNAs (circRNAs) are a unique class of endogenous RNA molecules characterised by their covalently closed-loop structure, generated through back-splicing of pre-mRNA, as reviewed by Liang et al. (2021) and Zhang et al. (2024). This circularity renders them resistant to exonucleases, making them more stable compared to linear RNAs. While endogenous circRNAs are primarily thought to be non-coding and act as miRNA sponges by sequestering miRNAs to modulate gene expression (Hansen et al., 2013), studies suggest that some circRNAs are capable of encoding peptides *via* cap-independent translation mechanisms (C. Chen & Sarnow, 1995; Y. Yang et al., 2017). In this

context, m^{6}A modifications appear to play an important role in regulating translation initiation of endogenous circRNAs (Y. Yang et al., 2017).

Therapeutically, circRNAs have been explored for their ability to act as miRNA sponges, particularly in diseases such as cancer, where they can inhibit miRNAs that are overexpressed in cancerous cells (X. Liu et al., 2022). However, there is growing interest in their potential as alternative mRNA therapeutics capable of encoding proteins, given their stability and potential for prolonged protein expression. To achieve this, therapeutic circRNAs typically require an IRES to facilitate translation, as they lack the 5' cap structure found in linear mRNA (Figure 1.7B).

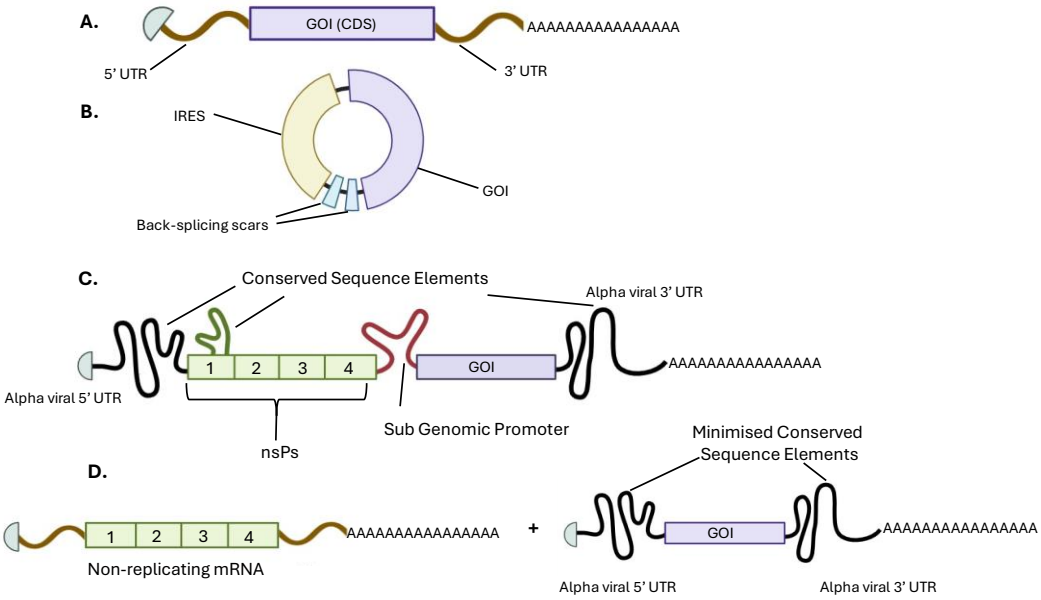


Figure 1.7. Schematic showing protein-encoding RNA modalities. (A) Linear non-replicating mRNA, where GOI stands for gene of interest. (B) CircRNA featuring an internal ribosome entry site (IRES) in yellow, the GOI in purple, and back-splicing junction scars. (C) Self-amplifying RNA, showing conserved sequence elements (CSE) within the 5' and 3' UTRs, and within the CDS of non-structural protein 1 (nsP1) that are required for RNA self-replication and amplification of the GOI. (D) Trans-amplifying RNA, with the non-structural proteins (nsPs) encoded in a separate, linear non-replicating mRNA, and the transreplicon containing the GOI and minimal CSEs for trans-replication. Schematic created using BioRender.com.

1.2.3.2. Self-amplifying and trans-amplifying RNA

Self-amplifying RNA (saRNA) is an innovative therapeutic platform typically derived from alphaviruses, as reviewed by Comes et al. (2023) and Silva-Pilipich et al. (2024). It consists of a large ssRNA molecule (10-12 kb) that encodes both a replicase complex and a transgene for the protein of interest. First, the RNA encoding the replicase complex, consisting of non-structural proteins (nsPs)

such as a capping enzyme (nsP1) a helicase (nsP2) and an RNA-dependent RNA polymerase (nsP4), is translated (Figure 1.8A). The replicase then anchors to membrane-bound structures called spherules and copies the long mRNA into a complementary negative strand RNA (Figure 1.8B). This negative strand RNA serves as a template for the replicase to generate more full-length saRNA, owing to conserved features in its UTRs. Simultaneously, the replicase recognises a subgenomic promoter (SGP) downstream of the replicase in the negative strand, leading to the synthesis of a greater proportion of subgenomic RNA compared to the full-length replicating RNA (Figure 1.8C). By replacing the subgenomic content with an RNA sequence encoding an antigen downstream of the SGP, this process facilitates the amplification of transgenic RNA, allowing for high levels of protein production from lower initial RNA doses.

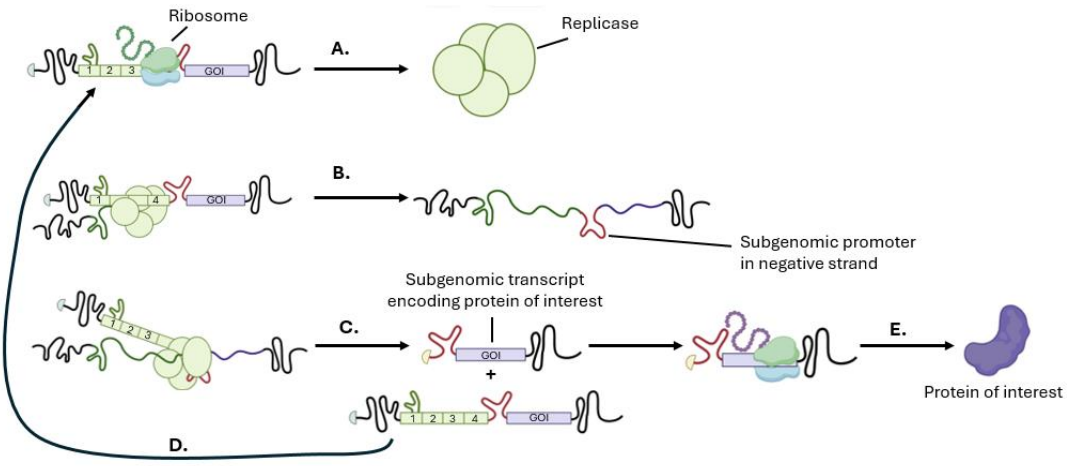


Figure 1.8. Schematic illustrating self-amplifying RNA amplification. (A.) Translation of non-structural proteins (nsP1-4) from the input saRNA leads to the formation of the alphavirus replicase complex. (B.) The alphavirus replicase transcribes the full-length positive strand RNA to generate a complementary negative single-stranded RNA (ssRNA). (C.) The conserved sequence elements (CSEs) and the subgenomic promoter within the negative strand are recognised by the replicase, which synthesises new full-length positive strand saRNA and a large amount of subgenomic RNA, encoding the protein of interest. (D.) The newly transcribed full-length saRNA is translated, continuing the cycle of replication. (E.) The newly transcribed subgenomic RNA is translated, leading to the production of the protein of interest. Image created using BioRender.com.

The effects of nucleoside modifications such as Ψ and $m^1Ψ$ also vary when incorporated into other RNA modalities, such as circRNA and saRNA. For example, m^6A-modified circRNAs exhibit minimal immune activation, while Ψ and $m^1Ψ$ tend to reduce translation efficiency in circRNA contexts (Y. G. Chen et al., 2019). Similarly, Ψ and $m^1Ψ$ impair the function of saRNA by interfering with

replicase activity on the transgene, preventing efficient amplification (Erasmus et al., 2020). In contrast, modifications such as m^5C and m^5U are well tolerated in saRNA therapeutics and help dampen immune responses without impairing translation (Azizi et al., 2024). Once saRNA is replicated, it will lose the initial modifications, meaning these modifications may be more beneficial for reducing TLR-mediated endosomal recognition during the early stages of delivery.

Trans-amplifying RNA (taRNA) represents an innovative approach that involves separating the replicase and transgene into two distinct RNA molecules, as reviewed by Yıldız et al. (2024). In this system, a non-replicating mRNA (nrRNA) encodes the replicase, while a transreplicon (TR) encodes the therapeutic transgene, which is amplified by the replicase through conserved features in its UTRs (Beissert et al., 2020). This design offers an advantage over saRNA in that $m^1\Psi$ can be incorporated into the nrRNA to enhance its stability and reduce immunogenicity without affecting the replication of the transgene. As the nrRNA is used in excess relative to the TR, most of the RNA mass of taRNA may benefit from the $m^1\Psi$ modification, potentially making it less immunogenic than saRNA. However, this approach introduces new challenges, such as ensuring that both RNA components are delivered to the same cell, as the delivery of the TR alone may not provide a therapeutically useful dose.

1.2.4. Current and Emerging Applications of RNA Therapeutics

The only mRNA therapeutics available on the global market are COVID-19 vaccines, with a total of seven approved products, as reviewed by Hu et al. (2024). Most of these vaccines encode the spike protein of SARS-CoV-2, while one specifically targets the receptor-binding domain. All non-replicating mRNA vaccines, with the exception of two developed and approved in China, have reported using m1Ψ to enhance mRNA stability and therapeutic efficacy (J. Lu et al., 2020; Tian et al., 2022; Xia, 2021). One of the seven approved vaccines is an saRNA COVID-19 vaccine, which received approval in Japan in 2023 (Hồ et al., 2024).

Beyond these approved COVID-19 vaccines, the landscape of mRNA-based therapies is rapidly expanding, with over 100 linear mRNA therapies currently in preclinical or phase trials, as reviewed by Parhiz et al. (2024) and Deyhimfar et al. (2024). These investigational therapies extend far beyond vaccination against infectious viral diseases, with promising strategies aimed at cancer treatment, genome editing, and protein replacement therapies.

For example, mRNA cancer vaccines are being developed to target tumour-associated antigens for colorectal cancer and glioblastoma, personalised neoantigens for melanoma, and enhancing chimeric antigen receptor-T cell responses *in vivo*. mRNA is also being explored for enhancing chimeric antigen receptor (CAR)-T cell responses *in vivo*. Beyond oncology, mRNA technologies are being explored for antibody therapies against toxins and for bacterial mRNA vaccines to combat antibiotic resistance. Additionally, Cas9 mRNA therapies are under investigation for genome editing, alongside mRNA-based protein replacement therapies for genetic disorders like methylmalonic acidaemia, cystic fibrosis, and haemophilia, reflecting the broad range of potential therapeutic applications.

Multiple saRNA candidates have entered pre-clinical testing as reviewed by Schmidt & Schnierle (2023), with vaccines developed against influenza virus, respiratory syncytial virus, rabies virus, Zika virus, Ebola virus, Venezuelan equine encephalitis virus, and HIV-1. Conversely, the development of trans-amplifying RNA (taRNA) has progressed more slowly. One taRNA vaccine candidate against influenza, which expresses the hemagglutinin of the virus, was able to induce protective immune responses with less antigenic RNA compared to an saRNA candidate (Beissert et al., 2020).

circRNA development has progressed more slowly than its linear counterparts and have yet to enter phase I trials. This slower advancement is primarily due to limitations in purification methods, which make it challenging to differentiate between nicked circularised RNA and circRNA (Loan Young et al., 2023). However, a study by Qu et al. (2022) compared a circular RNA SARS-CoV-2 receptor-binding domain vaccine with its mRNA counterpart, revealing that the

circRNA vaccine exhibited better shelf-life, higher and more durable antigen expression, and induced a stronger immune response compared to a more conventional mRNA vaccine. This suggests that while circRNA research faces challenges, its potential benefits may pave the way for future advancements in vaccine development.

1.3. Current challenges in mRNA therapeutics

Bringing mRNA therapeutics to market presents a range of challenges that span both scientific and logistical domains. These challenges include intrinsic hurdles related to the stability of RNA, difficulties in achieving targeted and efficient delivery, and the complexities of rapidly scaling up manufacturing processes while ensuring consistent product quality. Many of these factors are interdependent, making the development of mRNA-based therapies a multifaceted challenge.

1.3.1. Managing the instability of mRNA

2'-in-line hydrolysis involves the cleavage of the RNA backbone through a nucleophilic attack by the 2' OH group of ribose on the adjacent phosphodiester bond (Y. Li & Breaker, 1999). This reaction initially forms an unstable intermediate, a 2',3'-cyclic phosphate, along with a cleaved 5'-hydroxyl (Figure 1.9). The 2',3'-cyclic phosphate is susceptible to further nucleophilic attack, ultimately resulting in the formation of a stable 3'-phosphate and 2'-hydroxyl at the cleavage site. In-line cleavage renders mRNA non-functional and prevents the production of full-length protein. Although this degradation can occur spontaneously in solution, due to the presence of H_2O molecules, it is also the mechanism employed by specific RNA endonucleases (Thompson et al., 1994).

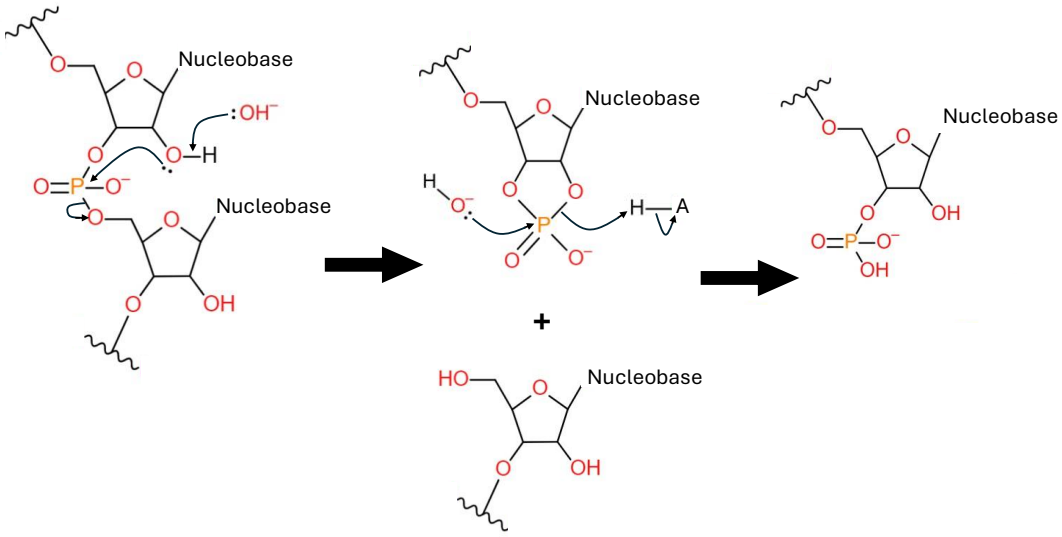


Figure 1.9. Schematic illustrating 2'-in-line hydrolysis of RNA. The reaction begins when a base (represented here as a hydroxide ion) abstracts a proton from the 2'-OH group on the ribose sugar,

enhancing the nucleophilicity of the 2'-oxygen atom. The activated 2'-oxygen atom then attacks the adjacent phosphate atom, resulting in cleavage of a P-O bond and producing a 3'-nucleotide with a 5'-OH group. In the subsequent step, a second hydroxyl ion donates a pair of electrons to the phosphate atom, while the 2'-oxygen atom abstracts a proton from a donor (denoted as A-H), thereby regenerating the 2'-OH group. Notably, either the 2'- or 3'-oxygen atom may act as the electron donor in this second step. Image made using chemdoodle.com.

Several sequence optimisation strategies to manage chemical instability largely overlooked in-line hydrolysis as a key determinant of RNA stability and instead, focus on another metric: the minimum free energy (MFE) (Mauger et al., 2019; Victor et al., 2019; H. Zhang et al., 2023). MFE refers to the predicted energy state of an RNA molecule's most stable secondary structure, assuming that this is also the most probable structure under thermodynamic equilibrium, as reviewed by Lorenz et al. (2016). Multiple algorithms have been developed that aim to find sequences that facilitate RNA secondary structures with the lowest MFE, considering nucleotide number, composition, and arrangement (Mathews et al., 1999; H. Zhang et al., 2023). However, RNA is a dynamic molecule, and it has been suggested that it exists in a range of sub-optimal structures (Zuker, 1989).

More recently, an alternate type of mRNA stability metric was developed with aim of predicting a sequence's susceptibility to in-line hydrolysis (Wayment-Steele et al., 2021). These metrics attempt to estimate the probability that each nucleotide is in an unpaired state, leaving its 3' phosphodiester linkage vulnerable to adopting the in-line attack conformation necessary for cleavage.

A new strategy developed by Leppek et al. (2022) tested several predictive algorithms in wet-lab experiments and found that sequence optimisation based solely on the metrics described above did not correlate with RNA stability in solution or in cells. They used a massively parallel reporter assay called Pooled Evaluation of mRNA in-solution Stability, and In-cell Stability and Translation RNA-seq (PERSIST-seq), to quantify IVT mRNA stability by detecting intact mRNAs across multiple timepoints, using RT-PCR and Illumina sequencing to determine mRNA half-lives in a high throughput manner. Through this, Leppek et al. (2022) identified that asymmetric internal loops ending with 3' uridine were the most susceptible to in-line hydrolysis. Optimising the sequence surrounding the site of 2' in-line hydrolysis proved to be more effective. Guanosine and

cytidine downstream of the vulnerable uridine were the most effective at preventing cleavage, followed by adenosine, whereas uridine directly downstream increased the risk of hydrolysis.

Where sequence optimisation is not possible to reduce the risk of 2’ in-line hydrolysis, nucleotide modifications such as Ψ and $m^1Ψ$ its derivatives can be introduced. Both Ψ and $m^1Ψ$ reduce the effects of sequence dependence on 2'-in-line hydrolysis, likely due to a combination of changed nucleophilicity at the site of substitution and enhanced base-stacking interactions which acts to stabilise RNA (Davis, 1995; Leppek et al., 2022).

Strategies aimed at decreasing deadenylation-dependent decay have also emerged to enhance mRNA stability. One effective method developed by the Wang lab involves the site-specific introduction of exonuclease-resistant modifications through 3' end ligation of chemically synthesised RNA featuring exonuclease-resistant elements at their respective 5' end (Aditham et al., 2022). These modifications, including secondary structures such as G4 quadruplexes and phosphorothioate modifications, significantly improve mRNA stability and thereby increase protein production in cells.

1.3.2. Production of full length IVT mRNA without contaminants

IVT using bacteriophage RNA polymerases, such as T7, T3, or SP6, is the standard method for synthesising mRNA for both research and large-scale GMP-grade biopharmaceutical production. Despite its widespread use, IVT is known to generate dsRNA by-products that can stimulate the immune system and reduce the efficacy of mRNA therapeutics (as reviewed by Lenk et al. (2024)). These dsRNA by-products arise from different mechanisms during the transcription process.

One primary source of dsRNA by-products is RNA hairpin-initiated extension, where the T7 RNA polymerase re-initiates on the transcribed RNA once it reaches the 3’ end of the double-stranded DNA, resulting in the production of dsRNA products, as illustrated in Figure 1.10A (Gholamalipour et al., 2018). The likelihood of this form of dsRNA being generated increases if the restriction

enzyme used to linearise the plasmid DNA leaves a 3' overhang (Schenborn & Mierendorf, 1985). This mechanism is also thought to be partly responsible for single-stranded RNAs (ssRNA) that are 3' end heterogeneous, typically by a few nucleotides, as a consequence of terminated loop-back extension (Gholamalipour et al., 2018).

Another common source of contamination comes from T7 polymerase initiation failing to transition into the stable transcription elongation structure (Dousis et al., 2023). This leads to the production of short ssRNA byproducts, known as abortive transcripts, which are approximately 10 nucleotides in length (Figure 1.10B).

The third reported mechanism involves antisense transcription that initiates from the promoter-less DNA end in a sequence specific manner (Mu et al., 2018; Wu et al., 2020). dsRNA by-products then arise from the negative RNA strand product annealing with the intended ssRNA product (Figure 1.10C) (Mu et al., 2018).

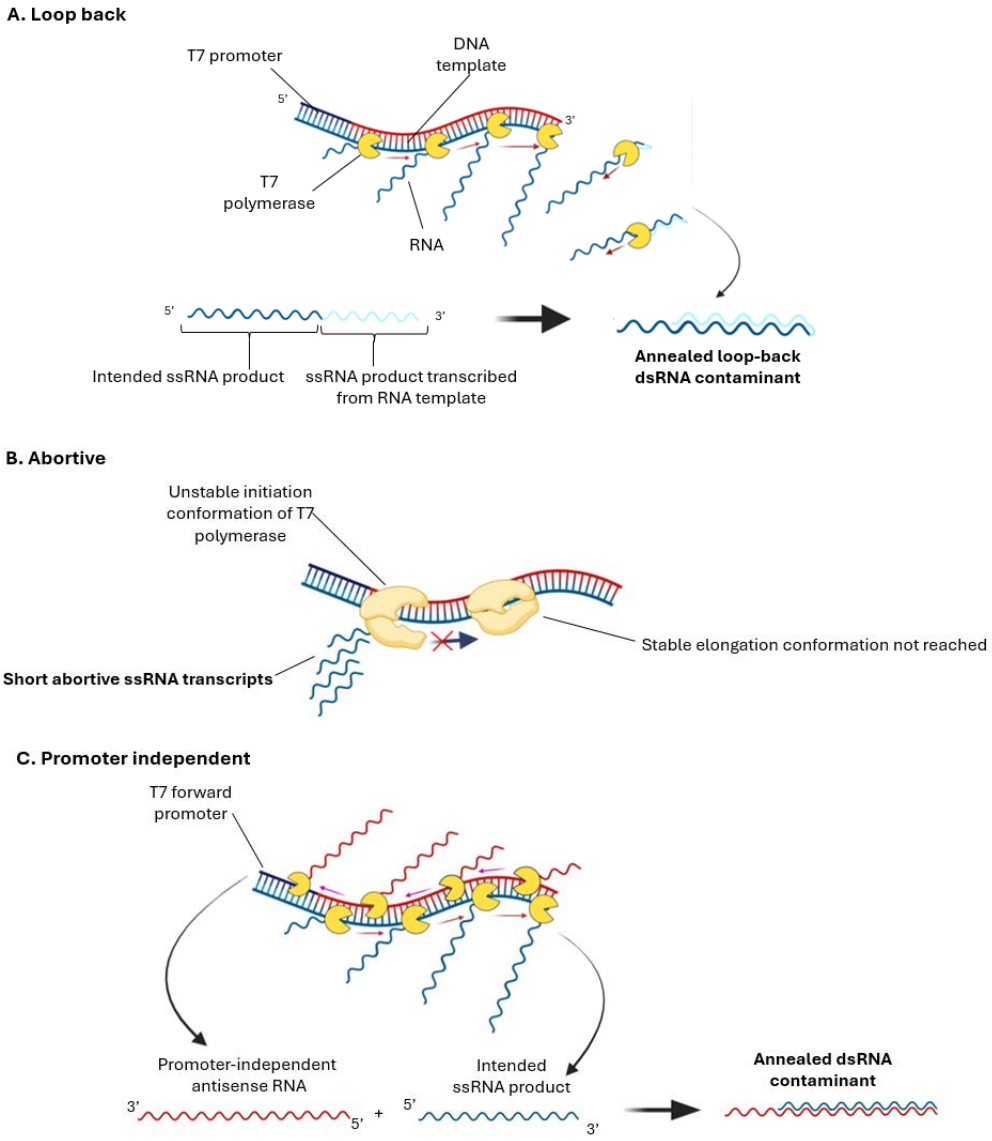


Figure 1.10. Illustration of various IVT by-products. (A) RNA-templated loop-back extension occurs when T7 RNA polymerase re-initiates at the 3' end of the DNA template due to RNA self-complementarity. (B) Abortive short ssRNA transcripts result from the failure of T7 RNA polymerase to transition into the elongation phase. (C) Promoter-independent transcription by T7 polymerase leads to full-length ssRNA using the non-coding strand as the template, which then anneals with the intended ssRNA product. Image made using Biorender.com.

Detection and quantification of dsRNA and other by-products is critical for assessing mRNA quality and safety, and several analytical techniques are commonly employed for this purpose. High-performance liquid chromatography is widely used to quantitatively assess dsRNA contamination in mRNA products (Karikó et al., 2011). The other, more qualitative, standard for detecting dsRNA involves the use of J2 or K1 antibodies in dot blot assays or enzyme-linked immunosorbent assays (ELISAs). These assays are effective for identifying dsRNA over 40 nucleotides long (Bonin et al., 2000; Schonborn et al.,

1991). Interferon-β assays conducted in immune cells can provide an indirect measure of the innate immune response triggered by dsRNA (Dousis et al., 2023; Karikó et al., 2005). Acridine orange staining of denaturing polyacrylamide gels can also be used to differentiate between double and single-stranded RNA and provide additional information about the length of dsRNA contaminants (Mu et al., 2018).

Efforts to mitigate the formation of dsRNA by-products have included optimising IVT reaction conditions and designing modified RNA polymerases. For example, reducing Mg^{2+} ion concentrations in the IVT buffer has been shown to decrease dsRNA production by template-independent transcription (Boman et al., 2024; Mu et al., 2018). The use of modified nucleosides like Ψ, $m^1Ψ$, and m^5C has also been reported to lead to lower dsRNA production (Mu & Hur, 2021). Additionally, altering the steady-state concentration of nucleotides, such as UTP or $m^1ΨTP$, through a fed-batch process has been found to reduce dsRNA levels of promoter-independent transcription of a DNA template encoding mRNA with a poly(A) tail (Ziegenhals et al., 2023). DNA template design has also proven effective: linearising templates with restriction enzymes that leave 5' overhangs (Mu et al., 2018; Schenborn & Mierendorf, 1985), or using DNA templates that are only double stranded in the T7 promoter region, can significantly reduce the formation of dsRNA by-products (Mu et al., 2018).

Post-IVT purification techniques are essential for removing dsRNA contaminants by employing reverse-phase high-performance liquid chromatography (Karikó et al., 2011) and cellulose-based columns (Baiersdörfer et al., 2019).

In a recent development, Moderna engineered a double-mutant T7 polymerase that produces substantially lower levels of dsRNA during IVT and resulted in IVT mRNA with increased 3’ end homogeneity (Dousis et al., 2023). They reported post-IVT purification of mRNA products made by this mutant T7 polymerase did not further reduce dsRNA associated interferon-beta (IFN-β) immune response in fibroblasts or *in vivo*.

1.4. Delivery of exogenous mRNA

From the early days of nucleic acid delivery, encapsulation methods have been investigated to enhance the stability and internalisation of RNA therapies. As reviewed by Sahin et al. (2014), the delivery of naked mRNA has shown limited efficacy, primarily due to its rapid degradation in circulation, immune activation, and inefficient cellular uptake. Additionally, the size and negative charge of mRNA molecule render its passive cellular entry extremely unlikely (Dowdy, 2017). Initial efforts in the field focused on plasmid DNA and mRNA, later expanding to include ASOs and siRNAs in the 2000s (Hou et al., 2021; Qin et al., 2023; Semple et al., 2001). A range of delivery materials has been explored, including lipids, polymers, proteins, and lipid-like compounds.

Encapsulated mRNA enters cells through endocytosis, the process by which the cell membrane engulfs extracellular material into intracellular vesicles called endosomes, as reviewed by Dowdy et al. (2022). The plasma membrane invaginates around the material, creating a vesicle that pinches off into the cytoplasm. These vesicles fuse with early endosomes, which maintain a slightly acidic environment with a pH of around 6.3, as reviewed by Pei & Buyanova (2019). As the endosomes mature into late endosomes, the pH decreases to around 5.5. Eventually, late endosomes merge with lysosomes, where the pH decreases further to around pH 4 to 5. mRNA must escape into the cytosol at some point during intracellular trafficking to avoid pH or RNAse-mediated degradation in later stages of the endolysosomal pathway and reach cytosolic ribosomes to be translated.

1.4.1. Lipid nanoparticles

In the 1990s, liposome-based drug delivery systems were investigated for genetic therapies, but their effectiveness was limited by the use of permanently charged cationic lipids, as reviewed by Buschmann et al. (2021) and Cullis & Hope (2017). These lipids caused significant toxicity and were rapidly cleared from circulation, prompting the need for more refined alternatives.

A key breakthrough in nucleic acid delivery was achieved when researchers at Inex Pharmaceuticals introduced ionisable cationic lipids to the field (Semple et al., 2001). These lipids can switch from a protonated state at low pH to a neutral state at physiological pH, a property that allows for the efficient complexing with negatively charged nucleic acids during production, while reducing the toxicity associated with permanently charged cationic lipids.

This spurred the development of more sophisticated lipid nanoparticles (LNPs), as illustrated in Figure 1.11, which became the foundation for RNA delivery systems (as reviewed by Hou et al. (2021) and Tenchov et al. (2021)). These LNPs are composed of four essential components: an ionisable cationic lipid for complexing RNA; a helper lipid such as dioleoylphosphatidylethanolamine (DOPE) to facilitate interactions with biological membranes; cholesterol to enhance the structural stability of the formulation, minimise interactions with blood proteins and minimise vehicle disassembly; and a poly(ethylene glycol) (PEG) conjugated lipid to ensure colloidal stability during production and also to potentially reduce opsonisation, the process whereby blood proteins mark NPs for immune clearance (Owensiii & Peppas, 2006).

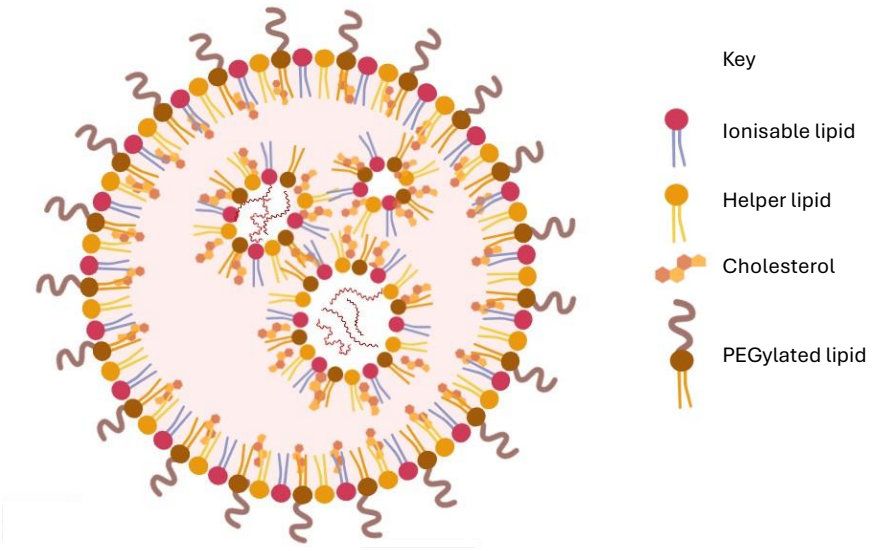


Figure 1.11. Schematic representation of a multilamellar lipid nanoparticle encapsulating mRNA. The structure of LNPs is thought to be influenced by various factors, including charge ratio, the chemical structure of individual components, the relative proportion of components, and the method of preparation. Image made using Biorender.com.

Substantial optimisation by Jayaraman et al. (2012) demonstrated that, for amino ionisable lipids with similar structures, *in vivo* hepatic gene silencing is significantly influenced by pK_a, with optimal values for effective siRNA delivery falling between 6.2 and 6.5. The potential of LNPs was realised with the approval of *Onpattro*, the first Food and Drug Administration (FDA)-approved LNP-encapsulated siRNA medicine in 2018 (Akinc et al., 2019). However, LNPs have since been adapted to better suit the delivery of mRNA. Relative to siRNA, mRNA molecules are significantly longer, carry more negative charges, and have been reported to differ in how they package into LNPs (Kulkarni et al., 2018; Yanez Arteta et al., 2018). These differences are thought to potentially affect the stages at which mRNA and siRNA escape from endosomes during endocytosis (Hunter et al., 2023; J. Shin et al., 2024).

To optimise LNPs for mRNA, rather than siRNA delivery, the ionisable lipids in LNPs have been redesigned, as reviewed by Hou et al. (2021) and Tang et al. (2023). The newer ionisable lipid (SM-102) used in the Moderna COVID-19 vaccine demonstrates improved *in vivo* delivery and faster clearance from the liver and plasma following intramuscular administration compared to MC3 (Hassett et al., 2019). In addition, Hassett et al. (2019) suggested the enhanced biodegradation of SM-102 increases the general tolerability of LNPs relative to MC3 LNPs, making them more suitable for repeated dosing while maintaining effective mRNA delivery.

While the widespread administration of LNP-based COVID-19 vaccines *via* IM injection has demonstrated their general safety, there have been reports of LNP administration causing a significant increase in anti-PEG antibodies (Ju et al., 2022). In some patients, these antibodies can lead to enhanced and rapid clearance of systemically delivered PEGylated NPs, potentially reducing therapeutic efficacy (B.-M. Chen et al., 2021). Alternatives to PEGylated lipids, such as Poly(2-ethyl-2-oxazoline) (PEOZ) lipids, have shown promise. Sanchez et al. (2024) showed that the use of PEOZ lipids in mRNA-LNPs facilitate similar levels of anchored VHH protein expression in liver tissue as PEGylated lipid-

containing LNPs while also reducing the incidence of anti-PEOZ antibodies relative to the incidence of anti-PEG antibodies.

The immunostimulatory properties caused by the lipid components have been suggested to restrict current LNP technology to applications like vaccines, where LNPs serve as adjuvants (Y. Lee et al., 2023). However, a recent Moderna clinical trial in humans demonstrated that repeated IV administration of ionisable LNPs containing PEGylated lipids for protein replacement therapy was well tolerated (Koeberl et al., 2024).

1.4.2. Alternative synthetic formulation vehicles

Cationic synthetic polymers, such as poly(ethylenimine) (PEI), complex with mRNA effectively through electrostatic interactions. By neutralising the negative charges of mRNA, they facilitate formation of complexes that promote cellular uptake *via* endocytosis and enhance endosomal escape, as reviewed by Lin et al. (2023). However, the positive charge required for encapsulation and, potentially for endosomal escape, is associated with increased toxicity, making it difficult to achieve a balance between efficiency and safety for *in vivo* applications. Furthermore, the electrostatic interactions between positively charged long polymers and negatively charged RNA can lead to overstabilisation, a phenomenon where the mRNA remains too tightly bound to the polymer, thereby reducing its overall performance (Bettinger, 2001).

To address these challenges, researchers have developed next-generation cationic polymers, such as poly(β-amino ester) (PBAE) (as reviewed by Yousefi Adlsadabad et al. (2024)). PBAE polymers present several advantages over PEI, such as reduced cytotoxicity and enhanced RNA delivery efficiency (Sunshine et al., 2012). Additionally, they exhibit significant potential for structural diversity, enabling the efficient generation of large libraries of structurally diverse polymers (Karlsson et al., 2020). PBAE polymers can also be co-formulated with lipid components such as PEGylated lipids, to create lipo-polyplexes (Kaczmarek et al., 2016). Co-formulation with PEGylated lipids has

been reported to promote PBAE nanoparticle stability in the presence of serum proteins (Eltoukhy et al., 2013).

One suggested limiting factor for efficient mRNA translation post-endosomal escape is the absence of bound PABP and eIF4E, both essential for cap-dependent mRNA translation. To address this, J. Li, He, et al. (2017) and J. Li, Wang, et al. (2017) developed polymer-based nanoparticles designed to co-deliver mRNA complexed with either PABP or eIF4E proteins. This co-delivery approach has demonstrated enhanced protein expression *in vitro*, highlighting the potential of polymer-based nanoparticles to deliver ribonucleoproteins and improve mRNA translation efficiency, thereby supporting therapeutic efficacy.

1.4.3. Biological delivery systems

Viruses have evolved sophisticated mechanisms to achieve endosomal escape, a critical step for delivering their genetic material into host cells, as reviewed by Staring et al. (2018). Enveloped viruses typically use fusion peptides that are activated by the acidic environment of the endosome. These peptides undergo conformational changes that enable the viral envelope to merge with the endosomal membrane, releasing the viral contents into the cytoplasm. In contrast, non-enveloped viruses employ a variety of peptides to facilitate their escape from the endosome, with many viruses utilising distinct mechanisms tailored to their composition and lifecycle (Staring et al., 2018).

Viral-like particles (VLPs) are generated by transfecting plasmids encoding packaging proteins, envelope proteins, and the mRNA of interest into host cells, which then package the mRNA into VLPs (Gupta et al., 2024; He et al., 2022; X. Wang et al., 2024). VLPs exploit fusogenic proteins, derived from enveloped viruses. One advantage of this fusogenic approach is its ability to bypass immune detection by endosomal TLRs. A notable example of a fusogenic protein used in VLP systems is the vesicular stomatitis virus envelope protein (VSV-G). However, the toxicity of VSV-G has prompted research into safer alternative fusogenic proteins, such as mouse syncytin-A protein (Segel et al., 2021; Coquin et al., 2019).

Despite their advantages, VLPs have a very low loading capacity for mRNA (Prel et al., 2015), which limits their effectiveness as delivery vehicles for therapeutic applications. MS2-VLPs have been developed to enhance the loading capacity of retroviral particles by fusing the viral MS2 coat protein (MCP) to packaging proteins and incorporating MS2 RNA aptamers into the delivered mRNA (B. Lu et al., 2019; Prel et al., 2015). While this strategy has been shown to improve reporter mRNA expression, concerns exist that MCP-packaging fusion proteins may unintentionally encapsulate random host cellular RNA (B. Lu et al., 2019).

Exosomes are small extracellular vesicles, typically between 50 and 100 nm in size, that are secreted by cells to facilitate intercellular communication by transferring genetic material, such as microRNAs and mRNAs, tRNAs and proteins (as reviewed by Ibba et al. (2021), Y. Lu et al. (2023) and Weng et al. (2022)). Owing to their biocompatibility, stability in systemic circulation, and natural ability to cross cell membranes, exosomes have emerged as promising mRNA delivery vehicles. For instance, exosomes have been used in xenograft models to deliver mRNA to arrest the growth of HER2+ breast cancer and to deliver Cas9 mRNA to cultured cells where they were able to repress expression levels of miRNAs associated with leukaemia (Forterre et al., 2020; Usman et al., 2018).

However, several challenges hinder the widespread use of exosomes for mRNA delivery, as reviewed by Y. Lu et al. (2023) and Yamashita et al. (2018). First, large-scale production of exosomes with high quality remains difficult, as yields from biological fluids or cell culture supernatants are typically low. This problem is compounded when loading large mRNAs into exosomes, as the insertion process can be inefficient (Z. Yang et al., 2019). Furthermore, concerns about the safety and immunogenicity of exosomes persist, especially regarding the potential toxicity from immortalised donor cells (Green et al., 2015; Nakamura et al., 2017).

1.4.4. Tissue targeting of delivery vehicles

Regardless of the therapeutic indication, whether vaccine or protein replacement, mRNA delivery and translation must occur within a specific set of cells for therapeutic efficacy. In the case of mRNA vaccines, this is thought to involve the uptake and translation of mRNA within APCs or directly into T cells following IM administration to stimulate a robust immune response (as reviewed by Gote et al. (2023)). For mRNA-based protein replacement therapies, successful translation must occur within specific tissues and cells that lack the target protein, as reviewed by Lin et al. (2023). However, achieving delivery beyond the liver following intravenous administration remains challenging with current LNP formulations, as reviewed by T. Zhang et al. (2024). This tendency toward hepatic tissue accumulation of LNPs has been suggested to be due to the neutral charge of the LNP, which facilitates interactions with highly expressed apolipoprotein E proteins on the cell surface of hepatocytes. Changing the net charge of LNPs is thought to alter the composition of the protein corona, which may play a crucial role in facilitating organ-specific targeting (Cheng et al., 2020; Dilliard et al., 2021).

The use of targeting ligands on LNP surfaces offer potential for enhancing receptor-mediated uptake in specific tissues, as reviewed by Lin et al. (2023). One example of receptor-targeted LNPs includes N-acetylgalactosamine (GalNAc)-conjugated LNPs, which specifically bind to the hepatocyte's asialoglycoprotein receptor (ASGPR) (Kasiewicz et al., 2023). Goswami et al. (2019) demonstrated that conjugating mannose to cholesterol prior to mRNA encapsulation facilitates targeting of the mannose receptor CD206 on APCs, both *in vitro* and *in vivo*. This approach led to increased antibody titres in APCs, and has been further validated by Lei et al. (2024). However, to date, no such ligand-targeted NP drug delivery system has been clinically approved for mRNA therapeutics.

To optimise tissue-specific targeting of nanoparticles, innovative screening methods have been developed. High-throughput DNA barcoding of nanoparticles involves the encapsulation of DNA barcodes alongside reporter

mRNA within LNPs, allowing for the simultaneous assessment of the biodistribution and efficacy of multiple formulations in the same animal (Huayamares et al., 2023; Paunovska et al., 2018; Sago et al., 2018). Another method developed by Lindsay et al. (2019) involves the comparison of mRNA biodistribution through positron emission topography/computed tomography (PET/CT) imaging, coupled with bioluminescence to track reporter protein expression.

1.5. Endosomal escape as a bottleneck of mRNA therapeutics

Following internalisation of therapeutic mRNA by endocytosis, efficient endosomal escape is essential for mRNA translation and therapeutic efficacy. Despite its critical role, the process remains highly inefficient, with the majority of delivered mRNA remaining trapped and degraded within endo-lysosomal compartments as reviewed by Chatterjee et al. (2024). This inefficiency limits the overall success of many mRNA delivery systems and poses a significant challenge in their development.

The mechanisms by which endosomal escape occurs remain poorly understood, as reviewed by Dowdy et al. (2022). One hypothesis is the "proton sponge" effect, where the acidification of the endosome causes the NP to absorb water, leading to an increase in volume and subsequent rupture of the endosome (Vermeulen et al., 2018). Sonawane et al. (2003) observed this phenomenon in DNA-PEI polyplexes, where they noted a 140% increase in endosomal volume. This phenomenon appears to be more applicable to polymer-based NPs and LNPs may escape the endosome *via* a different mechanism (Wittrup et al., 2015; Yanez Arteta et al., 2018). As the pH within the endosome drops, ionisable lipids in LNPs are thought to adopt a hexagonal phase structure, facilitating membrane fusion and endosomal escape (Yanez Arteta et al., 2018). However, variability in endocytic trafficking pathways across different cell types complicates understanding of this transient process (Hunter et al., 2023; Sayers et al., 2019), posing a bottleneck in the optimisation of mRNA delivery systems.

1.5.1. Microscopy-based approaches to probe endosomal trafficking and escape

A variety of microscopy-based techniques are employed to visualise mRNA trafficking during internalisation and endosomal escape. While these methods provide valuable insights into endosomal dynamics, they often lack the quantitative precision needed to assess how mRNA escapes into the cytoplasm. Below, key techniques, their limitations, and their combined use are discussed.

1.5.1.1. Visualisation of unlabelled mRNA

Fluorescence *in situ* hybridisation (FISH) is a commonly used technique to visualise the localisation of mRNA in fixed cells, where fluorescently labelled oligonucleotide probes bind to specific RNA sequences, allowing for mRNA detection (as reviewed by Young et al., (2020)). Single-molecule FISH (smFISH) enhances this by using multiple probes to increase signal intensity, enabling the visualisation of individual mRNA molecules (Haimovich & Gerst, 2018a). However, FISH techniques alone have not been sufficient to accurately quantify the subcellular localisation of delivered mRNA, and they typically need to be combined with immunofluorescence (IF) to incorporate reference markers (Paramasivam, Stöter, et al., 2022).

To address this, Paramasivam, Franke, et al. (2022) developed a combined smFISH and IF approach that enables simultaneous detection of unlabelled mRNA and endosomal markers. This method provided insight into the co-localisation of delivered mRNA with specific endosomal proteins and offered insights into the vicinity of mRNA relative to endosomes. However, dual smFISH:IF showed limitations in detecting mRNA encapsulated in LNPs within intact endosomes. Paramasivam, Franke, et al. (2022) suggested this may be due to the absence of accessible primary amines near encapsulated mRNA, which impairs effective cross-linking during fixation.

1.5.1.2. Visualisation of labelled mRNA

Fluorescent labelling of mRNA throughout the CDS can be used for live cell tracking of its intracellular localisation. However, this method has significant limitations. Ribosomes may stall when they encounter labelled nucleotides, resulting in impaired protein translation and faster degradation of the mRNA (Custer & Walter, 2017). Additionally, commonly used cyanine-based fluorophores have been reported to have intrinsic subcellular targeting tendencies, complicating the interpretation of whether mRNA localisation reflects the true intracellular distribution of the mRNA or the fluorophore (Negwer et al., 2017; W. J. Rhee & Bao, 2010). Furthermore, encapsulating

fluorophore labelled and unlabelled mRNAs can result in differences in NP size (Bernard et al., 2023), potentially influencing the route of endocytic uptake and overall performance. As a result, body-labelled mRNA should be cautiously used as a proxy for mRNA internalisation, as this alone has been reported to not be a reliable predictor of transfection efficacy (Paramasivam, Franke, et al., 2022; Sayers et al., 2019).

An alternative method for labelling mRNA, introduced by Winz et al. (2012) and then developed by Custer & Walter (2017), addresses the limitations associated with CDS-labelled mRNA by enzymatically incorporating N_3-dATPs into the poly(A) tail of IVT mRNA. This click-chemistry labelling technique has been used primarily for "in-cell" labelling, allowing mRNA to be tagged after it has entered cells or organisms, such as zebrafish (Westerich et al., 2020). Importantly, this method has been validated to not interfere with translation, preserving the functional integrity of the labelled mRNA (Anhäuser et al., 2019).

Multiply Labelled Tetravalent RNA Imaging Probes (MTRIPs) provide an alternative method for labelling mRNA, with minimal impact on protein expression (Kirschman et al., 2017). These oligonucleotide probes bind to the 3' UTR of IVT mRNA, enabling the tracking of mRNA within cells by live-cell imaging and by proximity ligation assays. While MTRIPs have been used to visualise mRNA uptake and trafficking, their application for studying endosomal escape and co-localisation with endocytic markers remains unexplored.

1.5.1.3. Visualisation of proteins as endocytic markers

To investigate mRNA trafficking and its association with endosomal compartments, fluorescently tagged endosomal markers, such as Rab5 and Rab7, are frequently used (C.-C. Chen et al., 2017; Johansson et al., 2024). These markers allow real-time tracking of early and late endosomes, respectively. However, overexpression of these markers, typically achieved through transient plasmid transfection, can lead to altered endosomal function and size (Bucci et al., 1992). Additionally, endosomal markers like Rab proteins

often localise to multiple endosomal subpopulations, making it challenging to delineate specific trafficking routes (Shearer & Petersen, 2019).

A more recent approach to probing endosomal escape involves the use of stable cell lines expressing fluorescently tagged galectin proteins, such as Gal3 and Gal9 (Beach et al., 2022; Joshi et al., 2020; Munson et al., 2021). These proteins are recruited to damaged endosomes, enabling visualisation of transient endosomal membrane rupture events, which indicate escape attempts. The number of Gal9 puncta was reported to generally correlates with both the mean intensity and proportion of GFP-positive cells following transfection with LNPs, suggesting it serves as an effective predictor of LNP efficacy (Munson et al., 2021). However, not all delivery vehicles induce galectin recruitment. For example, extra-cellular vesicles have demonstrated the ability to achieve endosomal escape without triggering Gal3-Azami Green puncta formation (Joshi et al., 2020). Furthermore, certain LNP components, such as hydroxycholesterol derivatives may act as mTOR inhibitors, leading to a decoupling of endosomal escape from successful translation (Munson et al., 2021).

1.5.2. Functional assays

Functional outputs following delivery of mRNA in cells are typically assessed using a variety of quantitative techniques, including luciferase assays for bioluminescent measurement and fluorescence-based methods for visualising protein expression (Dastgerdi et al., 2024). To mitigate the effects of signal amplification in following mRNA delivery, one strategy is to use mRNA constructs encoding intentionally destabilised reporter proteins, such as firefly luciferase fused to degradation-promoting C-terminal PEST (Luc2CP) or destabilised EGFP (d2eGFP) (Kijima et al., 2018; Orlandini von Niessen et al., 2019). These unstable proteins degrade rapidly, helping to reduce the compounding effects of repeated rounds of translation from each mRNA molecule.

Protein expression *in vivo* can be evaluated through delivery of mRNA encoding a glycosylphosphatidylinositol (GPI)-anchored camelid VHH antibody using cell-surface flow cytometry (Sago et al., 2018; Sanchez et al., 2024). For assessing selective delivery *in vivo*, options include monitoring bioluminescence following the delivery of firefly luciferase mRNA (Lindsay et al., 2019). Additionally, far-red fluorescence can be monitored after delivering cyclic recombinase (Cre) mRNA in genetically modified mice, which are engineered with a LoxP-flanked STOP cassette that prevents transcription of a far-red fluorescent protein variant (tdTomato) in all tissues (Kauffman et al., 2018).

Protein expression is a commonly used quantitative measure to assess mRNA delivery, but it remains an indirect indicator of success, introducing a significant time lag between nanoparticle internalisation and translation. Several variables can affect protein expression levels beyond nanoparticle delivery efficiency. These include how the mRNA interacts with the nanoparticle and influences its properties (Bernard et al., 2023) and the specific endo-lysosomal trafficking routes (Hunter et al., 2023). Therefore, functional assays should be complemented with label-free quantitative techniques that can differentiate between subcellular compartments and monitor mRNA localisation to obtain deeper insights into the mRNA delivery pathways and endosomal escape mechanisms.

1.6. Proximity Dependent Ligation

1.6.1. Development of proximity labelling techniques

The concept of proximity-based labelling was first introduced with DamID, a technique where DNA adenine methyltransferase (Dam) is fused to a DNA-binding protein of interest (Steensel & Henikoff, 2000). Dam methylates nearby DNA, enabling researchers to study chromatin interactions and has been used to as a tool to identify transcriptional targets of DNA binding proteins (Taverner et al., 2004).

This approach was later applied to protein labelling by use of an engineered mutant *E. coli* biotin ligase (BirA*) that was designed to be more promiscuous in biotinylating proteins beyond its natural substrate, biotin carboxyl carrier protein (Choi-Rhee et al., 2004). BirA* activates free biotin by coupling it with ATP, producing a highly reactive biotin-AMP ester. This ester reacts with nearby lysine sidechain amines, covalently tagging proteins within approximately 10 nm of the ligase (Choi-Rhee et al., 2004). Since biotin-AMP is short-lived, only proteins in close proximity to BirA* are labelled. BioID combines the proximity-dependent biotinylation of BirA* fusion proteins with streptavidin pull-down and liquid chromatography tandem mass spectrometry (LC-MS/MS) to isolate and identify neighbouring proteins near the ligase (Roux et al., 2012).

In 2013, APEX was introduced as another advancement in proximity labelling. APEX, an engineered ascorbate peroxidase (APX) with a molecular weight of 28 kDa, was modified to function as a monomer in reducing environments like the cytosol by removing cysteine residues (Martell et al., 2012). Initially developed for generating contrast in electron microscopy, this technique was then applied as a spatioproteomic tool (H.-W. Rhee et al., 2013). APEX catalyses the one-electron oxidation of biotin-phenol into a biotin-phenoxyl radical, which covalently tags nearby proteins (Figure 1.12). These biotinyl radicals are short-lived and highly reactive, labelling electron-rich amino acids such as tyrosine, tryptophan, cysteine, and histidine within approximately 20 nm of the enzyme's active site. The 'labelling radius' is not fixed but

decreases with distance depending on factors like antioxidant concentration (H.-W. Rhee et al., 2013).

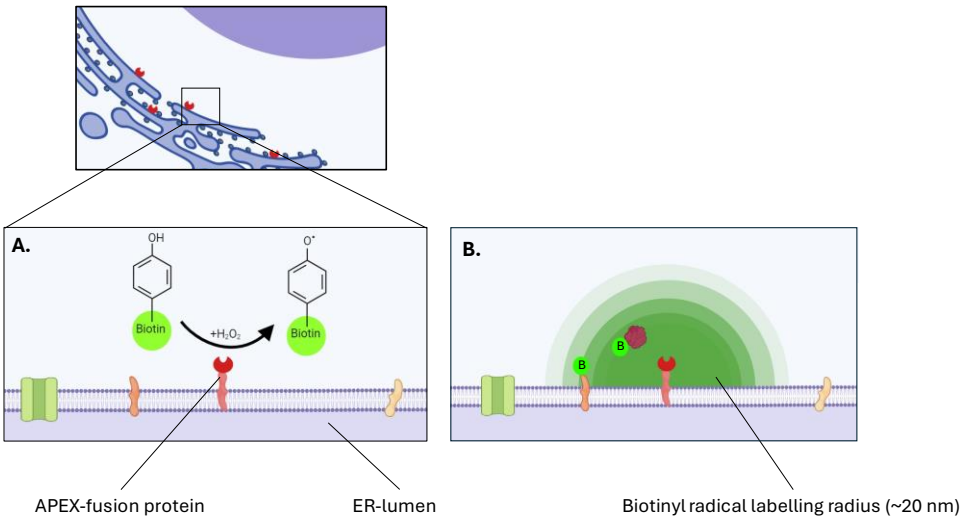


Figure 1.12. Schematic illustrating APEX labelling of proteins on the cytosol facing side of the ER. (A.) APEX catalyses the formation of biotinyl radicals. (B.) The typical labelling radius of the biotinyl radicals is shown, with biotinylated proteins indicated by a green "B." Image created with BioRender.com.

APEX2, an improved version of APEX, was engineered through a yeast phage display-driven mutation (A134P) to increase sensitivity and labelling efficiency. This mutation reduces susceptibility to H_2O_2-induced inhibition, making APEX2 more effective for spatioproteomics applications (Lam et al., 2014). APEX2 offers several advantages over BioID for proximity-based protein labelling, including faster labelling (1 minute versus 24 hours) and the ability to use a variety of aromatic substrates conjugated to biotin. However, a key drawback of APEX and APEX2 proximity-dependent labelling approaches is the exposure of cells to hydrogen peroxide, which can be toxic and can lead to off-target oxidative reactions that complicate results (Finnegan et al., 2010; Kang & Rhee, 2022). Since the introduction of APEX2, several engineered BirA biotin ligases have been developed to enhance BioID. These include TurboID, which accelerates labelling speed, and ultraID, which not only improves labelling speed but also reduces the size of the tag (Branon et al., 2018; Kubitz et al., 2022).

Both APEX2 and BioID are routinely used as spatial proteomics tools, with biotinylated proteins typically enriched using streptavidin beads, followed by

analysis *via* western blot or LC-MS/MS (Kang & Rhee, 2022). However, conventional mass spectrometry (MS) can misidentify non-biotinylated proteins that are enriched due to non-specific interactions with streptavidin (Kang & Rhee, 2022). Newer LC-MS/MS techniques such as Spot-BioID directly detect trypsinised biotinylated peptides, improving specificity and reducing false positive rates (S.-Y. Lee et al., 2019; S. Shin et al., 2024).

1.6.2. Application of proximity labelling to RNA

Proximity labelling has extended beyond protein studies to RNA research, offering new insights into spatial subcellular transcriptomics. One of the first approaches to use proximity labelling for RNA was APEX-RIP (RNA Immunoprecipitation), developed by the Ting lab (Kaewsapsak et al., 2017a). In this method, APEX2 is used to label proteins in close proximity to RNA, followed by cross-linking the RNA to these proteins. The labelled protein-RNA complexes are then isolated by streptavidin pulldown, enabling researchers to investigate RNA-protein interactions. However, APEX-RIP has been reported to have low accuracy in non-membrane-enclosed subcellular regions, limiting the technique's applicability to regions such as the lumen of the ER and mitochondria (Kaewsapsak et al., 2017).

Building upon these findings, the Ting lab developed APEX-seq in 2019, a direct RNA labelling method that bypasses the need for protein intermediaries (Fazal et al., 2019). APEX-seq uses APEX2 fused to a target protein, which, under standard biotinylation conditions, biotinylates nearby RNA. This RNA is then isolated *via* streptavidin affinity purification and analysed by RNA-sequencing. Ting's lab initially demonstrated the specificity of APEX-seq by labelling RNA within the mitochondrial matrix, a membrane-enclosed environment. This was followed by the creation of a spatial map of RNA across various organelles by using multiple APEX2 fusion cell lines that the Ting lab had previously developed for proteomics research. APEX-seq enabled the exploration of RNA trafficking mechanisms, including both RNA-dependent and independent pathways involved in mitochondrial localisation (Fazal et al., 2019).

Around the same time, the Ingolia lab also validated APEX-seq for spatial subcellular transcriptomics (Padrón et al., 2019). They used an *in vitro* APEX2 biotinylation technique to confirm RNA labelling by using streptavidin-horseradish peroxidase (HRP) RNA dot blot assays to detect biotinylated RNA. By fusing APEX2 to eIF4A1, which is known to be trafficked to stress granules under conditions of heat shock and hippuristanol-induced stress, they demonstrated the potential of APEX2 in characterising both RNA and protein components of stress granules (Padrón et al., 2019). In contrast to existing techniques for isolating core SG components, Padrón et al. (2019) reported that APEX2-seq facilitated detection of weaker, longer range interactions.

In an effort to improve upon the existing methods, the Zou lab explored alternative APEX2 biotin substrates beyond biotin-phenol (Y. Zhou et al., 2019). They sought substrates that would react with greater specificity towards RNA over proteins and DNA. Their research identified biotin-aniline as a promising substrate for preferential RNA labelling and biotin-naphthalene for DNA labelling through *in vitro* dot blot assays. Zou's work also explored whether specific RNA sequences or structures exhibited higher APEX labelling efficiency, with their findings suggesting that guanine bases were preferentially labelled, potentially due to guanine's lower redox potential compared to other nucleotides (Wang et al., 2020; Y. Zhou et al., 2019).

While proximity-dependent labelling has yet to be applied specifically to therapeutic mRNA or mRNA delivery in the literature, similar approaches have been employed to investigate the delivery of small drug-like molecules using cell lines stably expressing fusion HaloTag proteins (Peraro et al., 2018). HaloTag is a labelling protein that accepts chloroalkane as a substrate, which can be conjugated to various molecules, making the system particularly versatile (Los et al., 2008). Peraro et al. (2018) utilised fusion HaloTag cell lines, where HaloTag proteins localised to the outer mitochondrial membrane or the nucleus, to quantify the dose-dependent penetration of chloroalkane-linked small molecules. This assay provided valuable insights into the uptake of small drug-like molecules into the cytosol and nucleus.

1.7. Aims and objectives of project

The primary aim of this thesis is to address the current limitations in quantifying the subcellular localisation of delivered mRNA within cells. This project will utilise the APEX2-RT-qPCR technique to investigate the localisation of IVT mRNAs encoding both reporter and therapeutic proteins. The selection of APEX2-RT-qPCR is based on its predicted advantages over existing analytical techniques and other proximity biotinylation methods, including its potential for quantitative analysis of unlabelled RNA, rapid labelling times, and adaptability to different subcellular compartments. These features make it particularly well-suited for studying the intracellular fate of mRNA therapeutics.

The key objectives of this project are:

- To create and validate APEX2-expressing cell lines for the specific purpose of quantifying the subcellular localisation of mRNA within the cytosol and the ER.
- To explore how different RNA features, such as signal sequences, influence the localisation of delivered mRNA and the efficiency of APEX2 labelling.
- To use APEX2-RT-qPCR alongside functional assays to investigate the relationship between the localisation of delivered mRNA and protein expression, providing insights into how localisation affects translational efficiency.
- To investigate the potential of APEX2-RT-qPCR for quantifying escape from endosomes and delivery to sites of subcellular translation, by using various delivery vehicles that exhibit differing transfection efficiencies.

By addressing these objectives, this study aims to establish APEX2-RT-qPCR as a versatile tool for subcellular mRNA quantification. Its combination of quantitative precision and the ability to track RNA localisation without prior labelling of IVT mRNA has the potential to significantly advance the understanding and development of mRNA therapeutics

Chapter 2 - Materials and Methods

2.1. Materials

2.1.1. General Cell Culture Reagents

Table 1 Cell culture reagents.

Reagent	Composition	Supplier
Dulbecco's modified eagle medium (DMEM)	N/A	Sigma-Aldrich
Foetal bovine serum (FBS)	N/A	Gibco
OptiMEM	N/A	Gibco
Trypsin-EDTA (Trypsin)	N/A	Invitrogen
Lipofectamine 2000	N/A	Invitrogen
Fugene HD	N/A	Promega
Phosphate buffered saline (PBS)	4.3 mM Na_2HPO_4, 1.5 mM KH_2PO_4, 137 mM NaCl, 2.7 mM KCl, pH 7.4	In-house

2.1.2. Plasmids

Table 2. Plasmids Employed: Details include source, use and resistance gene for each plasmid. CLJ lab – Catherine Jopling Lab.

Plasmid Name	Source	Use
APEX2-ERM	Addgene (Cat no.: 79055)	Lentiviral transfer plasmid allowing for integration of APEX2 fusion protein expected to localise to the endoplasmic reticulum. Utilised for lentiviral transduction
APEX2-NES	Addgene (Cat no.: 92158)	Lentiviral transfer plasmid allowing for the integration of APEX2 fusion protein expected to be diffuse throughout cytosol. Utilised for lentiviral transduction.
PsPAX2	Addgene (Cat no.: 12260)	Lentiviral 2nd generation packaging plasmid. Utilised for lentiviral transduction.
PMD2.G	Addgene (Cat no.: 12259)	Lentiviral VSV-G envelope expressing plasmid. Utilised for lentiviral transduction.
NLucHCV	CLJ lab (backbone from pGL3 Luciferase Reporter Vectors Promega)	Utilised for classical restriction cloning to provide the vector backbone. This ended up being the one vector backbone utilised for all vector constructs generated in this project.
NlucSecHCV	CLJ lab (backbone from pGL3 Luciferase Reporter Vectors Promega)	Used for *in vitro* transcription to produce mRNA for transfection.
FlucHCV	CLJ lab (backbone from pGL3 Luciferase Reporter Vector Promega). (Roberts et al., 2011).	Utilised for *in vitro* transcription to generate spike-in RNA.
eGFP-C1	Clontech	Utilised for *in vitro* transcription to generate spike-in RNA and to generate eGFP CDS insert for NEBuilder HiFi Assembly cloning.
RLuc	pRL-SV40 (Promega)	Utilised for *in vitro* transcription to generate Renilla luciferase spike-in RNA.
mCher	David Heery group	Utilised to generate mCherry CDS insert for NEBuilder HiFi Assembly cloning.

Table 3. Plasmids Generated: Details include backbone, modifications/insertions.

Plasmid Name	Backbone	Insert/ modification	Cloning strategy
NlucHBB	NlucHCV *Made by Poppy Winlow whilst in the Jopling lab.*	Replaced HCV UTRs from with Hbb UTRs using a IDT GeneBlock of Nluc CDS flanked by HBB UTRs.	Classical restriction cloning
NlucSecHBB	NlucHCV *Made by Poppy Winlow whilst in the Jopling lab.*	Replaced HCV UTRs from with Hbb UTRs using a IDT GeneBlock from IDT of NlucSec CDS flanked by HBB UTRs	Classical restriction cloning
NlucaltSPCSHBB	NlucSecHBB	Substituted IL6 SPCS with PMEL SPCS	NEBuilder® HiFi DNA Assembly
PMELHBB	NlucHBB	Substituted Nluc CDS with PMEL CDS	NEBuilder® HiFi DNA Assembly
PMELaltSPCSHBB	PMELHBB	Substituted PMEL SPCS with IL6 SPCS	NEBuilder® HiFi DNA Assembly
CytoPMELHBB	NlucHBB	Substituted Nluc CDS with a modified PMEL CDS where the endogenous SPCS is removed.	NEBuilder® HiFi DNA Assembly
PMEL_mutant_HBB	PMEL_HBB	Altered 6 nucleotides in 3' end of coding sequence so encoded protein has LL>AA mutation and is predicted to be retained at cell surface as observed by Theos et al. (2006).	NEBuilder® HiFi DNA Assembly
CleanCap_NlucSec_ HBB	NlucSec_HBB	Replaced 3 guanines downstream of T7 promoter with AGG to facilitate the use of CleanCap AG co-transcriptional capping.	NEBuilder® HiFi DNA Assembly
CleanCap_NlucSec_ Opt_HBB	CC_NlucSec_HBB	Substituted nucleotides in 5' UTR of HBB for the established optimal Kozak sequence for improved translational efficiency.	NEBuilder® HiFi DNA Assembly
CC_mCher_Opt_ HBB	CC_NlucSec_Opt_HBB	Substituted the NlucSec CDS from CC_NlucSec_Opt_Hbb for mCher CDS.	NEBuilder® HiFi DNA Assembly
CC_eGFP_Opt_ HBB	CC_NlucSec_Opt_HBB	Substituted the NlucSec CDS from CC_NlucSec_Opt_Hbb for eGFP CDS.	NEBuilder® HiFi DNA Assembly

2.2. Molecular Biology Techniques

2.2.1. GeneBlock Design

The NlucHBB and NlucSecHBB sequences were designed using the CDS derived from standard Promega Nluc sequences. The haemoglobin subunit β (HBB) UTRs were selected for their therapeutic relevance, obtained from the NCBI transcript database (NM_000518.5) (Marotta et al., 1977; Proudfoot, 1977) and cross-referenced against sequences utilised by Wayment-Steele et al. (2021). Restriction enzyme recognition sites were incorporated into the Nluc GeneBlocks at the 5' end corresponding to HindIII and at the 3' end corresponding to EcoRI, with a random four-nucleotide leader sequence added to both ends to enhance cutting efficiency.

For the PMEL GeneBlocks, sequences were sourced from the NCBI transcript database (NM_006928.5) (Kwon et al., 1991; Maresh et al., 1994) with a minor modification at the 5' end of the CDS, directly downstream of the signal peptide cleavage site. This alteration changed the amino acid sequence in the final peptide at the N-terminus from KVP to KGP to enhance immunogenicity following van Stipdonk et al. (2009). Restriction sites were not added to the ends, as NEBuilder HiFi DNA Assembly was employed for cloning these constructs.

2.2.2. Restriction Digests

A total of 300 ng of GeneBlock, as listed in Supplementary or plasmid DNA (pDNA), as listed in Table 2, was added to 10 X CutSmart buffer (NEB) and digested with the restriction enzymes specified in Table 4. The digestion reaction was incubated in a water bath at 37 °C for 2 hours. For applications involving classical restriction enzyme cloning, restriction enzymes were heat-inactivated at 80 °C for 10 minutes. Following heat-inactivation, the restriction digests were separated by agarose gel electrophoresis, and the appropriate bands were extracted and purified from the gel as described in 2.2.4.

For *in vitro* transcription (IVT) of FlucHCV RNA and NlucHCV, the linearised pDNA was isolated using phenol-chloroform extraction followed by sodium

acetate precipitation, as outlined in 2.2.3. For construct verification, the digested samples were analysed by gel electrophoresis (2.2.4) without further purification.

Table 4. Restriction enzymes utilised to digest plasmids.

Plasmid	Amount	Enzymes	Task
APEX2-ERM	300 ng	StuI (NEB, R01875), SpeI-HF (NEB, R0133), EcoRI-HF (NEB, R31015)	Diagnostic digest
APEX2-NES	300 ng	XhoI (NEB, RO1465), Sac II (NEB, RO1575)	Diagnostic digest.
PsPAX2	300 ng	SpeI (NEB, RO133)	Diagnostic digest.
PMD2.G	300 ng	NCO1-HF (NEB, R31935)	Diagnostic digest.
FlucHCV (Roberts et al., 2011)	5 µg	EcoRI-HF (NEB, R3101S)	Linearisation to generate template for IVT.
NlucHCV	2 µg	EcoRI-HF (NEB, R3101S), HindIII (NEB, R0104S)	Digestion to prepare vector backbone for restriction enzyme cloning.
NlucSecHCV	5 µg	EcoRI-HF (NEB, R3101S)	Linearisation to generate template for IVT.

2.2.3. Phenol chloroform DNA extraction

To purify the linearised Fluc HCV and NlucSec HCV pDNA template for IVT, the restriction digest was first diluted with NF water to a total volume of 200 µL. An equal volume of phenol (pH 8, Sigma) and chloroform (1:1 ratio) was then added to the mixture. The solution was vigorously shaken for 1 minute and centrifuged at 17,000 x *g* for 5 minutes. The aqueous phase was carefully transferred to a new Eppendorf tube, and an equal volume of chloroform was added. After another round of vigorous shaking for 1 minute, the mixture was centrifuged at 17,000 x *g* for 5 minutes. The aqueous phase was transferred to a new tube, followed by the addition of 500 µL of 100% ethanol and 20 µL of 3M sodium acetate. This mixture was shaken vigorously and stored at -20 °C overnight. The next day, the samples were centrifuged at 17,000 x *g* for 15 minutes. The supernatant was then discarded, and the pellet was washed twice with 80% ice-cold ethanol, each followed by centrifugation at 17,000 x *g* for 5 minutes. Finally, the pellet was resuspended in 20 µL of nuclease-free water.

2.2.4. Agarose Gel Electrophoresis

1% agarose gels were prepared by dissolving 0.4 g of agarose (Sigma) in 40 mL of 1 X Tris/Borate/EDTA (TBE) buffer (0.13 M Tris, pH 7.6, 45 mM boric acid, 2.5 mM EDTA) by heating the solution in a microwave. Once the solution cooled to approximately 50 °C, 4 µL of SYBR Safe DNA gel stain (Invitrogen) was added and the mixture was poured into a gel casting tray. The gel was allowed to set for 20 minutes at room temperature (RT).

For DNA gels, samples containing pDNA, PCR products, or restriction digests were mixed with 6X DNA loading dye (NEB). GeneRuler 100 bp or 1 kb DNA ladder (Thermo Fisher Scientific) and the DNA samples were loaded into the wells of the gel. The gel was run in 1 X TBE buffer at 100 V for 1 to 1.5 hours, depending on the visible migration of the loading dye.

For RNA gels, 2 µL of *in vitro* transcribed mRNA (IVT mRNA), prepared as outlined in 2.2.6, was taken from samples both before and after enzymatic polyadenylation. Each RNA sample was mixed with an equal volume of RNA Gel Loading Buffer II (Invitrogen), which contained 95% formamide, 0.025% xylene cyanol, 0.025% bromophenol blue, 18 mM EDTA, and 0.025% SDS. The RNA samples in loading buffer were heated at 70 °C for 5 minutes, immediately placed on ice for 5 minutes, and then loaded into the wells of the 1% agarose gel. Electrophoresis for RNA samples was conducted under the same conditions as for DNA gels (in 1X TBE buffer at 100 V).

DNA and RNA bands were visualised and imaged using the GelDoc Go Gel Imaging System (Bio-Rad). For DNA band extraction, bands were visualised using blue light, and desired bands were excised using a clean scalpel and forceps. Excess agarose was trimmed, and DNA was extracted using the Wizard PCR Clean-Up Kit (Promega), which was used according to the manufacturer's instructions. The DNA was eluted in 20 µL of nuclease-free water and quantified using a Nanodrop OneC spectrophotometer.

2.2.5. PCR

Inserts for NEBuilder HiFi assembly cloning were generated by PCR amplification from pDNA or GeneBlocks, as listed in Supplementary Table 1 using primers detailed in Supplementary Table 2. Templates for IVT were produced by PCR using primers listed in Supplementary Table 3.

Amplification was carried out using 1 ng of plasmid or GeneBlock DNA in a master mix containing 1 µL of Phusion High-Fidelity DNA Polymerase (NEB), 10 µL of 1 X High-Fidelity Buffer, deoxynucleotide triphosphates (dNTPs) at a final concentration of 200 µM, and primers at a final concentration of 10 µM each. The final reaction volume was adjusted to 50 µL with nuclease-free water. PCR cycling was performed using standard Phusion conditions for 35 cycles with annealing temperatures as listed for individual primers in Supplementary Table 2.

2.2.6. *In vitro* transcription

In vitro transcription (IVT) was performed according to the manufacturer's instructions, using either 1 µg of linearised plasmid, prepared as described in 2.2.2 and purified as in 2.2.3 or 200 ng of PCR product per 20 µl reaction. The PCR product was generated using primers listed in Supplementary Table 3, as detailed in 2.2.5 and purified as outlined in 2.2.4. IVT was conducted using the T7 Megascript Kit for non-capped RNA for spike-in purposes and non-capped mRNA with HCV UTRs for transfection, the T7 mMessage mMachine Kit for ARCA-capped mRNA, or the HiScribe T7 CleanCapAG Kit (NEB) for CleanCapped mRNA. For the synthesis of IVT mRNA containing S^6GTP, the HiScribe T7 CleanCapAG IVT Kit (NEB) was used, with GTP partially replaced by S^6GTP (Jena GmbH) at ratios of 10 % S^6GTP to canonical GTP and 25 % S^6GTP to canonical GTP, achieving a final combined concentration of 5 mM.

Following 2-4 hours of transcription at 37 °C in a PCR machine (Alpha Laboratories), the newly synthesised RNA was treated with Turbo DNase (Thermo Fisher Scientific) at 37 °C for 15 minutes. For RNA intended for cell transfection, enzymatic A-tailing was performed using the Poly(A) Tailing Kit

(Invitrogen) at 37 °C for 45 minutes. All IVT RNA, regardless of whether it underwent enzymatic A-tailing, was subsequently purified with the RNA Clean & Concentrator-25 Kit (Zymo Research), reconstituted in 100 µL of RNase-free water, quantified, aliquoted, and stored at -80 °C for long-term preservation.

2.2.7. Ligations

The HbbNluc and HbbNlucSec GeneBlocks, listed in Supplementary Table 1, were digested with restriction enzymes as described in 2.2.2, followed by gel electrophoresis as in 2.2.4. The inserts were then ligated into the digested vector backbone derived from the NlucHCV plasmid, detailed in Table 2, using T4 DNA Ligase (NEB). A 3:1 molar ratio of insert to vector was used for the ligation reactions. Each ligation reaction was prepared in a final volume of 20 µL, consisting of 1 µL of T4 DNA Ligase (NEB), 2 µL of 10 X T4 Ligase Buffer, 50 ng of digested vector, 150 ng of digested insert DNA, and nuclease-free water to reach the final volume. Samples were incubated in a thermocycler at 16 °C for 16 hours.

2.2.8. NEBuilder HiFi DNA Assembly

NEBuilder HiFi DNA assembly was employed to generate plasmids, as detailed in Table 3, suitable for IVT. Gibson primers, as listed in Supplementary Table 2, were designed using the NEBuilder Assembly Tool (https://nebuilder.neb.com/#!/), with the vector backbone insertion sequence included. The resulting primers were ordered from Integrated DNA Technologies (IDT).

The Gibson primer pairs were used in PCR reactions (2.2.5), PCR products were run on an agarose gel as in 2.2.4 and the appropriate bands were extracted and cleaned up.

The NEBuilder HiFi DNA Assembly Cloning kit (NEB) was used for the assembly process. For optimal cloning efficiency, 50 ng of vector was combined with inserts at a 2:1 vector-to-insert molar ratio. For 4-6 fragment assemblies, a 3:1 molar ratio of each insert to the vector was used. The mixture included 10 µL of Gibson Assembly Master Mix and nuclease-free water to bring the total volume

to 20 µL. Samples were incubated in a thermocycler at 50 °C for 50 minutes. After incubation, samples were stored at –20 °C until transformation. 2 µL of the assembly reaction was used for the transformation as described in 2.2.9.

2.2.9. Transformations

Plasmids were transformed *via* heat shock into 50 µL of chemically competent DH5α™ cells (Invitrogen, Subcloning Efficiency™). Competent cells were defrosted on ice before either adding 1 ng of purified plasmid or 5 µL of ligation mixture. Alternatively, following use of NEBuilder HiFi Assembly cloning mix, 2 µL of the HiFi assembly ligation mix was added to NEB 5α Competent E. coli cells provided in the NEBuilder HiFi Assembly cloning kit. Following the addition of pDNA or ligation mixture, the cell mixture was incubated on ice for at least 30 minutes.

Cells were then heat-shocked by placing them in a 42 °C water bath for 60 seconds, followed by immediate transfer to ice for 2 minutes. For recovery after the heat-shock, 200 µL of SOC recovery media (Table 5) was added to the cells, and the mixture was incubated at 37 °C on a shaking rotator for 1 hour. Subsequently, the cells were streaked onto pre-warmed selective LB agar plates (Table 5) containing ampicillin (100 µg/mL) and incubated for 16 hours at 37 °C.

Table 5. Buffer compositions for growth of bacteria.

Buffer	Composition
Luria-Bertani (LB) medium	1 % w/v tryptone, 0.5 % w/v yeast extract and 1% w/v NaCl in diH2O, pH 7.0. For solid medium, 2 % w/v bacteriological agar was added. Medium was autoclaved and stored at room temperature until required.
SOC medium	2 % tryptone, 0.5 % yeast extract, 10 mM NaCl, 2.5 mM KCl, 10 mM MgCl2, 10 mM MgSO4, and 20 mM glucose. Purchased from NEB and stored at 4 °C
Ampicillin	1000 X stock solution made as 100 mg/mL in sterile 50 % diH2O and 50 % EtOH to prevent freezing and stored at -20 °C.
Selective LB plates	Ampicillin was added to melted LB agar medium that was starting to cool to make a final concentration of 1 X and poured onto Petri dishes in a laminar flow hood. Plates were stored at 4 °C for up to 1 month

2.2.10. Minipreps

Single bacterial colonies were isolated from a selective ampicillin LB agar plate (Table 5) using a sterile pipette tip and transferred to 10 mL of LB broth with 0.1 % (v/v) ampicillin. The cultures were inoculated for 16 hours at 37 °C in a shaking incubator. Cultures were then transferred to 15 mL Falcon tubes and centrifuged at 4000 x *g* for 10 minutes at 4 °C. The supernatant was discarded, and pDNA was isolated from the bacterial pellet using the Wizard Plus SV DNA Purification System by Promega, following the manufacturer's protocol, with the exception of eluting in 25 µL of nuclease-free water instead of 100 µL. DNA concentration was measured using a Nanodrop OneC spectrophotometer (ThermoFisher Scientific). Minipreps for classical restriction enzyme cloning were assessed by restriction digestion (2.2.2), followed by gel electrophoresis (2.2.4) to confirm that the bands were of the correct size. The validated samples were then sent for sequencing, as described in 2.2.12. Alternatively, new constructs generated using NEBuilder HiFi Assembly cloning were sent directly for sequencing (2.2.12).

2.2.11. Maxipreps

A single bacterial colony was selected and inoculated into 10 mL of LB broth containing 0.1% (v/v) ampicillin, then grown overnight at 37 °C with shaking. From this overnight culture, 500 µL was transferred to 300 mL of LB broth with 0.1% (v/v) ampicillin and incubated for an additional 24 hours at 37 °C with shaking. Following incubation, the culture was divided into six 50 mL Falcon tubes and centrifuged at 4000 x *g* for 10 minutes at 4 °C. The supernatant was discarded, and the pellets were recombined in buffer P1 during the resuspension step. pDNA was then isolated using the Qiagen Plasmid Plus Maxi Kit according to the manufacturer's instructions. The isolated pDNA was resuspended in 100 µL of nuclease-free water, quantified using a Nanodrop OneC, and stored at -20 °C.

2.2.12. Sequencing of plasmids

5 µL of pDNA, at a concentration of 100 ng/µL, and 5 µL of primer, at a concentration of 3.2 pmol/µL, were sent to Source Bioscience Ltd for Sanger Sequencing using primers listed in Supplementary Table 4.

2.3. Cell Culture

2.3.1 Cell maintenance

HEK293T cells (supplied by ATCC) and A549 cells (supplied by ATCC) were maintained at 37 °C with 5 % CO2 in a T75 flask containing Dulbecco's Modified Eagle Medium (Sigma, cat no 6429) supplemented with 10 % foetal bovine serum (FBS). HEK293T cells were passaged by washing once with PBS (Table 1) and then incubating in trypsin (Table 1) for 2 minutes at room temperature. In contrast, A549 cells were incubated with trypsin (Table 1) for 5 to 7 minutes at 37 °C with 5 % CO2. After trypsinisation, the volume of fresh DMEM added and the volume of cells discarded were adjusted to achieve the desired seeding confluency. Cells were routinely passaged at 70 % confluency, with a 1:12 split ratio, on Mondays, Wednesdays, and Fridays.

2.3.2 Freezing cells

Cells were first seeded in 10 cm plates 24 hours before freezing and grown to an approximate confluency of 70 %. The following day, cells were trypsinised and resuspended in 10 mL of media. Cells were pelleted by centrifugation for 5 min at 200 x *g* before resuspension in 1 mL of 9:1 FBS:DMSO (PanReac AppliChem). The resuspended cells were then transferred to 1 mL cryovials and the vials then transferred temporarily to a CoolCell® cell freezing container (Corning) at -80 °C for at least 24 hours before long term storage in liquid nitrogen.

2.3.3 Thawing cells

Cells were rapidly thawed in a pre-heated 37 °C water bath following removal from liquid nitrogen tanks and then re-suspended in 10 mL of pre-warmed media. To remove the DMSO, cells were pelleted by centrifugation at 200 x *g* for 5 min at room temperature and resuspended in media before adding the cell suspension to T75 flasks.

2.3.4 Blasticidin kill curve

Wild-type (WT) cells were seeded into 6-well plates 24 hours before blasticidin treatment and grown to an approximate confluency of 40 %. The following day, the media was removed and replaced with fresh media containing blasticidin. Concentrations of blasticidin spanned 0 to 10 µg/mL and increased in increments of 2 µg/mL. Cells were observed for confluence and signs of death (rounded or floating cells) using a light microscope for 7 days, with media containing the same amount of blasticidin replaced on day 4. The concentration of blasticidin chosen for the selection of APEX-transduced cells was established by determining the lowest concentration of blasticidin able to cause 100 % wild-type cell death by days 3 to 5.

2.3.5 Lentivirus production

WT HEK293T and WT A549 cells were seeded in 10 cm plates 24 hours before plasmid transfection and grown to approximately 40 % confluency. The following day, 4.6 µg of either APEX2-ERM plasmid or APEX2-NES plasmid, along with 3.2 µg of psPAX2 plasmid and 1.2 µg of pMD2G plasmid, were mixed in 500 µL of Opti-MEM (Gibco). To this solution, 25 µL of Fugene HD Transfection Reagent (Promega) was added, and the mixture was gently vortexed before incubating at room temperature 15 minutes. The resulting transfection mixture was then applied to the cells and left to incubate for 48 hours. After incubation, the media from the transfected cells was filtered through a 0.45 µm Millex HV filter (Millipore) to isolate the lentiviral stock, which was then stored at -80 °C prior to lentiviral transduction.

2.3.6 Lentivirus transduction

Lentiviral transduction was employed to establish stable HEK293T and A549 cell lines expressing ER-localised APEX2 (APEX2-ERM) and cytoplasmic APEX2 (APEX2-NES). For APEX2-ERM transduction, cells were seeded in 6 cm plates 24 hours before transduction and grown to approximately 50 % confluency. The following day thawed lentiviral stock was added to the cells, and Polybrene (Millipore) was included at a final concentration of 8 µg/mL to enhance

transduction efficiency. 24 hours after lentivirus application, the infected media was replaced with fresh DMEM containing 10 % FBS. Cells expressing APEX2 were selected by adding 6 µg/mL of blasticidin (Sigma) to the medium and monitoring cell survival over the subsequent week, as in 2.3.4. Cells were treated with blasticidin for 1.5 weeks, and the surviving cells were maintained as a mixed population. 6 µg/mL of blasticidin was routinely added to APEX2 expressing cell lines every 1-2 weeks, thus preventing overgrowth by cells that had stopped expressing these genes.

For APEX2-NES transduction, the process was similar, except that WT HEK293T and A549 cells were seeded in 6-well plates. The lentiviral stock was diluted in Opti-MEM (Gibco) to final dilutions of 1:2, 1:4, 1:10, and 1:200 in a total volume of 2 mL before being added to cells. Following transduction, the infected media was replaced with fresh DMEM containing 10 % FBS, and selection was performed with 6 µg/mL of blasticidin, as described for APEX2-ERM.

2.3.7 APEX2-mediated labelling

The APEX2 labelling protocol was carried out as in Fazal et al. (2019). APEX2 expressing cells were seeded into 6-well plates or 10 cm-plates 24 hours before biotinylation and grown to 80 % confluency. The following day, cells were treated with 500 µM Biotin Aniline (Iris Biotech, Cat no. L53970) or Biotin Phenol (Sigma Aldrich, Cat no.= SML2135). Cells were then incubated for 30 min at 37 °C 5% CO_2. 1 mM H_2O_2, freshly diluted in NF water from 30 % H_2O_2 (v/v) (Sigma), was added to the pre-existing media in each well and incubated for 1 min at room temperature with gentle swirling throughout. To quench the biotinylation reaction, the media was aspirated and replaced with a solution containing 5 mM Trolox (Sigma Aldrich), 10 mM sodium ascorbate (Sigma Aldrich) and 10 mM sodium azide (Sigma Aldrich) in NF PBS and incubated for 1 minute. The quenching solution was then replaced, and the cells were washed 3 times with an azide-free solution consisting of 10 mM sodium ascorbate and 5 mM Trolox in PBS. For RNA work, total RNA was extracted using TriReagent and processed as described in 2.4.5. This RNA is referred to as the “input RNA”. For SIM imaging, cells were processed as described in 2.5.6.

2.3.8 Membrane-cytoplasm biochemical fractionation

The biochemical membrane fractionation protocol was carried out using an adaptation of the protocol used by Child et al. 2021.

HEK293T APEX2-ERM cells were seeded into 10 cm plates 24 hours before fractionation and grown to an approximate confluency of 80 %. 4 hours post-transfection of IVT mRNA, as detailed in 2.4.1, media was aspirated, and cells were scraped into 5 mL of chilled PBS. The cell suspension was then centrifuged at 1500 x *g* for 3 minutes at 4 °C, and the supernatant was discarded. The resulting pellet was resuspended in 1 mL of chilled PBS.

200 µL of this suspension was set aside for both total protein and total RNA extraction. For total protein extraction, 100 µL of the sample was pelleted at 15000 x *g* for 3 minutes at 4 °C, supernatant was aspirated, the pellet was then resuspended in 32 µL of 1X SDS loading buffer and protein extracted as described in 2.5.2. For total RNA extraction, the 100 µL of the sample was pelleted at 1500 x *g* for 3 minutes at 4 °C. The supernatant was discarded, and the pellet was mixed with 400 µL of TRI reagent and kept on ice.

The remaining cell pellet was resuspended in 200 µL of 1st step lysis buffer (Table 6) and pipetted gently 25 times with a 200 µL tip. After a 10-minute incubation on ice, the suspension was centrifuged at 2000 x *g* for 5 minutes at 4 °C. The supernatant was collected as the cytosolic fraction and split for protein and RNA analysis. For protein extraction, 70 µL of the cytosolic fraction was mixed with 18 µL of 4X SDS loading buffer and kept on ice. For RNA extraction, 130 µL of the cytosolic fraction was combined with 400 µL of TRIzol LS and kept on ice.

The cell pellet was washed with 100 µL of 2nd step lysis buffer (Table 6), pipetted 10 times, and immediately centrifuged at 2000 x *g* for 5 minutes at 4 °C. The resulting lysate, referred to as the “cytosolic wash”, was set aside for protein extraction by taking 70 µL and adding it to 18 µL of 4X SDS loading buffer.

For ER fraction collection, the pellet was resuspended in 200 µL of 3rd step lysis buffer (Table 6) and pipetted 10 times, followed by a 10-minute incubation on

ice. After centrifugation at 2400 x *g* for 5 minutes at 4 °C, the supernatant was collected as the ER fraction. For protein analysis, 70 µL of the ER fraction was combined with 18 µL of 4X SDS loading buffer, while for RNA analysis, 130 µL of the ER fraction was mixed with 400 µL of TRIzol LS.

The final pellet, containing nuclear material, was resuspended in 250 µL of 1X SDS loading buffer and protein then extracted as described in 2.5.2.

0.1 ng of elution spike-in eGFP RNA, generated as described in section 2.2.6, was added to all RNA fractions in TRIzol LS and TRI-reagent prior to RNA extraction, as outlined in section 2.4.5.

Table 6. Buffers used in biochemical membrane fractionation and their compositions.

Buffer	Composition	Storage
1 X lysis buffer	10 mM Tris pH 7.4 75 mM NaCl 2.5 mM $MgCl_2$ 1 cOmplete mini EDTA free protease inhibitor cocktail tablet (Roche) added per 10 mL. 1 mM DTT	Stored as 5 X lysis buffer (Tris, NaCl, $MgCl_2$) at -20 °C. cOmplete mini tablet (Roche) and DTT added on the day of the experiment.
First-Step Lysis buffer	0.0015 % Digitonin (w/v) in 1 X lysis buffer, supplemented with 0.002 % (v/v) Ribolock RNAse inhibitor (Thermo Fisher Scientific)	Made on the day of the experiment.
Second-Step Lysis buffer	0.004 % Digitonin (w/v) in 1 X Lysis Buffer, supplemented with 0.002 % (v/v) Ribolock RNAse inhibitor (Thermo Fisher Scientific)	Made on the day of the experiment.
Third-Step Lysis buffer	2 % n-Dodecyl B-D-maltoside (DDM) in 1 X lysis buffer, supplemented with 0.002 % (v/v) Ribolock RNAse inhibitor (Thermo Fisher Scientific)	Made on the day of the experiment.

2.4. RNA Techniques

2.4.1 Transfection of mRNA

Twenty-four hours before transfection, cells were seeded into 6-well plates or 10 cm plates and grown to an approximate confluency of 80 %. A master mix was prepared for each mRNA to be transfected. For each well of a 6-well plate, 0.75 µg of IVT mRNA was combined with 3.8 µL of Lipofectamine 2000 (Invitrogen) in a final volume of 250 µL of OptiMEM (Invitrogen). For each 10 cm plate, 3.75 µg of IVT mRNA was mixed with 22 µL of Lipofectamine 2000 in a final volume of 750 µL of OptiMEM.

Following the manufacturer's instructions, the RNA was first added to OptiMEM and Lipofectamine 2000 was then added to a separate OptiMEM solution and incubated for 5 minutes. The RNA solution was then added to the Lipofectamine solution, and the mixture was incubated for 20 minutes to allow for complex formation. The resulting transfection mixture was added dropwise to the cells to ensure even distribution and minimise disturbance of the nanoparticles. The cells were incubated with the transfection media for up to 8 hours to facilitate mRNA uptake. 4 hours post-transfection, cells were typically used for APEX2 biotinylation experiments as described in 2.3.7, or, for biochemical membrane fractionation, as described in 2.3.8. If reporter mRNA encoding fluorescent proteins was transfected, the cells were imaged as detailed in 2.5.8.

2.4.2 Between Body and Tail labelling of *in vitro* transcribed mRNA

The between body and tail (BBT) labelling approach used was adapted from Anhäuser et al. (2019). NlucSec HBB IVT RNA, generated using the T7 mMessage mMachine kit as described in 2.2.6, was purified using Zymo RNA Clean & Concentrator 25 columns before polyadenylation. To incorporate 2-azido dATP into the region of the poly(A) tail closest to the 3' UTR, 1 µg of RNA was treated with yeast poly(A) polymerase. The reaction mixture, prepared with the yeast poly(A) polymerase kit (Invitrogen), included
5 µL of 5 X yPAP buffer, 0.2 µM IVT RNA, 2-azido dATP (Jena) at a final concentration of 1 mM, and 0.5 µL of yeast poly(A) polymerase, adjusted to a

final volume of 25 µL with nuclease-free water. The mixture was incubated for 15 minutes with the aim of achieving limited polyadenylation. Following incubation, the RNA was purified again using the RNA Clean & Concentrator Kit 25 (Zymo Research) and eluted in 20 µL of nuclease-free water. The RNA was then poly(A)-tailed using the Poly(A) Tailing Kit (Invitrogen) as detailed in 2.2.6, purified with the RNA Clean & Concentrator-25 Kit (Zymo Research), and reconstituted in 20 µL of RNase-free water. The RNA was quantified, aliquoted, and stored at -80 °C for long-term preservation.

To label the 2-azido dATP with fluorophores, strain-promoted azide-alkyne cycloaddition was performed by incubating 2-azido dATP-labelled mRNA with fluorophores conjugated to Dibenzylcyclooctyne (DBCO). A 20 µL sample of the 2-azido dATP-labelled mRNA was combined with Aurora Fluor 488 DBCO (Biosynth Laboratories) to a final concentration of 100 nM and 50µM respectively. Additionally, 5 µL of 10 X nuclease-free PBS (Invitrogen) was added to achieve a final concentration of 1 X PBS, and the total volume was adjusted to 50 µL with nuclease-free water. The reaction was incubated at 37 °C for 1.5 hours to allow the click chemistry reaction to proceed. After incubation, the RNA was purified once more using the RNA Clean & Concentrator Kit 25 (Zymo Research) and eluted in 20 µL of nuclease-free water.

2.4.3 pCp-biotin labelling of *in vitro* transcribed RNA

To generate a 3' end biotinylated spike in RNA, a ligation technique using ssRNA ligase was adapted from the method described by Kore et al. (2009). This involved pCp biotin labelling of an IVT FlucHCV RNA that was non-capped and non-polyadenylated, generated as described in 2.2.6.

The labelling reaction was prepared by combining the following components in a final volume of 30 µL: 10 % T4 RNA Ligase Reaction Buffer (NEB), 10 % dimethyl sulfoxide (DMSO), and 15 % polyethylene glycol 8000 (NEB). The mixture also contained 1 mM ATP, 0.1 mM pCp-Biotin (Jena GmBh), 1 µg of IVT RNA, and 10 units of T4 RNA Ligase 1 (NEB). After thoroughly mixing these components at room temperature, the reaction was incubated overnight at 16

°C to allow for effective ligation. Following the incubation, the RNA was purified using a Zymo Clean & Concentrator-25 column to remove any unreacted components and isolate the biotinylated RNA.

2.4.4 *In vitro* biotinylation of S^6GTP *in vitro* transcribed mRNA

To validate the incorporation of S^6GTP by T7 polymerase, 1 µg of IVT mRNA containing S^6GTP, generated as described in 2.2.6, was biotinylated following Downie Ruiz Velasco et al. (2024). A negative control without S^6GTP was performed for comparison. Biotinylation was carried out by mixing the RNA with 25 µL of biotinylation buffer (10 X stock: 100 mM HEPES, pH 7.5, and 10 mM EDTA) and 50 µL of MTSEA-XX-Biotin (Cambridge Biotech) at a concentration of 100 µg/mL in dimethylformamide. The final volume was adjusted to 250 µL with nuclease-free water. The reaction was incubated in the dark at room temperature for 30 minutes on a rotator to ensure thorough mixing and effective biotinylation. Post-reaction, RNA was purified using chloroform extraction. An equal volume of chloroform was added to the mixture, which was shaken vigorously and then centrifuged at 17,000 x *g* for 15 minutes at 4 °C. The aqueous phase was decanted, and 500 µL of isopropanol was added to precipitate the RNA. The sample was incubated overnight at -20 °C, then centrifuged at 17,000 x *g* for 15 minutes at 4 °C. The RNA pellet was washed twice with ice-cold 75 % ethanol, then re-suspended in 20 µL of nuclease-free water. The purified RNA was used for further analyses, including dot blot assays to visualise biotinylation (2.4.7) or streptavidin pulldown RT-qPCR (2.4.8.2) for quantitative assessment of biotinylation.

2.4.5 RNA isolation using TRI reagent

Total RNA for mRNA analysis by RT-qPCR and input RNA for pulldown experiments was extracted using TRI-reagent (Sigma). RNA was extracted from HEK293T and A549 cells grown in 6-well or 10 cm plates. The medium was removed upon harvesting, and the wells were washed twice with PBS. For 6-well plates, 1 ml of TRI-reagent was added, while 2 ml and 3 ml were used for 6 cm and 10 cm plates, respectively, at room temperature. The TRI-reagent was mixed

by repeated pipetting to ensure complete cell lysis followed by a 10-minute incubation at room temperature. Trizol-LS was used at a 3:1 ratio of Trizol-LS to the RNA-containing solution for samples in solution, such as those from fractionation or post-streptavidin pulldown. The TRI-reagent solution was transferred to microfuge tubes, and 200 µl of chloroform was added per 1 ml of TRI-reagent. The tubes were shaken vigorously for 30 seconds and incubated for 5 minutes at room temperature. Samples were then centrifuged at 17,000 x *g* for 15 minutes at room temperature. The aqueous layer was carefully transferred to new microfuge tubes, and 500 µl of filtered isopropanol (Sigma) along with 1 µl of co-precipitant GlycoBlue (Invitrogen) was added. The tubes were inverted to mix and left to precipitate overnight at -20 °C. The following day, the samples were centrifuged at 17,000 x *g* for 15 minutes at 4 °C. The supernatant was discarded, and the RNA pellet was washed twice with 500 µl of 75 % chilled ethanol. After the final centrifugation at 17,000 x *g* for 5 minutes at 4 °C, the supernatant was removed, and the pellet was allowed to air dry for 1 minute before being resuspended in 15-40 µL of nuclease-free water.

2.4.6 Reverse Transcription Quantitative Polymerase Chain Reaction

Reverse transcription was performed using the GoScript Reverse Transcription System (Promega, cat. no. A2800) with random primers, in accordance with the manufacturer's instructions. For this process, 100 ng of RNA was used as the template. In APEX2 experiments, either 100 ng of input RNA or 8 µL of the 16 µL output pulldown RNA was utilised. The input RNA for the streptavidin pulldown was normalised by mass, while the output RNA used for reverse transcription was normalised by volume. For post-fractionation experiments, reverse transcription was carried out using 1 µL of RNA re-suspended in 30 µL of nuclease-free water. Quantitative PCR (qPCR) was performed in technical triplicates with a total reaction volume of 10 µL. Each qPCR reaction consisted of 0.5 µL of cDNA, GoTaq mastermix at a final concentration of 1X, primers (listed in Supplementary Table 5) at a final concentration of 0.25 µM and nuclease-free water (Promega) to a total of 10 µL The reactions were conducted

on a Rotor-Gene Q instrument (Qiagen), following the cycling conditions specified in Table 7. Each sample was run in triplicate, and primer efficiency was assessed through serial dilution.

Table 7. qPCR Cycling Parameters

Step	Number of cycles	Temperature	Time
Enzyme activation	1	95 °C	2 min
Denaturation	40	95 °C	10 secs
Annealing		56 °C	1 min

2.4.7 Streptavidin-HRP RNA dot blot assay

To confirm the biotinylation of total RNA from APEX labelling experiments, 3' end biotinylated spike-in RNA, and IVT mRNA containing S^6GTP were analysed using streptavidin HRP dot blot assays, following the protocol optimised by Makino et al. (2021). 2 µL of APEX2-mediated biotinylated RNA, as described in 2.3.7, or biotinylated S^6GTP IVT mRNA, as detailed in 2.4.4 was spotted onto an Amersham Hybond membrane (GE Healthcare). For the 3' biotinylated spike-in RNA, prepared as described in 2.4.3, the RNA was serially diluted, and 2 µL of each dilution was spotted onto the membrane. The membrane was allowed to dry, then ultraviolet (UV)-crosslinked using a Stratalinker (Strategene) and incubated with 10 X blocking buffer (125 mM NaCl, 9 mM Na_2HPO_4, 7 mM NaH_2PO_4, and 10 % SDS) for 20 minutes at room temperature. Streptavidin-HRP (Abcam) was diluted in 10 X blocking solution to a final concentration of 1:2500 and the membrane incubated with Streptavidin-HRP for 10 minutes. The membrane was subsequently transferred to 1 X blocking buffer for 10 minutes, followed by two 5-minute washes with wash buffer (10 mM Tris-Base, 10 mM NaCl, 1.05 mM $MgCl_2$, pH 9.5).

2.4.8 Streptavidin Affinity Purification

2.4.8.1. Streptavidin pulldown optimisation using spike-in RNA

During the optimisation of the streptavidin pulldown procedure, several modifications were made to the washing steps prior to bead incubation with RNA, the inclusion of a blocking step before incubating beads with input RNA,

the washing steps post-incubation with RNA, and the methods of elution of biotinylated RNA from streptavidin beads. In doing so, elements of APEX2-labelled RNA streptavidin pulldown protocols from Ting (Fazal et al., 2019), Ingolia (Padrón et al., 2019), and Zou (Li et al., 2022) were tested as detailed in Table 8.

Table 8. Overview of different streptavidin affinity protocols tested.

Protocols / Step	Ting protocol (Fazal et al., 2019)	Ingolia protocol (Padrón et al., 2019)	Zou protocol (Li et al., 2022)
Washing of beads pre-incubation with RNA	Tested – Used in final protocol.	Tested – Not used in final protocol.	Tested – (Same as Ting protocol).
Blocking of beads pre-incubation with RNA	N/A	Tested – Not used in final protocol.	Tested – Used in final protocol.
Washing post-blocking	N/A	Tested – Not used in final protocol.	Tested – Used in final protocol.
Incubation with RNA	Tested – Combined with Zou protocol in final protocol (2 hour end-over-end rotation at 4 °C).	Tested – Not used in final protocol.	Tested – Combined with Ting protocol in final protocol (Binding solution).
Washing post-incubation with RNA	Tested – Combined with Ingolia and Zou in final protocol (3 X wash with B&W buffer)	Tested – Combined all 3 in final protocol (use of 0.1 % SDS in PBS).	Tested – Same as Ting protocol.
Elution of biotinylated RNA from beads	Tested – Used in final protocol.	Tested – Not used in final protocol.	Not tested.

2.4.8.2. Optimised Streptavidin Affinity Purification Protocol

The streptavidin pulldown procedure used 0.5 ng of 3' end biotinylated spike-in RNA, as described in 2.4.3. Additionally, 0.1 ng of unlabelled Nluc RNA was used as a second spike-in and was added before RNA extraction to account for variations in RNA extraction efficiency and normalisation of RT-qPCR data. The final optimised protocol was subsequently applied for the isolation of APEX2-mediated biotinylated RNA from total RNA and to validate S^6GTP incorporation by T7 polymerase during IVT post *in vitro* biotinylation (2.4.4).

Pierce Streptavidin Magnetic beads (Thermo Fisher Scientific) were washed 3 times with Bead Washing buffer (Table 9), washed twice in a solution consisting of 0.1 M NaOH (Sigma) and 0.05 M NaCl (Invitrogen) and then washed once in a solution of 0.1 M NaCl (Invitrogen). The beads were then incubated in blocking buffer (Table 9) on a rotator, for end-over-end rotation, for 2 hours at room temperature. Each 10 µL of streptavidin beads was incubated in a final volume of 200 µL of blocking buffer. After the two hours of block, the beads were washed three times with 4 M NaCl Washing buffer (Table 9) and then washed three times in 1 X binding buffer (Table 9) before being resuspended in 100 µL of 2 X binding buffer. At this point 0.1 ng of biotinylated spike in RNA and 500 ng of the total (input) RNA was added to 100 µL of beads in 2 X binding buffer and diluted to a total volume of 200 µL to reach a concentration of 1 X binding buffer. Alternatively, when validating biotinylation of S^6GTP IVT mRNA, RNA was diluted 1 in 30 and then 1 µL was added to the beads and binding buffer. The streptavidin magnetic beads were then incubated with the diluted RNA on a rotator, for end-over-end rotation, for 2 hours at 4 °C

Following incubation, beads were placed on a magnet and the supernatant was discarded. The streptavidin magnetic beads were washed once with Bead Washing buffer, twice with 0.1 % SDS in PBS and then twice more with Bead Washing buffer before being resuspended in 54 µL nuclease free water. 33 µL of 3 X proteinase digestion buffer (Table 9), 10 µL of Proteinase K (Thermo Fisher Scientific) and 3 µL of RNasin ribonuclease inhibitor (Promega) were added to the solution and incubated at 42 °C for 1 hour, following the method of biotinylated RNA extraction from streptavidin magnetic beads described by Fazal et al. (2019). The beads were then heated to a 55 °C in a ThermoMixer (Eppendorf) for 1 hour, shaking at 400 RPM throughout. Following the second incubation, 0.1 ng of unlabelled eGFP or Renilla Luciferase spike-in RNA was added to the solution, to allow for normalisation of RNA isolation efficiency in subsequent RT-qPCR analysis. The beads were then placed on a magnet, the supernatant containing the eluted RNA extracted and 25 µL of NF water was added to the supernatant. Next, 375 µL of Trizol LS (Thermo Fisher Scientific)

was added to the solution, RNA was isolated as described in 2.4.5 and then resuspended in 16 µL of NF water. This RNA is referred to as the "output RNA".

Table 9. Buffer compositions for optimised streptavidin affinity purification protocol.

Reagent	Composition	Source
Bead Washing buffer	5 mM Tris-HCl pH = 7.5, 0.5 mM EDTA, 1 M NF NaCl (Invitrogen), 0.1 % TWEEN 20 (Sigma Aldrich)	Zou lab (Li et al., 2022)
Blocking buffer	2 mg/ml Yeast tRNA (Invitrogen), 50 µg/ml Glycogen (Invitrogen)	Zou lab (Li et al., 2022)
4 M NaCl Washing buffer	0.1 M Tris-HCl pH = 7.5, 4 M NF NaCl (Invitrogen), 10 mM EDTA, 0.2 % Tween-20	Zou lab (Li et al., 2022)
1 X Binding buffer	0.1 M Tris-HCl pH = 7.5, 1 M NF NaCl, 10 mM EDTA, 0.2 % Tween-20	Zou lab (Li et al., 2022)
2 X Binding buffer	0.2 M Tris-HCl pH = 7.5, 2 M NF NaCl, 20 mM EDTA, 0.4 % Tween-20	Zou lab (Li et al., 2022)
0.1% SDS in PBS	0.1 % SDS (w/v) in NF PBS (Invitrogen)	Inspired by Ingolia lab (Padrón et al., 2019)
3 X Protein Digestion buffer	30 % v/v 10 X PBS (Invitrogen), 30 % v/v of 20 % (w/v) N-laurylsarcosine sodium (Sigma), 32.5 % (v/v) of NF water, 6 % NF 0.5 M EDTA (Invitrogen), 1.5 % (v/v) 1 M DTT (Invitrogen)	Ting lab (Fazal et al., 2019)

2.5. Protein Techniques

2.5.1. Antibodies and buffers

Table 10. Buffers utilised for western blotting and their compositions.

Buffer	Composition
RIPA	50 mM Tris-HCl, pH 8.0, 150 mM NaCl, 1.0 % NP-40, 0.5 % w/v sodium deoxycholate, 0.1% w/v SDS.
4 X SDS loading dye	0.5 M Tris HCl pH 6.8, 40 % v/v glycerol, 0.8 % w/v SDS, 0.0025 % v/v bromophenol blue, 5 % β-mercaptoethanol.
1 X SDS running buffer	25 mM Tris-HCl pH 8.3, 192 mM glycine, 0.1 % w/v SDS.
Blotting buffer for semi-dry transfer	50 mM Tris-HCl pH 8.3, 192 mM glycine, 20 % v/v methanol.
Blotting buffer for Turbo-Blot transfer	Bio-Rad Propriety.
1 X TBS-T	10 mM Tris HCl, 15 mM NaCl, 0.1 % v/v Tween 20 (Sigma).
5 % milk blocking solution	5 % w/v dried skimmed milk powder (Marvel) in 1 X TBS-T.

Table 11. Antibodies used for western blotting.

Antibody	Dilution	Incubation conditions	Species of origin	Type	Supplier
Anti-V5 antibody	1:50,000	45 mins, RT	Mouse	Primary	Invitrogen (46-0705)
Anti-FLAG M2 antibody	1:5000	Overnight, 4 °C	Mouse	Primary	Sigma (F1804)
Anti-GP100 antibody	1:1000	Overnight, 4 °C	Rabbit	Primary	Abcam (a137078)
Anti-β-tubulin antibody	1:5000	1 hour, RT	Rabbit	Primary	Abcam (ab6046)
Anti-calnexin antibody	1:5000	1 hour, RT	Rabbit	Primary	Abcam (ab22595)
Anti-mouse IgG-HRP	1:1000	1 hour, RT	Goat	Secondary	Sigma (A6154)
Anti-rabbit IgG-HRP	1:1000	1 hour, RT	Goat	Secondary	Dako

Table 12. Reagents used for direct and indirect immunofluorescence imaging.

Reagents	Dilution	Species of origin	Type	Supplier
Anti-V5 antibody	1:400	Mouse	Indirect (primary)	Invitrogen (46-0705)
Anti-RCN2	1:200	Rabbit	Indirect (primary)	Proteintech (10193-2-AP)
Anti-FLAG M2 antibody	1:200	Mouse	Indirect (primary)	Sigma (F1804)
Anti-mouse secondary antibody Alexa Fluor 488	1:500	Chicken	Indirect (secondary)	Invitrogen (A28175)
F(ab') Anti-rabbit Secondary Antibody, Alexa Fluor™ 546	1:500	Goat	Indirect (secondary)	Invitrogen (A-11071)
Alexa Fluor 546 Phalloidin	1:500	N/A	Direct	Invitrogen (A22283)
Streptavidin AlexaFluor 488	1:500	N/A	Direct	Invitrogen (S11223)
Streptavidin AlexaFluor 647	1:500	N/A	Direct	Invitrogen (S21374)

2.5.2. Protein extraction

Cells were seeded into 10 cm plates and grown to approximately 80 % confluency to generate protein extracts for western blotting. After removing the media, cells were washed twice with chilled PBS and incubated in 300 µL of RIPA buffer (Table 10) on ice for 10 minutes. The cells were then scraped using a cell lifter (Thermo Fisher Scientific) and sonicated on ice using the Diagenode Bioruptor, with two 5-minute rounds of sonication at 30-second intervals. Following sonication, the cell debris was pelleted by centrifugation at 16,000 x *g* for 10 minutes at 4 °C, and the supernatant was transferred to -20 °C for short-term storage or -80 °C for long-term storage.

2.5.3. Bradford Assay

Total protein concentration was determined by measuring the absorbance at 595 nm using an Eppendorf photometer or on a Nanodrop OneC spectrophotometer relative to a standard curve created by measuring the absorbance of series of known dilutions containing bovine serum albumin (BSA) (Sigma) between the concentrations of 2 µg/mL and 10 µg/mL, as shown below in Table 13.

Table 13. Bradford assay components.

Component	Sample	Blank	Standard 1	Standard 2	Standard 3	Standard 4	Standard 5
ddH_2O	798 µL	798 µL	797 µL	796 µL	795 µL	794 µL	793 µL
Bradford reagent (Bio-Rad)	200 µL	200 µL	200 µL	200 µL	200 µL	200 µL	200 µL
RIPA buffer	-	2 µL	2 µL	2 µL	2 µL	2 µL	2 µL
BSA (2 mg/mL)	-	0	1 µL (2 µg/mL)	2 µL (4 µg/mL)	3 µL (6 µg/mL)	4 µL (8 µg/mL)	5 µL (10 µg/mL)
Total protein extracted	2 µL	-	-	-	-	-	-

2.5.4. SDS-polyacrylamide gel electrophoresis

SDS-PAGE gel electrophoresis was utilised to separate proteins based on their molecular weight, allowing for the resolution and identification of specific protein bands.

SDS-PAGE gels were cast, the proportions of which are listed in Table 14. The amount of total protein lysate loaded varied between 10 µg and 40 µg and differences between samples were accounted for by adding NF water and then 4 X SDS PAGE dye (Table 10) to 1 X final concentration. Samples were first heated to 95 °C for 10 minutes, before briefly cooling on ice for 5 minutes and spinning each sample down in a small centrifuge to minimise the loss of sample due to condensation. Samples were loaded alongside 5 µL of Color Prestained Protein Standard, Broad Range (NEB) ladder into either a 10 % or 12 % SDS-PAGE gel. The gel was immersed in 1X SDS running buffer (Table 10) and electrophoresed at 120 V until the bands in the ladder were adequately separated and the bromophenol blue dye had just migrated off the gel, which typically took approximately 90 minutes.

Table 14. Components of SDS-polyacrylamide gels.

Resolving gel components	**Volume of components in 10 % resolving gel**	**Volume of components in 12 % resolving gel**	**Stacking gel components**	**Volume of components in stacking gel**
30 % acrylamide/ 0.8 % bis-acrylamide	6 mL	6 mL	**30 % acrylamide/ 0.8 % bis-acrylamide**	650 µL
1.5 M Tris-HCl pH 8.8	4.5 mL	3.75 mL	**1M Tris-HCl pH 6.8**	625 µL
10 % SDS	180 µL	150 µL	**10 % SDS**	50 µL
ddH_2O	7.32 mL	6.9 mL	**ddH_2O**	3.625 mL
10 % APS	120 µL	100 µL	**10 % APS**	50 µL
TEMED	24 µL	20 µL	**TEMED**	10 µL

2.5.5. Western blotting

Protein was transferred from the gel onto a PVDF membrane (Thermo Fisher Scientific) pre-soaked in methanol by semi-dry transfer at 70 mA for 90 min at RT in 1 X blotting buffer for semi-dry transfer (Table 10). Alternatively, protein was transferred from the gel onto a PVDF membrane (Thermo Fisher Scientific) using a Trans-Blot® Turbo™ Transfer System (BioRad), Trans-Blot Turbo 1 X Transfer buffer (BioRad) (Table 10), Trans-Blot Turbo Mini-size Transfer Stacks on the 1.5 mm program (12 minutes).

The PVDF membrane was then blocked for 1 hour at room temperature in 5 % milk blocking solution (Table 10). Then the PVDF membrane was transferred to 5 % milk blocking solution (Table 10) containing the primary antibody (Table 11) and rotated either for 45 minutes at room temperature, 1 hour at room temperature or overnight at 4 °C. Following primary antibody incubation and three washes in TBS-T (Table 10), the membrane was incubated with the appropriate secondary antibody (Table 11) in 5 % milk solution for 1 hour at room temperature, washed in TBS-T a further 3 times, and the blot exposed to Pierce ECL Plus Western Blotting solution (Thermo Fisher Scientific). The PVDF membrane was then imaged using an ImageQuant LAS-500.

For stripping and re-blotting, membranes were treated with Re-Blot Plus Strong Solution (Millipore) for up to 15 minutes. Subsequently, they were blocked again in 5 % milk blocking solution for 1 hour at room temperature before being re-probed with the desired primary antibodies.

2.5.6. Immunofluorescence staining and imaging

1.5 H precision coverslips (Marienfeld) were incubated overnight in a 1:1 solution of Poly-D-lysine (0.1 mg/mL) (Gibco) and sterile PBS at room temperature in a 24-well plate. The coverslips were then washed three times by adding PBS and rocking the 24-well plate for 5 minutes and then allowed to dry completely in a tissue culture hood. The coverslips were then directly used or stored at 4 °C for up to one week before use.

Cells were seeded in 24-well plates containing the PDL coated coverslips and grown to approximately 80 % confluency. For co-localisation between APEX2 enzymatic activity and the relevant subcellular marker, cells were first treated as in 2.3.7. When conducting immunofluorescence (IF) for co-localisation between the APEX2 epitope tag and the relevant subcellular marker, treatment with the biotin-phenol, H_2O_2 and quenching solutions were omitted.

Cells were fixed in 4 % paraformaldehyde (PFA) in PBS for 10 minutes at room temperature, followed by three washes with chilled PBS. Permeabilisation was achieved using 0.5 % Saponin from Quillaja Bark (Sigma) in PBS for 10 minutes at room temperature, followed by three additional washes with ice-cold PBS. Cells were then blocked with 0.1% Saponin and 3 % BSA in PBS for 1 hour at room temperature to prevent non-specific interactions.

Primary antibodies, as detailed in Table 12, were diluted in a solution containing 0.1% Saponin and 3 % BSA in PBS. A total of 200 µL of this solution was added to each well of the 24-well plate, and the plate was incubated overnight at 4 °C. The next day, cells were incubated with secondary antibodies or Streptavidin AlexaFluor conjugates, as specified in Table 12, at room temperature for 1 hour. Three PBS washes were performed between antibody staining steps and following the removal of the secondary antibodies. If required, DAPI (Invitrogen) was added to a final concentration of 0.1 µg/mL in PBS and incubated with the cells for 15 minutes in the dark. Following this, the cells were washed three times with ice-cold PBS.

Prior to mounting coverslips, 1 µL of 1:100 diluted 0.1 µm Tetraspeck beads (Invitrogen) in ddH_2O was added to 8 µL of Fluoromount G (Thermo Fisher Scientific) on each glass slide. Then, coverslips were dipped in ddH_2O to remove any residual salts from PBS washes and mounted onto the slide. Coverslips were allowed to harden overnight at room temperature and stored in the dark before imaging using the Elyra PS1 Super Resolution system (Zeiss). A 63X objective lens was used with 30 °C oil to match the refractive index of the mounting solution.

2.5.7. Secreted luciferase assays

APEX2-ERM expressing HEK293T and A549 cells were seeded in a 6-well plate and grown to approximately 80 % confluency. To measure secreted nanoluciferase, cells transfected with NlucSec mRNAs were assayed using the Nano-Glo® Luciferase Assay System (Promega). At each time point, 30 µL of cell culture media containing the secreted NanoLuc luciferase was collected and added to a well of a white round-bottom 96-well plate. An equal volume (30 µL) of Nano-Glo® Luciferase Assay Buffer was added, bringing the total volume to 60 µL per well. The plate was then wrapped in aluminium foil and gently rocked at 60 RPM for 5 minutes to ensure thorough mixing. All reagents and media were equilibrated to room temperature before use. Luciferase activity was measured in triplicate using a GloMax Navigator luminometer (Promega) with a 1-second integration time, and the luminescence was recorded as relative light units.

2.5.8. Fluorescent cell imaging

APEX2-ERM HEK293T cells were seeded into 6-well plates 24 hours before transfection and grown to approximately 80% confluency. The following day, the media was removed, and the cells were incubated with Hoechst 33342 (Invitrogen) at a working concentration of 2 µg/mL in PBS for 10 minutes at room temperature. After incubation, the cells were washed once with room-temperature sterile PBS, and OptiMEM was added before mRNA transfection, as detailed in 2.4.1. The cells were incubated for 8 hours post-transfection and imaged at 1 hour, 2 hours, 4 hours, and 8 hours using a Zoe Fluorescent Cell Imager (BioRad).

For live-cell imaging experiments using the Gal9-mCherry HEK293 cell line, cells were seeded to 60% confluency in a Greiner glass-bottom 96-well plate. On the day of imaging, cells were stained with Hoechst 33342 following the same procedure as described above. Next, either 150 ng of fluorophore-labelled Nluc mRNA, complexed with 0.2 µL Lipofectamine MessengerMax (Invitrogen) as outlined in 2.4.1, or 150 ng of naked eGFP mRNA (TriLink BioTechnologies) was introduced to the cells. Cells were then

maintained at 5% CO_2 and 37 °C within a Celldiscoverer 7 microscope (Zeiss) and imaged once per hour over an 18-hour period.

2.6 Data Analysis

All statistical analyses were conducted using GraphPad Prism. To determine the appropriate statistical tests, dataset normality was first assessed using the Shapiro-Wilk test. Parametric tests were applied to normally distributed data, while non-parametric tests were used for datasets that showed significant deviation from a Gaussian distribution. Standard deviation was calculated using GraphPad Prism, with values derived using the *n*-1 method. When the same statistical test was applied to multiple comparisons within an experiment, p-values were adjusted using Bonferroni's correction to determine statistical significance. Adjusted p-values are indicated in the respective figure legends.

2.6.1. Analysis of RT-qPCR data

APEX2 mRNA expression data was analysed by the $2^{-\Delta\Delta Ct}$ method, first normalising each cell line to 18S rRNA as a housekeeping gene. The resulting values were then presented relative to the APEX2-ERM cell line within the corresponding cell type.

Prior to streptavidin affinity purification optimisation, biotinylated RNA enrichment post-streptavidin affinity purification was analysed by comparing output RNA relative to input RNA, using the $2^{-\Delta\Delta Ct}$ method where ΔCt = Ct(output) – Ct(input) for each gene individually and ΔΔCt = ΔCt(+biotin-aniline) – ΔCt(no biotin-aniline negative control). It should be noted that this enrichment analysis did not make use of housekeeping genes.

During streptavidin affinity purification optimisation, biotinylated RNA enrichment post-streptavidin affinity purification was analysed using the $2^{-\Delta\Delta Ct}$ method, where ΔCt = Ct(Fluc spike in RNA) – Ct(elution spike in) and ΔΔCt = ΔCt(biotinylated Fluc spike in RNA) – ΔCt(non-biotinylated spike in RNA).

Post streptavidin affinity purification optimisation, biotinylated RNA enrichment post-streptavidin affinity purification was analysed by first normalising to the

biotinylated spike-in RNA and showing data relative to the no-biotin-aniline negative control, using the $2^{-\Delta\Delta Ct}$ method, where ΔCt = Ct(RNA of interest) – Ct(Biotinylated spike in Fluc RNA) and ΔΔCt = ΔCt(+biotin-aniline) – ΔCt(-biotin-aniline control).

Statistical analyses to assess enrichment post-streptavidin pulldown, relative to the no-biotin-aniline negative control, were conducted using either a one-sample t-test or a Wilcoxon signed-rank test, depending on whether the data significantly deviated from a Gaussian distribution. Cases where Wilcoxon signed-rank tests were used are indicated in the respective figure legends. Comparisons between experimental conditions were performed using a Student's t-test with Welch's correction.

In the analysis of S^6G incorporation following *in vitro* biotinylation *via* streptavidin pulldown-RT-qPCR, the data were first normalised to the biotinylated spike-in RNA and expressed relative to the 25% S^6G condition. To evaluate the enrichment of the 0% S^6G condition compared to the 25% S^6G condition, a one-sample Student's t-test was conducted whereas to assess the statistical significance between the 0% and 10% S^6G conditions, a Student's t-test was utilised.

In the analysis of membrane fractionation experiments, RNA expression was analysed using the $2^{-\Delta\Delta Ct}$ method, where ΔCt = Ct(Gene of interest in fraction of interest) – Ct(spike-in RNA in fraction of interest), and ΔΔCt = ΔCt(Gene of interest in membrane bound fraction) – ΔCt(Gene of interest in cytosolic fraction). This method allowed for the identification of the gene's predominant localisation. Statistical analyses to assess enrichment in the membrane-bound fraction, relative to the cytosolic fraction, were conducted using a one-sample t-test.

2.6.2. Quantitative analysis of structured illumination microscopy images

Structured illumination microscopy (SIM) images were analysed primarily by calculating Mander's co-occurrence coefficient in Zen Black software to assess

the degree of overlap between the two channels, providing insight into the extent to which signals from each channel coincide. Pearson's correlation scores were also calculated in Zen Black (Supplementary Figure 11), although the assumption of a linear relationship between pixel intensities cannot be made due to the absence of direct interaction between the two proteins (Dunn et al., 2011; Pike et al., 2017). No other statistical tests were applied to the SIM data.

Co-occurrence analysis was conducted on images captured within the general focal plane of the nuclear mid-section. Regions of interest were manually delineated around the nuclei for analysis. Before quantitative assessment, the minimum threshold was adjusted based on negative controls. This was achieved by adjusting the minimum threshold based on the highest intensity signal observed in either the no primary antibody control or the no biotin-aniline negative control when analysing streptavidin signals. For studies in APEX2-NES HEK293T cells assessing co-occurrence with DAPI, the minimum threshold for the DAPI channel was set based on the highest intensity signal observed in the cytosolic regions.

2.6.3. Statistical analysis of secreted luciferase data

Statistically significant differences in the delivered nanoluciferase mRNA featuring distinct characteristics were assessed using Student's t-tests with Welch's correction at each time point. In the analysis of secreted nanoluciferase assays comparing IVT mRNA with varying proportions of S^6G, Student's t-tests with Welch's correction were performed to evaluate differences in luciferase activity between the 0% S^6G and 25% S^6G conditions at each time point.

Chapter 3 - Generation and characterisation of APEX2-expressing cells

3.1. Introduction

The primary aim of the work in this chapter was to establish and validate cell lines expressing the APEX2 enzyme targeted to two distinct subcellular locations: the cytoplasm and the ER. Translation is typically initiated in the cytoplasm but the presence of SPCS leads to re-localisation of ribosomes translating mRNA encoding secreted or membrane proteins to the ER for continued translation (Keenan et al., 2001).

HEK293T cells, derived from female human embryonic kidney cells, were selected due to their established use in APEX2 RNA studies (Fazal et al., 2019; Padrón et al., 2019). These cells are suitable due to their ability to permit the diffusion of biotin-phenol (Lam et al., 2014) and biotin-aniline (Zhou et al., 2019) and their high transfectability, which facilitates the efficient introduction of exogenous genetic material (E. Tan et al., 2021).

A549 cells, derived from a male lung epithelial carcinoma, were also chosen for their therapeutic relevance. Lung epithelial cells represent a valuable model for studying therapeutic mRNA delivery, particularly given the interest in the use of inhalation-based administration of therapeutic mRNA for lung-localised conditions such as cystic fibrosis (Jiang et al., 2024). Additionally, the APEX2 system has not yet been reported in A549 cells, making this an important addition to the field.

APEX2-ERM and APEX2-NES systems were used in this study because they have been extensively used by the Ting lab for proximity-dependent labelling of proteins and RNA within the cytosol and ER (Fazal et al., 2019; Hung et al., 2017; Lam et al., 2014). Stable APEX2-fusion cell lines were generated in HEK293T and A549 cells using a lentiviral (LV) transduction approach. Investigating whether APEX2 was localised and functional at the expected

subcellular sites was crucial for determining the level of confidence in subsequent experiments exploring the localisation of delivered mRNA.

This chapter first details the generation of the APEX2-ERM and APEX2-NES cell lines, followed by analysis of their enzymatic activity using RNA dot blot assays and immunofluorescence. The findings demonstrate successful targeting of APEX2 to the intended subcellular compartments, with the ER-targeted APEX2 (APEX2-ERM) overlapping with an ER marker protein and the cytosolic APEX2 (APEX2-NES) localising primarily to the cytoplasm. The chapter concludes by discussing additional validation methods, such as testing different antibodies, assessing ER function with phenotypic assays, and considering alternatives to LV transduction.

3.2. Generation of APEX2-ERM cell lines by lentiviral transduction

First, cell lines with APEX2 targeted to the ER membrane were developed using a LV transduction approach. The LV transfer plasmid containing APEX2 fused to the N-terminal signal anchor domain from rabbit CYP450 2C1 (Figure 3.1) (Ahn et al., 1993) is designed to integrate the transgene into the genome and localise the APEX2 fusion protein to the cytosolic-facing side of the ER membrane. This APEX2 fusion system has been previously utilised by the Ting laboratory (Fazal et al., 2019; Lam et al., 2014) and the Ingolia laboratory (Padrón et al., 2019) in HEK293T cells.

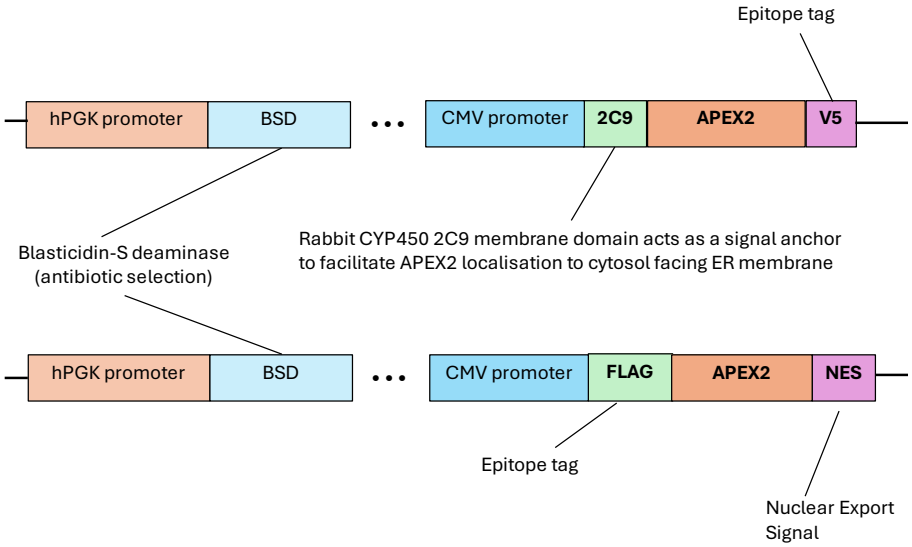


Figure 3.1. Plasmid schematics illustrating APEX2 coding sequences, antibiotic resistance genes and their corresponding promoters. Plasmids are engineered for integration into the host genome via lentiviral transduction.

Blasticidin kill curves were performed on HEK293T and A549 cell lines to determine the minimum antibiotic concentration necessary for effective selection of transduced cells. It was determined that a concentration of 6 µg/mL blasticidin in the growth medium was the minimum required to induce cell death within 3 to 5 days in both cell lines. The APEX2-ERM lentiviral transfer plasmid, psPAX2 packaging plasmid, and PMD2.G envelope plasmid were co-transfected into HEK293T cells to generate LV stock (Figure 3.2).

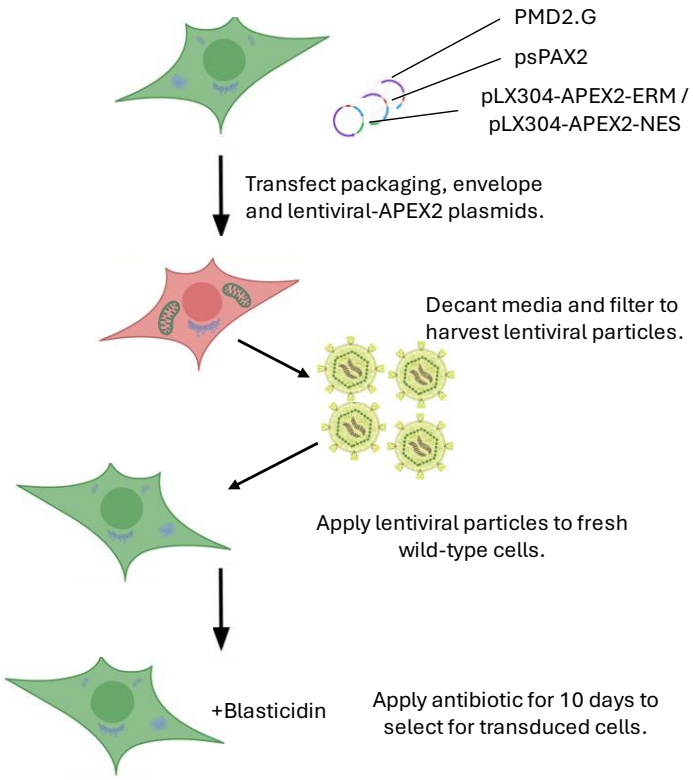


Figure 3.2. Schematic of lentiviral transduction workflow.

Fresh HEK293T cells were subsequently transduced with the APEX2-ERM LV stock and selected with 6 µg/mL blasticidin for one week, resulting in approximately 20% cell survival. These transduced cells were then maintained as a mixed cell population.

To verify the stable expression of APEX2-ERM in transduced HEK293T and A549 cells, western blots using an antibody against the V5 epitope tag were performed across passages 3 to 15 post-transduction. A prominent band at 28 kDa was observed in the APEX2-ERM HEK293T cells but not in the WT HEK293T cells, confirming successful LV transduction (Figure 3.3). A clear band was visualised using an antibody specific to the V5 epitope at a highly diluted concentration (1:50,000), compared to the manufacturer's recommended 1:1,000 dilution, suggesting that APEX2-ERM was highly expressed in both the A549 and HEK293T cell lines.

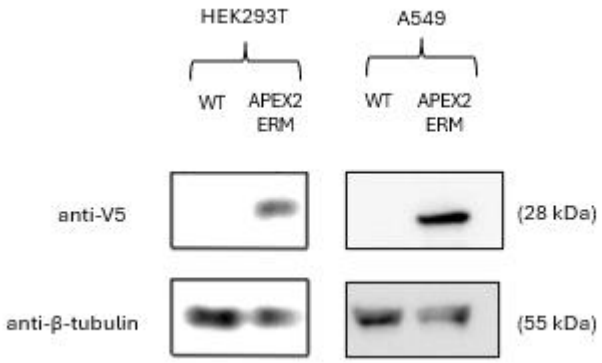


Figure 3.3. Western blots demonstrating stable expression of APEX2-ERM in transduced HEK293T and A549 cells. Blots were first probed with anti-β-Tubulin and then stripped and re-probed with an anti-V5 antibody to detect the V5-tagged APEX2-ERM construct. Images shown are representative of at least three independent experiments.

3.3. Generation of APEX2-NES cell lines by lentiviral transduction

To generate cells with APEX2 localised to the cytoplasm, APEX2-NES HEK293T and APEX2-NES A549 cell lines were produced using a LV system where APEX2 was fused to a nuclear export signal (NES), as illustrated in Figure 3.1. Multiple APEX2 fusion protein cell lines were generated by reducing the proportion of LV particles to fresh Opti-MEM during transduction to provide a level of control over APEX2-NES expression levels in transduced cells. Both HEK293T and A549 cells exposed to 10%, 25%, 50%, and 100% APEX2-NES lentivirus survived blasticidin selection. However, HEK293T cells exposed to 10% APEX2-NES LV particles were excluded from further analysis due to slower growth rates. To investigate whether reducing the amount of LV particles corresponded to lower APEX2-NES expression, two APEX2-NES cell lines from each cell type were selected based on the assumption that cell lines exposed to lower proportions of LV particles in fresh Opti-MEM would exhibit reduced APEX2 expression. HEK293T lines exposed to 25% and 50% lentivirus, and A549 lines exposed to 10% and 25% lentivirus were chosen, as these cell lines were exposed to the lowest concentrations of LV and survived blasticidin selection while maintaining normal growth rates.

To verify the stable expression of APEX2-NES in transduced HEK293T and A549 cells, western blots using an antibody against the FLAG epitope tag were

performed across passages 3 to 15 post-transduction. Bands at 28 kDa were detected in APEX2-NES 25% and 50% HEK293T cells, and in APEX2-NES 10% and 25% A549 cells, but not in WT cells, confirming successful transduction (Figure 3.4.). In both cell types, the FLAG band detected by western blot in cells exposed to a lower proportion of LV in fresh Opti-MEM was thinner compared to those exposed to a higher LV proportion. This suggests that exposure to reduced lentivirus levels corresponded to lower APEX2 protein expression in both cell types.

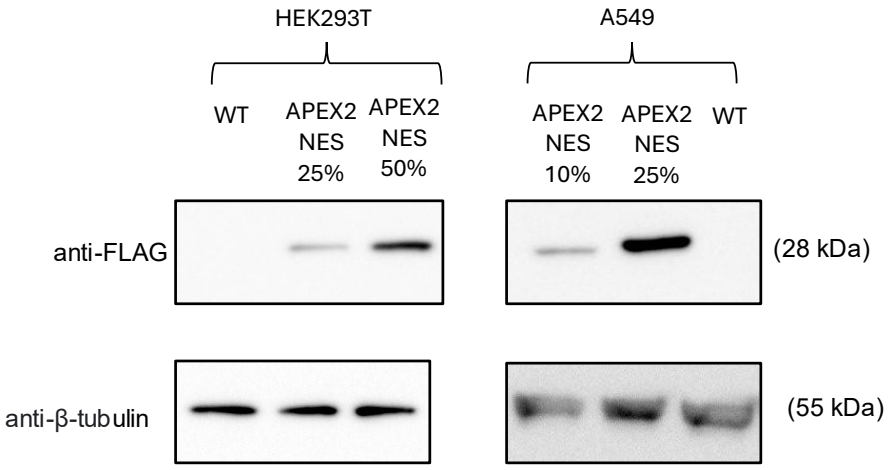


Figure 3.4 Western blot showing stable expression of APEX2-NES in transduced HEK293T and A549 cells. The APEX2-NES constructs include a FLAG epitope tag, which was detected with a specific anti-FLAG antibody. Blots were then stripped and re-probed with an anti-β-tubulin antibody. NES 10%, NES 25%, and NES 50% refer to the ratio of Opti-MEM containing LV particles to fresh Opti-MEM used during LV transduction. Images shown are representative of at least three independent experiments.

Since protein-level comparison was not feasible due to the differing epitope tags, RT-qPCR was used to quantify and directly compare APEX2 mRNA expression between the NES and ERM cell lines. The RT-qPCR results for APEX2 mRNA in APEX2-NES cell lines were normalised to 18S rRNA and are presented relative to APEX2-ERM in each cell type. In HEK293T cells, APEX2 mRNA levels in the APEX2-NES 25% cell line were comparable to those in the APEX2-ERM cells (Figure 3.5A). In contrast, APEX2 mRNA levels in the APEX2-NES 50% HEK293T cell line were approximately 2-fold higher than those in the APEX2-NES 25% cell line, which is consistent with the western blot results (Figure 3.4). APEX2 mRNA levels in WT HEK293T cells were significantly lower than in APEX2-ERM, confirming the specificity of the assay. The similarity in mRNA

levels between APEX2-ERM and APEX2-NES 25% cell lines, combined with the fact that APEX2-NES 10% HEK293T had slower growth rates compared to HEK293T WT cells suggested that APEX2-NES 25% was the most suitable variant to use in subsequent experiments. Therefore, from this point forward, "APEX2-NES HEK293T" will refer to the APEX2-NES 25% HEK293T cell line.

In A549 cells, APEX2 mRNA levels in the APEX2-NES 25% cell line were also comparable to those in APEX2-ERM cells (Figure 3.5B). In the APEX2-NES 50% cell line, APEX2 mRNA levels were nearly three-fold higher than in APEX2-ERM. Additionally, APEX2 mRNA levels in the APEX2-NES 10% cell line were significantly lower than those in APEX2-ERM, while mRNA levels in APEX2-NES 100% were similar to those in the APEX2-NES 50% cell line. Although this suggests potential saturation above the 50% level, the lack of statistical significance in NES 100% relative to APEX2-ERM precludes a definitive conclusion. Overall, in both HEK293T and A549 cells, the strategy ensured that the minimum amount of LV was used to establish the NES transduced cell lines, thereby minimising the chances of individual cells ending up with excessive copies of APEX2.

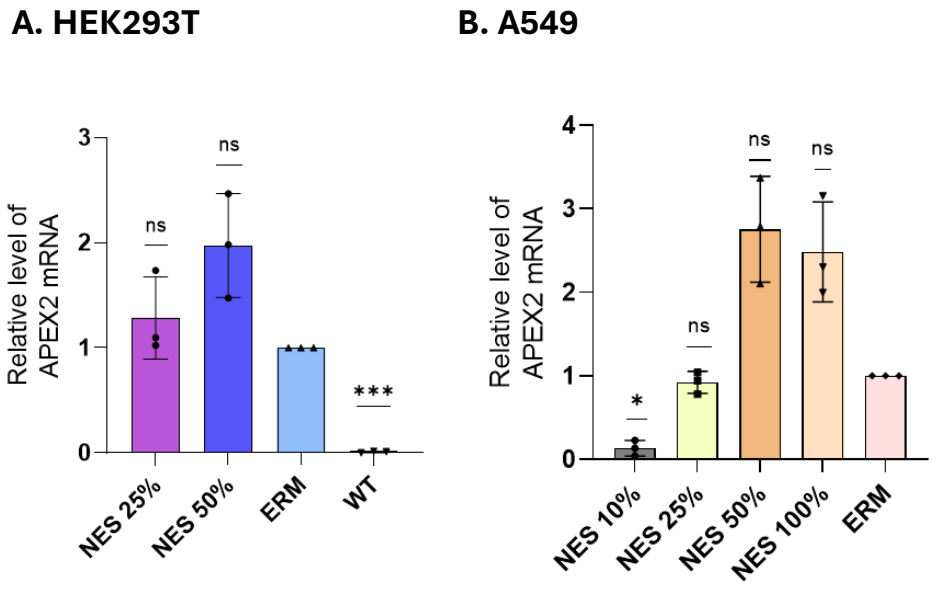


Figure 3.5. Relative expression of APEX2 mRNA in various lentivirally transduced cell lines, analysed by RT-qPCR. (A) Expression in HEK293T cell lines, and (B) expression in A549 cell lines. APEX2 mRNA levels were normalised to 18S rRNA and are presented relative to APEX2-ERM in each cell type. Bars represent the mean of three biological replicates, with error bars indicating standard deviation (SD). Statistical significance in A is denoted as *** p < 0.00016 ˙ and ns = not significant. Statistical significance in B is denoted as * = p < 0.0125 and ns = not significant.

3.4. APEX2 enzymatic activity on RNA in cells imaged by RNA dot blot assay

To determine whether APEX2 in the cell lines described above was able to biotinylate RNA: APEX2-NES HEK293T, APEX2-ERM HEK293T, and APEX2-ERM A549 cell lines were incubated with biotin-aniline, an APEX2 substrate chosen for its greater selectivity for RNA labelling over proteins compared to biotin-phenol (Zhou et al., 2019). Cells were treated with biotin-aniline for 30 minutes at 37 °C, followed by a 1-minute exposure to H_2O_2 to facilitate APEX2's enzymatic activity on spatially proximal RNA prior to total RNA isolation. A negative control, in which biotin-aniline was omitted, was included to assess specificity of the streptavidin-HRP RNA dot blot assay. After RNA isolation, the samples were UV-crosslinked onto a membrane and incubated with streptavidin-horseradish peroxidase (HRP), followed by the addition of ECL substrate for detection of biotinylated RNA. Streptavidin-HRP signals were observed only when both biotin-aniline and H_2O_2 were present, confirming that RNA biotinylation occurred specifically under these experimental conditions (Figure 3.6). Work was not conducted in A549 APEX2-NES cell lines at this stage, as the focus was on validating the labelling in the HEK293T cell lines to ensure the method was successful before applying it to a more experimental cell line such as A549.

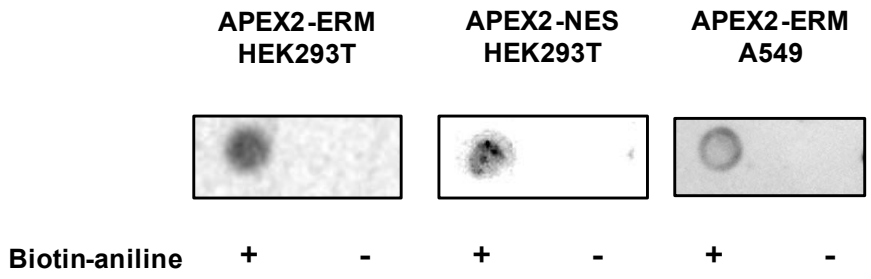


Figure 3.6. Streptavidin-HRP dot blot assays show APEX2 is enzymatically active towards RNA in various APEX2 cell lines. APEX2 expressing cells were incubated with or without biotin-aniline before exposure to H_2O_2. Total RNA was extracted and dotted onto a membrane, and the blot probed with streptavidin-HRP. The images for HEK293T APEX2-ERM are representative of three biological repeats (N=3), while those for HEK293T APEX2-NES and A549 APEX2-ERM are from a single experiment (N=1).

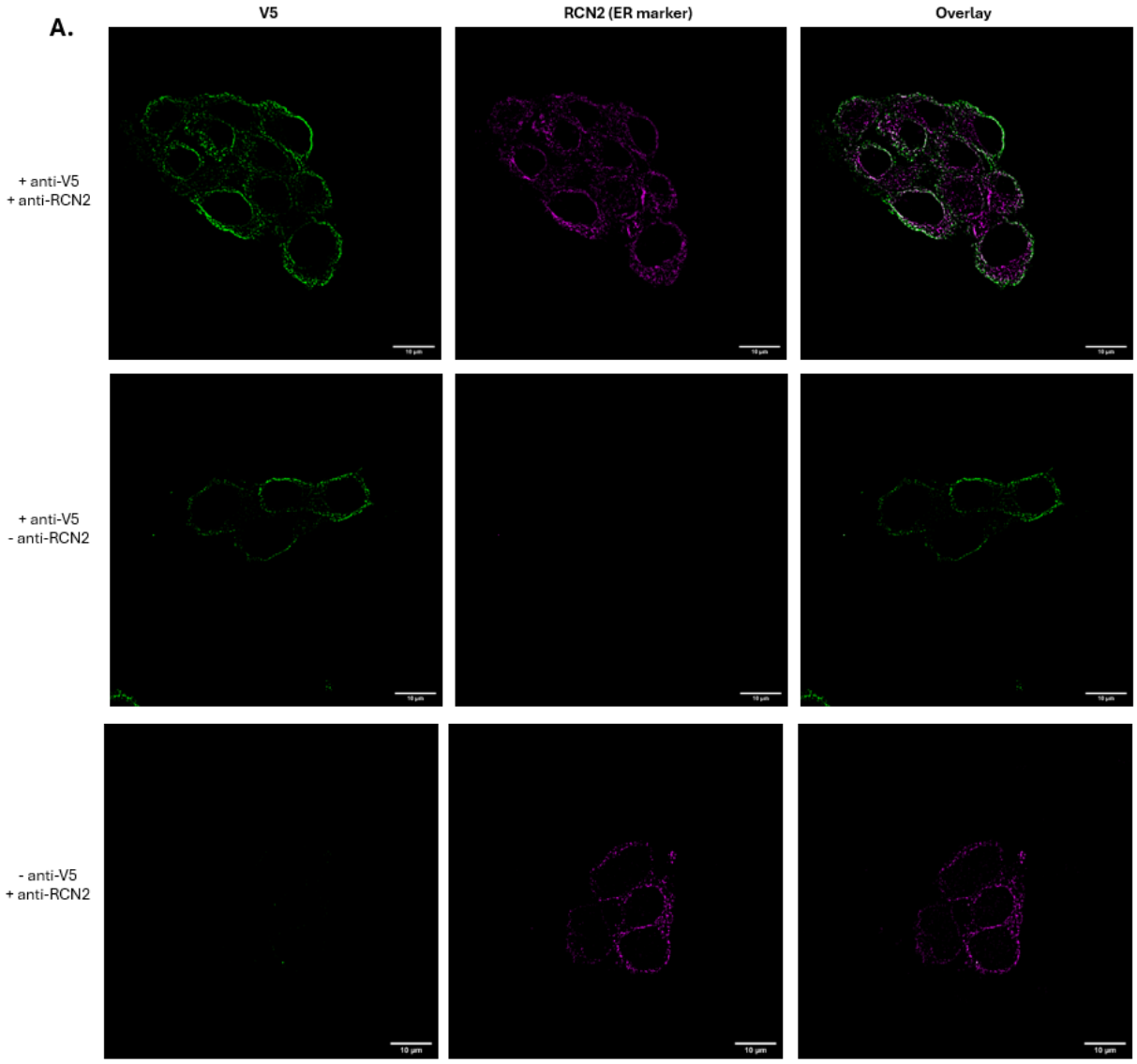
3.5. APEX2-ERM HEK293T epitope tag co-localises with the ER

To gain insight into the localisation of APEX2 in the APEX2-ERM cell line, structured illumination microscopy (SIM) was utilised. SIM was selected for its super-resolution capabilities, which allow for the precise visualisation of subcellular structures beyond the limits of conventional light microscopy (as reviewed by Chen et al. (2023)). The first SIM experiment was designed to assess co-occurrence of the V5 epitope tag of APEX2-ERM in HEK293T with RCN2, a calcium-binding protein localised to the lumen of the endoplasmic reticulum (ER), which serves as a well-established ER marker. This approach, based on the methodology of Hung et al. (2017), allowed for the initial probing as to the extent of APEX2-ERM's ER localisation.

The SIM images were analysed by calculating Mander's coefficient to evaluate the extent of overlap between the two channels, offering insight into how closely the signals from each channel coincide. Co-occurrence analysis was performed on images captured within the general focal plane of the nuclear mid-section. Regions of interest were manually delineated around the nuclei, and in this experiment, RCN2 and V5 signals were used as references to identify peri-nuclear regions of cells.

In APEX2-ERM HEK293T cells, the V5 epitope tag of APEX2-ERM was mainly localised to the ER, though it tended to distribute to regions of the ER more distal from the nucleus (Figure 3.7A). This is supported by weak to moderate co-occurrence scores (M1 = 0.6233 ± 0.04986, M2 = 0.6701 ± 0.08678) (Figure 3.7B). Importantly, the V5 epitope signal was consistently detected in all

HEK293T cells expressing APEX2-ERM, indicating expression across the mixed cell population.



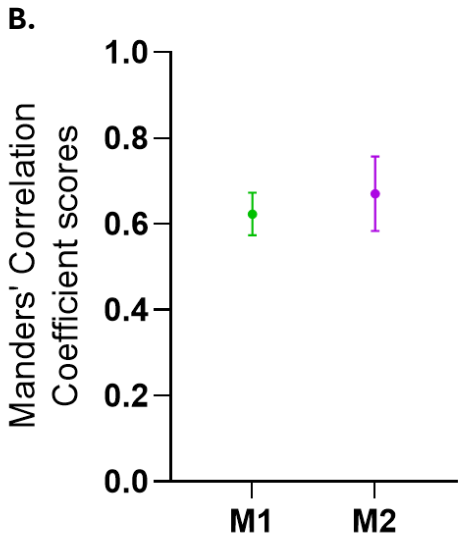
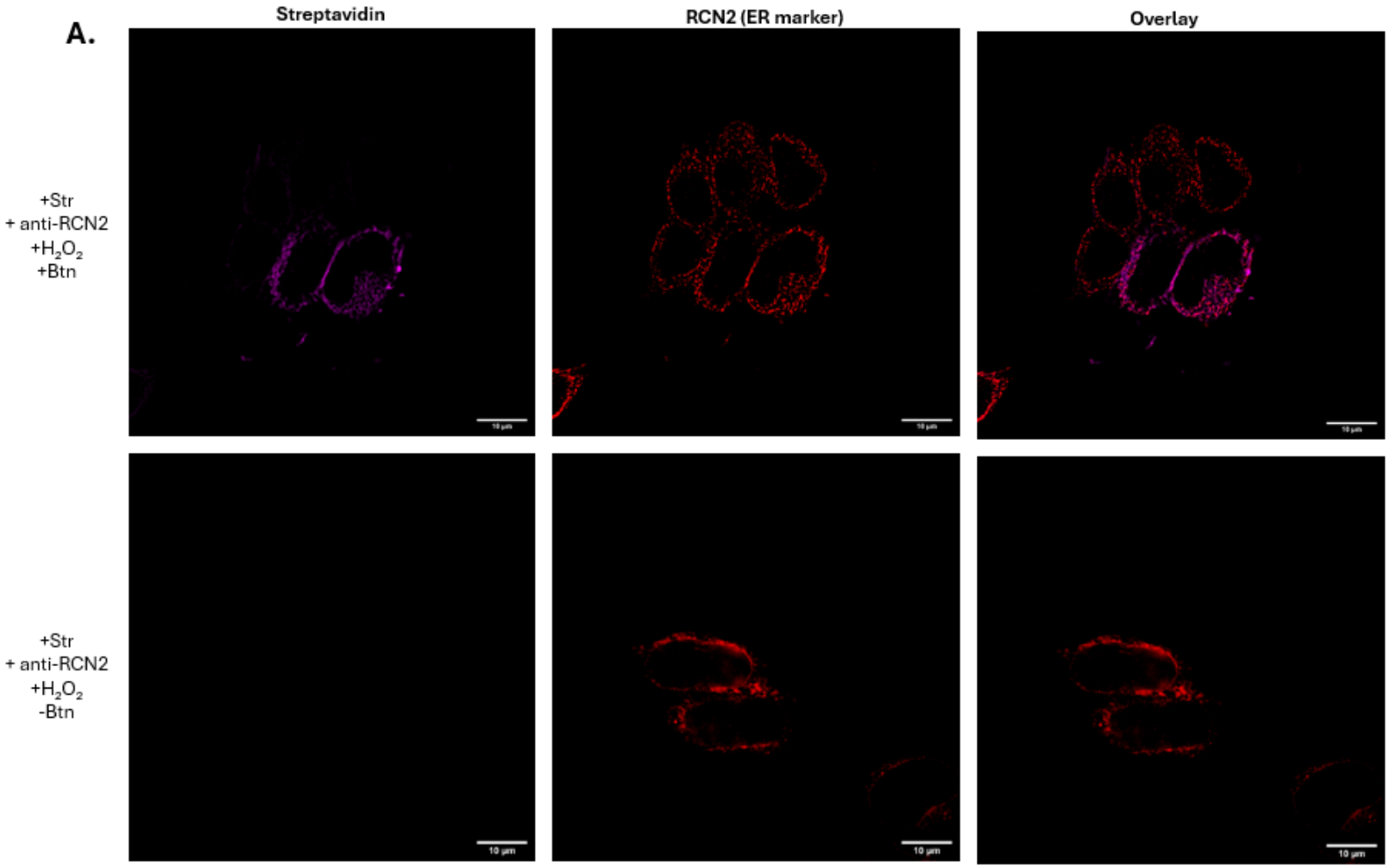


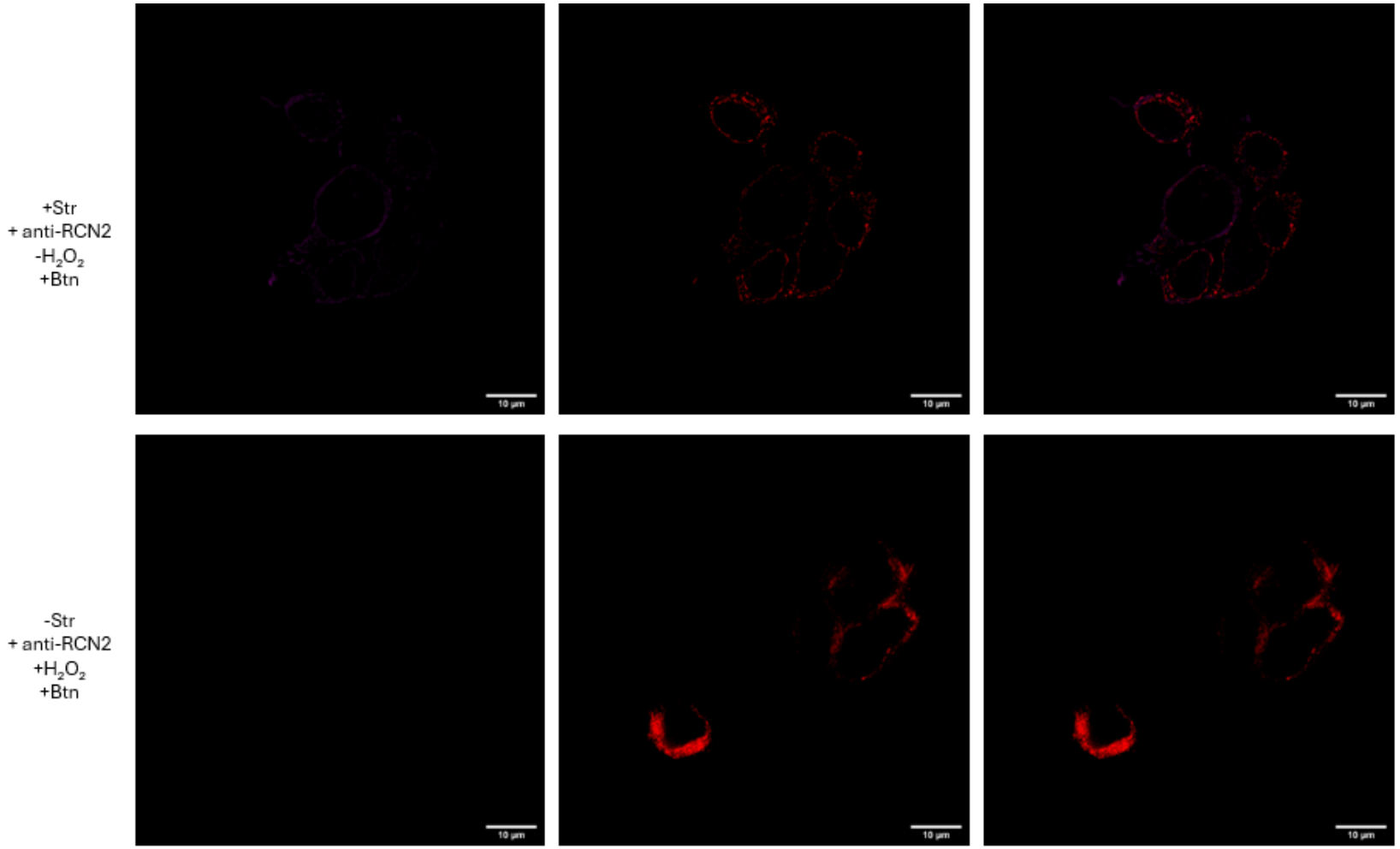
Figure 3.7. Representative structured illumination microscopy (SIM) images illustrating the co-occurrence of the V5 epitope tag of APEX2-ERM (green) with RCN2, an endoplasmic reticulum (ER) membrane protein (purple), in APEX2-ERM HEK293T cells. (A) Cells were fixed and stained with either anti-V5, anti-RCN2, or both to visualise the distribution of APEX2-ERM relative to the ER membrane. Scale bar: 10 µm. Magnification: 63X. (B) Mander's coefficient scores of manually delineated objects from images taken at the mid-section of the nucleus. M1 indicates the degree of overlap of the V5 channel with the RCN2 channel, while M2 represents the extent of RCN2's overlap with the V5 channel. The coefficient scores were calculated from 17 cells, taken from 1 biological repeat (N=1), with error bars indicating SD.

3.6. APEX2-ERM HEK293T enzymatic activity co-localises with the ER

While the SIM data presented in Figure 3.7 confirmed that APEX2-ERM generally localised to the ER membrane, it was also important to determine whether APEX2 labelling was occurring in proximity to the ER. To achieve this, APEX2-ERM HEK293T cells were incubated with biotin-phenol, which is considered more effective for APEX2-mediated protein labelling than biotin-aniline (Zhou et al., 2019), prior to exposure to H_2O_2. This method provides a comprehensive view of the spatial distribution of APEX2-mediated biotinylation of biological molecules. Co-occurrence analysis was performed on images captured within the general focal plane of the nuclear mid-section. Regions of interest were manually delineated around the nuclei, using RCN2 as a reference marker. In this study, M1 reflects the degree of overlap of the Streptavidin channel onto the RCN2 channel, while M2 represents the degree of overlap of the RCN2 channel onto the Streptavidin channel. Cells were counted to determine the proportion of cells exhibiting detectable enzymatic activity.

In cells with detectable biotinylation above background levels (Figure 3.8A), co-occurrence scores were high, indicating substantial spatial overlap between the signals (M1 = 0.8467 ± 0.1846, M2 = 0.9521 ± 0.06747) (Figure 3.8B). This suggests APEX2's enzymatic activity in these cells is closely associated with the ER membrane.


A.
Streptavidin
RCN2 (ER marker)
Overlay
+Str
+ anti-RCN2
+H2O2
+Btn
+Str
+ anti-RCN2
+H2O2
-Btn
10 µm


+Str
+ anti-RCN2
-H_2O_2
+Btn
-Str
+ anti-RCN2
+H_2O_2
+Btn
10 µm

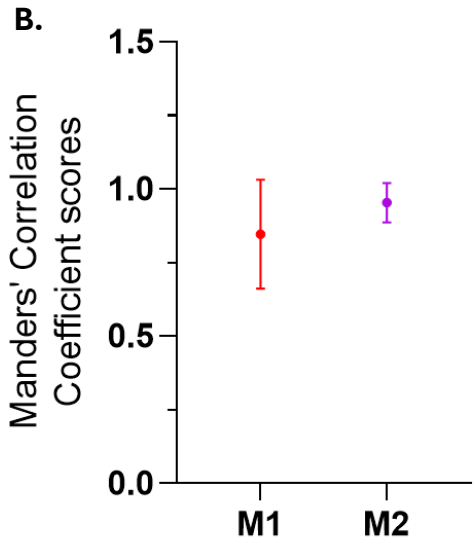


Figure 3.8. Representative structured illumination microscopy (SIM) images illustrating the co-occurrence of APEX2-ERM enzymatic activity (red) with RCN2 (purple) in APEX2-ERM HEK293T cells. (A) Cells were incubated with or without biotin-phenol (Btn) prior to exposure to H_2O_2, which catalyses APEX2-mediated biotinylation. Following this, cells were stained with Streptavidin AlexaFluor (Str), anti-RCN2, or both, allowing for the visualisation of APEX2-ERM enzymatic activity relative to the ER. Scale bar: 10 µm. Magnification: 63X. (B) Mander's co-occurrence scores of manually delineated objects from images taken at the mid-section of the nucleus. M1 indicates the degree of overlap of the Streptavidin channel with the RCN2 channel, while M2 represents the extent of RCN2's overlap with the Streptavidin channel. The coefficient scores were calculated from 24 cells, taken from 1 biological repeat (N=1), with error bars indicating SD.

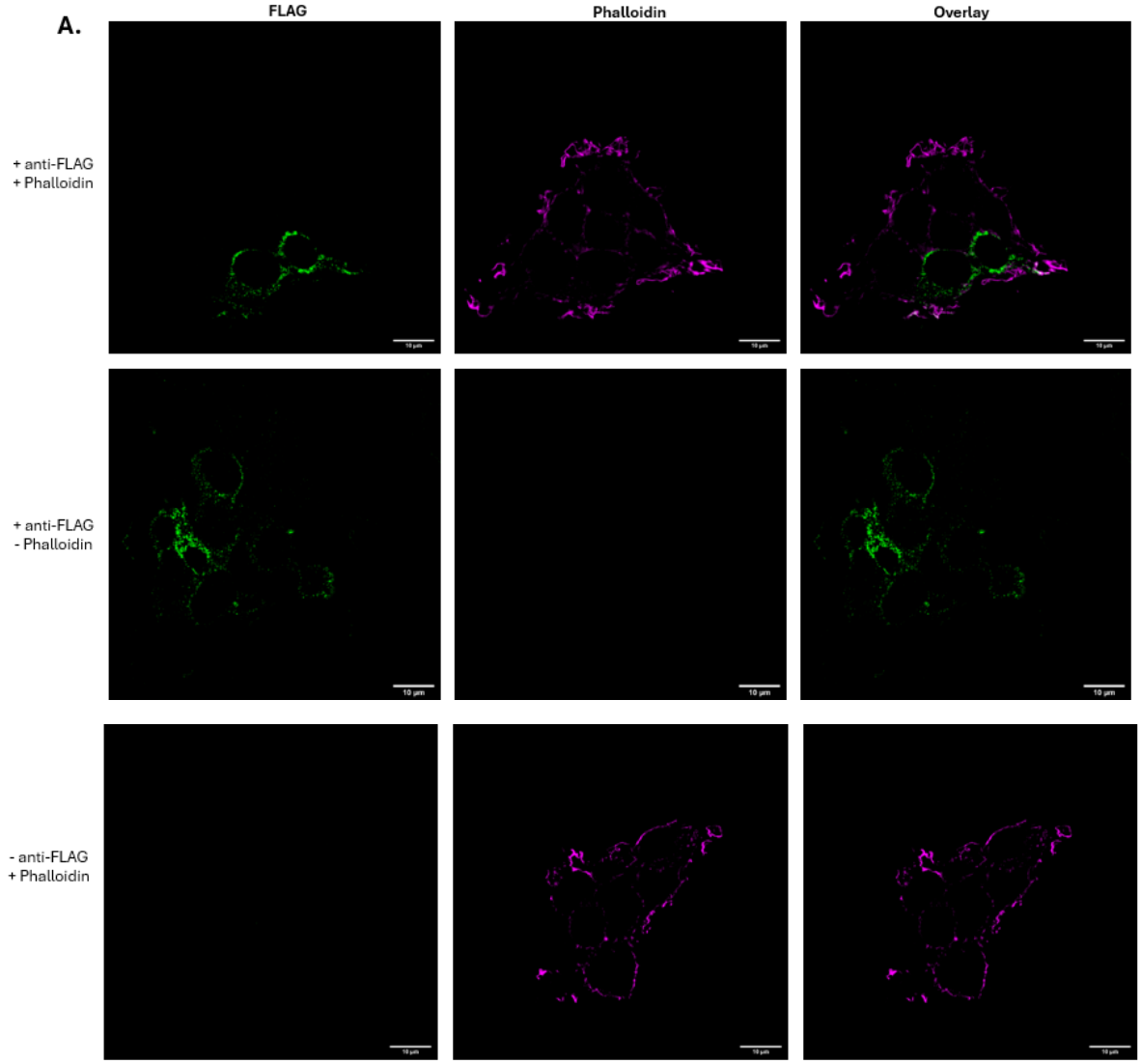
3.7. APEX2-NES HEK293T epitope tag overlaps with cytosolic marker

To investigate the localisation of APEX2 in the APEX2-NES HEK293T cell line, SIM was employed to investigate the overlap between the FLAG epitope tag of APEX2-NES and phalloidin, a marker for actin filaments (Figure 3.9A). Co-occurrence analysis was performed on images captured within the general focal plane of the nuclear mid-section. Regions of interest were manually delineated around the nuclei, using phalloidin as a reference to identify peri-nuclear regions of cells.

The M1 score reflects the overlap of the FLAG signal onto the phalloidin signal, indicating the extent to which APEX2-NES is localised within the cytosol. The relatively high M1 score (0.6989 ± 0.1855) suggests that the FLAG-tagged APEX2-NES is broadly distributed in the cytosol where actin filaments are present (Figure 3.9B). In contrast, the M2 score represents the overlap of phalloidin signal onto the FLAG signal. The very low M2 score (0.1120 ± 0.1141) indicates that most of the phalloidin signal does not overlap with the FLAG

signal. While APEX2-NES protein expression is expected to be less extensive than that of actin, the spatial restriction of the APEX2-NES FLAG tag within the cytosol was notable. Furthermore, the amount of FLAG epitope detected varied significantly across cells and 52 % of APEX2-NES HEK293T cells were positive for FLAG signal (Figure 3.9C)

Consequently, further investigation was necessary to understand the localisation of APEX2-NES enzymatic activity to clarify these findings.



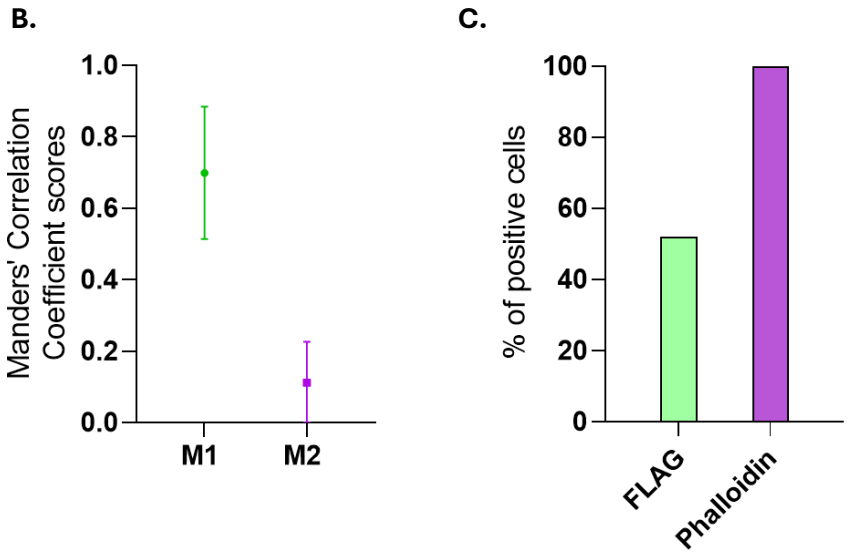
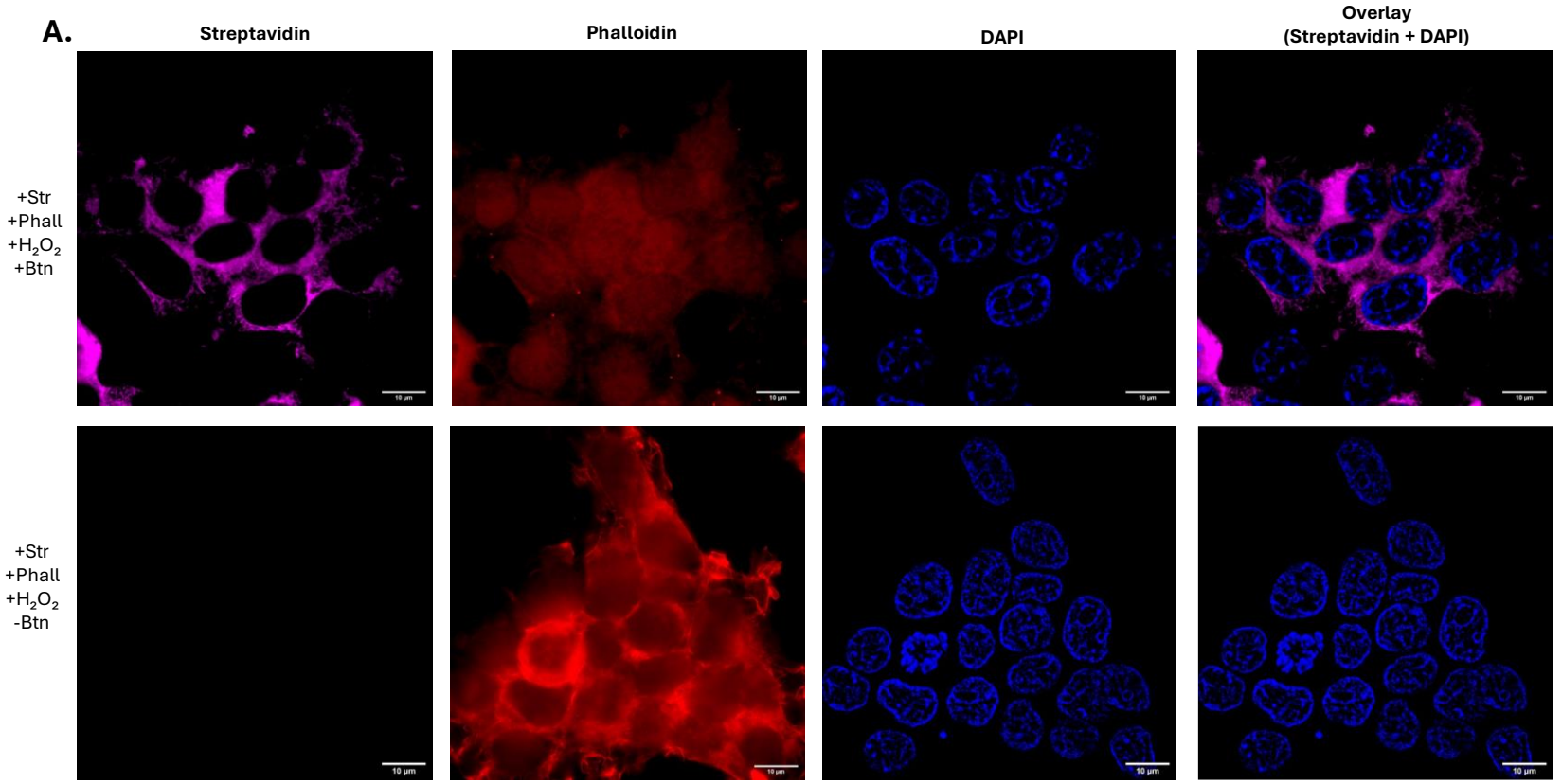


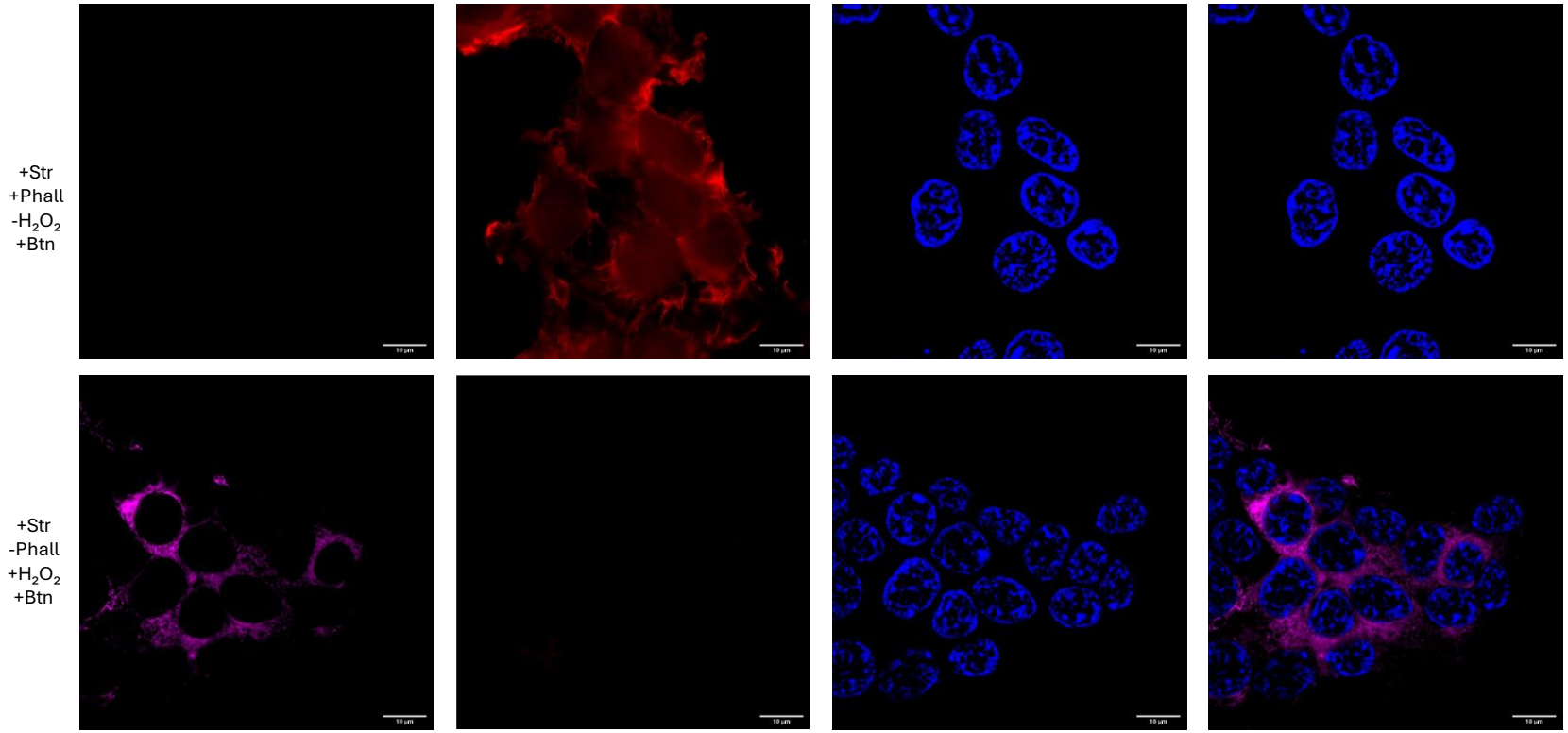
Figure 3.9. Representative structured illumination microscopy (SIM) images illustrating the co-occurrence of the FLAG epitope tag of APEX2-NES (green) with phalloidin (purple) in APEX2-NES HEK293T cells. (A) Cells were stained with either anti-FLAG, phalloidin, or both, enabling the visualisation of the distribution of APEX2-NES relative to the actin cytoskeleton. Scale bar: 10 µm. Magnification: 63X. (B) Mander's co-occurrence co-existence scores of manually delineated objects from images taken at the mid-section of the nucleus. M1 indicates the degree of overlap of the FLAG channel with the phalloidin channel, while M2 represents the extent of phalloidin's overlap with the FLAG channel. The Mander's co-occurrence coefficient were calculated from 15 cells, taken from 1 biological repeat (N=1), with error bars indicating SD.

3.8. APEX2-NES HEK293T enzymatic activity is restricted from the nucleus

To examine the localisation of APEX2 enzymatic activity in the APEX2-NES HEK293T cell line, SIM was employed to investigate the localisation of biotin adducts resulting from APEX2-NES enzymatic activity relative to subcellular markers. A biotinylation reaction was carried out by incubating APEX2-NES HEK293T cells with biotin-phenol for 30 minutes followed by exposure to H_2O_2, as described previously for APEX2-ERM HEK293T cells in Section 3.6. Unanticipated overlap between phalloidin and diamidino-2-phenylindole (DAPI) signals was observed, suggesting the presence of the cytosolic marker within nuclei (Figure 3.10A). Consequently, phalloidin was excluded from the co-occurrence analysis, and regions of interest were manually delineated around the nuclei using DAPI as a reference. In this experiment, M1 reflects the degree of overlap of the Streptavidin channel onto the DAPI channel, while M2 represents the extent of DAPI channel's overlap with the Streptavidin channel.

Co-occurrence scores (M1 = 0.07563 ± 0.07382, M2 = 0.2336 ± 0.2388) between Streptavidin signal and DAPI signal indicate that APEX2-NES enzymatic activity was not present in the nucleus (Figure 3.10B).


A.
Streptavidin
Phalloidin
DAPI
Overlay
(Streptavidin + DAPI)
+Str
+Phall
+H2O2
+Btn
+Str
+Phall
+H2O2
-Btn


+Str
+Phall
-H_2O_2
+Btn
+Str
-Phall
+H_2O_2
+Btn

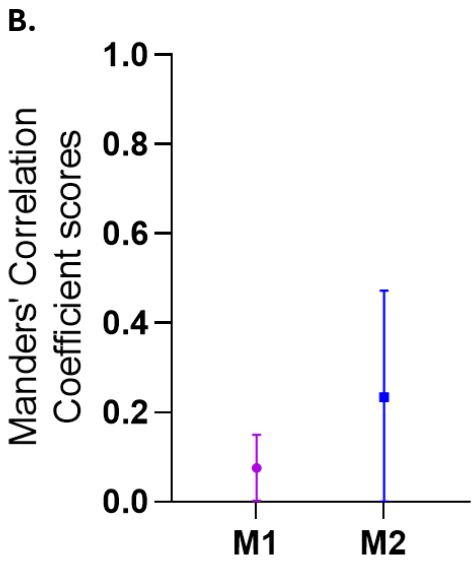
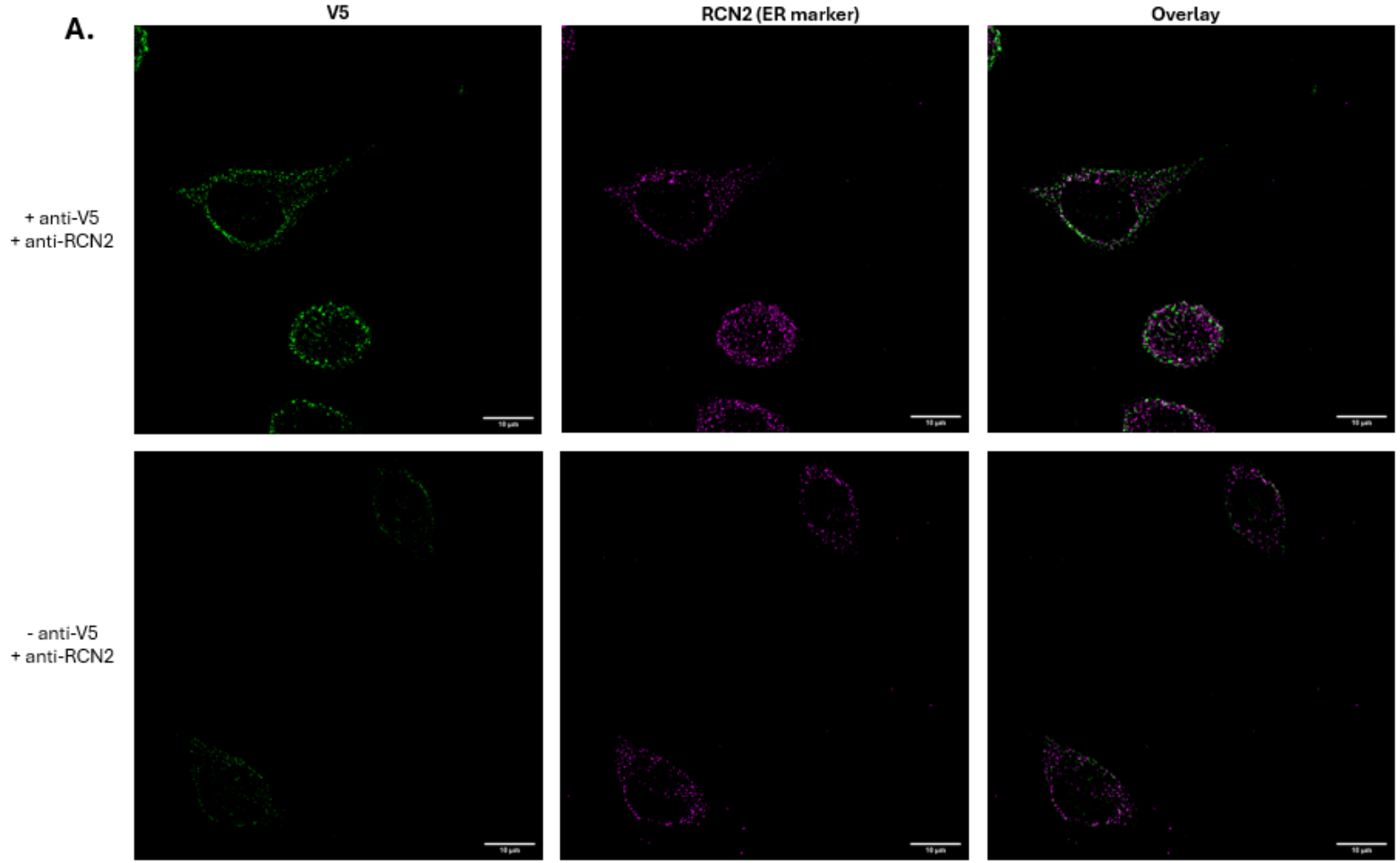


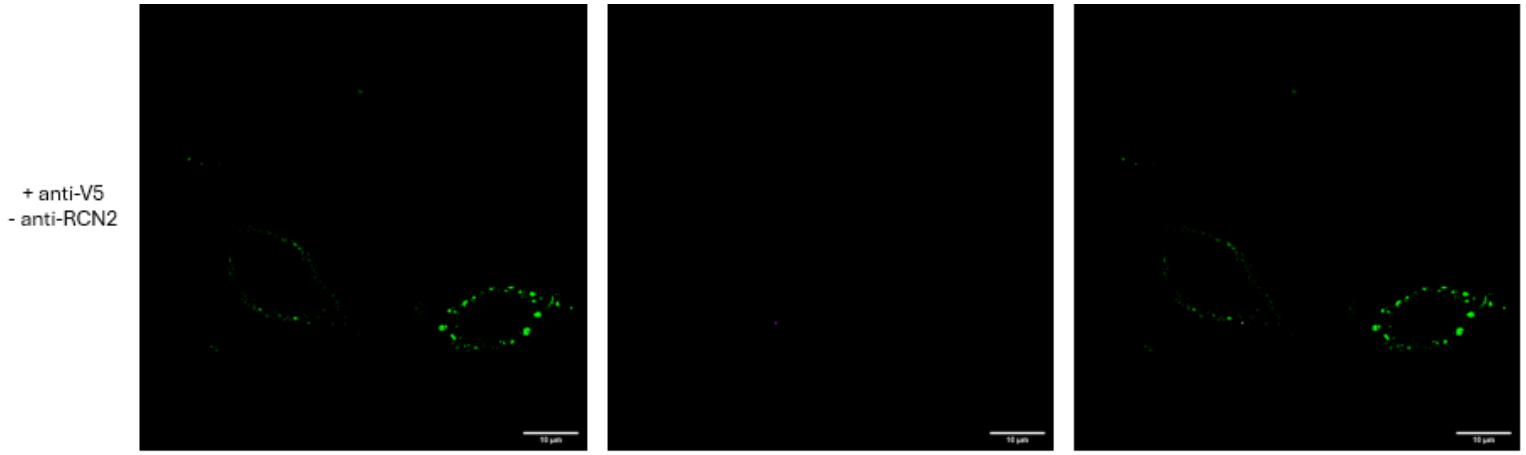
Figure 3.10. Representative structured illumination microscopy (SIM) images illustrating the co-occurrence of APEX2-NES enzymatic activity (purple) with DAPI (blue) in APEX2-NES HEK293T cells. (A) Cells were incubated with or without biotin-phenol (Btn) prior to exposure to H_2O_2, which catalyses APEX2-mediated biotinylation. Following this, cells were stained with various combinations of Streptavidin AlexaFluor (Str), Phalloidin and DAPI, enabling the assessment of APEX2-NES' enzymatic activity distribution relative to the nucleus. Scale bar: 10 µm. Magnification: 63X. (B) Mander's co-existence scores of manually delineated objects from images taken at the mid-section of the nucleus. M1 indicates the degree of overlap of the streptavidin channel with the DAPI channel, while M2 represents the extent of DAPI's overlap with the streptavidin channel. The coefficient scores were calculated from 57 cells, taken from 1 biological repeat (N=1), with error bars indicating SD.

3.9. Lack of Overlap Between ER Protein RCN2 and the APEX2-ERM Epitope Tag in APEX2-ERM A549 Cells

Having confirmed that the APEX2-ERM epitope tag is generally localised to the ER in APEX2-ERM HEK293T cells, the next step was to determine whether APEX2-ERM A549 cell lines exhibited similar localisation patterns. SIM was employed to assess the overlap between the V5 epitope tag of APEX2-ERM and RCN2 in A549 cells (Figure 3.11A). In this experiment, M1 reflects the degree of overlap of the V5 channel onto the RCN2 channel, while M2 represents the degree of overlap of the RCN2 channel onto the V5 channel. The co-occurrence scores (M1 = 0.1194 ± 0.1097, M2 = 0.3903 ± 0.1898) indicated that the V5 epitope tag was generally not confined to the ER membrane (Figure 3.11B). The higher M2 score compared to the M1 score suggests that a larger proportion of the ER membrane overlaps with the V5 epitope tag than *vice versa*. This may suggest that the epitope tag is not restricted to the ER. The anti-V5 negative control displayed a notable level of signal, albeit much lower than that

observed in the experimental conditions. This required a higher minimum threshold for co-occurrence analysis to account for and exclude non-specific signals.


A.
V5
RCN2 (ER marker)
Overlay
+ anti-V5
+ anti-RCN2
- anti-V5
+ anti-RCN2


+ anti-V5
- anti-RCN2

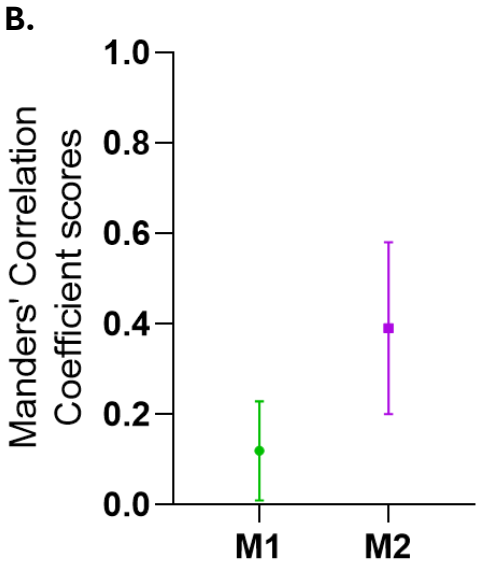
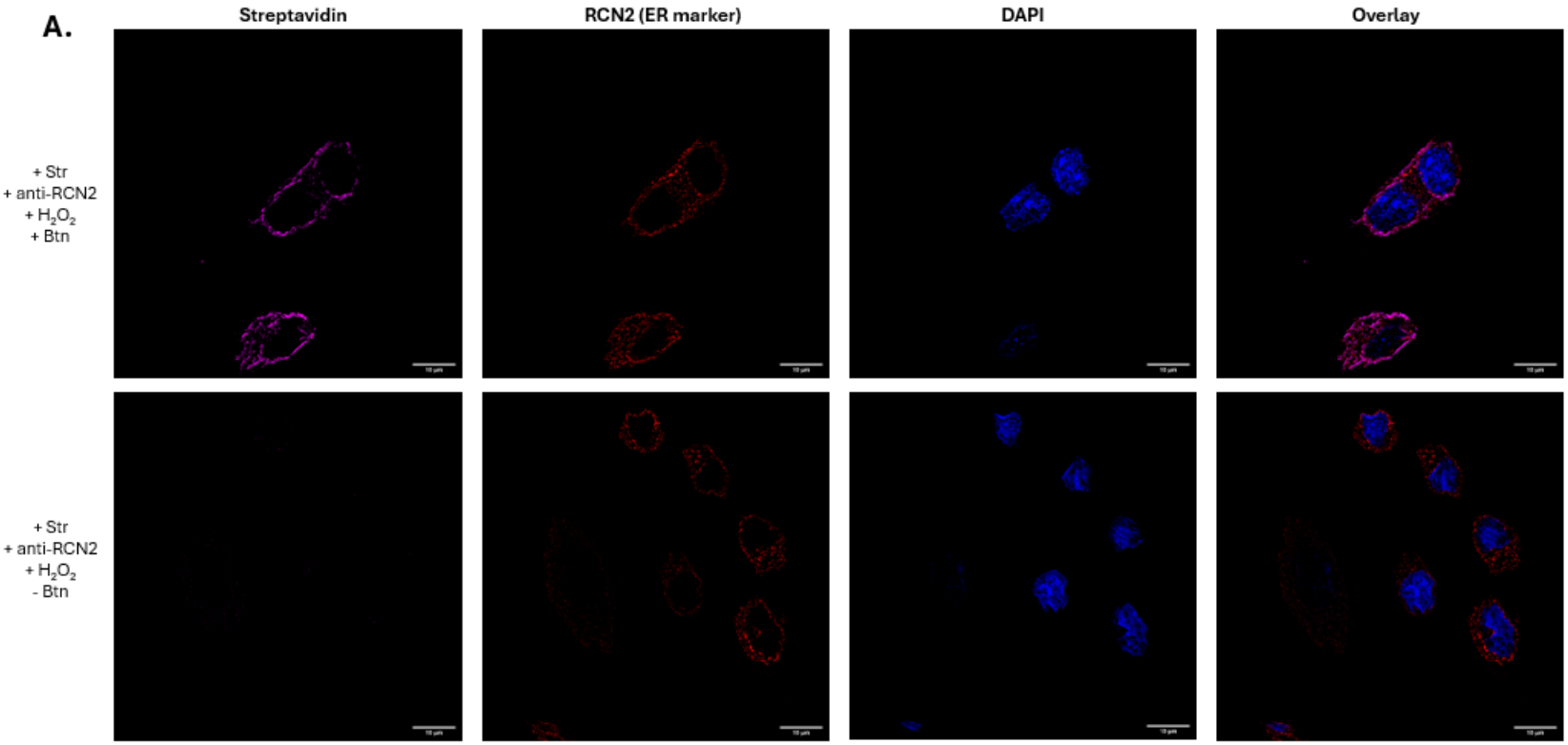

B.
Manders' Correlation Coefficient scores
1.0
0.8
0.6
0.4
0.2
0.0
M1
M2

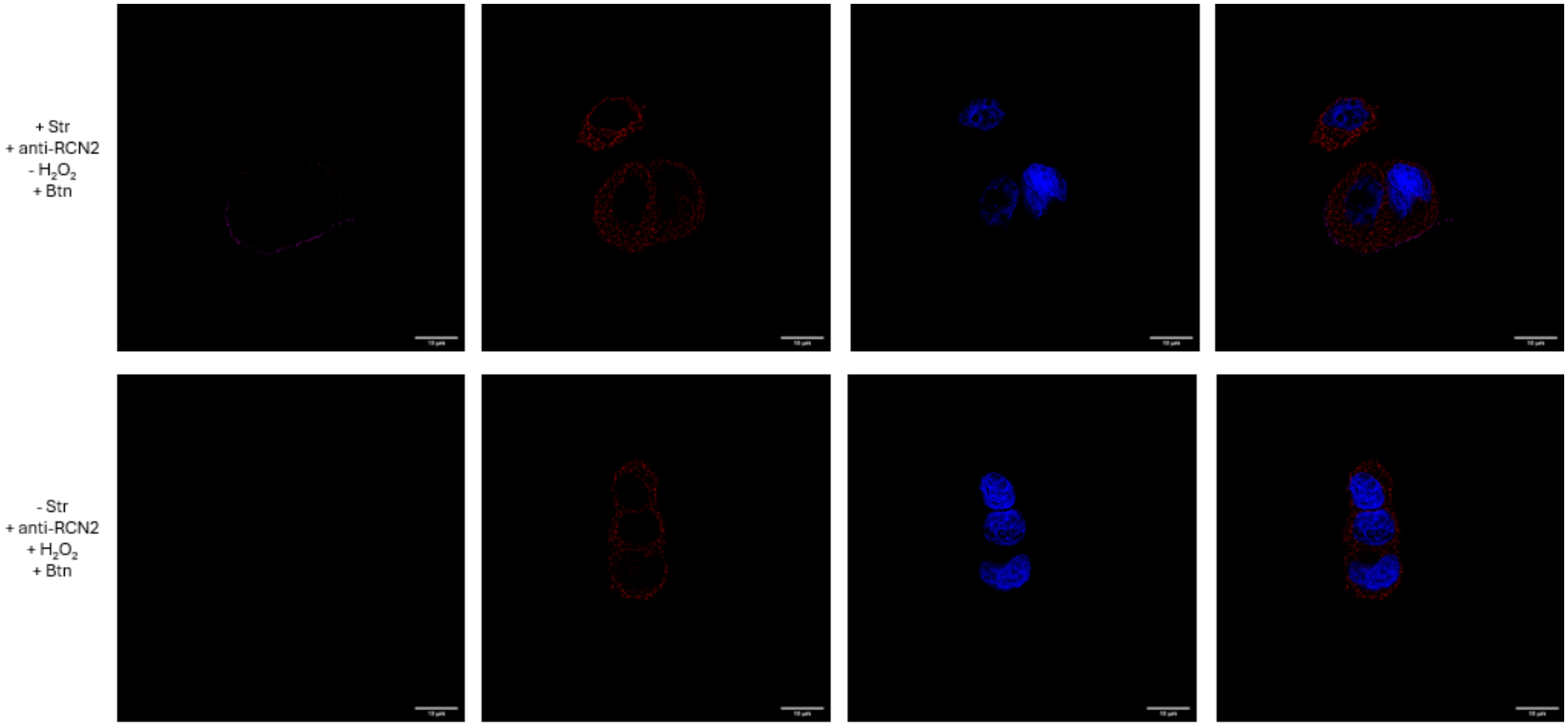
Figure 3.11. Representative structured illumination microscopy (SIM) images illustrating the co-occurrence of the V5 epitope tag of APEX2-ERM (green) with RCN2, an endoplasmic reticulum (ER) membrane protein (purple), in APEX2-ERM A549 cells. (A) Cells were fixed and stained with either anti-V5, anti-RCN2, or both to visualise the distribution of APEX2-ERM relative to the ER. Scale bar: 10 µm. Magnification: 63X. (B) Mander's co-existence scores of manually delineated objects from images taken at the mid-section of the nucleus. M1 indicates the degree of overlap of the V5 channel with the RCN2 channel, while M2 represents the extent of RCN2's overlap with the V5 channel. The coefficient scores were calculated from 26 cells, taken from 1 biological repeat, with error bars indicating SD.

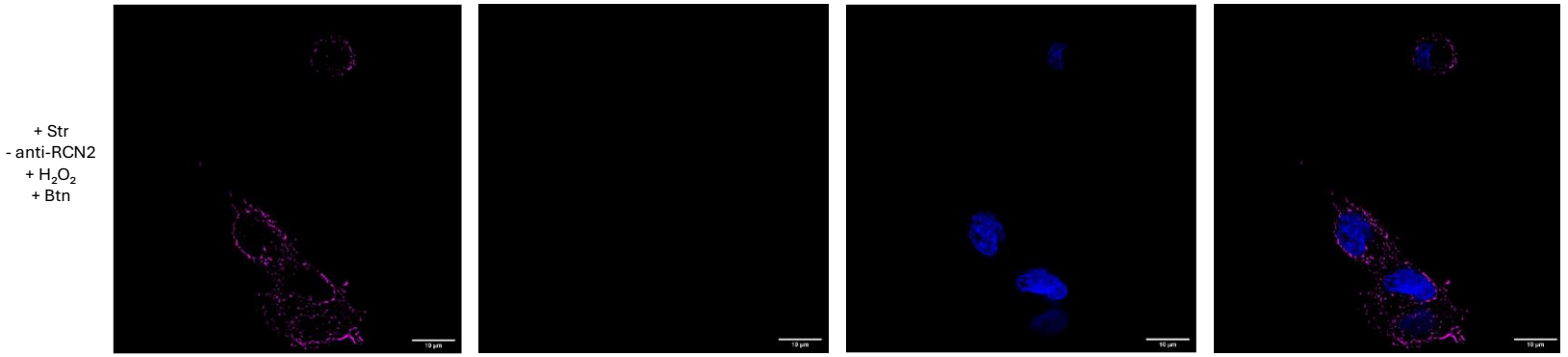
3.10. Overlap of APEX2 enzymatic activity with ER protein marker RCN2 in APEX2-ERM A549 cells

Given that the SIM data presented in Figure 3.11 revealed weak Manders co-occurrence scores for APEX2-ERM's V5 epitope and RCN2, it was necessary to investigate the overlap between APEX2-ERM enzymatic activity and the ER membrane in APEX2-ERM A549 cells. A biotinylation reaction was carried out by incubating APEX2-ERM A549 cells with biotin-phenol followed by exposure to H_2O_2, as described previously for APEX2-ERM HEK293T cells in Section 3.6. In this experiment, M1 reflects the degree of overlap of the Streptavidin channel onto the RCN2 channel, while M2 represents the degree of overlap of the RCN2 channel onto the Streptavidin channel. The cells were quantified to determine the proportion exhibiting detectable enzymatic activity.

Analysis of SIM images (Figure 3.12A) indicated striking differences in M1 and M2 scores, with M1 = 0.5584 ± 0.05264 and M2 = 0.8908 ± 0.08908 (Figure 3.12B). The streptavidin signal was generally not restricted to the ER, as indicated by the weak M1 score. However, the majority of the RCN2 marker appeared to overlap with the streptavidin signal, as reflected by the strong M2 score. This discrepancy may indicate that APEX2-mediated biotinylation was occurring both at the ER and outside the ER, pointing to overexpression of APEX2-ERM in these cells.


A.
Streptavidin
RCN2 (ER marker)
DAPI
Overlay
+ Str
+ anti-RCN2
+ H_2O_2
+ Btn
+ Str
+ anti-RCN2
+ H_2O_2
- Btn


+ Str
+ anti-RCN2
- H_2O_2
+ Btn
- Str
+ anti-RCN2
+ H_2O_2
+ Btn


+ Str
- anti-RCN2
+ H_2O_2
+ Btn

B.

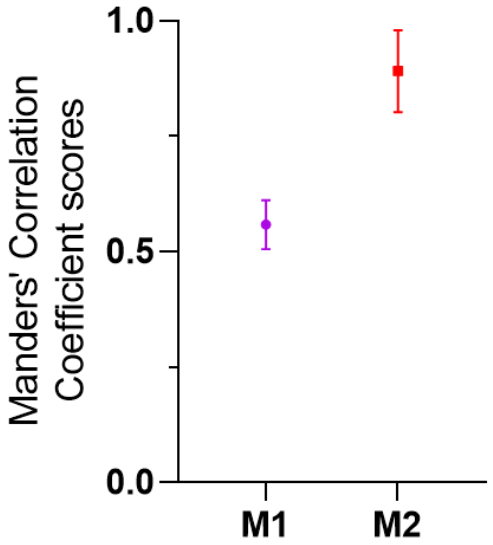

Manders' Correlation
Coefficient scores
1.0
0.5
0.0
M1
M2

Figure 3.12. Representative structured illumination microscopy (SIM) images illustrating the co-occurrence of APEX2-ERM enzymatic activity (red) with RCN2 (purple) in APEX2-ERM A549 cells. (A) Cells were incubated with or without biotin-phenol (Btn) prior to exposure to H_2O_2, which catalyses APEX2-mediated biotinylation. Following this, cells were stained with Streptavidin AlexaFluor (Str), anti-RCN2, or both, allowing for the visualisation of APEX2-ERM enzymatic activity relative to the ER. Scale bar: 10 µm. Magnification: 63X. (B) Mander's co-occurrence coefficient of manually delineated objects from images taken at the mid-section of the nucleus. M1 indicates the degree of overlap of the streptavidin channel with the RCN2 channel, while M2 represents the extent of RCN2's overlap with the streptavidin channel. The coefficient scores were calculated from 25 cells, taken from 1 biological repeat, with error bars indicating SD.

3.11. SIM imaging reveals differential biotinylation efficiency in APEX2-expressing cell lines

As detailed in Sections 3.6, 3.8 and 3.10: APEX2-ERM HEK293T, APEX2-NES HEK293T, and APEX2-ERM A549 cells were incubated with biotin-phenol and H_2O_2 to assess APEX2 enzymatic activity. SIM images revealed considerable variation in biotinylation signals within each mixed cell population, with only a subset of cells displaying streptavidin-positive signals once the minimum intensity threshold for Manders' co-occurrence coefficient analysis had been applied. The proportion of streptavidin-positive cells also differed across the three cell lines (Figure 3.13): APEX2-NES HEK293T showed the highest proportion at 85%, followed by APEX2-ERM A549 at 64%, and APEX2-ERM HEK293T at 45%.

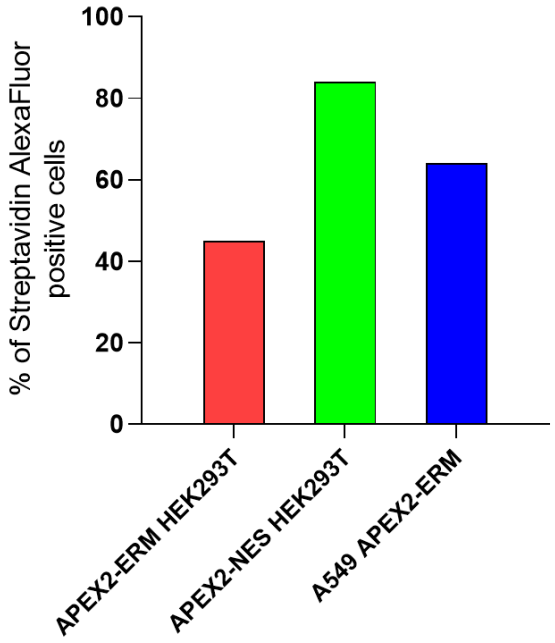


Figure 3.13. Comparison of Streptavidin-Positive Cell Proportions Across APEX2 Cell Lines. APEX2-ERM HEK293T, APEX2-NES HEK293T, and APEX2-ERM A549 were incubated with biotin-phenol and H_2O_2 prior to paraformaldehyde fixation and then cells were stained with Streptavidin-Alexa Fluor to detect biotinylation. The proportion of streptavidin-positive cells was quantified across the three cell lines,

employing the same minimum intensity thresholds used for Manders' co-occurrence coefficient score analysis. For APEX2-ERM HEK293T, 51 cells were counted; for APEX2-NES HEK293T, 107 cells were counted; and for APEX2-ERM A549, 39 cells were counted. Data were derived from one independent experiment for each cell line.

3.12. Discussion

This chapter details the successful generation of stable HEK293T and A549 cell lines expressing APEX2 targeted to either the ER membrane (APEX2-ERM) or the cytosol (APEX2-NES) using LV transduction. The expression levels of APEX2 were confirmed through western blotting and RT-qPCR, while the enzymatic activity of APEX2 on RNA was validated using RNA dot blot assays. Super-resolution imaging was employed to examine the localisation of the APEX2 epitope tag and the distribution of APEX2-mediated biotinylation in relation to the intended subcellular compartment. APEX2-ERM HEK293T exhibited the most favourable co-occurrence scores when investigating both overlap between APEX2-ERM's epitope tag and the ER as well as overlap between APEX2's enzymatic activity and the ER. This led to the decision to focus on this cell line in the following chapters.

The high levels of APEX2-ERM protein expression observed by western blot and SIM co-occurrence analysis of APEX2-ERM A549 raised concerns as to the potential impact on ER function due to APEX2-ERM overexpression. To investigate the possibility that ER function is disrupted in APEX2-ERM cell lines, functional assays or higher resolution imaging could be employed. For instance, single molecule fluorescence in situ hybridisation combined with selective permeabilisation using digitonin could be used to assess whether ER-targeted mRNAs remain anchored to the ER (Wu & Palazzo, 2020). Alternatively, transmission electron microscopy imaging could be employed to assess APEX2-ERM overexpression by providing insights into morphological changes, such as excessive folding of the ER membrane, that are indicative of APEX2 protein overexpression at the ER (Lam et al., 2014; Martell et al., 2017).

To mitigate overexpression concerns, a controlled dilution of LV stock was used to generate APEX2-NES cell lines. The lowest effective concentration of LV particles was determined based on successful APEX2 expression (Figure 3.4

and Figure 3.5), antibiotic resistance, and appropriate growth rates. Interestingly, the lowest effective concentration of APEX2-NES LV particles resulted in similar APEX2 mRNA expression levels to those observed in APEX2-ERM cell lines by RT-qPCR. Finer control over gene transfer levels could be attained by titrating LV particles using quantitative real-time PCR or digital droplet PCR (Corre et al., 2022; Scherr et al., 2001). However, even if transduction efficiency were to be more tightly controlled, there remains a possibility that certain cells receive several copies of the APEX2 gene, while others receive fewer copies, leading to inconsistent expression levels (Woods, 2003). An alternative approach would be to employ CRISPR knock-in techniques to either tag endogenous proteins with APEX2 at their native loci or introduce genetic material encoding APEX2 fusion proteins at well-established loci, such as the AAVS1 loci. The latter approach would ensure a controlled maximum number of copies per cell, which could vary depending on the chromosomal structure, given that tissue-cultured cells often do not maintain the typical diploid number of chromosomes. Both strategies could provide tighter control over expression levels and therefore might influence the extent to which APEX2 localisation is restricted to the intended subcellular region.

In the APEX2-ERM A549 cell line, the no V5 negative control displayed high background that was not observed when investigating the APEX2-ERM HEK293T cell line. While this background was eliminated for co-occurrence analysis of the experimental condition, it is possible that low but genuine signals were lost which may have affected the Manders co-occurrence analysis. The comparison between the proportion of FLAG-positive APEX2-NES cells detected during the co-occurrence analysis with phalloidin (52%) and the proportion of streptavidin-positive APEX2-NES HEK293T cells observed in the SIM imaging of APEX2's enzymatic activity with Streptavidin AlexaFluors (85%) is striking. Phalloidin staining appeared diffuse when APEX2-NES HEK293T cells were exposed to H_2O_2 (Figure 3.10), raising concerns about the validity of phalloidin as a cytosolic marker under these experimental conditions. Exploring

alternative cytosolic markers, such as anti β-actin, and alternative anti-FLAG antibodies could potentially provide more valuable insights.

SIM imaging revealed a stronger overlap between APEX2-mediated enzymatic activity and the ER membrane compared to the overlap between the APEX2-ERM epitope tag and the ER membrane in both APEX2-ERM cell lines. Interestingly, imaging of APEX2-mediated enzymatic activity using streptavidin-AlexaFluors raised the possibility that APEX2-mediated biotinylation may not be restricted to the ER membrane in APEX2-ERM A549 cells (Figure 3.12). While this observation could potentially be attributed to biotin adducts diffusing away from APEX2's active site during the labelling or the subsequent 10-minute fixation period, the strong co-occurrence observed in APEX2-ERM HEK293T cells under the same experimental conditions (Figure 3.8) makes this explanation unlikely. A more plausible reason is that the overexpression of APEX2-ERM in APEX2-ERM A549 may explain this observed phenomenon. Further investigations comparing the overlap between APEX2-labelled molecules and APEX2's epitope tag would provide valuable insights into the extent of biotinylated molecule diffusion during labelling and fixation.

To draw more definitive conclusions about the extent to which APEX2 localises to its intended subcellular compartment, additional controls for SIM imaging should be implemented. These could include wild-type cells as negative controls to assess background expression of the epitope tag, the use of immunofluorescence markers for both the ER and cytoplasm in both APEX2-ERM and APEX2-NES cell lines to improve compartment distinction, and the use of an ER immunofluorescence marker specific to the rough ER, the site of SRP-dependent mRNA targeting and translation.

The variation in the proportion of APEX2-expressing cells found to be streptavidin-positive, following treatment with biotin-phenol and H_2O_2, across the three APEX2 cell lines during SIM imaging (Figure 3.13) is noteworthy. This inconsistency may be due to insufficient cell membrane permeability of biotin-phenol, as suggested by previous studies (Hwang & Espenshade, 2016; B. Tan et al., 2020). Alternatively, this variation may reflect differences in the

proportion of APEX2 that is enzymatically active across the cell lines. To address this issue, further independent replicates are necessary. Intracellular flow cytometry using Streptavidin AlexaFluor conjugates could be used to optimise APEX2 labelling conditions for future APEX2 studies.

A limitation of the SIM experiments was the impact of poly-D-lysine (PDL) treatment on cell morphology. In this study, HEK293T cells required PDL treatment to prevent detachment from glass coverslips. Although A549 cells could adhere without PDL, they were easily dislodged during H_2O_2 treatment, which was necessary to assess the localisation of APEX2's enzymatic activity. To maintain consistency across both epitope and enzymatic co-occurrence studies, PDL treatment was applied in all SIM experiments. PDL-treated cells tend to flatten (Mazia et al., 1975), which complicates Z-plane imaging and limited the assessment of the full three-dimensional structure of the ER, that extends above and below the nucleus. As a result, whilst multiple Z planes of cells were imaged, only one Z plane was for image analysis, which might not be representative of the whole ER.

The cell lines generated in this chapter lay the groundwork for further studies into APEX2 labelling of endogenous and exogenous RNA in the cytosol and ER. The localisation of APEX2's epitope tag and enzymatic activity in the APEX2-ERM HEK293T cell line provided the greatest confidence compared to other APEX2 cell lines, establishing it as the primary focus for exploring APEX2 in assessing mRNA delivery.

Chapter 4 - Optimisation of streptavidin affinity pulldown for APEX2-RT-qPCR

4.1. Introduction

Following the generation of APEX2-ERM and APEX2-NES cell lines in HEK293T and A549, as detailed in Chapter 3, the next objective was to evaluate the efficacy of APEX2 in labelling delivered mRNA within these cells.

The pioneering work of Fazal et al. (2019) from the Ting lab and Padrón et al. (2019) from the Ingolia lab advanced the field of subcellular spatial transcriptomics by utilising APEX2 for direct RNA biotinylation. The Ting lab had previously developed an approach known as APEX-RIP (RNA immunoprecipitation), where APEX2 is used to label proximal proteins that are then cross-linked to their associated RNA (Kaewsapsak et al., 2017). RNA linked to labelled proteins are then isolated from the RNA:protein complexes by streptavidin pulldown. However, Kaewsapsak et al. (2017) noted limitations with APEX-RIP, particularly when APEX2 was fused to bait proteins localising to non-membrane enclosed regions.

Building on this, Fazal et al. (2019) demonstrated the first successful use of APEX2 for direct labelling of RNA. APEX2, localised to the mitochondrial matrix (APEX2-mito), was employed under standard biotinylation conditions to biotinylate RNA. The labelled RNA was then isolated using streptavidin pulldown and analysed *via* RT-qPCR to validate the spatial proximity of APEX2 RNA labelling. Fazal et al. (2019) carried out APEX-seq in various APEX2 fusion cell lines to map the spatial distribution of RNA across multiple organelles. APEX2-seq data from the APEX2-NES cell line was used as a normalisation control for both the APEX2-RT-qPCR data obtained from APEX2-mito cells and all subsequent APEX-seq data in the various APEX2-fusion cell lines (Fazal et al., 2019).

In contrast, the Ingolia lab focused on validation of APEX-seq using an APEX2-GFP (cytosolic) cell line and an APEX2-ERM cell line (Padrón et al., 2019). They

first employed an *in vitro* APEX2 biotinylation technique to validate direct RNA labelling *via* streptavidin dot blot, and bypassed cellular validation of direct APEX2-mediated RNA labelling by RT-qPCR. Instead, they utilised APEX-seq to confirm APEX2-mediated biotinylated RNA enrichment above background levels in open cellular compartments. During the validation stage of their APEX-seq protocol, Padrón et al. (2019) plotted the correlation between biological replicate samples from the APEX2-seq experiments, demonstrating good reproducibility for both the APEX2-GFP and APEX2-ERM cell lines. APEX-seq data from the APEX2-GFP cell line was used as a normalisation control both during APEX-seq validation in the APEX2-ERM cell line and when addressing more biologically complex questions.

Both studies (Fazal et al., 2019; Padrón et al., 2019) utilised similar normalisation methods by relying on cytosolic APEX2 labelling data for comparison. Fazal et al. (2019) noted that cytosolic APEX2 labelling was less accurate than anticipated, which likely influenced its use as a control for normalisation. However, neither study addressed potential biases introduced by this normalisation approach, nor did they consider alternatives such as normalisation to no biotin-phenol or no H_2O_2 negative controls.

Zhou et al. (2019) sought to enhance APEX2-mediated RNA labelling by exploring alternative APEX2 biotinyl substrates beyond the conventional biotin-phenol compound, aiming to identify substrates that react with less selectivity towards electron-rich biological molecules. Using an *in vitro* assay, they discovered that biotin-aniline preferentially labels RNA over proteins and DNA, in contrast to biotin-phenol, which selectively labels proteins over other biological molecules. Building on this, Zhou et al. (2019) then worked to improve APEX2-RT-qPCR and APEX-seq protocols by utilising biotin-aniline as the APEX2 substrate in a mitochondrial APEX2 cell line. Their methodology involved normalising data relative to pre-enrichment levels and employing no biotin-aniline negative controls, illustrating a different approach to account for the recovery of non-biotinylated RNA post-streptavidin pulldown (Zhou et al., 2019).

In this chapter, initial attempts at APEX2 labelling followed by RT-qPCR did not demonstrate the expected specific enrichment of mRNA. Despite reproducing published methods, variability in the localisation of both endogenous RNA and exogenously delivered mRNA anticipated to co-localise with APEX2 became apparent. These inconsistencies underscored the limitations of the existing streptavidin pulldown protocols for isolating APEX2-labelled mRNA and the necessity for refinement.

To address these challenges, this chapter then details the development of an improved streptavidin affinity pulldown protocol, refined by comparing the enrichment of a 3'-end biotin-labelled spike-in RNA to that of a non-biotinylated spike-in RNA, with detection of enrichment post-streptavidin pulldown achieved through RT-qPCR. Incorporating the 3' end biotin-labelled spike-in RNA in the streptavidin pulldown enables internal normalisation for RT-qPCR. This advancement represents a significant improvement to the confidence in reproducibility of APEX2-mediated RNA labelling experiments.

4.2. Variability in APEX2-RT-qPCR results for endogenous RNA using existing streptavidin pulldown methods

The APEX2 labelling protocol developed by the Ting lab (Fazal et al., 2019) was employed to determine whether APEX2-ERM HEK293T and APEX2-NES HEK293T cells can effectively label endogenous mRNA localised to the cytosolic facing surface of the ER and cytoplasm, respectively. This method utilised the APEX2 streptavidin pulldown technique developed by Fazal et al. (2019) to evaluate the enrichment of selected endogenous mRNAs, comparing their localisation as detected by RT-qPCR, to the predicted subcellular localisation reported by Fazal et al. (2019), as detected by RNA-seq.

Biotin-aniline was used as the APEX2 RNA labelling substrate, as the Zou lab demonstrated it to be superior for RNA labelling compared to other biotin-conjugated aromatic substrates (Zhou et al., 2019). A negative control, where biotin-aniline was omitted, was also included. SSR2, SFT2D2, and TMX1 were selected as ER-associated mRNA markers, having been identified as enriched

in APEX2-ERM HEK293 through APEX-seq (Fazal et al., 2019). Similarly, FAU1 was included as a cytosolic RNA, as it was found to be enriched in the APEX2-NES HEK293 cell line through APEX2-seq (Fazal et al., 2019). MT-CO2, a mitochondrial transcript encoded by mitochondrial DNA (Anderson et al., 1981), was reported to be enriched after APEX2 labelling in APEX2-mito HEK293 cells, where APEX2 is fused to a bait protein that directs the APEX2 fusion protein to the mitochondrial matrix (Fazal et al., 2019). Therefore, MT-CO2 served as a negative control, as it was not expected to show enrichment after APEX2 labelling in either the APEX2-ERM HEK293T or APEX2-NES HEK293T cell lines. Lastly, 18S rRNA was included as a positive control, as it is anticipated to be enriched relative to the no-biotin-aniline negative controls in both APEX2-ERM and APEX2-NES cell lines.

To analyse the data, RT-qPCR was performed on both the RNA post streptavidin pulldown (referred to as the "output RNA"), and 2% of the RNA pre streptavidin pulldown (referred to as the "input RNA"). The output qPCR data for each mRNA were then normalised to the input RNA, expressed as fold change relative to the no biotin-aniline negative control.

The RT-qPCR analysis of ER-associated RNAs identified by Fazal et al. (2019), revealed varying degrees of enrichment in the APEX2-ERM HEK293T cells (Figure 4.1A). Among these, SSR2 exhibited the most notable enrichment, although this result was not statistically significant due to the small sample size (N=2). SFT2D2 showed a small but detectable level of enrichment. TMX1, despite being predicted to be enriched, did not display enrichment relative to the no-biotin-aniline control.

The data exhibited considerable variability, with 18S rRNA showing inconsistent enrichment levels relative to the no-biotin control. This variability suggests potential issues with the reproducibility of the methodology. Additionally, unexpected enrichment relative to the no-biotin control was observed for MT-CO2 mRNA. This surprising result, combined with the variability seen with 18S rRNA, raises concerns about the specificity and reliability of the experiment.

The same approach was subsequently tested in APEX2-NES HEK293T cells, where APEX2 is expected to label cytosolic RNA (Figure 4.1B). In this context, most RNAs examined by RT-qPCR exhibited varying degrees of enrichment over the no-biotin aniline control. This result was observed regardless of whether the RNAs were predicted to localise to the cytosol or the ER, contrasting with the findings from the APEX2-ERM cells.

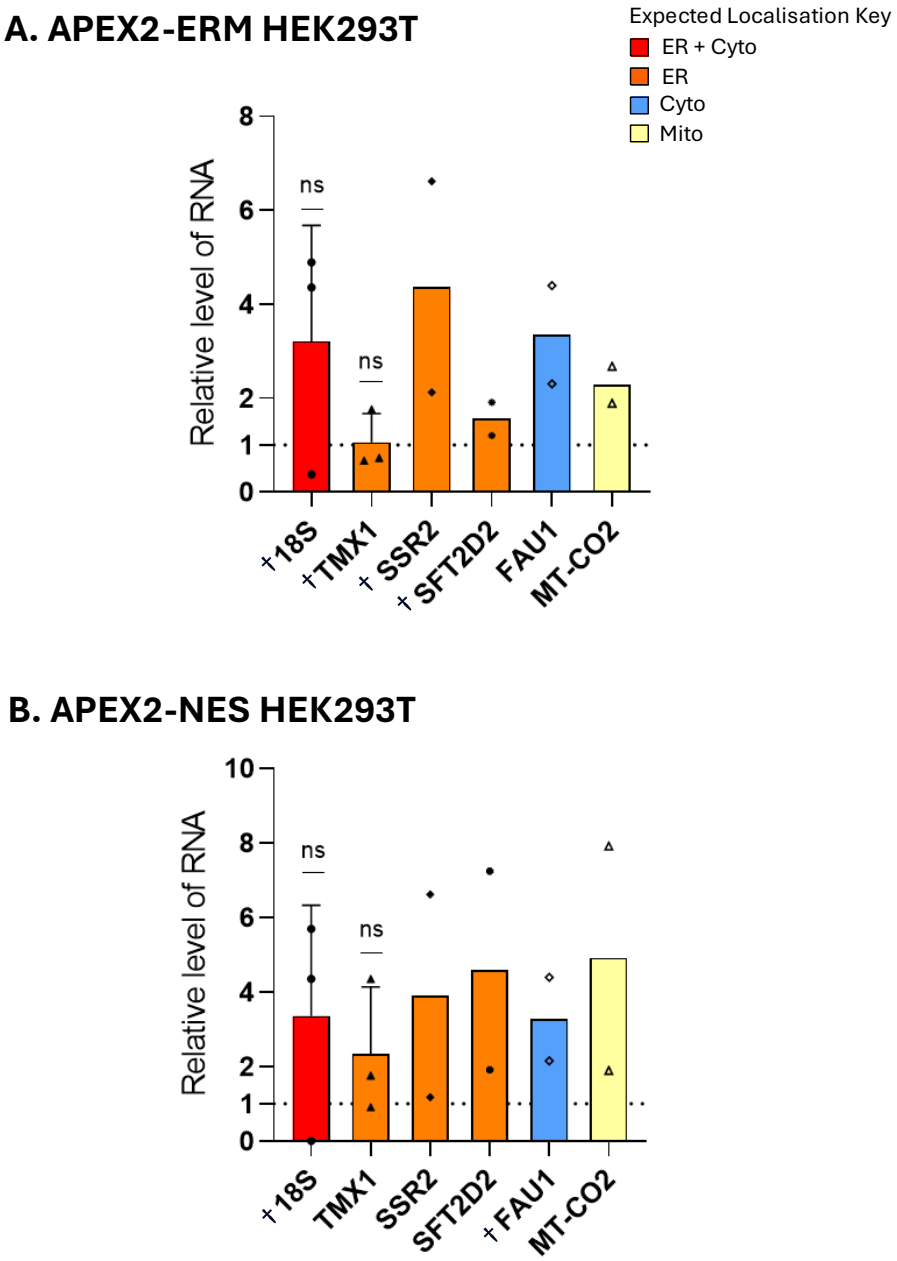


Figure 4.1. APEX2-RT-qPCR results for the enrichment of endogenous genes in APEX2 HEK293T cells using the Ting lab protocol (Fazal et al., 2019). (A.) shows the enrichment of mRNA in the APEX2-ERM HEK293T cell line, which is expected to localise to the cytosolic-facing side of the ER. (B.) shows the enrichment of mRNA in the APEX2-NES HEK293T cell line, which is expected to be diffused throughout the

cytosol. Streptavidin enrichment for each RNA was normalised to the input RNA and is presented relative to the no biotin-aniline control. The key indicates the predicted subcellular localisation of each transcript, and † denotes that the transcript is expected to be enriched in the respective cell line. Where N ≥ 3, bars represent the mean of three independent experiments with error bars showing standard deviation (SD). Where N = 2, bars represent the mean of two experiments. Statistical significance is denoted as ns = not significant.

The observed inconsistencies in APEX2-RT-qPCR results for endogenous mRNA in the APEX2-ERM and APEX2-NES HEK293T cells suggest potential issues with the protocol developed by Fazal et al. (2019). Interestingly, Chapter 3 demonstrated that APEX2 and its enzymatic products correctly localise to their intended subcellular regions in both HEK293T cell lines. Therefore, it is plausible that the protocol may be ineffective in consistently isolating biotinylated RNA. Consequently, the following section will investigate the APEX2 protocol developed by Ingolia lab (Padrón et al., 2019) to examine whether it offers a more consistent approach to selective enrichment of APEX2-mediated biotinylated RNA.

Ingolia's APEX2 protocol (Padrón et al., 2019) differs from the Ting protocol (Fazal et al., 2019) primarily in the streptavidin pulldown method, particularly in the cleaning and elution processes. Rather than using a NaOH solution to remove potential RNase or residual RNA from streptavidin magnetic beads before incubation with input RNA, the method described by Padrón et al. (2019) employs a viscous solution of $MgCl_2$, SDS and sodium deoxycholate. Additionally, the protocol developed by Padrón et al. (2019) incorporates a blocking solution of Poly(I:C) and Denhardt's solution, composed of Ficoll (type 400), polyvinylpyrrolidone, and BSA, both before and during the incubation of input RNA with strepavidin beads. For elution of biotinylated RNA post-streptavidin pulldown, Padrón et al. (2019) applied Trizol directly to the beads instead of enzymatically digesting the streptavidin using Proteinase K, as was carried out by Fazal et al. (2019).

The enrichment of transcripts tested by RT-qPCR in APEX2-ERM HEK293T cells using the protocol developed by Padrón et al. (2019) displayed substantial variability across biological replicates (Figure 4.2A). For instance, 18S rRNA showed an enrichment of 5.8 in one experiment compared to the no-biotin-aniline control but was lower than the negative control in three other replicates.

Notably, FAU1 mRNA, which encodes a ribosomal protein, was the most enriched transcript, despite being used as a cytosolic control transcript by Fazal et al. (2019).

Similarly, enrichment results for genes tested by RT-qPCR in APEX2-ERM HEK293T cells using the protocol developed by Padrón et al. (2019) varied significantly across biological repeats (Figure 4.2B). Mitochondrial RNA exhibited high enrichment in some replicates, contradicting earlier results obtained with the protocol. FAU1 showed moderate enrichment in one of two replicates, while TMX1, an ER-associated RNA, displayed no enrichment. Although these results are generally consistent with expectations, they diverge from previous findings using the original protocol.

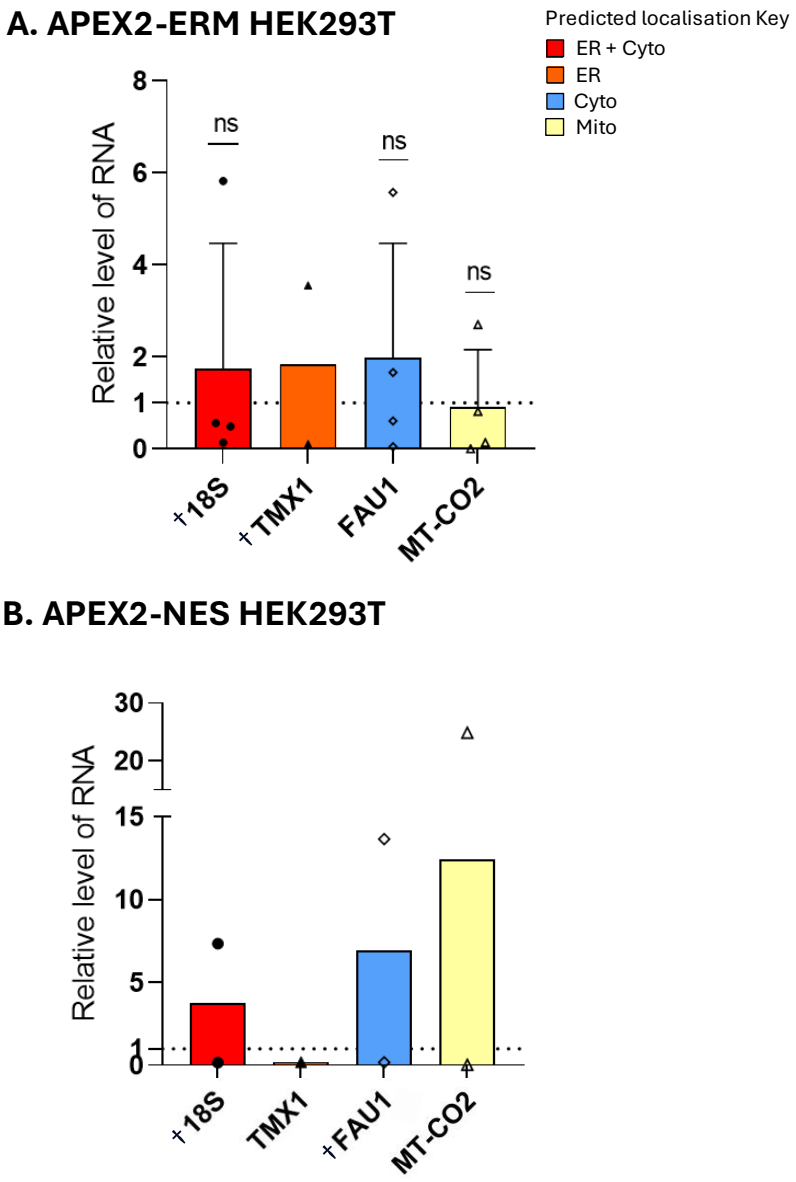


Figure 4.2. APEX2-RT-qPCR results for the enrichment of endogenous genes in HEK293T cells using the Ingolia lab protocol (Padrón et al., 2019). (A) shows the enrichment of mRNA in the APEX2-ERM HEK293T cell line, which is expected to localise to the cytosolic-facing side of the ER. (B) shows the enrichment of mRNA in the APEX2-NES HEK293T cell line, which is expected to be diffused throughout the cytosol. Streptavidin enrichment for each RNA was normalised to the input RNA and is presented relative to the no biotin-aniline control. The key indicates the predicted subcellular localisation of each transcript, and † denotes whether the transcript is expected to be enriched in the respective cell line. Where N ≥ 3, bars represent the mean of three independent experiments with error bars showing SD. Where N = 2, bars represent the mean of two experiments. Statistical significance is denoted as ns = not significant.

4.3. Variability in APEX2-RT-qPCR results for delivered mRNA using existing streptavidin pulldown methods

Given the variability observed with existing streptavidin pulldown protocols for endogenous RNA, described in section 4.2, efforts were made to test whether more reliable results could be obtained by transfecting IVT mRNA specifically

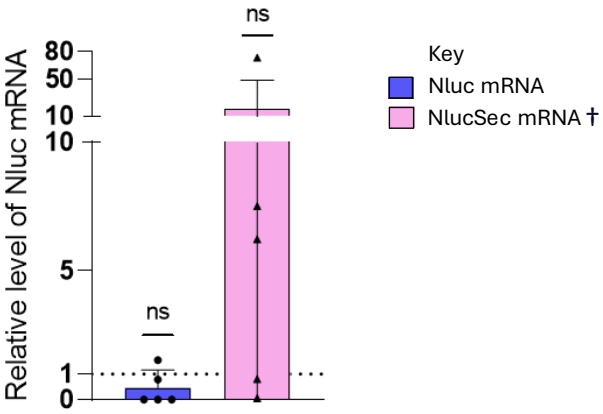
designed to localise to either the ER or the cytoplasm, while maintaining a shared coding sequence. Two plasmids were generated (Table 2), both encoding Nluc downstream of a T7 promoter and flanked by human β-haemoglobin UTRs, selected for their therapeutic relevance. The key difference between the plasmids was the inclusion of a signal peptide coding sequence from interleukin-6 (IL6), which directs the NlucSec mRNA to the ER *via* the signal recognition particle (SRP) pathway (Keenan et al., 2001), resulting in the secretion of the nanoluciferase protein. IVT was carried out to generate ARCA-capped, polyadenylated mRNA, which was then delivered into APEX2-ERM and APEX2-NES HEK293T cells *via* lipofectamine transfection (as detailed in 2.4.1). Four hours post-transfection, cells were treated with biotin-aniline for 30 minutes, treated with H_2O_2 for 1 minute and then total RNA was extracted. In this section, the streptavidin affinity purification protocol developed by Padrón et al. (2019) was used to isolate APEX2-labelled mRNA. As in Figure 4.2, a no biotin-aniline negative control was included and data was analysed performing RT-qPCR on both the output RNA and 2% of the input RNA before streptavidin pulldown. The output qPCR data for each mRNA were then normalised to the input RNA, expressed as fold change relative to the control.

NlucSec mRNA showed inconsistent enrichment in the streptavidin pulldown after APEX2 labelling in APEX2-ERM HEK293T cells compared to the no biotin-aniline control (Figure 4.3A). In some experiments, no enhancement was observed, while in others, an enrichment of up to 60-fold was detected. In contrast, cytosolic Nluc mRNA in APEX2-ERM HEK293T cells did not exhibit any enrichment in the streptavidin pulldown following APEX2 labelling when compared to the no biotin-aniline control.

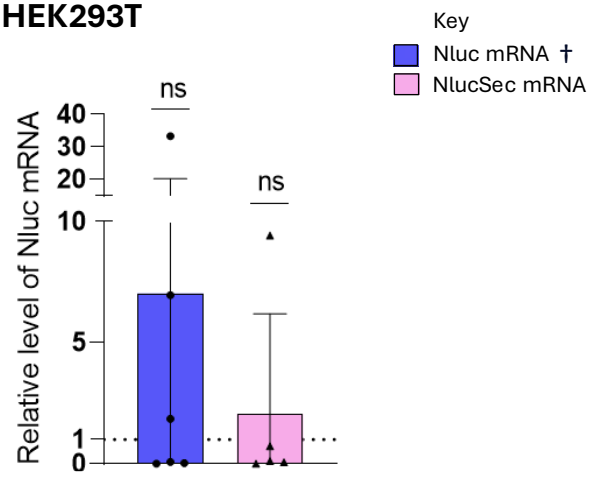
Both cytosolic Nluc and NlucSec mRNA showed selective enrichment post-streptavidin pulldown in the APEX2-NES HEK293T cell line relative to the no biotin-aniline negative control, though with notable variability across biological replicates (Figure 4.3B). Interestingly, cytosolic Nluc mRNA exhibited stronger enrichment in some biological repeats compared to NlucSec, aligning with its expected cytosolic localisation. Despite this, the enrichment of NlucSec in

APEX2-NES HEK293T cells remained high in multiple biological repeats, raising concerns about the specificity of the APEX2-NES system. This resulted in a shift in focus towards the use of the APEX2-ERM system to better investigate ER-specific delivered mRNA trafficking. Nevertheless, the variability across biological repeats once again highlighted the importance of refining the pulldown protocol to improve the consistency of the APEX2-RT-qPCR technique.

A. APEX2-ERM HEK293T



B. APEX2-NES HEK293T



C.

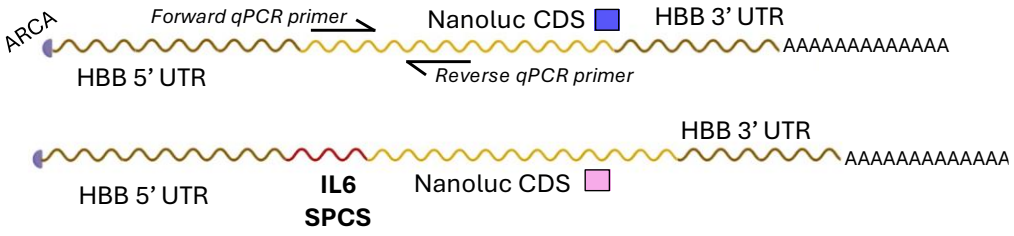


Figure 4.3. APEX2-RT-qPCR results for the enrichment of transfected Nluc mRNA in HEK293T cells using the streptavidin pulldown protocol developed by the Ingolia lab (Padrón et al., 2019) and analysed by RT-qPCR. IVT mRNA was delivered into APEX2-ERM HEK293T cells by lipofectamine transfection and incubated for 4 hours before exposure to biotin-aniline and H_2O_2. Nluc and NlucSec mRNA levels post-streptavidin pulldown were quantified by RT-qPCR, normalised to the input RNA and are presented relative to the no biotin-aniline control. (A) Shows the enrichment of Nluc mRNA in the APEX2-ERM cell line, which is expected to localise to the cytosolic-facing side of the ER. (B) Shows the enrichment of Nluc mRNAs in the APEX2-NES HEK293T cell line. Bars represent the mean of at least three independent experiments, with error bars showing SD. Statistical significance is denoted as ns = not significant, based on Wilcoxon signed-rank tests. † in the key denotes whether the delivered mRNA is expected to be enriched in the respective cell line. (C) Shows an illustration of the two Nluc mRNAs.

4.4. Generation and validation of 3' end biotinylated spike-in RNA

Following the investigation of existing streptavidin pulldown protocols for isolating APEX2-mediated biotin-labelled mRNA, efforts were redirected towards developing a more consistent streptavidin pulldown method. Incorporating a biotinylated spike-in RNA was considered an effective strategy to optimise the streptavidin pulldown. The use of *in vitro* biotinylated spike-in RNA helps eliminate variability introduced by the efficiency of RNA labelling by APEX2 in cells. Additionally, *in vitro* biotinylation is expected to be more efficient than APEX2 labelling in cells, and using the same stock of spike-in RNA for each pulldown can be expected to enhance reproducibility across optimisation experiments.

To minimise potential interference with reverse transcription and ensure that the processivity of the reverse transcriptase remains unaffected by modified nucleotides, a 3’ end biotinylation approach using T4 ligase 1 was chosen. This method also allows precise control over the biotinylation process, limiting the biotin modification to a single site per RNA molecule, thereby standardising the assay.

To achieve this, an *in vitro* transcribed firefly luciferase (Fluc) RNA, with hepatitis C virus (HCV) UTRs flanking the coding sequence, was 3' end biotinylated using pCp-biotin and T4 ligase 1 following an adaptation of the approach described by Kore et al. (2009) (Figure 4.4A). The plasmid encoding Fluc was readily available in the lab (Roberts et al., 2011) and chosen

specifically because Fluc is not endogenously expressed in HEK293T cells, ensuring no interference from native RNA during the pulldown experiments.

Following the biotinylation reaction, the RNA was serially diluted in nuclease-free water and assessed by streptavidin-HRP dot blot to confirm successful biotinylation. As shown in Panel B of Figure 4.4, the 3' end biotinylation of Fluc RNA was confirmed. Additionally, this RNA dot blot assay helped estimate the appropriate amount of biotinylated spike-in RNA for use in streptavidin pulldown optimisation. Based on the results, it was determined that 1-0.1 ng of biotinylated RNA would be suitable input for streptavidin pulldown optimisation.

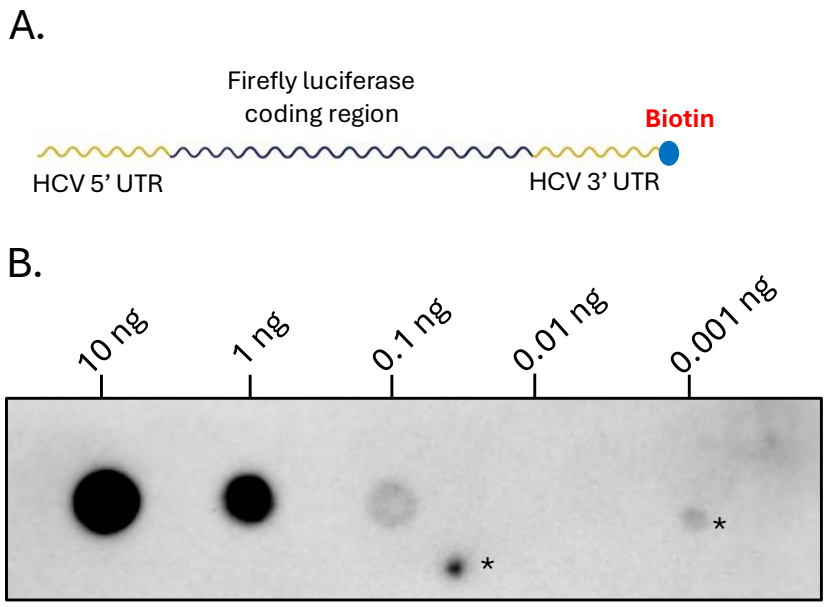


Figure 4.4. Generation and validation of 3' end biotinylated spike-in RNA. (A) Schematic illustration of 3' end biotinylated spike-in RNA where HCV stands for Hepatitis C Virus. (B) Streptavidin-HRP dot blot assay validating the biotinylation of spike-in Fluc IVT RNA, which was serially diluted in nuclease free water prior to analysis. * indicates cross-reactivity. N=1.

4.5. Optimisation of streptavidin pulldown using 3' end biotinylated spike-in RNA

To optimise the streptavidin pulldown, the 3' end biotinylated spike-in Fluc RNA was utilised as the input RNA instead of APEX2-mediated labelled RNA. This approach aimed to bypass the variability in RNA labelling efficiency of APEX2 and enable optimisation of the pulldown in a system free from background noise associated with unlabelled RNA. To evaluate the effectiveness of this new

method, the biotinylated spike-in RNA was compared to a negative control condition, where the same RNA without 3' end biotinylation was used (Figure 4.5).

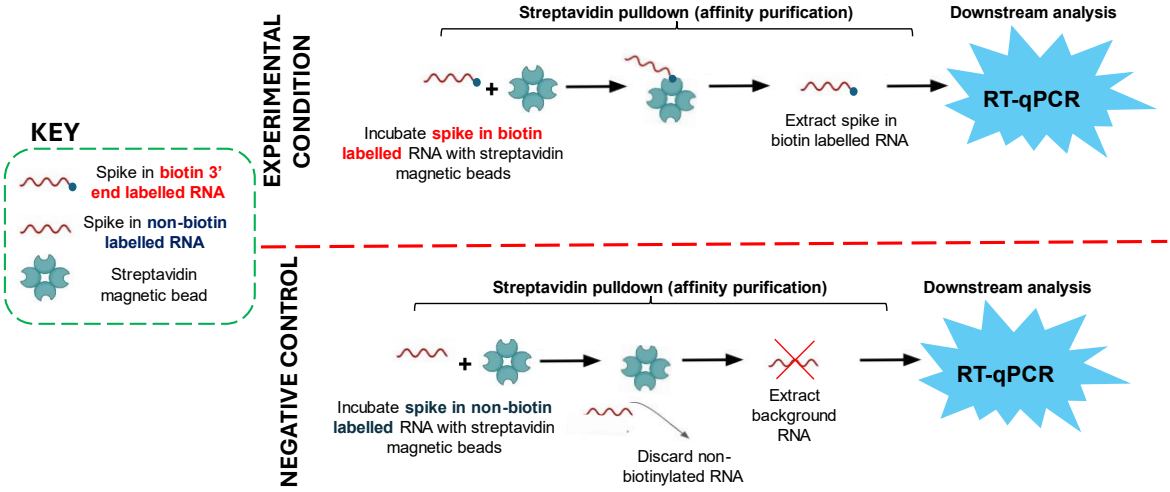


Figure 4.5. Schematic showing workflow for streptavidin affinity purification optimisation by use of a synthetically biotinylated RNA.

Initially four protocols were tested in parallel to evaluate their effectiveness in enriching 3' end biotinylated spike-in RNA relative to the non-biotinylated control. Based on the streptavidin dot blot assay results (Section 4.4), either 0.1 ng of biotinylated spike-in Fluc HCV RNA or 0.1 ng non-biotinylated spike-in RNA were added to streptavidin magnetic beads. The first protocol, following the unmodified Ting lab method described by Fazal et al. (2019), as referenced in Section 2.4.8.1, did not achieve selective enrichment of the biotinylated spike-in RNA relative to the unlabelled spike in RNA (Figure 4.6, Protocol 1).

In contrast, the second protocol, which incorporated additional washing steps utilised by Padrón et al. (2019) into the Fazal et al. (2019) protocol, yielded improved results (Figure 4.6, Protocol 2). The level of biotinylated spike-in RNA recovered post-pulldown was higher than the negative control, suggesting that enhanced washing conditions may have reduced non-specific interactions. This improvement indicates that more rigorous washing, beyond the recommendations of Fazal et al. (2019), can enhance specificity of the pulldown.

The third protocol, which involved the addition of a blocking step in a solution of yeast tRNA and nuclease free BSA prior to RNA incubation with streptavidin magnetic beads, as carried out by Li et al. (2022), along with additional washing steps, inspired by Padrón et al. (2019), proved to be the most effective. This method achieved the lowest level of non-specific enrichment in the non-biotinylated spike-in RNA negative condition (Figure 4.6, Protocol 3). The blocking step seemed particularly successful in minimising non-specific recovery of unlabelled RNA and enhancing the accuracy of the pulldown.

The fourth protocol, which combined the blocking step with the stringent washing steps from the second protocol, did not perform as well as the third protocol alone (Figure 4.6, Protocol 3). While Protocol 4 was more effective than the Protocol 1, the inclusion of both blocking and stringent washing steps led to reduced performance compared to using the blocking step alone.

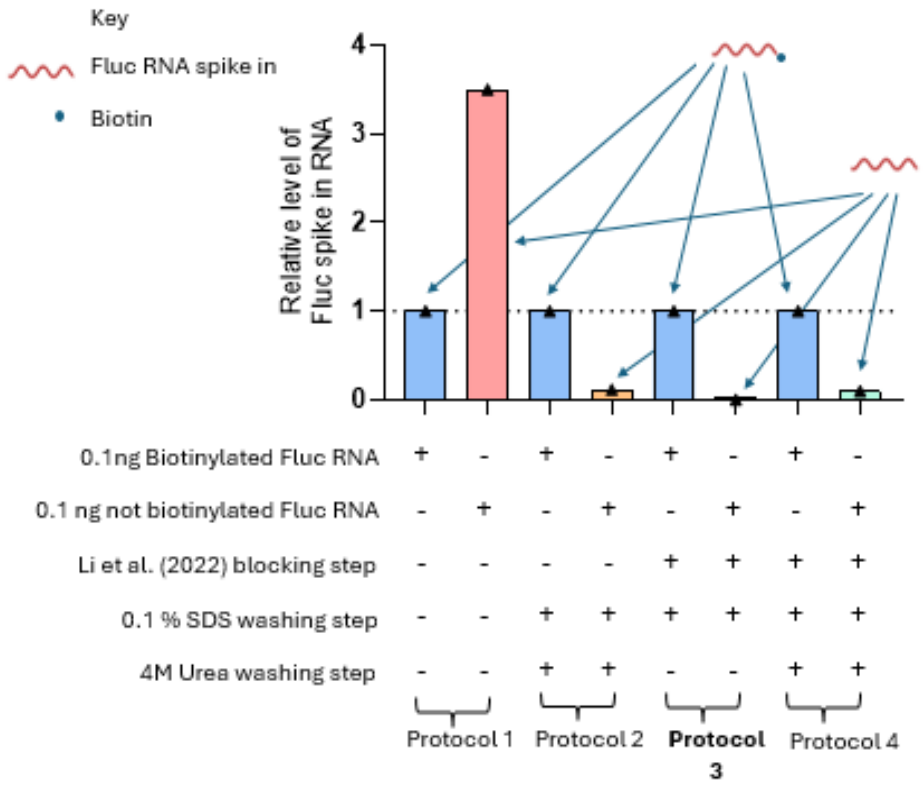
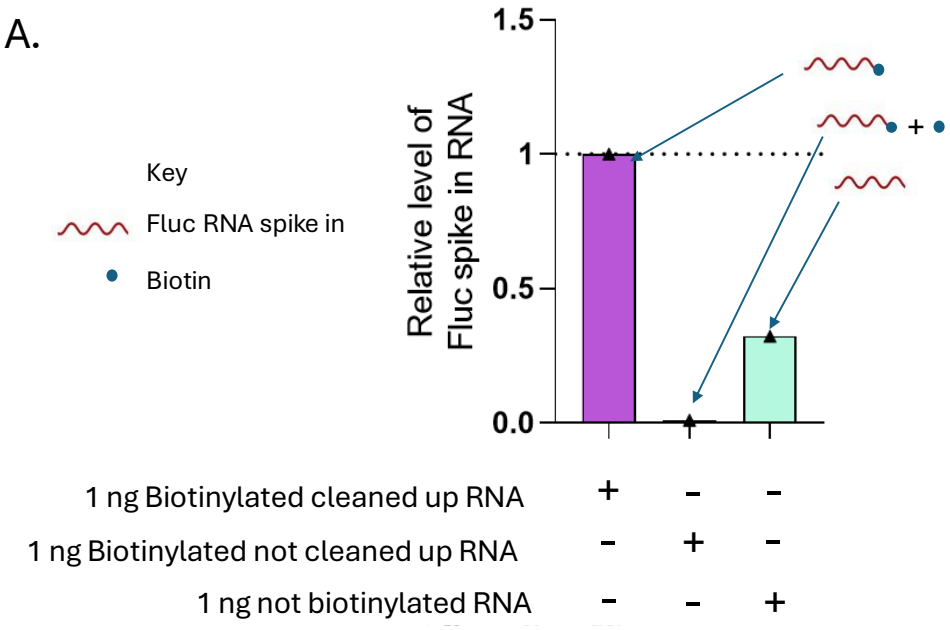


Figure 4.6. Comparison of streptavidin affinity purification techniques using spike-in IVT RNA, with and without *in vitro* biotinylation, analysed by RT-qPCR. Four workflows were tested in parallel, all loosely based on the Ting lab protocol developed in Fazal et al. (2019):
Protocol 1. The unmodified protocol.
Protocol 2. Additional washing steps following RNA incubation with streptavidin magnetic beads, inspired

by Padrón et al. (2019).
Protocol 3. Addition of a blocking step from Li et al. (2022) before RNA incubation with streptavidin magnetic beads and more stringent washing (0.1% SDS).
Protocol 4. Incorporating the Li et al. (2022) blocking step before RNA incubation with streptavidin magnetic beads and washing steps inspired by Padrón et al. (2019).
Streptavidin enrichment for each RNA was normalised to an elution spike-in Nluc IVT RNA and is presented relative to the non-biotinylated spike-in Fluc RNA. Bars represent one independent experiment.

The biotinylated Fluc spike-in RNA had previously been purified to remove excess free pCp-biotin using a cleanup procedure involving ethanol precipitation and spin column purification. This step ensured that unincorporated biotin, which could compete with biotinylated RNA for streptavidin binding sites, was removed.

In this experiment, protocol 2 described above (Figure 4.6) was tested with both cleaned and non-cleaned biotinylated spike-in Fluc RNA. Two quantities of cleaned biotinylated Fluc RNA, 1 ng and 0.1 ng, were used as input to determine the optimal concentration for pulldown efficiency. In addition, 1 ng and 0.1 ng of non-cleaned biotinylated RNA were also utilised as input to determine the extent to which unincorporated biotin affects the streptavidin pulldown. The results indicated that the streptavidin pulldown was more effective with 0.1 ng of biotinylated RNA compared to 1 ng (Figure 4.7), suggesting that there may be an optimal range of input RNA for effective pulldown. Using too much biotinylated RNA may reduce efficiency by potentially overwhelming the binding capacity of streptavidin. Additionally, when non-purified RNA was used, no selective enrichment over the non-biotinylated spike-in RNA was observed. This result highlights the critical importance of removing free biotin prior to the pulldown, as the excess pCp-biotin competes with biotinylated RNA for streptavidin binding, leading to a reduction in pulldown efficiency.



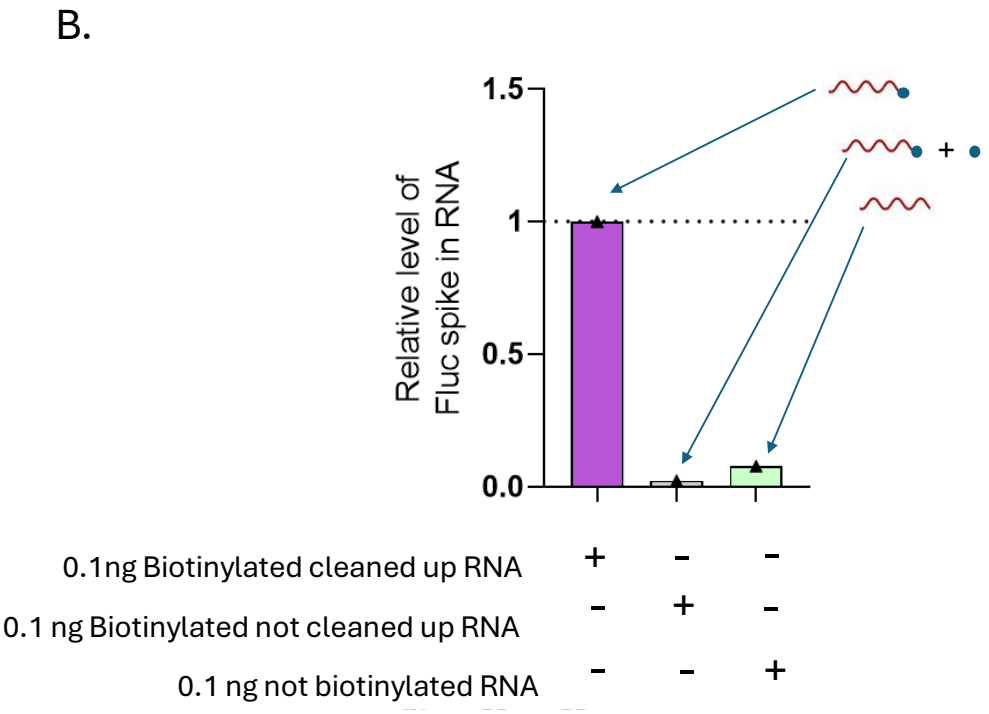


Figure 4.7. Comparison of the combined protocols based off Fazal et al. (2019) and Li et al. (2022) using two different concentrations of biotinylated spike-in Fluc RNA as analysed by RT-qPCR. Where (A) 1 ng and (B) 0.1 ng of input RNA. The figure also includes data for spike-in biotinylated RNA before purification to remove excess free pCp-biotin. Streptavidin enrichment for each RNA was normalised to an elution spike-in Nluc IVT RNA and is presented relative to the non-biotinylated spike-in Fluc RNA. Bars represent one independent experiment.

Given the strong binding affinity between biotin and streptavidin (Sano & Cantor, 1995), it was hypothesised that this bond might hinder the efficient release of labelled RNA during the extraction process. Breaking the biotin-streptavidin

interaction without damaging the RNA is particularly challenging, and this inefficiency may result in a bias towards recovery of unlabelled RNA, which is anticipated to be less tightly bound to the streptavidin beads. The next step involved testing different RNA extraction methods to determine whether the strong binding affinity between biotin and streptavidin limited pulldown efficiency. In parallel, an additional condition was tested to assess whether the use of a binding buffer, as utilised by Li et al. (2022) from the Zou lab, would enhance the binding of biotin-labelled RNA to streptavidin compared to the use of 0.1 M NaCl, as utilised by Fazal et al. (2019). The results revealed little difference between the two RNA extraction methods in terms of pulldown efficiency (Figure 4.8). Consequently, the Trizol LS method was chosen for future experiments due to its lower cost. The comparison of the binding buffers, specifically the buffer utilised by Li et al. (2022) containing Tween, NaCl, EDTA, and Tris versus the RNA binding buffer Fazal et al. (2019) consisting of 0.1 M NaCl, suggested a notable difference in relative enrichment. The RNA binding buffer developed by Li et al. (2022) appeared to facilitate greater selective enrichment of the biotinylated spike-in Fluc RNA compared to the non-biotinylated Fluc spike-in RNA. Based on these observations, the RNA binding buffer developed by Li et al. (2022) was selected for use in subsequent experiments.

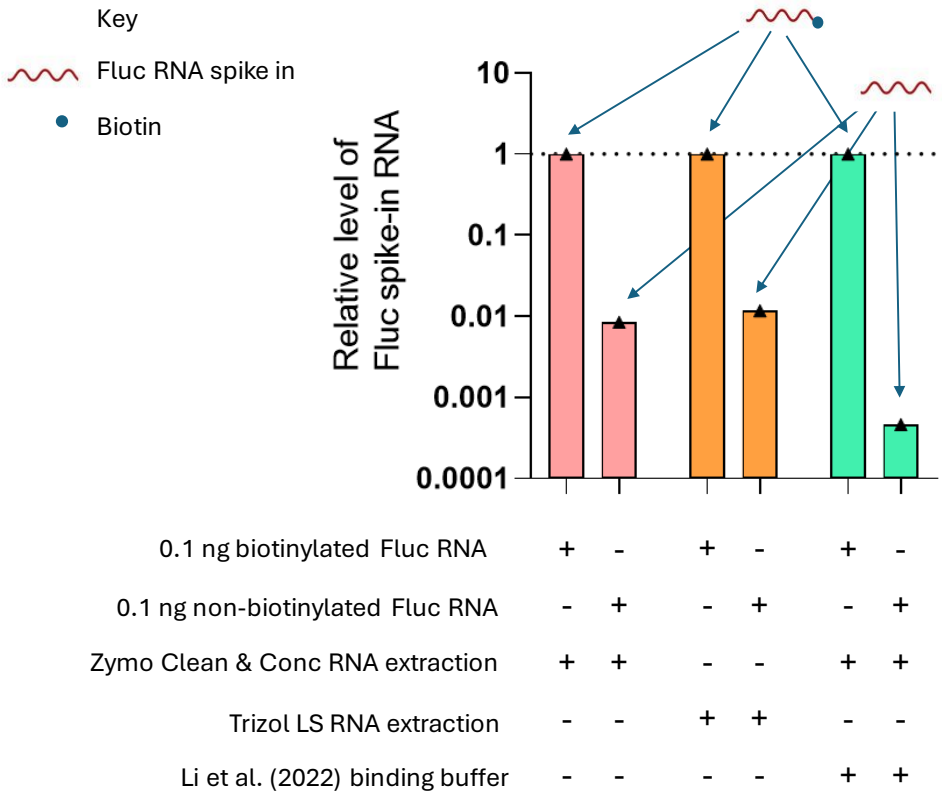


Figure 4.8. Comparison of two RNA extraction methods following streptavidin affinity purification using spike-in IVT RNA analysed by RT-qPCR. Additionally, a binding buffer utilised by Li et al. (2022) from the Zou lab to enhance the binding of biotinylated RNA to streptavidin magnetic beads was also evaluated. Streptavidin enrichment for each RNA was normalised to an elution spike-in Nluc IVT RNA and is presented relative to the non-biotinylated spike-in Fluc RNA. Bars represent one independent experiment.

To better simulate the conditions encountered with APEX2-mediated biotinylated RNA, where only a small fraction of RNA is expected to be labelled, total RNA extracted from wildtype HEK293T cells was added to the streptavidin pulldown (Figure 4.9). This allowed for the assessment of the optimised streptavidin pulldown protocol's ability to selectively enrich biotinylated RNA in the presence of excess non-biotinylated RNA.

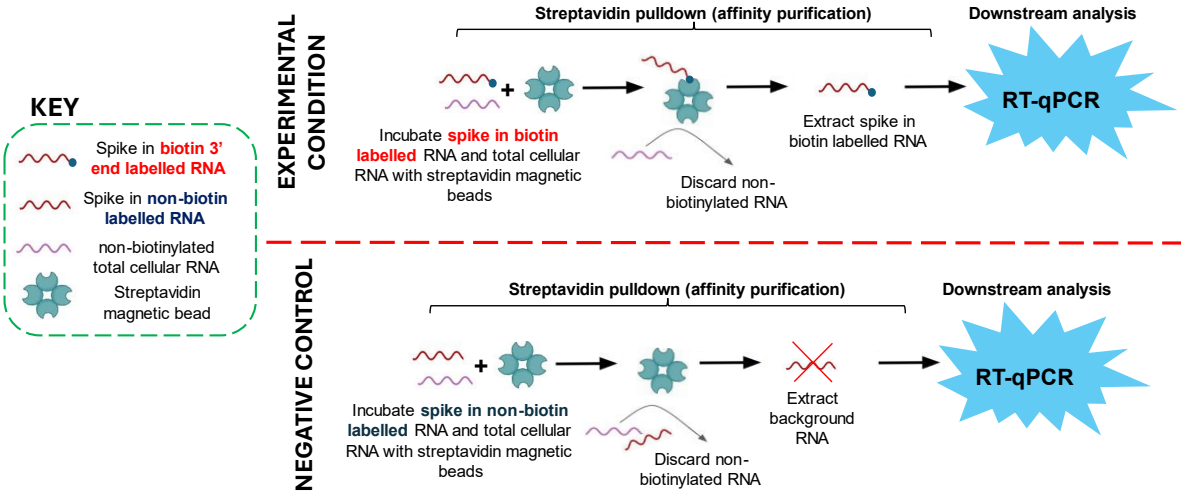


Figure 4.9. Schematic showing workflow for streptavidin affinity purification optimisation by use of a synthetically biotinylated RNA with the inclusion of background noise in the form of non-biotinylated RNA extracted from WT HEK293T cells.

The optimised pulldown protocol continued to perform effectively, even when substantial amounts of non-labelled RNA were present (Figure 4.10). The ability to successfully enrich biotinylated RNA under these conditions further validated the robustness of the optimised streptavidin pulldown protocol, demonstrating its ability to capture biotinylated RNA in more biologically relevant conditions, where non-biotinylated RNA would be abundant.

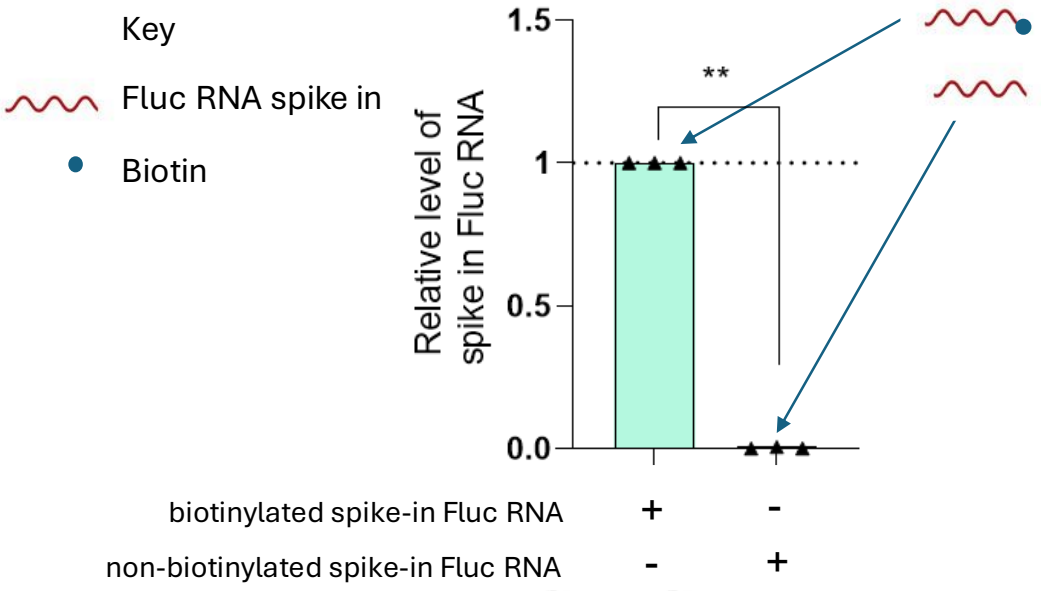


Figure 4.10. Optimised streptavidin affinity purification protocol selectively enriches biotinylated spike in Fluc RNA in the presence of excess unlabelled total RNA isolated from HEK293T wildtype cells, as analysed by RT-qPCR. Streptavidin enrichment for each RNA was normalised to an elution spike-in Nluc IVT RNA and is presented relative to the non-biotinylated spike-in Fluc RNA. Bars represent the mean of three independent experiments, with error bars showing SD. Statistical significance is denoted as ** = $P < 0.005$.

4.6. Initial challenges with optimised streptavidin pulldown in selectively enriching APEX2-mediated biotinylated RNA

To evaluate whether the optimised streptavidin pulldown protocol could effectively isolate APEX2-mediated biotinylated RNA, APEX2-ERM HEK293T cells were transfected with NlucSec mRNA, incubated for four hours, and then treated with biotin-aniline, followed by exposure to H_2O_2 to induce biotinylation. A negative control where biotin-aniline was omitted was included and data was analysed performing RT-qPCR on the output RNA using primers specific to Nluc mRNA and the biotinylated spike-in Fluc RNA. The Nluc mRNA qPCR data was normalised to the biotinylated spike-in Fluc RNA qPCR data and expressed as fold change relative to the no biotin-aniline control. The relative level of Nluc mRNA recovered from the experimental condition was not statistically different from that recovered from the negative control when 5 µg of input RNA was used in the streptavidin pulldown (Figure 4.11).

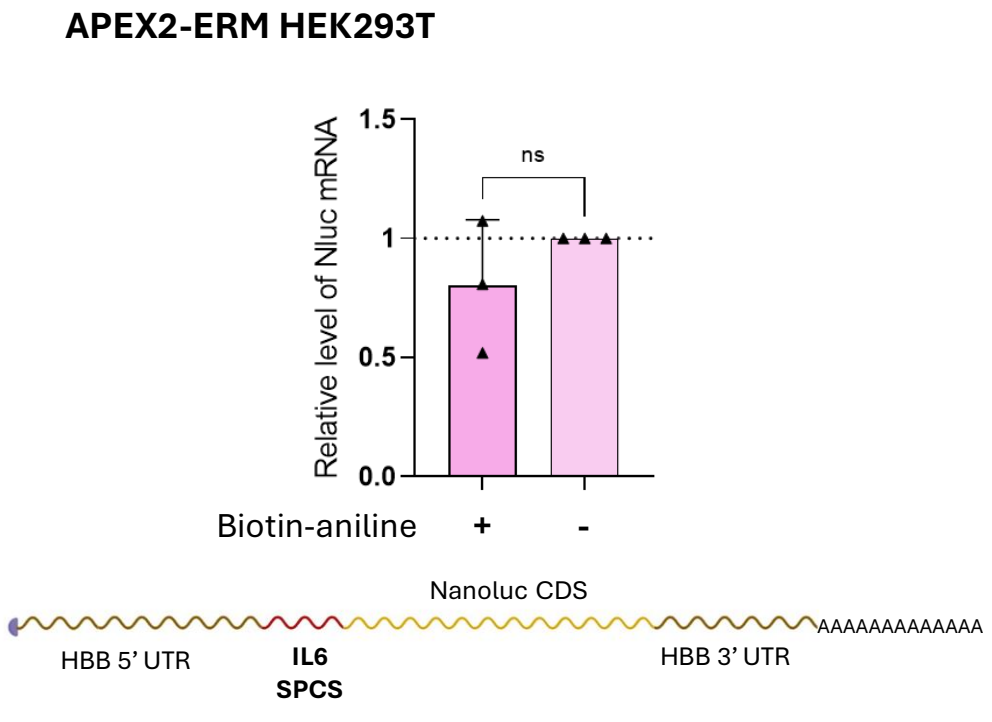


Figure 4.11. Evaluation of selective enrichment of APEX2-mediated biotinylated RNA in APEX2-ERM HEK293T cells using the optimised streptavidin affinity purification protocol, as analysed by RT-qPCR. NlucSec mRNA was delivered into APEX2-ERM HEK293T cells by lipofectamine transfection and incubated for 4 hours before exposure to biotin-aniline and H_2O_2. NlucSec mRNA levels post-streptavidin pulldown were quantified by RT-qPCR, normalised to the biotinylated spike-in Fluc RNA, and are presented relative to the no-biotin-aniline control. Bars represent the mean of three independent experiments, with

error bars showing SD. Statistical significance is denoted as ns = not significant. The schematic below the graph shows the transfected mRNA.

A similar experiment was performed in APEX2-NES HEK293T cells, this time transfecting cytosolic Nluc mRNA to assess whether the issue was specific to the ER-localised APEX2 HEK293T system. The results mirrored those of the APEX2-ERM experiment, with the relative level of Nluc mRNA recovered from the experimental condition showing no difference compared to the negative control (Figure 4.12). These findings together indicate that, in its current form, the optimised protocol was unable to effectively distinguish between biotinylated and unlabelled RNA in the context of APEX2-RT-qPCR experiments.

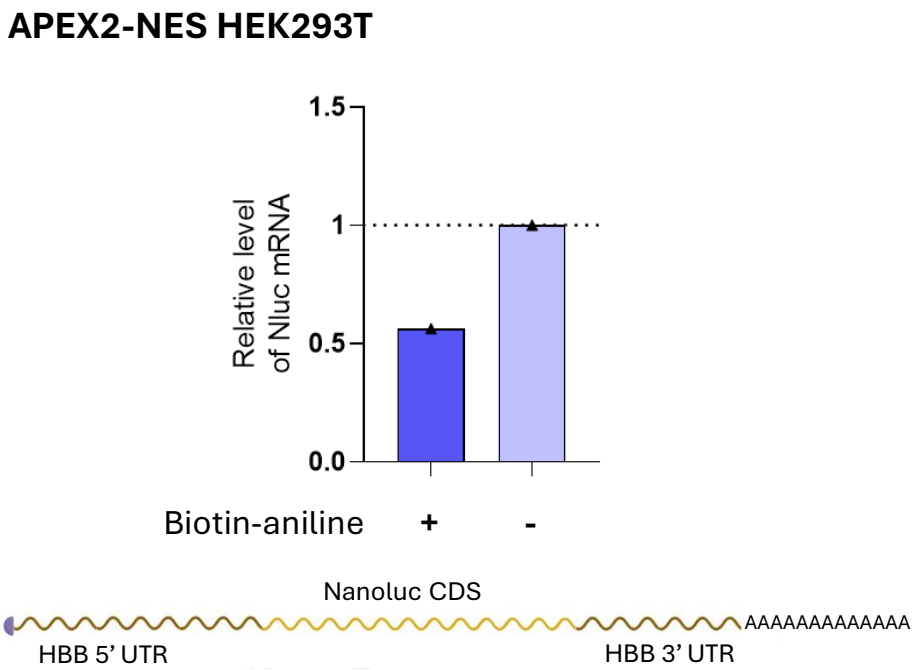


Figure 4.12. Evaluation of selective enrichment of APEX2-mediated biotinylated RNA in APEX2-NES HEK293T cells using the optimised streptavidin affinity purification protocol, as analysed by RT-qPCR. Nluc mRNA was delivered into APEX2-NES HEK293T cells by lipofectamine transfection and incubated for 4 hours before exposure to biotin-aniline and H_2O_2. Nluc mRNA levels post-streptavidin pulldown were quantified by RT-qPCR, normalised to the biotinylated spike-in Fluc RNA, and are presented relative to the no-biotin-aniline control. Bars represent one independent experiment. The schematic below the graph shows the transfected mRNA.

4.7. Successful enrichment of APEX2-mediated biotinylated RNA with optimised streptavidin pulldown and reduced input RNA

Given the high background noise observed when utilising the optimised streptavidin pulldown protocol, as discussed in Section 4.6, it was hypothesised that reducing the amount of input RNA could enhance the

selective enrichment of APEX2-mediated biotinylated RNA over non-labelled RNA. This hypothesis was based on earlier findings in Section 4.5, which suggested that the quantity of biotinylated RNA could impact the efficiency of the streptavidin pulldown. To test this, a streptavidin pulldown was performed using the same source of input RNA but with varying RNA amounts: 5 µg in one condition and 500 ng in another. This adjustment successfully led to the selective enrichment of APEX2-mediated biotinylated RNA over the no biotin-aniline control using the optimised streptavidin pulldown protocol for the first time (Figure 4.13A). The experiment was then repeated in six additional biological replicates with 500 ng input RNA, confirming that APEX2-RT-qPCR can effectively enrich IVT mRNA relative to a negative control in a statistically significant manner.

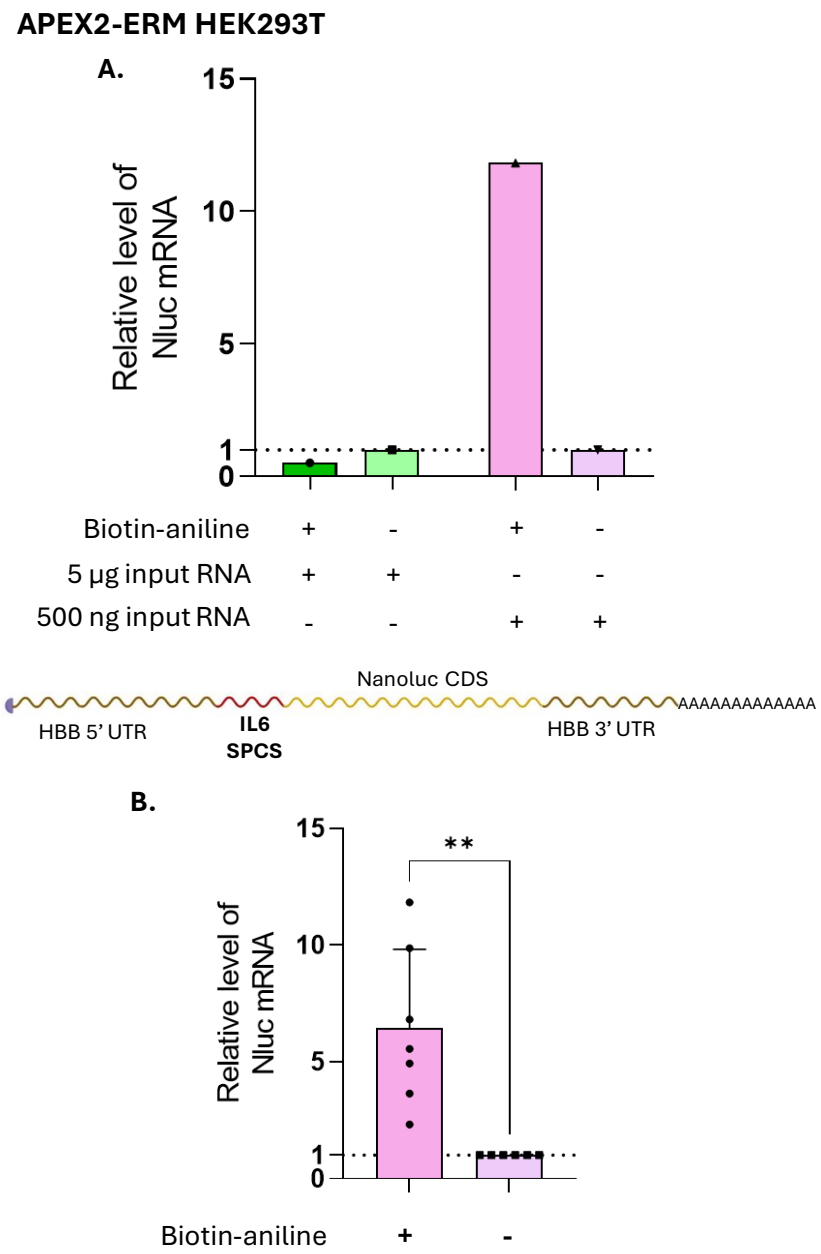


Figure 4.13. Comparison of different input amounts of APEX2-mediated biotinylated RNA in APEX2-ERM HEK293T using the optimised streptavidin affinity purification protocol, as analysed by RT-qPCR. NlucSec mRNA was delivered into APEX2-ERM HEK293T cells by lipofectamine transfection and incubated for 4 hours before exposure to biotin-aniline and H_2O_2. NlucSec mRNA levels post-streptavidin pulldown were quantified by RT-qPCR, normalised to the biotinylated spike-in Fluc RNA, and are presented relative to the no-biotin-aniline control. (A) Input amounts of 5 µg or 500 ng of APEX2-mediated biotinylated RNA extracted from APEX2-ERM HEK293T cells were tested, with a negative control where biotin-aniline was omitted to assess background noise. Bars represent the one independent experiment. (B) APEX2-RT-qPCR was performed on six biological repeats where the input RNA amount was dropped to 500 ng. Bars represent the mean of at least three independent experiments, with error bars showing SD. Statistical significance is denoted as ** = $P < 0.005$. The schematic between the graphs shows the transfected mRNA.

4.8. Evaluation of S^6GTP IVT mRNA to enhance APEX2 labelling efficiency of delivered mRNA

Once previous optimisations had led to more consistent enrichment of APEX2-labelled RNA, further enhancement of labelling efficiency was explored using a recently published methodology. Li et al. (2022) introduced a technique called MERR-APEX-seq, which demonstrated that supplementing APEX2-expressing cells with *S6*-Thioguanine (S^6G) could significantly increase APEX2's labelling efficiency of nascent endogenous RNA. To investigate whether this method could enhance labelling efficiency in the system, the MERR-APEX approach was adapted by performing IVT with varying levels of substitution of GTP for S^6GTP.

Previously, an ARCA co-transcriptional capping IVT kit was used in this study to generate capped-IVT mRNA with therapeutically relevant UTRs. While this kit simplifies the production of ARCA-capped RNA by combining all nucleotides in a single mixture, it requires GTP depletion to ensure proper capping. Moreover, it was understood that S^6GTP might not be incorporated by T7 RNA polymerase as efficiently as canonical GTP (Sergiev et al., 1997). This dual requirement to reduce non-capped RNA and account for varying incorporation efficiencies would likely necessitate IVT reagent optimisation when using the ARCA kit. To avoid this complication, plasmids were modified to replace the GGG sequence downstream of the T7 promoter with AGG, making the DNA template compatible with co-transcriptional capping with CleanCap AG, as shown in Figure 4.14. CleanCap AG co-transcriptional capping bypasses the need for GTP depletion (Henderson et al., 2021) during IVT, preserving the RNA capping process and enabling the incorporation of S^6GTP without additional optimisation of NTP concentrations.

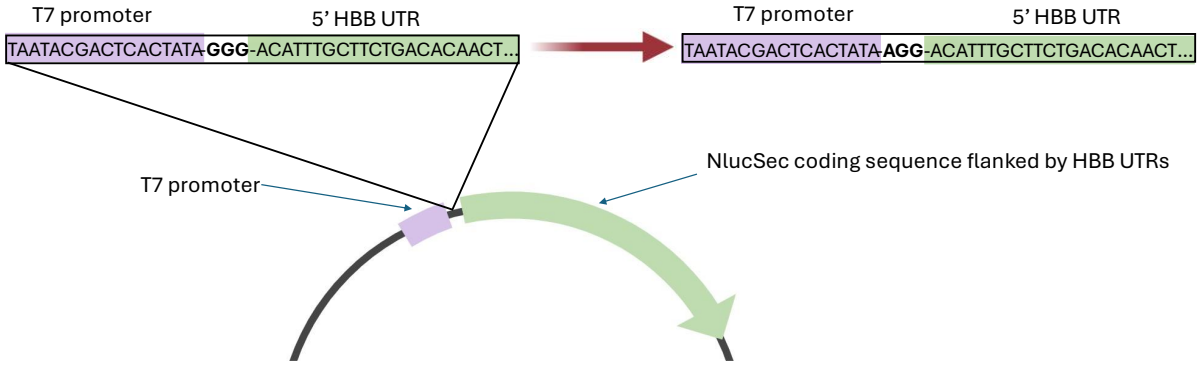


Figure 4.14. Schematic showing cloning of plasmid encoding NlucSec suitable for co-transcriptional ARCA capping (excerpt of sequence shown on the left) to a plasmid suitable for co-transcriptional CleanCap AG capping (excerpt of sequence shown on the right).

An *in vitro* biotinylation assay was employed to investigate whether S^6GTP was incorporated into the RNA during *in vitro* transcription. This *in vitro* biotinylation assay, originally developed to study nascent RNA transcription following 4-thiouridine (4^SU) labelling (Duffy et al., 2015), involves incubating RNA extracted from 4^SU-labelled cells with MTSEA-XX-Biotin. The MTSEA-XX-Biotin selectively forms covalent bonds with chemically modified thionucleobases.

CleanCapped IVT NlucSec mRNAs synthesised with varying ratios of S^6GTP to GTP (0%, 10% or 25%) were exposed to MTSEA-XX-Biotin *in vitro*. The RNA samples were subsequently analysed using a streptavidin-HRP RNA dot blot assay and subjected to streptavidin affinity purification, following the optimised protocol detailed in Section 4.5. Dot blotting of the *in vitro* biotinylated RNA revealed a visible signal in samples synthesised with 10% or 25% S^6GTP, confirming successful biotinylation, while no signal was detected in biotinylated RNA samples synthesised without S^6GTP (Figure 4.15A). Moreover, the extent of biotinylation increased with higher ratios of S^6GTP.

Additionally, the *in vitro* biotinylated S^6G-labelled RNA was subjected to streptavidin pulldown and analysed by RT-qPCR (Figure 4.15B). The same negative control, consisting of IVT RNA not synthesised with S^6GTP (0% S^6GTP) but exposed to biotin-XX-MTSEA, was included. The data were analysed by performing RT-qPCR on the output RNA using primers specific to Nluc mRNA and the biotinylated spike-in Fluc RNA, with results shown relative to the 25% S^6G condition. This analysis revealed significantly higher mRNA recovery in

samples containing S^{6}GTP compared to the control, with the highest recovery observed in the 25% S^{6}GTP samples. These results confirm that T7 polymerase effectively incorporates S^{6}GTP during transcription.

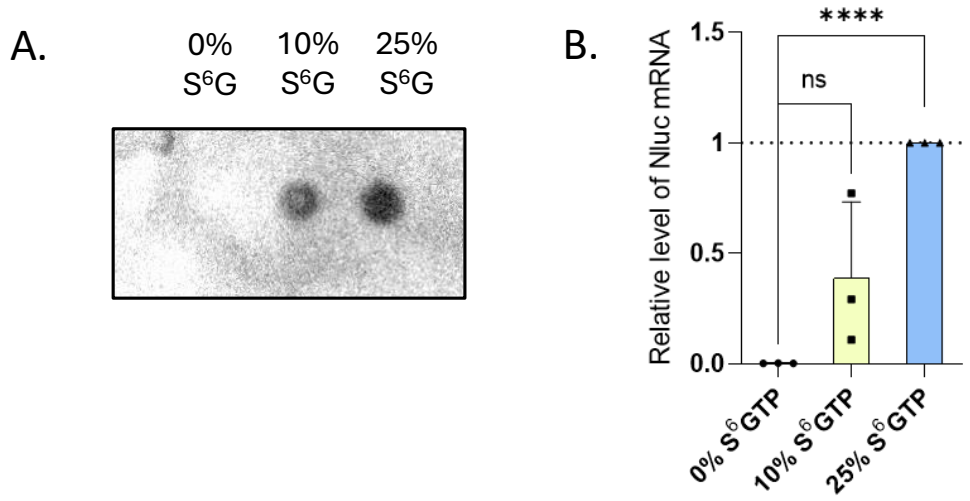


Figure 4.15. Validation of S^{6}GTP incorporation by streptavidin-HRP dot blot assay and streptavidin pulldown. IVT mRNA expected to contain varying proportions of S^{6}GTP was biotinylated *in vitro* using Biotin-XX-MTSEA, which reacts with sulphur-containing nucleotides. The mRNA was then extracted by chloroform extraction and analysed by a streptavidin-HRP dot blot assay (A) or subjected to the optimised streptavidin affinity purification protocol followed by RT-qPCR analysis (B, N=3). For (A), the image is representative of three independent experiments. For (B), data was normalised to an elution spike-in RNA (eGFP) and shown relative to the 25% S^{6}GTP condition. Bars represent the mean of three independent experiments, with error bars showing SD. Statistical significance is denoted as **** = P < 0.000025; ns = not significant.

Next, to assess whether S^{6}GTP incorporation would affect the translation of the mRNA, secreted luciferase activity was measured in APEX2-ERM HEK293T cells transfected with CleanCapped NlucSec mRNA containing varying ratios of S^{6}GTP to GTP (0%, 10%, or 25%). Secreted nanoluciferase activity was assessed at several time points post-transfection, prior to the addition of biotin-aniline at 4 hours and H_2O_2 for the APEX2 biotinylation reaction in the same cells. Statistical comparisons of luciferase activity between the 0% and 25% S6GTP conditions revealed no significant differences (Figure 4.16), indicating that S^{6}GTP did not significantly impair translation efficiency.

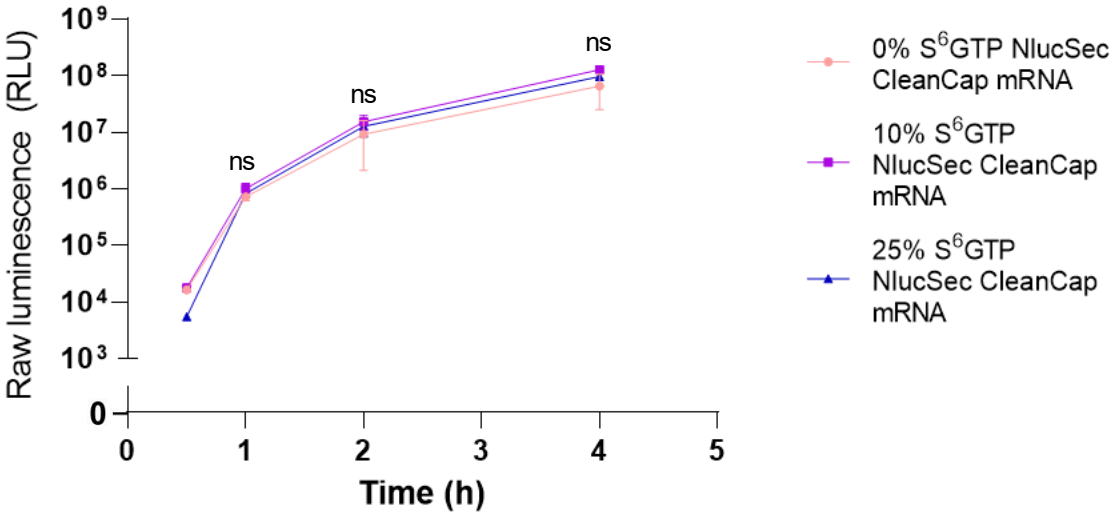


Figure 4.16. Comparison of luciferase activity of IVT NlucSec mRNA with varying proportions of S⁶GTP by secreted luciferase assay. mRNA was delivered into APEX2-ERM HEK293T cells by lipofectamine 2000 transfection, and luminescence in relative light units (RLU) was measured at different time points after media removal. At each time point, 3 technical repeats were taken (n=3). Measurements were taken at 30 minutes (N=1), 1 hour (N=3), 2 hours (N=3), and 4 hours (N=3), with errors bars showing SD. Statistical significance between the raw luminescence values of the 0% S⁶GTP and 25% S⁶GTP conditions is indicated, with "ns" representing non-significance.

After the 4 hour luciferase time course, APEX2-ERM HEK293T cells were treated with biotin-aniline for 30 minutes, , and were then exposed to H_2O_2 to induce APEX2-mediated biotinylation. A negative control, in which biotin-aniline was omitted, was included for normalisation purposes. Total RNA extracted from these cells was subjected to streptavidin pulldown and analysed by RT-qPCR. Both CleanCapped 25% S⁶G NlucSec mRNA and CleanCapped 0% S⁶G NlucSec mRNA demonstrated selective enrichment in the streptavidin pulldown compared to the no biotin-aniline control in some biological repeats (Figure 4.17). However, the enrichment relative to their corresponding negative controls was not statistically significant. Additionally, when comparing the enrichment of CleanCapped 25% S⁶G NlucSec mRNA to CleanCapped 0% S⁶G NlucSec mRNA, after normalising both to their respective no biotin-aniline controls, no statistically significant difference was observed. Therefore, in this experiment, the incorporation of S⁶GTP did not enhance the labelling efficiency of APEX2 in APEX2-ERM HEK293T cells.

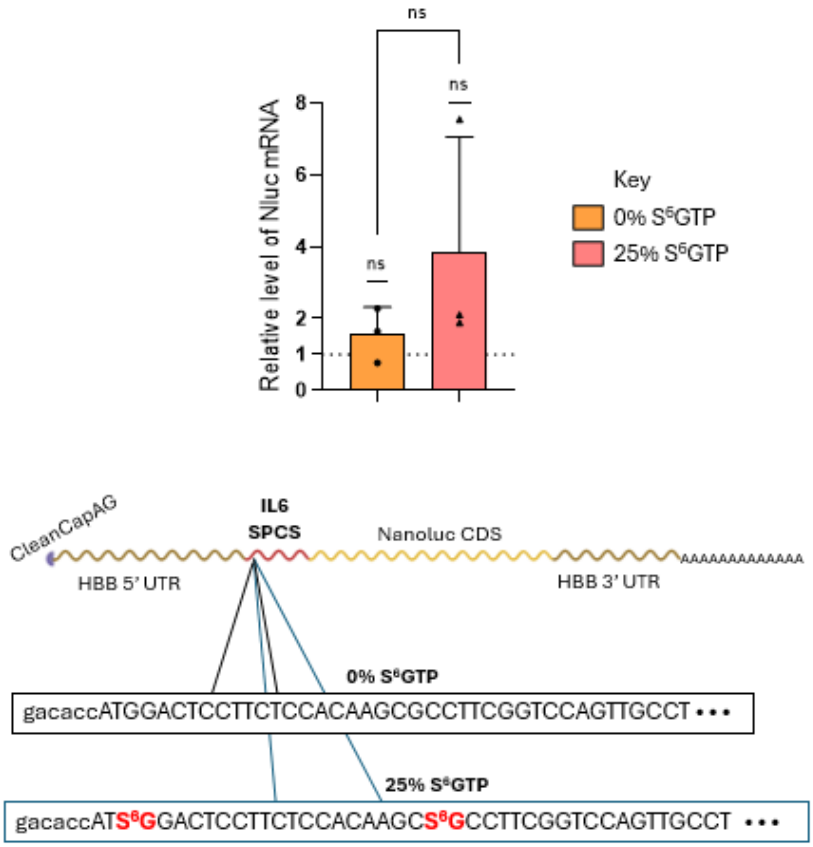


Figure 4.17. APEX2-RT-qPCR analysis of the relative enrichment of transfected NlucSec mRNA with varying proportions of S^{6}GTP following streptavidin affinity purification. NlucSec mRNA, synthesised by IVT with 0%, 10%, or 25% of standard GTP substituted with S^{6}GTP, was delivered into APEX2-ERM HEK293T cells by lipofectamine transfection and incubated for 4 hours before exposure to biotin-aniline and H_2O_2. Data was normalised to the biotinylated spike-in Fluc RNA and is shown relative to the no-biotin aniline control. Bars represent the mean of three independent experiments, with error bars showing SD. Statistical significance is denoted as ns = not significant. The schematic below the figure illustrates CleanCapped NlucSec mRNA and CleanCapped NlucSec mRNA synthesised with 25% S^{6}GTP.

4.9. Discussion

4.9.1. Streptavidin pulldown optimised for improved APEX2-RT-qPCR consistency

This chapter aimed to utilise APEX2-RT-qPCR to detect the localisation of delivered IVT mRNA to either the ER or cytoplasm. Issues arose with the published protocols (Fazal et al., 2019; Padrón et al., 2019), as they produced inconsistent results in this study, with enrichment in the experimental condition often being lower than in the negative control, where no APEX2 labelling was expected. These inconsistencies led to the optimisation of the streptavidin pulldown process by comparing enrichment of a 3' end biotinylated spike-in RNA relative to a negative control spike-in RNA. The optimisation process

involved testing various steps from three different protocols, ultimately refining the method by adding a blocking step prior to incubation of input RNA with beads and employing more stringent washing steps post-incubation of input RNA with streptavidin beads. Once the input RNA amount was reduced from 5 µg to 500 ng, the optimised streptavidin pulldown successfully enabled APEX2-mediated labelling to be detected above background noise. Furthermore, the continued use of the 3' end biotinylated spike-in RNA alongside the input RNA provides a novel normalisation method that may help reduce inconsistencies between experiments for both APEX2-RT-qPCR and APEX-seq techniques in the future.

Further optimisation of the streptavidin pulldown could also involve integrating steps from the protocol developed by Li et al. (2022), particularly their method of eluting biotinylated RNA from beads, which was not attempted in this study. Additionally, no efforts have been made to test the saturation point of the system with respect to biotin-labelled RNA. This could be investigated through serial dilution experiments to assess the binding capacity and recovery efficiency of the pulldown.

In an effort to enhance APEX2 labelling efficiency, the Metabolic Incorporation of Electron-Rich Ribonucleoside (MERR-APEX-seq) technique developed by Li et al. (2022) was adapted for delivered mRNA by incorporating S^6GTP during IVT. Secreted luciferase assays were conducted to determine whether S^6GTP incorporation into delivered mRNA impacted translation, as NlucSec localisation to the ER depends on successful translation. These assays confirmed that S^6GTP did not adversely affect translation. However, APEX2-RT-qPCR analysis showed no significant difference in labelling efficiency between mRNA synthesised with varying S^6GTP ratios and unmodified mRNA, indicating that S^6GTP did not improve APEX2's labelling activity for delivered mRNA in APEX2-ERM HEK293T cells. Streptavidin-HRP RNA dot blot assays and streptavidin pulldowns demonstrated that S^6GTP was effectively incorporated into the IVT mRNA, ruling out poor incorporation by T7 polymerase during IVT as the reason for the lack of improved labelling efficiency.

Li et al. (2022) demonstrated that APEX2 labelling efficiency is improved by the metabolic incorporation of S^6G *in vitro* using streptavidin-biotin dot blot analysis. Additionally, they evaluated recovery rates by comparing Ct values of post-streptavidin pulldown (output RNA) and input RNA samples from MERR-APEX-Seq and APEX-Seq, finding that the MERR-APEX approach increased labelling efficiency in HEK293T cell lines where APEX2 is localised to the mitochondrial matrix, as well as in a cell line where APEX2 localises to the nuclear lamina (Li et al., 2022). However, their MERR-APEX-Seq studies using APEX2-ERM and APEX2-NES cell lines did not include a comparison of recovery yields relative to conventional APEX2-Seq data. Moreover, they presented MERR-APEX2-ERM sequencing data normalised to MERR-APEX2-NES sequencing data, whereas their studies with APEX2-mitochondria and APEX2-nuclear lamina cell lines normalised enriched samples relative to input RNA (Li et al., 2022). This inconsistency may be due to the close relationship between the cytosolic-facing side of the ER membrane and the cytosol. Alternatively, it is possible that MERR-APEX2-Seq is more suitable for RNA not engaged in active translation, or that the MERR approach does not enhance APEX2 labelling efficiency in non-membrane-enclosed regions.

In experiments involving the adaptation of MERR-APEX-seq priniciples to IVT mRNA (Section 4.8), a 1:4 ratio of S^6GTP to GTP was used during IVT mRNA generation. This ratio was based on Sergiev et al. (1997), who found that using a 1:1 molar ratio of S^6GTP to GTP in IVT led to approximately 10% incorporation of S^6GTP. When Li et al. (2022) developed MERR-APEX-seq, they observed that S^6G incorporation into nascently transcribed RNA accounted for only 0.52% of total RNA 5 hours post-labelling. Thus, to avoid excessive incorporation in the IVT mRNA, a lower 1:4 ratio was selected to better approximate incorporation levels observed by Li et al. (2022). Future work could explore higher concentrations of S^6GTP during IVT to assess whether increased incorporation enhances labelling efficiency without compromising RNA integrity or function. Since Li et al. (2022) found that the MERR-APEX-seq technique favours the labelling of RNA with higher GC content, codon-optimising the IVT mRNA to increase its GC content

could help facilitate more efficient APEX2 labelling. NlucSec, which includes the IL6 signal protein coding sequence along with UTRs from human β-haemoglobin has a GC content of 51.3%, which is within the acceptable range for APEX2 labelling. Increasing the GC content of Nluc, however, to 55% or 60% might facilitate enhanced APEX2 labelling of IVT mRNA containing S^6GTP.

4.9.2. Optimising Conditions for APEX2 Labelling

Despite the successful optimisation of the streptavidin pulldown, inconsistencies in APEX2-RT-qPCR results persisted across experiments. For example, the enrichment observed post-pulldown in Figure 4.13B and Figure 4.17 varied significantly, despite the similarity of the mRNA constructs. These mRNA constructs differed only in the presence of ARCA versus CleanCap and were transfected into the same cells at similar passages (P4-P15 post-transduction). It is possible that this variation in enrichment between experiments was due to an apparent decline in the effectiveness of the APEX2-RT-qPCR experiments, which was often found in this study. At the end of the studies, it was found that using fresh H_2O_2, in particular, as well as fresh biotin-aniline apparently reversed the decline. Therefore, a key area of focus to improve the consistency of APEX2-RT-qPCR experiments would be to investigate the stability of biotin-aniline and H_2O_2 over time, in order to more accurately determine when reagent replacement is required. This could be assessed through stability time course assays, either by direct mass spectrometry of the reagents or by using APEX2 labelling followed by RNA dot blot assays for downstream analysis. Once the stability of the APEX2 labelling reagents is better understood, other critical variables, such as the incubation time of biotin-aniline with cells and the duration of H_2O_2 exposure, should be systematically tested in each cell type to enhance reproducibility.

In conclusion, the incorporation of a 3' end biotinylated spike-in RNA has proven to be a valuable tool for both optimising the streptavidin pulldown process and normalising data across APEX2 experiments. This approach has established conditions that facilitate using APEX2 to investigate mRNA delivery and localisation in future experiments.

Chapter 5 - Investigation of mRNA delivery

5.1. Introduction

Following the optimisation of the streptavidin pulldown technique to isolate APEX2-labelled RNA, as detailed in Chapter 4, the next objective was to employ APEX2-RT-qPCR to quantify the localisation of delivered reporter and therapeutically relevant mRNA in APEX2-expressing cells.

In published literature, direct APEX2 labelling of RNA using the APEX2-ERM and APEX2-NES cell lines has primarily been directed towards providing an overview of endogenous mRNA localisation at either subcellular region by RNA-seq (Fazal et al., 2019; Padrón et al., 2019; Zhou et al., 2019). Here, a novel approach is taken by focusing on the quantification of delivered mRNA using APEX2 labelling coupled with RT-qPCR, whilst applying the optimised streptavidin pulldown method developed in Chapter 4 to ensure that APEX2-labelled mRNA is quantified above background levels.

This study investigates the role of the Signal Peptide Coding Sequence (SPCS) in RNA delivery to the ER as a proof of principle for applying the APEX2-RT-qPCR technique to study mRNA delivery. The SPCS, located at the 5' end of the CDS of secreted proteins, plays a crucial role in directing mRNA and ribosomes to the ER in a co-translational manner via the Signal Recognition Particle pathway (Keenan et al., 2001). Recent studies highlight a growing awareness of the importance of the SPCS as a key feature in therapeutic mRNA design, not only for facilitating the secretion of proteins encoded by these mRNAs (Cheng et al., 2023), but also as a crucial component in the immune response elicited by COVID-19 mRNA vaccines (Mo et al., 2023). However, further understanding is needed regarding how the choice of signal peptide influences the subcellular trafficking kinetics of delivered mRNA.

APEX2-RT-qPCR was first applied to compare the subcellular distribution of exogenous reporter mRNAs, designed with or without a SPCS, and thus expected to localise either to the cytosol or to the ER. This approach assessed

the capability of APEX2-expressing cell lines and the APEX2-RT-qPCR system to differentiate between distinct subcellular localisations of delivered mRNA. Notably, the APEX2-ERM HEK293T cell line yielded the most robust results and became the main focus for subsequent experiments. This chapter then centres on the role of SPCS in reporter and therapeutically relevant mRNA. The use of the APEX2-ERM HEK293T and A549 cell lines facilitates almost simultaneous comparison between localisation of delivered mRNA at the ER, assessed by APEX2-RT-qPCR, and the resulting reporter protein activity, measured through secreted nanoluciferase assays. Understanding how the kinetics of co-translational ER targeting of delivered mRNA might relate to functional protein output introduces a new dimension in the design of therapeutic mRNA. To further understand the subcellular localisation of delivered mRNA, the use of a biochemical membrane fractionation technique was explored as an alternative method for studying both endogenous and delivered mRNA distribution. In addition, a novel IVT mRNA labelling technique, known to not interfere with translation (Anhäuser et al., 2019; Westerich et al., 2020) but that has not yet been tested in the context of mRNA delivery, was briefly investigated.

5.2. Investigation of selective APEX2 labelling and isolation of mRNA delivered to the cytosol and ER using the optimised streptavidin affinity purification method

To investigate the selectivity of the APEX2-RT-qPCR system in enriching ER-localised IVT mRNA over cytosolic mRNA and vice-versa, APEX2-expressing cells were transfected with either Nluc or NlucSec IVT mRNA using Lipofectamine 2000, as detailed in Section 4.3. This section and the subsequent APEX-RT-qPCR experiments presented in Chapter 5 maintain several consistent methodological elements: a 4-hour incubation period post-transfection, a subsequent 30 minute incubation with biotin-aniline prior to H_2O_2 treatment for APEX2-mediated RNA labelling, and inclusion of a negative control without biotin-aniline treatment. Following exposure to H_2O_2, total RNA was extracted, and biotinylated RNA was isolated using the optimised streptavidin affinity purification method established in Chapter 4. This

optimised method includes the addition of a biotinylated spike-in Fluc RNA to the input RNA sample prior to purification.

Quantification of the relative enrichment of the delivered mRNA of interest by RT-qPCR was performed by normalising to the biotinylated spike-in, with data shown relative to the no-biotin-aniline negative control. The consistent methodology allows for direct comparisons across the APEX2-RT-qPCR results presented throughout the chapter.

5.2.1. APEX2-ERM HEK293T

In the APEX2-ERM HEK293T system, transfection of cytosolic Nluc mRNA did not result in significant enrichment in the streptavidin pulldown compared to the no-biotin-aniline negative control (Figure 5.1). This indicates that Nluc mRNA, which is expected to localise to the cytoplasm, is not selectively pulled down when APEX2 labelling takes place at the ER. In contrast, NlucSec mRNA showed a significant 5-fold enrichment over its no-biotin-aniline negative control. These results demonstrate that the APEX2-ERM HEK293T system can effectively differentiate both delivered mRNA at the ER from its negative control and distinguish between delivered mRNA expected to localise to the ER from delivered mRNA expected to localise to the cytosol.

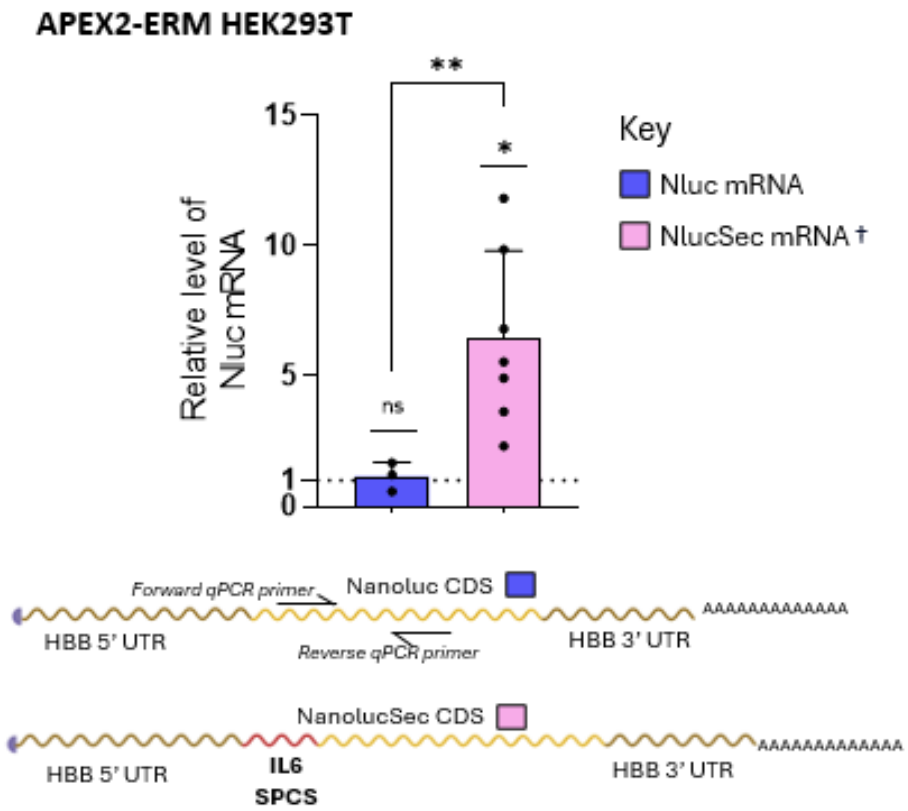


Figure 5.1. APEX-RT-qPCR selectively detects ER-localised Nluc mRNA in APEX2-ERM HEK293T cells. IVT mRNA was delivered into APEX2-ERM HEK293T cells by lipofectamine transfection and incubated for 4 hours. Cells were then treated with biotin-aniline for 30 minutes before exposure to H_2O_2. Nluc or NlucSec mRNA levels post-streptavidin pulldown were quantified by RT-qPCR, normalised to the biotinylated spike-in Fluc RNA, and are presented relative to the no-biotin-aniline control. Bars represent the mean of at least three independent experiments, with error bars showing standard deviation. Statistical significance is denoted as * $p < 0.025$ for one-sample t-tests following Bonferroni correction, while for unpaired t-tests comparing experimental conditions, significance is indicated as ** $p < 0.005$. Non-significant results are labeled as "ns". † in the key denotes whether the delivered mRNA is expected to be enriched in the respective cell line. The schematic below the graph shows the transfected mRNAs.

5.2.2. APEX2-NES HEK293T

The experimental approach, as described in Section 5.2, was then applied to APEX2-NES HEK293T cells. Nluc mRNA is expected to be translated in the cytosol meaning APEX2 labelling of Nluc but not NlucSec mRNA was expected to be detected.

The APEX2-RT-qPCR data collected 4.5 hours post-transfection demonstrated that the Nluc mRNA exhibited approximately 8-fold increase in the streptavidin pulldown from APEX2-labelled cells compared to the no-biotin control (Figure 5.2). However, the results from repeat experiments showed substantial variation, and as a consquence the difference between the biotin and no-biotin

conditions was not statistically significant. Similarly, the NlucSec mRNA showed around 4.5-fold enrichment compared to the no-biotin control, suggesting some level of selectivity in the APEX-NES biotinylation of cytosolically localised delivered mRNA. However, this result was also variable and not statistically significant. In contrast, results from the APEX2-ERM HEK293T cell line (Figure 5.1.) displayed a more consistent pattern of selective biotinylation at the intended subcellular location than results from the APEX2-NES HEK293T cell line (Figure 5.2). Therefore, subsequent experiments focused primarily on the APEX2-ERM cell line.

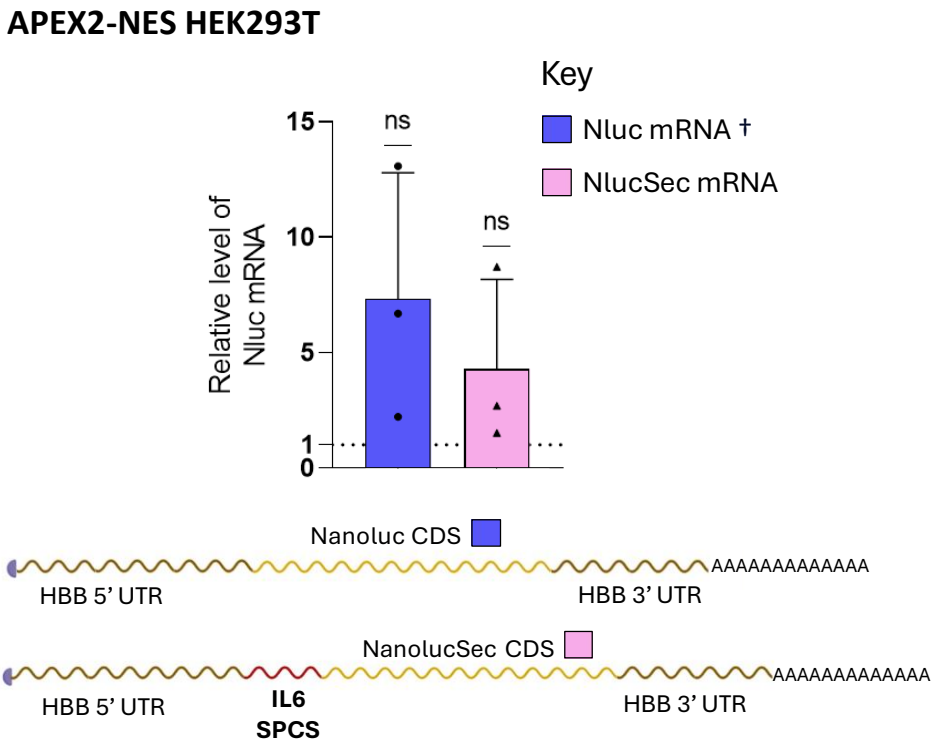


Figure 5.2. APEX2-RT-qPCR analysis of the relative enrichment of transfected Nluc and NlucSec mRNA following streptavidin affinity purification. IVT mRNA was transfected into APEX2-NES HEK293T cells using lipofectamine 2000 and incubated for 4 hours. Cells were then treated with biotin-aniline for 30 minutes before exposure to H_2O_2. Nluc or NlucSec mRNA levels post-streptavidin pulldown were quantified by RT-qPCR, normalised to the biotinylated spike-in Fluc RNA, and are presented relative to the no-biotin-aniline control. Bars represent the mean of three independent experiments, with error bars showing SD. Statistical significance is denoted as ns = not significant. The schematic below the figure illustrates the differences between the two mRNA utilised in this experiment.

5.2.3. APEX2-ERM A549

To investigate whether APEX2-ERM in a different cell line could selectively label ER-localised delivered mRNA, APEX2 labelling was performed in APEX2-ERM

A549 cells following Nluc and NlucSec transfection, using the same conditions as described in Section 5.2.

Almost no difference in the amount of Nluc mRNA recovered in the streptavidin pulldown between biotin-treated cells and the no-biotin control was detected by APEX2-RT-qPCR (Figure 5.3). This suggests that Nluc mRNA, expected to localise to the cytoplasm, was not selectively labelled by APEX2-ERM in this system. NlucSec mRNA showed some enrichment in the biotin-treated condition relative to the no-biotin control; however, the high level of experimental variability limited confidence in this result. These results suggest that the APEX2-ERM A549 system does not robustly differentiate between ER-localised and cytosolic mRNA under these experimental conditions.

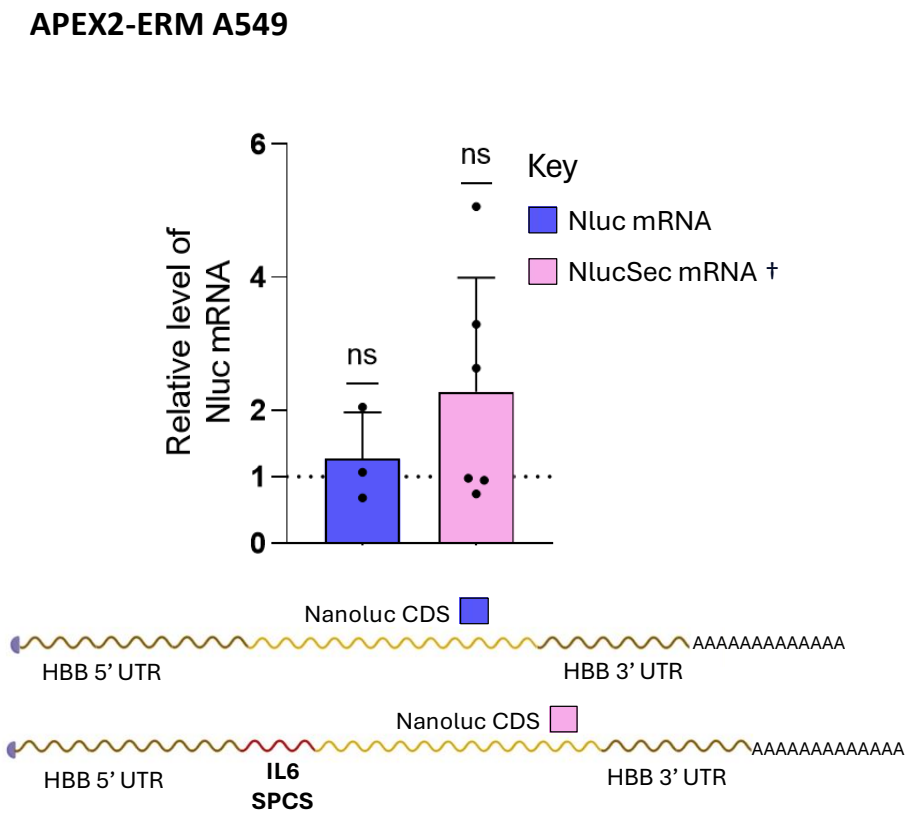


Figure 5.3. APEX2-RT-qPCR analysis of the relative enrichment of transfected Nluc and NlucSec mRNA following streptavidin affinity purification. IVT mRNA was transfected into APEX2-ERM A549 cells using lipofectamine 2000 and incubated for 4 hours. Cells were then treated with biotin-aniline for 30 minutes before exposure to H_2O_2. Nluc or NlucSec mRNA levels post-streptavidin pulldown were quantified by RT-qPCR, normalised to the biotinylated spike-in Fluc RNA, and are presented relative to the no-biotin-aniline control. Bars represent the mean of at least three independent experiments, with error bars showing SD. Statistical significance is denoted as ns = not significant. The schematic below the figure illustrates the two mRNA utilised in this experiment.

5.3. Comparison of effects of different signal peptide coding sequences on Nluc mRNA delivery to the endoplasmic reticulum using APEX2-RT-qPCR

5.3.1. APEX2-ERM HEK293T

Having established that APEX2 labelling can be used to specifically detect ER-localised mRNA in APEX2-ERM HEK293T cells, the method was used as a tool to investigate the effect of different signal peptide coding sequences (SPCS) on ER localisation in the APEX2-ERM HEK293T cell line. To do so, the plasmid encoding NlucSec with the interleukin 6 (IL6) SPCS was modified to replace the SPCS with one derived from the human PMEL gene (Figure 5.4). PMEL mRNA, featuring a native SPCS, encodes a melanoma-specific antigen and has been utilised as a mRNA cancer vaccine in an animal model of disease (Ben-Akiva et al., 2023).

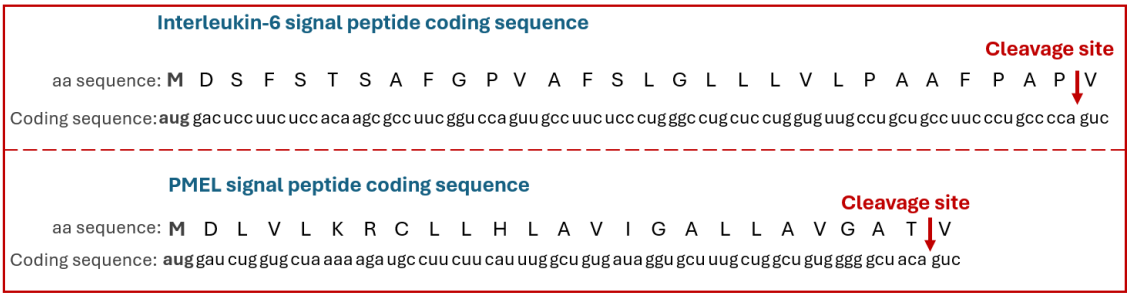


Figure 5.4. Comparison of two reporter mRNA with differing SPCS tested by APEX2-RT-qPCR and secreted nanoluciferase assays. The figure illustrates the amino acid (aa) sequence, the corresponding coding sequence and highlights the signal peptide cleavage site.

Two different NlucSec mRNAs, differing only in their SPCS, were transfected into APEX2-ERM HEK293T cells. Functional protein output was assessed using a secreted nanoluciferase assay, with measurements taken at 1, 2, and 4 hours post-transfection, prior to exposure to biotin-aniline and H_2O_2.

APEX2-RT-qPCR data collected 4.5 hours post-transfection demonstrated that the NlucSec IL6 SPCS Nluc mRNA exhibited approximately 5-fold enrichment, while the NlucSec PMEL SPCS mRNA showed around 7-fold enrichment compared to their respective no-biotin-aniline control (Figure 5.5A). This suggests that the PMEL SPCS facilitated more efficient mRNA trafficking to the ER at this time point, although the difference was not statistically significant

due to experimental variability. The NlucSec IL6 SPCS showed significant enrichment in APEX2-ERM HEK293T cells treated with biotin-aniline relative to its negative control, whereas the NlucSec PMEL SPCS mRNA did not achieve statistical significance owing to data variability. The NlucSec system as a secreted reporter protein system provided a means to quantify the translation of delivered mRNA in parallel with localisation studies. Secreted nanoluciferase activity varied between the two mRNAs 1 hour post-transfection, with NlucSec IL6 SPCS showing approximately 5-fold higher raw luminescence compared to NlucSec PMEL SPCS (Figure 5.5B – note that the graph is plotted on log scale). However, this difference diminished over time. Taken together with the APEX2-RT-qPCR data, these results suggest that the PMEL signal peptide may be more efficient at targeting NlucSec mRNA to the ER. However, as these differences were not statistically significant, further investigation, including additional time points, would be necessary to confirm this trend.

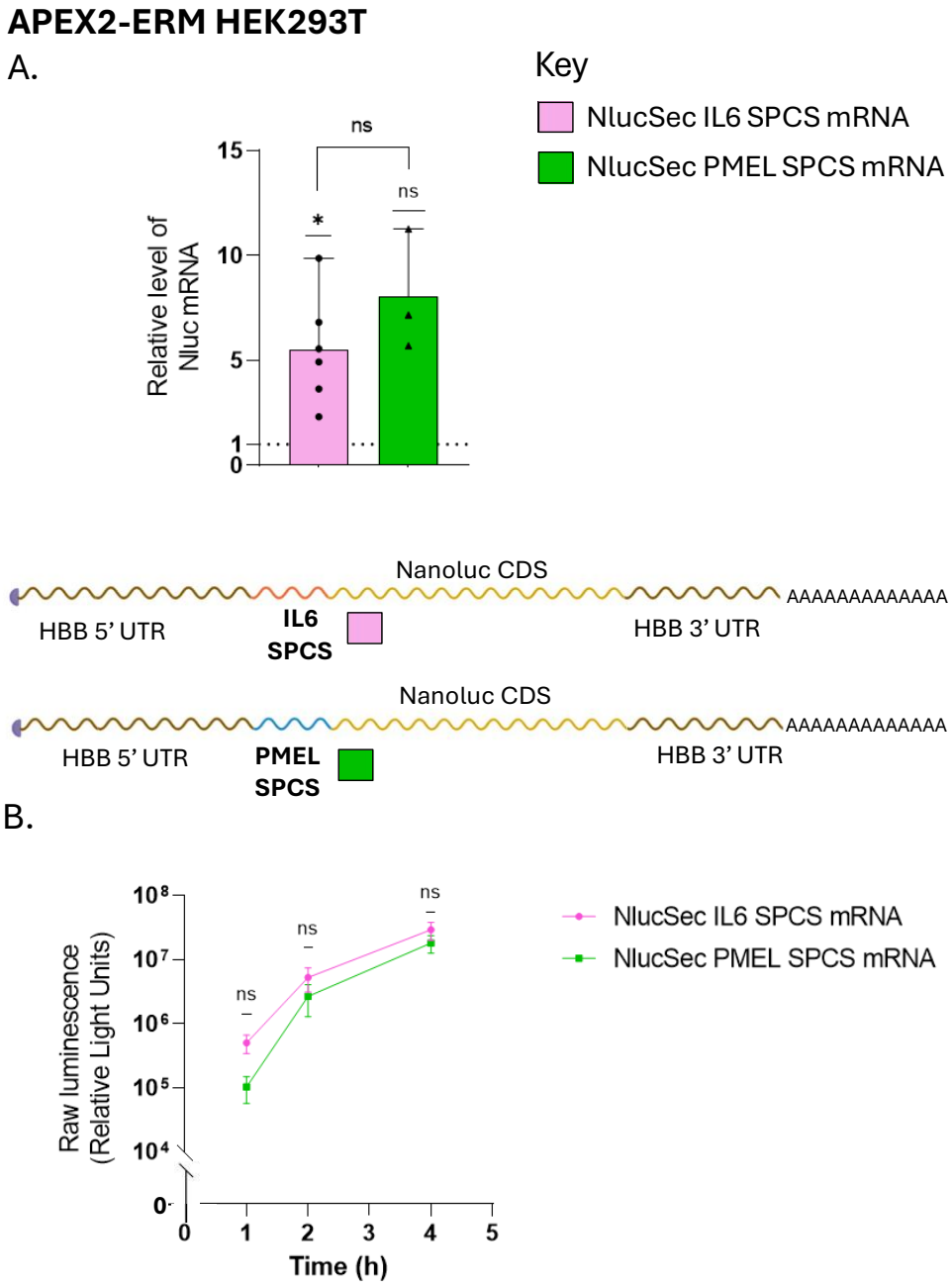


Figure 5.5. APEX-RT-qPCR analysis of delivered Nluc mRNA with different SPCS to the ER and secreted luciferase assays to assess functional protein output in APEX2-ERM HEK293T cells. (A.) APEX2-RT-qPCR analysis of the relative enrichment of transfected NlucSec (IL6 SPCS) and NlucSec (PMEL SPCS) mRNA following streptavidin affinity purification. IVT mRNA was transfected into APEX2-ERM HEK293T cells by lipofectamine 2000 and incubated for 4 hours. Cells were then treated with biotin-aniline for 30 minutes before exposure to H_2O_2. NlucSec IL6 SPCS or Nluc PMEL SPCS mRNA levels post-streptavidin pulldown were quantified by RT-qPCR, normalised to the biotinylated spike-in Fluc RNA, and are presented relative to the no-biotin-aniline control. Data represent mean of three independent experiments, with error bars showing SD. The schematic below the figure illustrates the two mRNA utilised in this experiment. (B.) Functional protein output assessed by secreted nanoluciferase activity prior to incubation of cells with biotin-aniline. Measurements were taken at 1 hour (N=3), 2 hours (N=3), and 4 hours (N=3), with three technical repeats at each time point. Bars represent the mean of at least three independent experiments, with error bars showing SD. Statistical significance in A is denoted as * = P < 0.025. Non-significant results are labeled as "ns".

5.3.2. APEX2-ERM A549

In order to establish whether there are cell type-specific differences in SPCS targeting to the ER, APEX2-ERM A549 cells were transfected with two different NlucSec mRNAs, differing only in their SPCS, using Lipofectamine 2000. Functional protein output was assessed by secreted nanoluciferase assay at 1, 2, and 4 hours post-transfection, prior to incubation of cells with biotin-aniline and H_2O_2 treatment.

APEX2-RT-qPCR analysis of the streptavidin pulldown showed that NlucSec IL6 SPCS mRNA was approximately 2-fold higher than its no-biotin-aniline negative control, though this enrichment was not statistically significant (Figure 5.6A). In contrast, NlucSec PMEL SPCS mRNA exhibited a 6-fold enrichment compared to its no-biotin-aniline control. Furthermore, the relative level of NlucSec PMEL SPCS mRNA was significantly higher than that of NlucSec IL6 SPCS mRNA. Overall, the APEX2-RT-qPCR data in APEX2-ERM A549 cells shows a similar trend to that observed in APEX2-ERM HEK293T, with the differences in NlucSec PMEL SPCS mRNA enrichment relative to NlucSec IL6 SPCS mRNA appearing more pronounced in the A549 cell line.

While secreted nanoluciferase activity following NlucSec PMEL SPCS transfection was consistently lower than that of NlucSec IL6 SPCS mRNA at all time points, this difference was not statistically significant (Figure 5.6B). The largest difference in luciferase activity occurred at the 4-hour mark; however, due to variability in raw luminescence across the three independent experiments, this difference did not reach significance. In summary, in the APEX2-ERM A549 system, the activity of secreted nanoluciferase was somewhat lower after the transfection of NlucSec PMEL SPCS mRNA compared to that of NlucSec IL6 SPCS mRNA. This finding was unexpected, especially considering that the APEX2-RT-qPCR data indicated a higher relative level of NlucSec PMEL SPCS mRNA at the ER at the 4.5-hour time point. However, it is important to note that the difference in luminescence was not statistically significant.

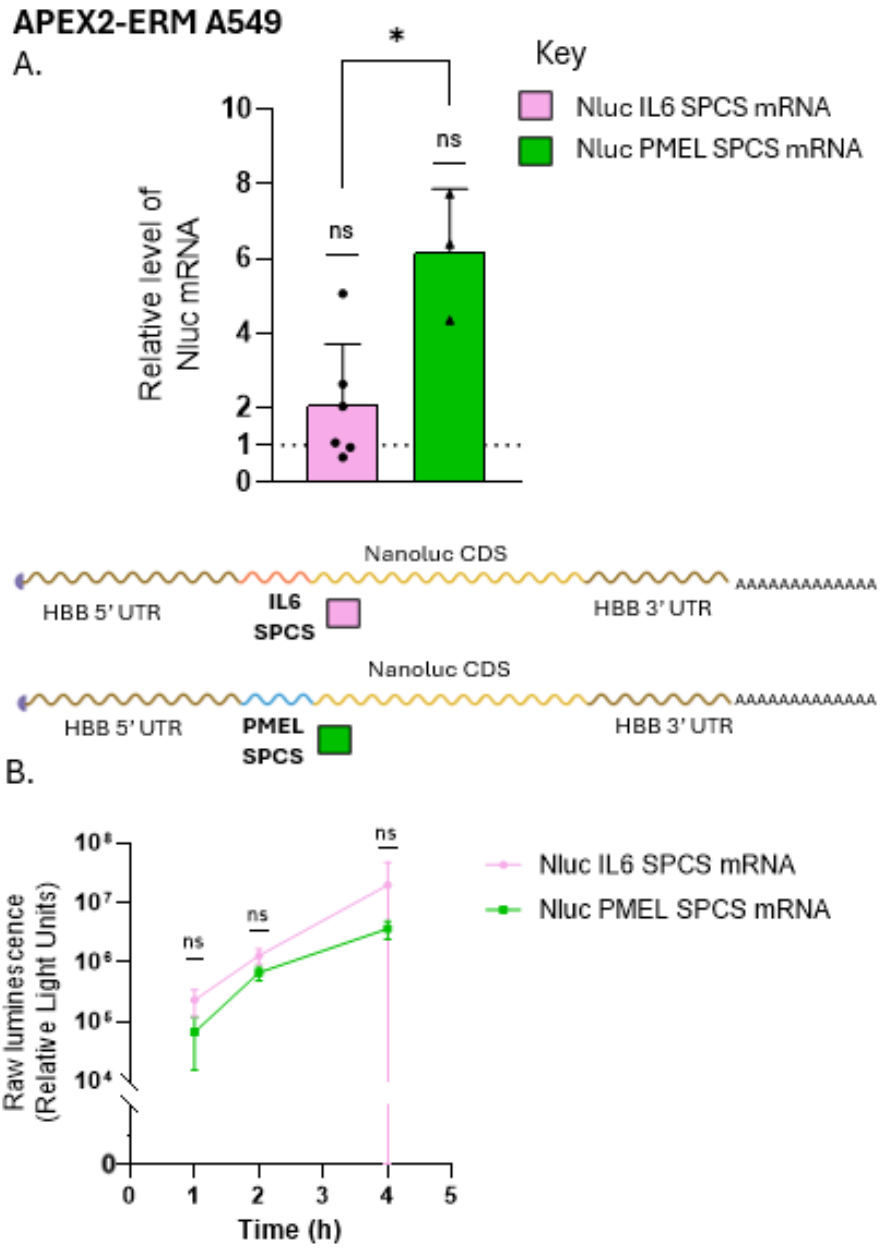


Figure 5.6 APEX-RT-qPCR analysis of delivery of Nluc mRNA with different SPCS to the ER and secreted luciferase assays to assess functional protein output in APEX2-ERM A549 cells. (A.) APEX2-RT-qPCR analysis of the relative enrichment of transfected NlucSec (IL6 SPCS) and NlucSec (PMEL SPCS) mRNA following streptavidin affinity purification. IVT mRNA was transfected into APEX2-ERM A549 cells by lipofectamine 2000 transfection and incubated for 4 hours. Cells were then treated with biotin-aniline for 30 minutes before exposure to H_2O_2. Nluc mRNA levels post-streptavidin pulldown were quantified by RT-qPCR, normalised to the biotinylated spike-in Fluc RNA, and are presented relative to the no-biotin-aniline control. Bars represent the mean of three independent experiments, with error bars showing SD. The schematic below the figure illustrates the differences between the two mRNA utilised in this experiment. (B.) Functional protein output assessed by secreted luciferase activity prior to incubation of cells with biotin-aniline. Measurements were taken at 1 hour (N=3), 2 hours (N=3), and 4 hours (N=3), with three technical repeats at each time point (n=3). Data represent mean of at least three independent experiments, with error bars showing SD. Statistical significance is denoted as * $P< 0.05$; ns = not significant.

5.4. Investigation of signal peptide coding sequence as a feature in therapeutic mRNA by APEX2-RT-qPCR

To explore the use of APEX2-RT-qPCR to detect subcellular localisation of delivered mRNA encoding therapeutically relevant proteins, three PMEL constructs were cloned into plasmids. The wildtype human PMEL CDS purchased as GeneBlocks with its native SPCS (WT PMEL); a modified construct where the PMEL SPCS was replaced with the IL6 SPCS (PMEL IL6 SPCS) and a PMEL construct purchased as a GeneBlock with the SPCS removed, which is anticipated to prevent co-translational trafficking to the ER (Cytosolic PMEL). To confirm successful expression, the three PMEL mRNAs were transfected into APEX2-ERM HEK293T cells and incubated for 4 hours before protein extraction and western blot analysis. In parallel, NlucSec (IL6 SPCS) mRNA was transfected to confirm transfection efficiency, and secreted luciferase activity was measured, showing similar levels to those observed in previous NlucSec transfection experiments (data not shown). PMEL protein expression was detected in total cell lysate of cells transfected with the PMEL mRNA, PMEL IL6 SPCS mRNA, and Cytosolic PMEL mRNA by western blot (Figure 5.7). There were no major differences in PMEL protein expression across the three mRNAs; however, PMEL protein levels following transfection with cytosolic PMEL mRNA appeared slightly lower than those following transfection with the wildtype human PMEL mRNA, while the PMEL IL6 SPCS mRNA produced approximately two-fold less PMEL protein than the wildtype human PMEL mRNA.

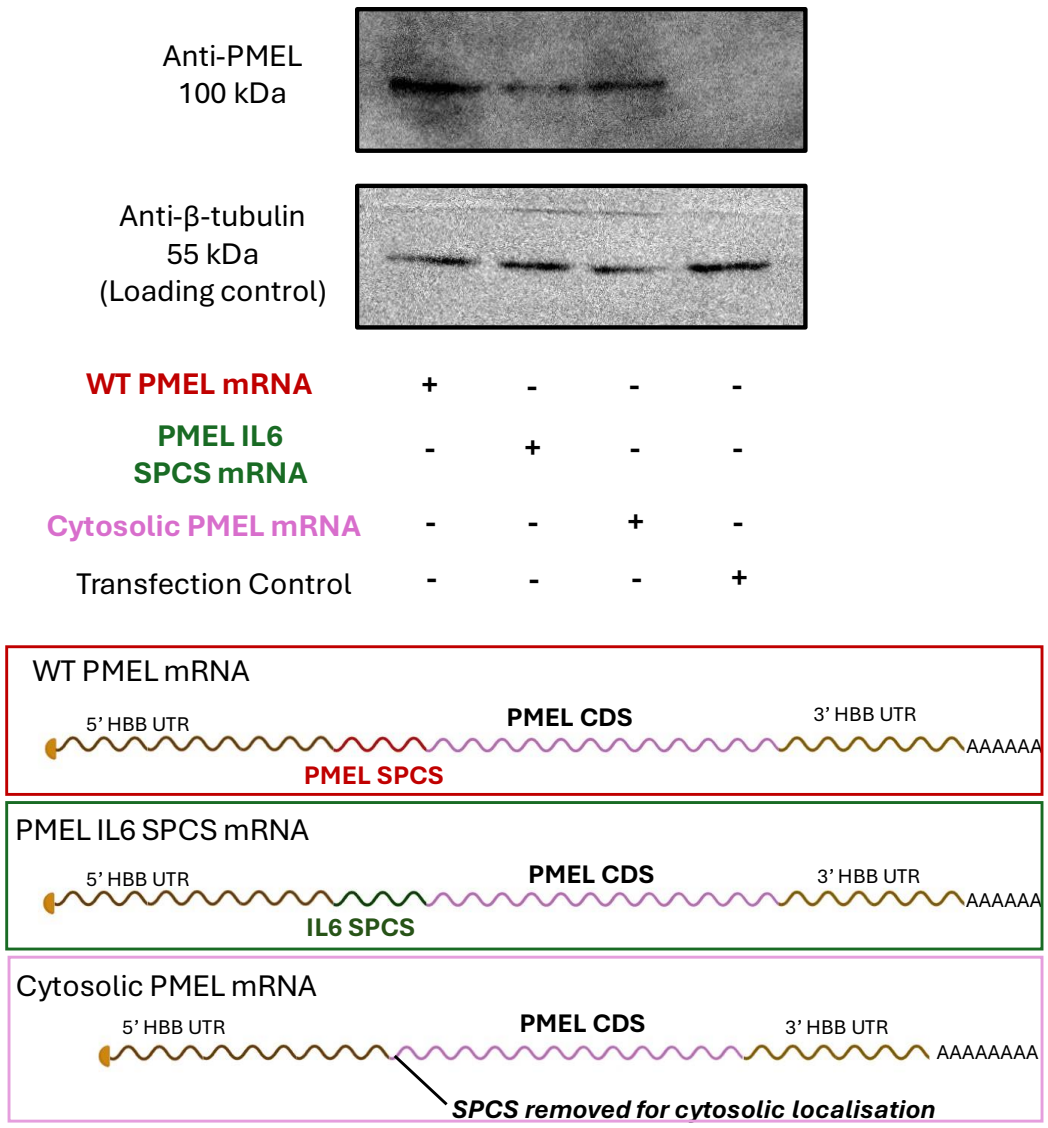


Figure 5.7. Western blot analysis of PMEL protein expression following transfection of different PMEL mRNAs into APEX2-ERM HEK293T cells. IVT mRNA was transfected into APEX2-ERM HEK293T cells using Lipofectamine 2000, followed by a 4-hour incubation prior to protein extraction. A negative control, where APEX2-ERM HEK293T cells were transfected with NlucSec mRNA, was included. The image is representative of three independent experiments. The schematic below the figure illustrates the sequences of the three mRNA used in this experiment.

To assess how substituting the SPCS affects PMEL mRNA localisation as detected by APEX2-RT-qPCR, PMEL mRNA with its native SPCS and PMEL IL6 SPCS mRNA were transfected into APEX2-ERM HEK293T cells and the APEX2 biotinylation reaction carried out 4.5 hours post-transfection.

APEX2-RT-qPCR data collected at 4.5 hours post-transfection showed that approximately 8-fold more PMEL mRNA was detected post-streptavidin pulldown in cells transfected with the PMEL native SPCS mRNA, while the PMEL IL6 SPCS mRNA showed a more modest 5-fold enrichment relative to their

respective negative controls (Figure 5.8). Although the difference in relative levels between these two conditions was not statistically significant due to experimental variability, this trend suggests that the PMEL SPCS may facilitate more efficient delivery of PMEL mRNA to the ER than the IL6 SPCS. This mirrors the general trend observed in the Nluc experiments in Section 5.3 and mirrors the pattern of PMEL protein production from these two mRNAs as detected by western blot (Figure 5.7).

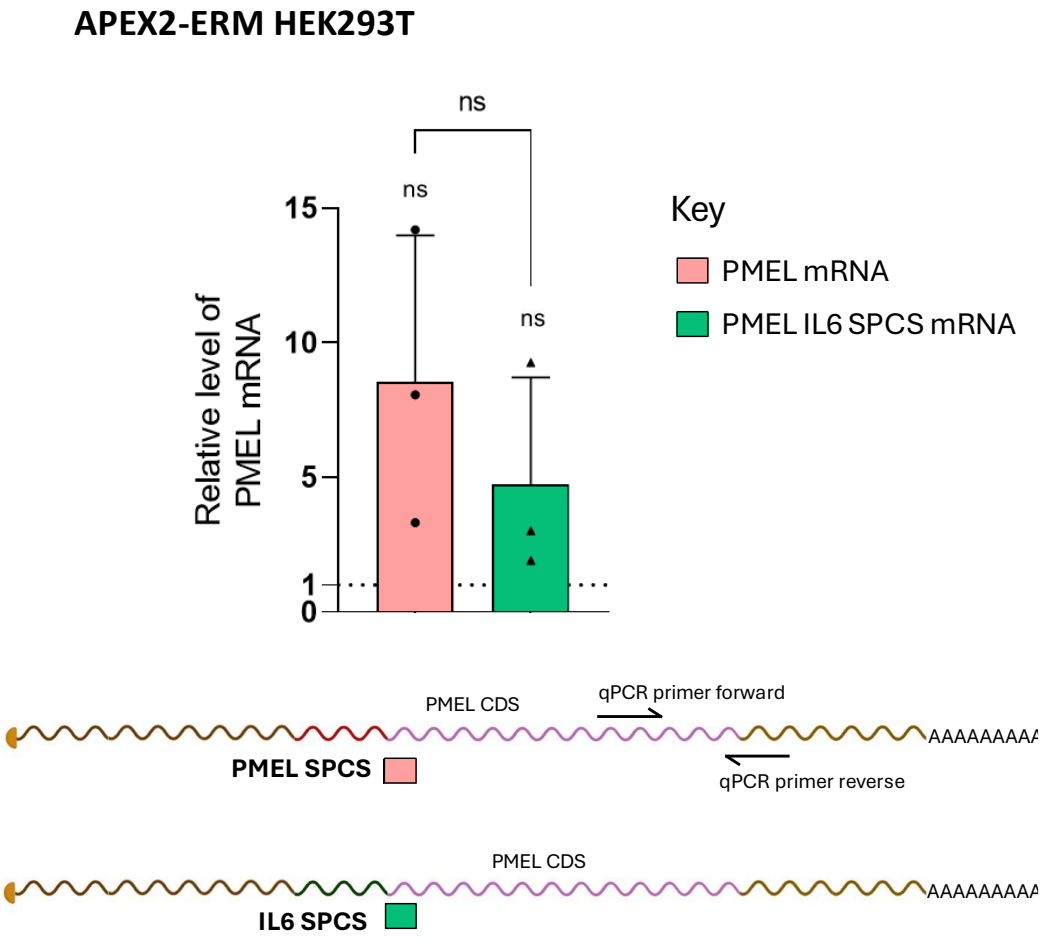


Figure 5.8. APEX-RT-qPCR analysis of delivery of PMEL mRNA with different SPCS to the ER in APEX-ERM HEK293T cells. PMEL IVT mRNA with either its native SPCS or IL6 SPCS were delivered into APEX2-ERM HEK293T cells by Lipofectamine 2000 transfection and incubated for 4 hours. Cells were then treated with biotin-aniline for 30 minutes before exposure to H_2O_2. PMEL mRNA levels post-streptavidin pulldown were quantified by RT-qPCR, normalised to the biotinylated spike-in Fluc RNA, and are presented relative to the no-biotin-aniline control. Bars represent the mean of three independent experiments, with error bars showing SD. Statistical significance is denoted as ns = not significant. The schematic below the figure illustrates the two mRNA utilised in this experiment.

5.5. Investigation of effect of puromycin on APEX2-RT-qPCR sensitivity

Active translation could obstruct APEX2 labelling, as ribosomes may physically block biotinyl radicals from reaching the mRNA, potentially causing these radicals to react with the ribosome itself rather than with the intended mRNA target. To assess whether active translation might interfere with APEX2 labelling, puromycin was tested as a potential method to increase the accessibility of APEX2-generated biotinyl radicals to nearby RNA. Puromycin works as an aminoacyl tRNA analogue and its incorporation into the nascent peptide causes ribosomes to dissociate from mRNA (Aviner, 2020).

Cytosolic PMEL mRNA was introduced into APEX2-NES HEK293T cells using Lipofectamine 2000. After a 4 hour incubation, the cells were treated with biotin-aniline for 30 minutes, followed by a two-minute exposure to puromycin before initiating the APEX2 labelling reaction with H_2O_2. Control conditions included the omission of puromycin and a separate no-biotin-aniline negative control for comparison and normalisation, respectively.

Efficient APEX2 labelling of cytosolic PMEL mRNA was observed, with APEX2-NES HEK293T cells treated with only biotin-aniline and H_2O_2 showing just over 15-fold enrichment relative to the no-biotin-aniline control (Figure 5.9). Cells exposed to both puromycin and biotin-aniline exhibited over 200-fold enrichment, suggesting that puromycin may enhance APEX2 labelling in this cell line. However, this is based on one independent experiment, and further investigation is required for validation.

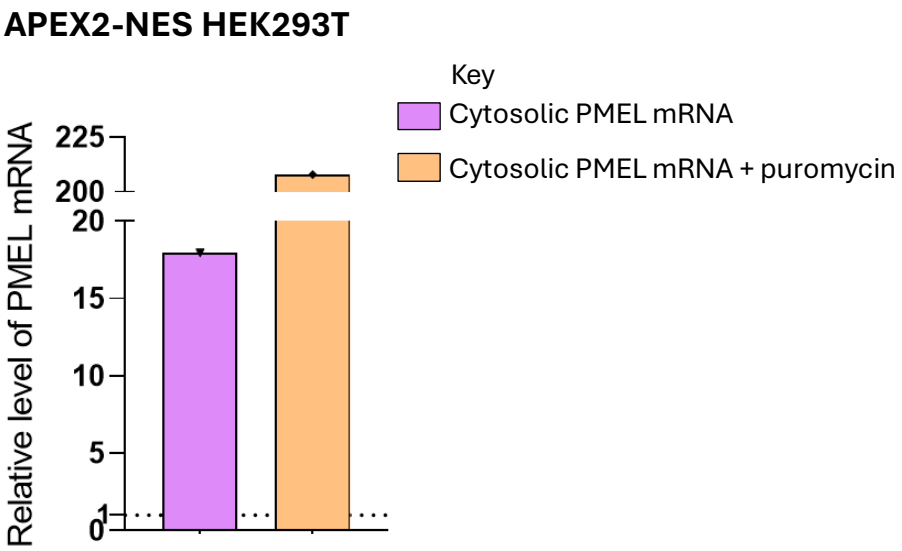


Figure 5.9. APEX-RT-qPCR analysis of puromycin's effect on APEX2 labelling following cytosolic PMEL mRNA delivery. Cytosolic PMEL mRNA was transfected into APEX2-NES HEK293T cells using Lipofectamine 2000 and incubated for 4 hours. Cells were then treated with biotin-aniline for 30 minutes, followed by a 2-minute incubation with puromycin prior to H_2O_2 exposure to initiate APEX2 labelling. PMEL mRNA levels post-streptavidin pulldown were quantified by RT-qPCR, normalised to the biotinylated spike-in Fluc RNA, and are presented relative to the no-biotin-aniline control. Bars represent one independent experiment.

5.6. Biochemical membrane fractionation experiments as an alternative to APEX2-RT-qPCR

5.6.1. Assessment of biochemical membrane fractionation by western blot and endogenous RT-qPCR analysis

Biochemical membrane fractionation was explored as an alternative method for investigating the subcellular localisation of both endogenous and delivered mRNA in APEX2-ERM HEK293T cells. APEX2-ERM HEK293T cells were fractionated into membrane, cytosolic, and nuclear components following NlucSec mRNA delivery *via* Lipofectamine 2000, enabling simultaneous investigation of protein localisation and the distribution of both endogenous and delivered RNA.

To validate this technique, western blots were conducted post-fractonation to qualitatively confirm the extent to which proteins known to localise to the ER which should form part of the membrane-bound fraction (Calnexin), cytosol (β-tubulin), and nuclear membrane (Lamin A/C) were restricted to their respective compartments.

The fractionation technique appeared to exhibit varying levels of cross-contamination across different fractions, as analysed by Western blot (Figure 5.10A). The nuclear fraction contained lamin A/C, as anticipated, along with the ER membrane protein calnexin, suggesting potential contamination from the membrane-bound fraction. The cytosolic fraction displayed the expected presence of β-tubulin but also contained calnexin, again indicating contamination from the membrane-bound fraction. The membrane-bound fraction showed a much stronger calnexin signal compared to the weaker contaminant bands observed in the nuclear and cytosolic fractions. Moreover, very little β-tubulin and lamin A/C were detected in the membrane-bound fraction, suggesting minimal cross-contamination of proteins expected to be present in the cytosolic and nuclear fractions.

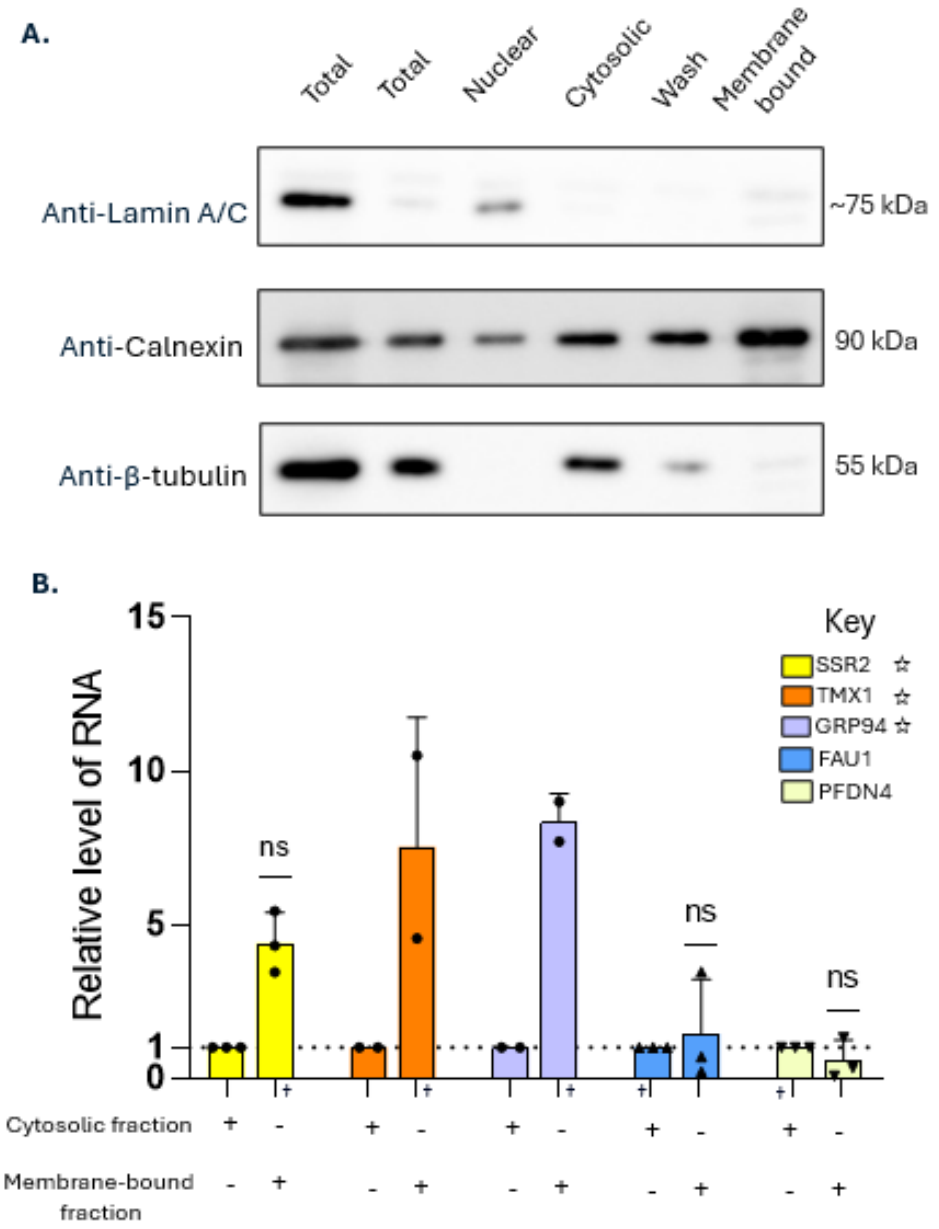


Figure 5.10. Analysis of protein distribution and transcript levels across fractions following biochemical membrane fractionation. (A) Western blot analysis shows the presence of proteins across cytosolic, membrane-bound (e.g., ER-associated), and nuclear fractions. "Total" refers to the lysate prior to fractionation, while the "wash" fraction represents proteins extracted from a wash step between the cytosolic and membrane-bound fractions. Lamin A/C, β-tubulin and calnexin serve as markers for the nucleus, cytoplasm and ER respectively. The image is representative of three independent experiments. (B) Relative levels of transcripts predicted to be associated with the cytosol or the ER across cytosolic and membrane-bound fractions, analysed by RT-qPCR. Data were normalised to a spike-in eGFP RNA and are shown relative to the cytosolic fraction. SSR2, GRP94, and TMX1 are expected to be ER-associated and are denoted with ☆ in the key, while FAU1 and PFDN4 are expected to be cytosolic. qPCR data for GRP94, and TMX1 represent the mean of two independent experiments. qPCR data for SSR2, FAU1 and PFDN4 represent the mean of three independent experiments, with error bars showing standard deviation. Statistical significance is denoted as “ns” = not significant.

Before investigating the localisation of delivered mRNA using the biochemical fractionation technique, it was necessary to first assess how effectively the fractionation process could separate endogenous RNA expected to be either

cytosolic or membrane-bound. SSR2, SFT2D2, and TMX1 were selected as ER-associated mRNA markers, having been identified as enriched in the APEX2-ERM cell line through APEX-seq by Fazal et al. (2019). GRP94 was also included as an ER-associated mRNA marker based on its enrichment in the APEX2-ERM cell line detected through APEX-seq by Padrón et al. (2019). FAU1 and PFDN4 served as cytosolic mRNA markers, as these transcripts were identified as enriched in the APEX2-NES cell line through APEX-seq by the Ting lab (Fazal et al., 2019) and in the Ingolia lab (Padrón et al., 2019) respectively.

RT-qPCR data were normalised using a spike-in RNA, added to each fraction in equal amounts before RNA extraction, allowing for comparison across fractions. Each transcript is presented relative to the cytosolic fraction.

With regards to ER-assocated mRNAs, SSR2 showed approximately four-fold enrichment in the membrane-bound fraction compared to the cytosolic fraction, with this difference reaching statistical significance (Figure 5.10B). Both GRP94 and TMX1 transcripts showed a general trend of enrichment in the membrane-bound fraction compared to the cytosolic fraction but were not tested for statistical significance due to the limited number of replicates. This pattern aligns with their expected ER localisation, though additional replicates are needed to confirm this trend.

Next, FAU1 and PFDN4 were analysed across both the cytosolic and membrane-bound fractions. For both transcripts, the relative amount of mRNA in the cytosolic fraction was not significantly higher than in the membrane-bound fraction (Figure 5.10B). However, FAU1 and PFDN4 were not enriched in the membrane-bound fraction, allowing for clear distinction from the ER-associated transcripts. This indicates that while the biochemical membrane fractionation technique may have limitations in separating cytosolic-associated mRNA, it is effective in isolating ER-associated endogenous transcripts.

5.6.2. Enrichment of delivered reporter mRNA in fractionated cells

The biochemical membrane fractionation technique was then utilised to investigate the subcellular localisation of delivered Nluc and NlucSec mRNA. APEX2-ERM HEK293T cells were transfected with either Nluc or NlucSec mRNA, followed by a 4 hour incubation period before fractionation and subsequent analysis by RT-qPCR.

Both Nluc and NlucSec mRNA showed approximately 3-fold to 4-fold enrichment in the membrane-bound fraction relative to the cytosolic fraction (Figure 5.11). However, NlucSec did not show statistically significant enrichment, and the result for Nluc was based on only one repeat. These findings suggest that this method is not effective for reliably discriminating the subcellular localisation of IVT RNA.

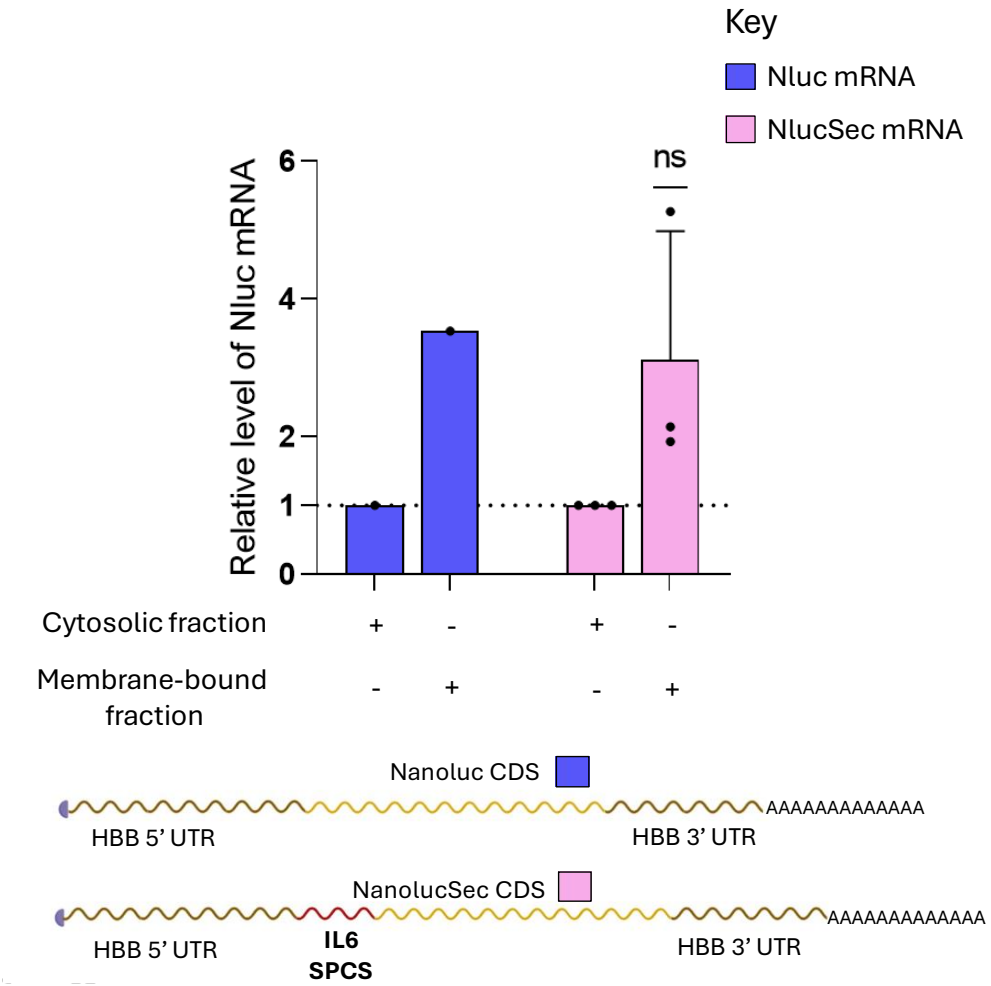


Figure 5.11. Relative levels of transfected mRNA across cytosolic and membrane-bound fractions following biochemical membrane fractionation, analysed by RT-qPCR. IVT mRNA was delivered into APEX2-ERM HEK293T cells via lipofectamine 2000 transfection and incubated for 4 hours before membrane fractionation. Nluc or NlucSec mRNA levels post-fractionation were quantified by RT-qPCR, normalised to an elution spike in Fluc RNA, and are presented relative to the cytosolic fraction. Bars represent the mean of three independent experiments for NlucSec and one independent experiment for Nluc, with error bars indicating SD. Statistical significance is denoted as "ns" = not significant. The schematic below the graph illlustrates the transfected mRNAs.

5.7. Investigation of between body and tail labelling of mRNA

To assess subcellular localisation of delivered mRNA, methods for directly visualising mRNA that could complement APEX2-RT-qPCR were explored. Recent advances in mRNA labelling techniques have provided new ways to visualise mRNA without compromising its translational efficiency. Commonly used labelling methods, which incorporate fluorophores into the CDS during IVT can cause ribosome stalling and thus alter how the mRNA interacts within cells, making them unsuitable for studying IVT mRNA dynamics in live cells (Custer & Walter, 2017). An alternative method, Between Body and Tail (BBT) labelling,

involves selectively labelling the beginning of the poly(A) tail (Figure 5.12). This approach uses yeast poly(A) polymerase to incorporate N_3-dATPs in a short initial round of A-tailing, followed by purification, further polyadenylation by E. Coli poly(A) polymerase and then addition of fluorophores by strain-promoted-azide-alkyne-cycloaddition (SPAAC) click-chemistry. By labelling the start of the poly(A) tail, canonical adenines in the rest of the poly(A) tail are free to bind to poly(A) binding proteins (PABPs), which is crucial for cap-dependent translation (as reviewed by Passmore & Coller (2022)). As a result, BBT labelling minimises interference with translation and may be more suitable for live-cell imaging of mRNA delivery, with studies showing no negative impact on translational output (Anhäuser et al., 2019; Westerich et al., 2020). However, BBT labelling has yet to be applied for visualising the subcellular localisation of delivered mRNA.

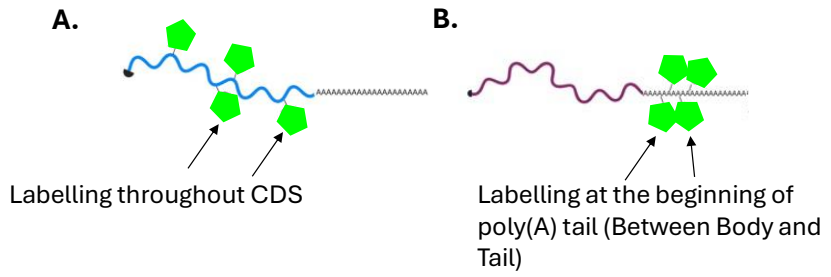


Figure 5.12. Schematic illustrating fluorophore labelled mRNAs. (A.) Shows (green) fluorophore labelling throughout the CDS whereas (B.) shows between body and tail labelling of mRNA.

To evaluate whether BBT labelling affects the translation of IVT mRNA, a secreted nanoluciferase assay was conducted comparing BBT-labelled NlucSec mRNA with unlabelled NlucSec mRNA containing UTRs derived from Hepatitis C Virus (HCV). APEX2-ERM HEK293T cells were transfected with either NlucSec mRNA using lipofectamine 2000 and a 4-hour time course was conducted to measure secreted nanoluciferase activity. The results from the secreted nanoluciferase assay indicated that BBT-labelled mRNA exhibited higher translational output than the NlucSecHCV mRNA (Figure 5.13). Although it was not determined in this study the extent to which the difference in translational output was due to the different UTRs, the different method of translation initiation or the labelling of the poly(A) tail increasing mRNA stability in cells, the findings suggest that BBT-labelled mRNA is translatable within

cells. This supports previous reports (Anhäuser et al., 2019; Custer & Walter, 2017; Westerich et al., 2020) indicating that this labelling strategy does not significantly impact translational efficiency.

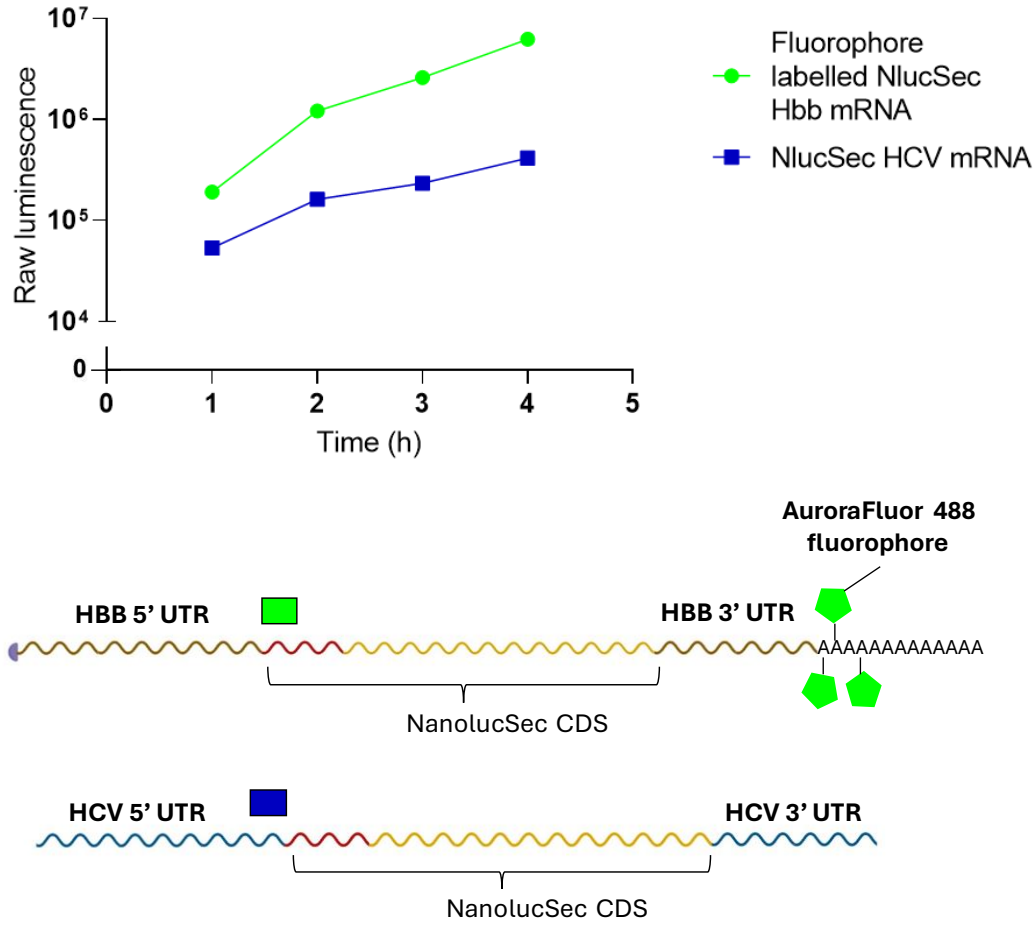
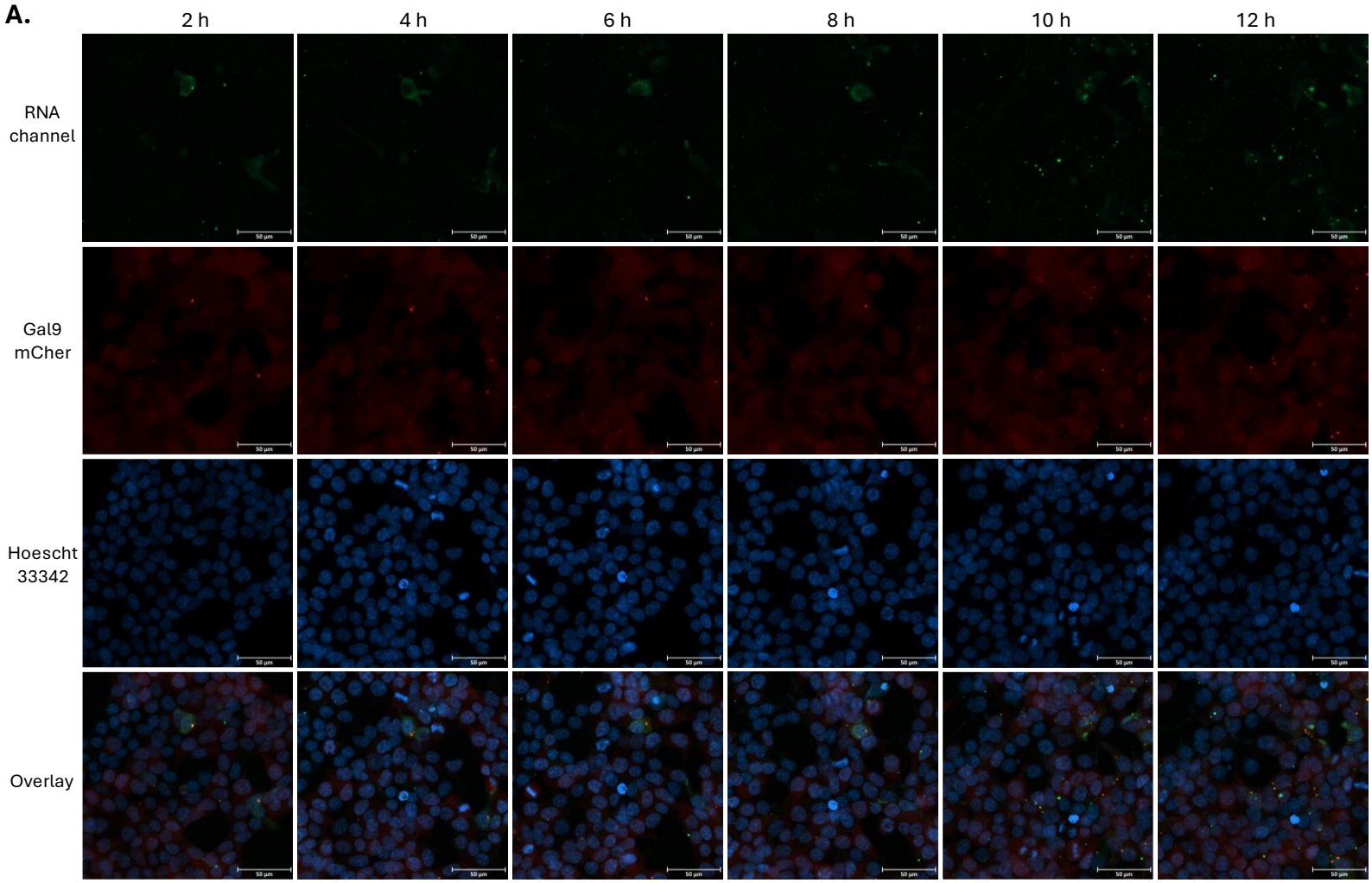


Figure 5.13. Comparison of secreted luciferase activity between BBT-labelled NlucSec mRNA and NlucSec HCV mRNA. APEX2-ERM HEK293T cells were transfected with either NlucSec mRNA labelled using the BBT approach with AuroraFluor 488 DBCO or unlabelled NlucSec mRNA using Lipofectamine 2000 (N=1). Luminescence was measured at multiple time points following media removal, with three technical replicates taken at each time point. The schematic below the figure shows the two mRNAs utilised in this experiment.

Having confirmed that BBT labelling allows efficient translation, live-cell imaging was employed to investigate the intracellular localisation of BBT-labelled mRNA. HEK293T cells stably expressing Galectin-9-mCherry (Gal9-mCherry), a system that allows for the visualisation of damaged endosomes due to its recruitment to the inner part of ruptured endosomal membranes (Munson et al., 2021), were transfected with either AuroraFluor 488-labelled NlucSec mRNA with Lipofectamine MessengerMax or a negative control where Gal9mCher cells were incubated with non-labelled naked mRNA. Imaging

revealed small puncta in the green channel, corresponding to the labelled mRNA (Figure 5.14A), some of which co-localised with red puncta that are indicative of endosomal damage. This suggests that a portion of the labelled mRNA may be retained within endosomes, while the remaining puncta may reflect mRNA that has successfully escaped into the cytoplasm. The negative control shows no puncta in the green channel (Figure 5.14B). Interestingly, the background mCherry signal in the negative control (Figure 5.14B) seems to get stronger throughout the time course, which could potentially be due to cell proliferation.


A.
2 h
4 h
6 h
8 h
10 h
12 h
RNA channel
Gal9 mCher
Hoescht 33342
Overlay

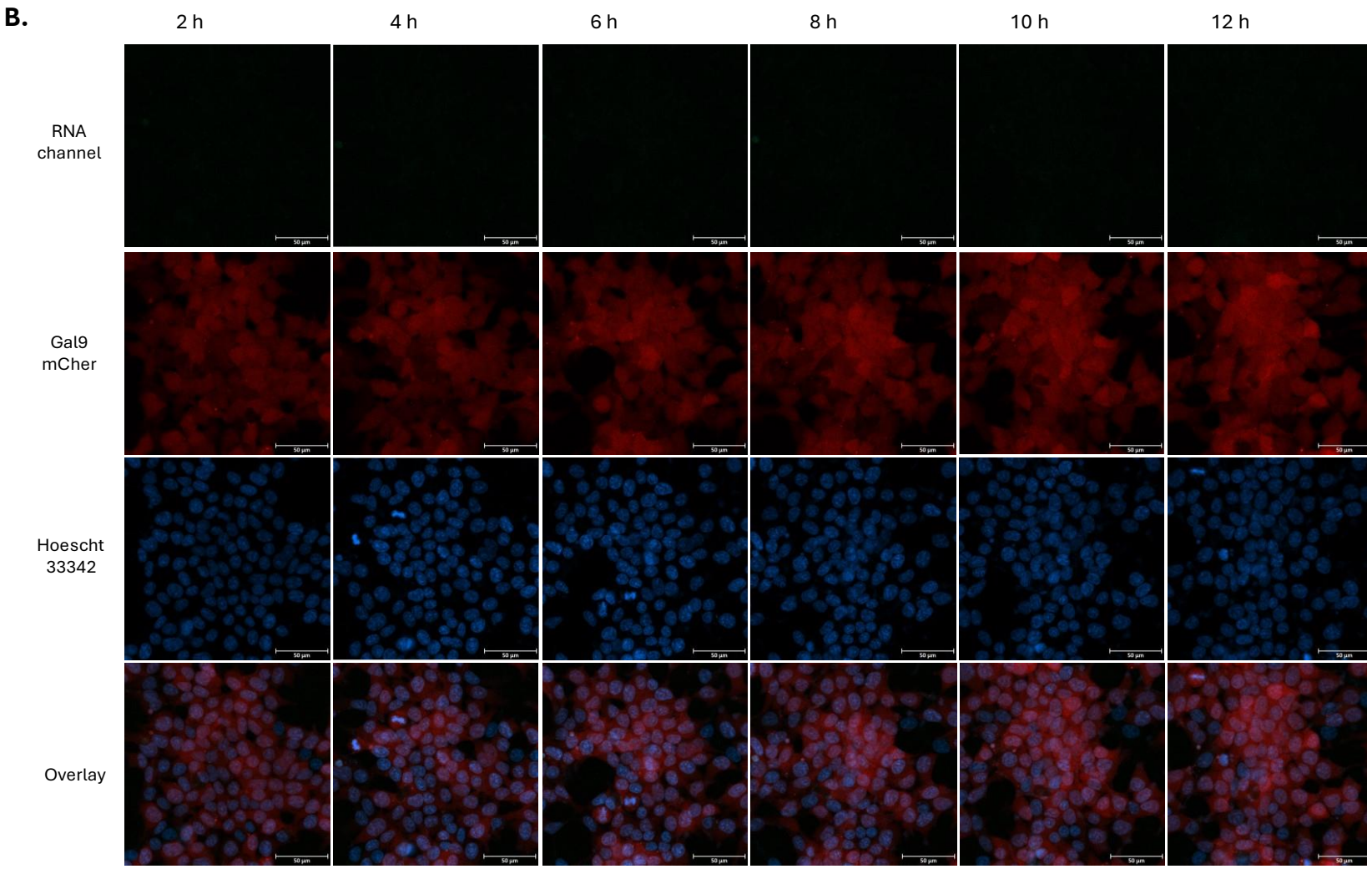


Figure 5.14. Live-cell time course imaging of delivered BBT-labelled mRNA in Gal9-mCherry HEK293T cells. NlucSec mRNA labelled using the BBT approach with AuroraFluor 488 (shown in A) or HEK293T Gal9-mCherry cells were incubated with naked eGFP mRNA (B.). Cells were maintained at 37°C with 5% CO_2 and imaged every hour for 12 hours to visualise the fluorophore-labelled mRNA. Images were captured at 10X magnification, with scale bars representing 50 µm. The images shown are representative of a single independent experiment and were taken by Sal Jones.

5.8. Pulse chase experiments with reporter mRNA encoding for fluorescent proteins.

To establish a comprehensive toolkit for investigating mRNA delivery using microscopy alongside APEX2-RT-qPCR and the BBT labelled approach, the Nluc CDS was replaced with mCherry or eGFP in plasmids used as templates for IVT (Figure 5.15). These CDSs did not have an SPCS, thereby enabling future comparisons of mRNA delivery using the APEX2-NES HEK293T system.

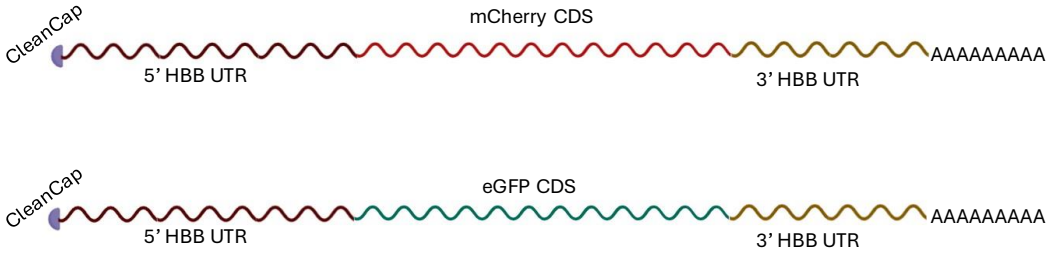
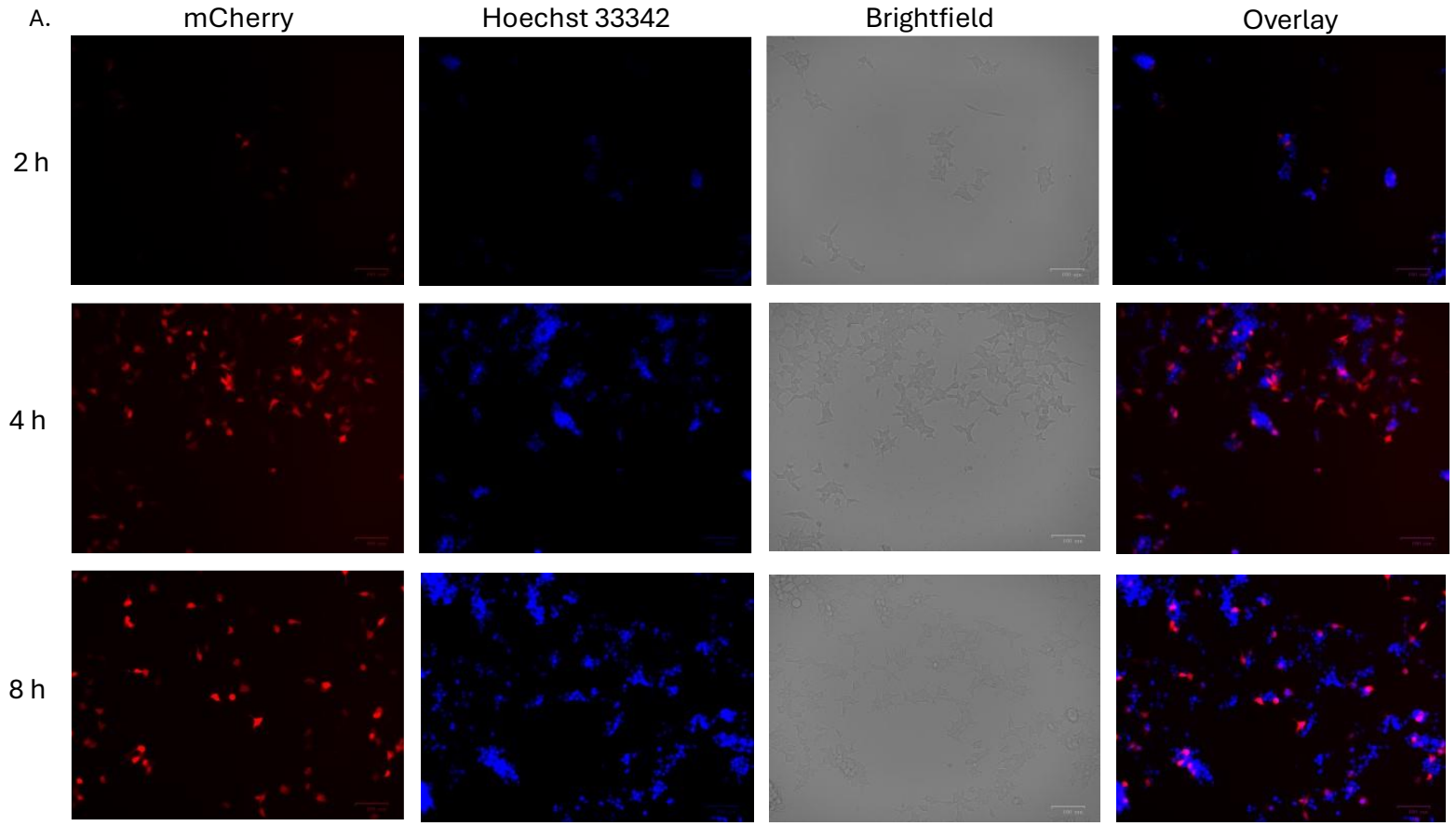


Figure 5.15. Schematic showing mRNAs encoding mCherry and eGFP flanked by human β-haemoglobin (HBB) UTRs and capped with CleanCap AG.

To gain insights into the dynamic processes of mRNA uptake and trafficking of mRNA encapsulated in Lipofectamine 2000 nanoparticles, a pulse-chase approach was employed with APEX2-ERM HEK293T cells. Prior to incubation with Lipofectamine complexes, cells were labelled with Hoechst 33542 to normalise transfection efficiency relative to cell number and to facilitate the quantification of differences in protein expression across various pulse conditions. However, APEX2-ERM HEK293T cells exhibited adverse reactions to PBS washing post-Hoechst treatment, leading to significant cell stacking and detaching during the washing process. As a result, accurate cell counting was compromised, necessitating a qualitative rather than quantitative assessment of experimental results. The pulse phase involved incubating the cells with Opti-MEM containing IVT reporter mRNA encapsulated in Lipofectamine 2000, while the chase phase utilised fresh Opti-MEM. The duration of the pulse varied, with the shortest set at 20 minutes, the longest at 60 minutes, and a control group where no chase was added.

mCherry protein expression was first observed 2 hours post-chase, consistent with the expected time lag between delivery and detectable reporter protein expression. Interestingly, the 20-minute pulse was sufficient to yield some reporter protein expression, although this was qualitatively slightly lower than the levels observed in the no-chase condition (Figure 5.16). These results suggest that a significant amount of mRNA internalisation occurs during the initial 20 minutes of exposure to mRNA formulated with Lipofectamine 2000, facilitating endosomal escape and subsequent translation within the 2 hours post-pulse.


A.
mCherry
Hoechst 33342
Brightfield
Overlay
2 h
4 h
8 h

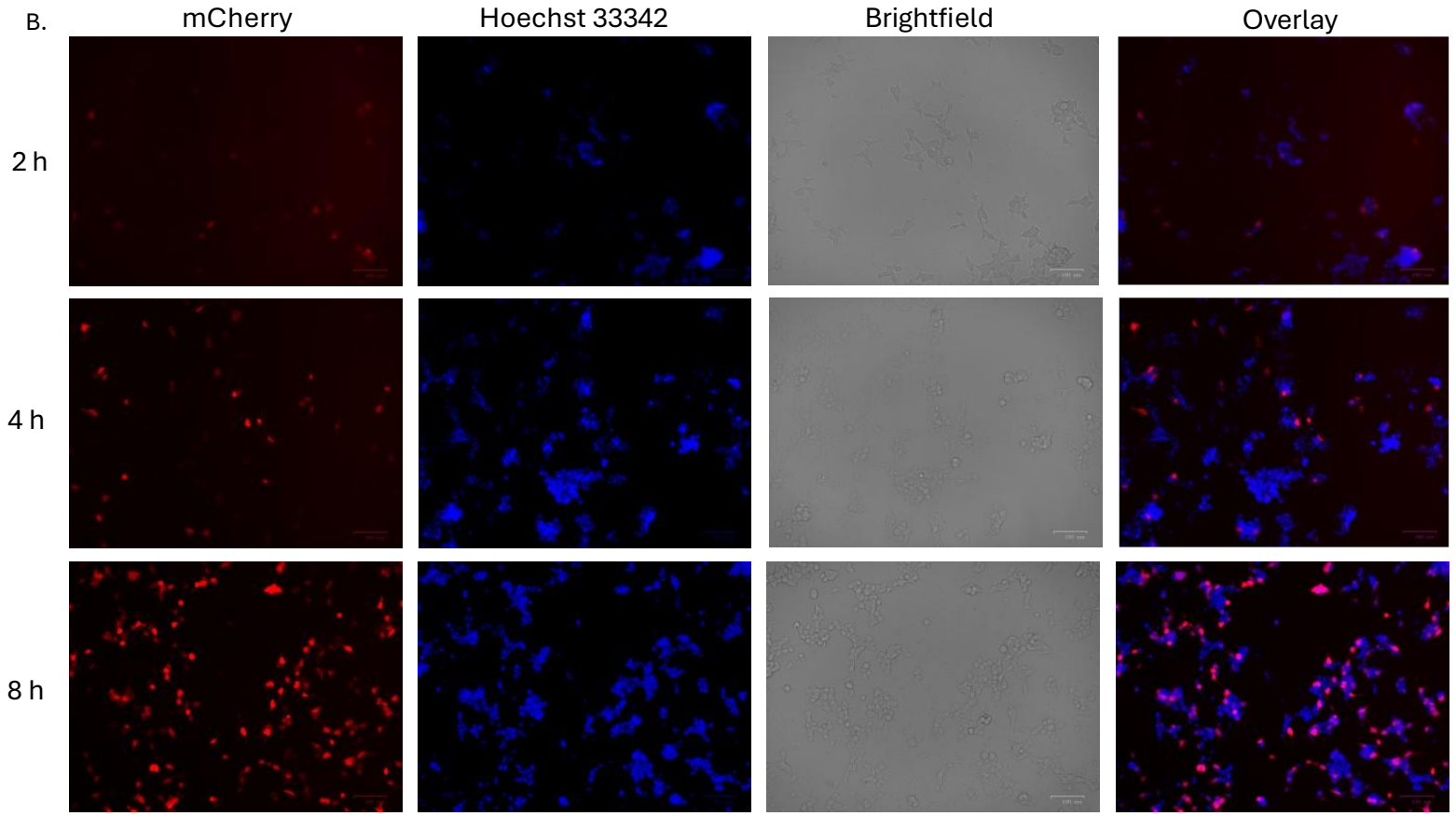


Figure 5.16. Representative live-cell images from a pulse-chase experiment in APEX2-ERM HEK293T cells, where cells were transfected with mCherry mRNA using Lipofectamine 2000. Panel A illustrates the time course of a 20-minute pulse of mCherry mRNA formulated with Lipofectamine 2000, while Panel B depicts the time course of a no-pulse chase condition. Images are from one independent biological experiment, with scale bars representing 100 µm and captured using a 4X magnification lens.

5.9. Discussion

In this chapter, APEX2-RT-qPCR was employed to assess the localisation of cytosolic and secreted Nluc mRNA constructs across three distinct APEX2 cell lines. The APEX2-ERM HEK293T line yielded the most reliable data, consistently showing enrichment of NlucSec mRNA at the expected ER localisation following APEX2 labelling. By contrast, the APEX2-NES HEK293T and APEX2-ERM A549 lines produced less consistent results in isolating mRNA expected to co-localise with the same subcellular region as APEX2. Building on these findings, further APEX2-RT-qPCR analyses and parallel reporter assays in APEX2-ERM HEK293T and APEX2-ERM A549 cells examined two Nluc mRNAs differing only in their SPCSs. In both cell lines, NlucSec mRNA with the PMEL SPCS showed greater enrichment in APEX2-RT-qPCR compared to NlucSec mRNA with the IL6 SPCS, though these differences were not consistently statistically significant. While secreted luciferase assays indicated a slight decrease in luciferase activity following transfection with Nluc PMEL SPCS mRNA relative to Nluc IL6 SPCS mRNA, this trend was not significant. Interestingly, when a different coding sequence (CDS) was tested, PMEL mRNA with its native SPCS showed a similar trend of enhanced ER enrichment relative to PMEL IL6 SPCS mRNA. Lastly, complementary approaches to APEX2-RT-qPCR were explored, offering potential tools for future studies to gain a more comprehensive understanding of IVT mRNA delivery and localisation.

An inherent limitation of the use of RT-qPCR downstream of the streptavidin pulldown to quantify APEX2-labelled RNA is the inability to distinguish between intact and partially degraded mRNA, which could impact the interpretation of localisation data. Analysis by RNA-seq would provide insight into both the level and the integrity of trafficked mRNA at the intended subcellular location..

APEX2-RT-qPCR exhibited inconsistencies across some biological replicates. A possible explanation for this is potentially due to the instability of reagents such as biotin-aniline and H_2O_2, as results tended to improve with the use of freshly prepared reagents. While replacing H_2O_2 often improved consistency, further investigation is needed to determine the shelf life of biotin-aniline at -80°C, the

stability of APEX2-biotinylated RNA at -20°C, and the impact of H_2O_2 degradation from air exposure on labelling efficiency. Preliminary RNA dot blot assays were conducted to assess biotinylated RNA stability at -20°C, but no chemiluminescence signal was detected possibly due to biotin-aniline degradation (data not shown). While RNA dot blots with streptavidin-HRP can provide an initial indication of stability, mass spectrometry could offer a more quantitative method for detecting biotinylated RNA. However, current mass spectrometry techniques for biotinylated RNA are not yet sensitive enough, owing to the low efficiency of APEX2 labelling and high background noise (Zhou et al., 2019). Beyond reagent stability, observations from SIM imaging in Chapter 3 highlighted additional factors that may affect consistency. In all three APEX2 cell lines, APEX2 enzymatic activity appeared uneven across cell populations, despite widespread epitope tag expression in the APEX2-ERM cell lines and consistent resistance due to the antibiotic selection marker introduced during lentiviral transduction. This variability suggests the need for further work to improve labelling uniformity across mixed cell populations.

Puromycin treatment prior to APEX2 labelling showed potential to increase APEX2 labelling efficiency and this warrants further investigation. Future studies should aim to optimise the incubation time of puromycin, as this may vary depending on the experimental context. For instance, when studying SPCS substitutions, prolonged treatment of cells with puromycin prior to APEX2 labelling could cause ER-targeted mRNA to diffuse away from the ER due to the co-translational trafficking nature of the SRP pathway. In addition, the amount of transfected mRNA was initially selected by measuring luciferase activity to ensure it fell within the linear range of luminometer detection (data not shown), though this may not be optimal for APEX2 labelling. Conducting dose-response experiments in each APEX2 cell line could help determine if a specific range of transfected mRNA doses maximises APEX2 labelling efficiency.

Once the optimal labelling conditions, the optimal dose range of IVT mRNA identified, and the effect of puromycin on APEX2 labelling have been determined, APEX2-RT-qPCR should be used to assess NlucSec and PMEL

mRNA localisation at earlier time points. A suitable timeframe for this assessment in APEX2-ERM HEK293T might be one hour post-transfection or earlier, given that differences in secreted luciferase activity (Figure 5.5B) were observed at the one-hour time point.

The use of PMEL mRNA allowed for the study of APEX2-RT-qPCR with a therapeutically relevant mRNA, as well as the testing of whether the length of delivered mRNA would impact APEX2 labelling efficiency. However, since only two CDSs were compared, it would be beneficial to conduct a more systematic investigation into the impact of mRNA length and structure. This could involve delivering multiple IVT mRNAs in parallel to the same APEX2 cells and employing APEX-seq to determine how mRNA length and structure influence the APEX2-RT-qPCR technique.

While APEX2-RT-qPCR was successfully employed alongside protein output measurements, there were several limitations associated with the assessment of protein output. One such limitation of reporter protein assays is the maturation time required for luminescent proteins, meaning that these assays reflect functional protein activity rather than translational output. A further potential limitation of the nanoluciferase assays presented in this chapter is the lack of normalisation, which affected how representative and interpretable the data was. Common normalisation methods for luciferase assays, such as co-transfection with a second reporter luciferase gene were not suitable for this study due to its focus on mRNA delivery. Alternative protein-based normalisation techniques, like Pierce BSA or Bradford assays, were also considered but are prone to variability in protein extraction efficiency. Normalising to total Nluc mRNA or using a spike-in RNA during extraction alongside a housekeeping gene may provide a more sensitive alternative. Additionally, the measurement of PMEL protein output was limited to a single time point and was not quantitative. To address these limitations in future studies, an intracellular flow cytometry technique could be applied to assess PMEL mRNA transfection efficiency. This approach would enable a quantitative comparison of functional protein output between different mRNAs.

Comparison of APEX-RT-qPCR with biochemical membrane fractionation and RT-qPCR gave insight into the advantages and disadvantages of both methods. Whilst the biochemical membrane fractionation approach used in this study effectively enriched ER-associated endogenous mRNA in the membrane-bound fraction, it was generally less successful at enriching endogenous mRNAs expected to localise to the cytosol in the cytosolic fraction. In addition, the biochemical fractionation technique did not significantly enrich the delivered NlucSec mRNA in the membrane-bound fraction. The variability observed in the RT-qPCR data likely arises from cross-contamination between the two fractions. However, interpreting this contamination is complicated by the fact that certain mRNAs encoding cytosolic proteins have been reported to be translated at the ER (Reid & Nicchitta, 2015), and mRNAs that rely on SPCS for co-translational ER targeting will initiate translation in the cytosol (Keenan et al., 2001). Transcripts assessed by RT-qPCR post-fractionation were selected based on APEX-seq data (Fazal et al., 2019; Padrón et al., 2019). A broader post-fractionation analysis, using either RT-qPCR or RNA-seq, that includes mRNA species encoding cytosolic proteins reported to localise to either the ER or cytosol through non-proximity-based labelling methods may provide a better understanding of how effectively this fractionation technique isolates cytosolic mRNA.

The BBT (Between Body and Tail) labelling strategy for mRNA showed promise as an alternative to more commonly used direct fluorophore labelling approaches, with promising implications as a means of visualisation of delivered mRNA or future studies investigating mRNA transfection. Preliminary experiments showed that BBT-labelled mRNA could be translated efficiently, and the attached fluorophores could be visualised in live cell imaging. However, heterogeneity in N_3-dATP incorporation by yeast poly(A) polymerase and the efficiency of fluorophore attachment by SPAAC click chemistry were not assessed in this study. Both factors could contribute to false-positive results, where mRNA observed to be functional in the luciferase assay represented a different subset from those detected in live-cell imaging. Future investigations

of BBT-labelled mRNA could explore the development of a direct RNA sequencing technique to quantify the proportion of mRNA containing N_3-dATP in the poly(A) tail and assess the labelling efficiency. BBT labelling, in conjunction with live-cell stains such as ER-tracker, could serve as an alternative approach to assess subcellular localisation following the delivery of labelled mRNA with different SPCS, facilitating corroboration with APEX-RT-qPCR data. Additionally, by leveraging live-cell imaging, the BBT approach may help identify optimal time points for subsequent APEX2 experiments.

In addition to the BBT labelling strategy, plasmids encoding fluorescent reporter proteins downstream of T7 promoters were developed as complementary molecular tools for investigating mRNA delivery. This toolkit could prove to be a valuable resource for future studies conducted alongside APEX2-RT-qPCR.

In summary, APEX2-RT-qPCR was developed as a method that can be used to investigate determinants of effective mRNA delivery to the ER at both the mRNA and delivery vehicle level and was successfully integrated with reporter protein assays. Furthermore, a toolkit of complementary approaches has been established for use in future studies.

Chapter 6 - Discussion

6.1. Summary

The primary aim of this thesis was to leverage the APEX2-RT-qPCR technique to investigate the subcellular localisation of delivered IVT mRNA with varied features to the two major sites of mRNA translation, the cytosol and the endoplasmic reticulum (ER). Additionally, this work aimed to advance APEX2-RT-qPCR as a tool with potential applications in assessing the efficacy of mRNA delivery vehicles.

To achieve this, stable cell lines expressing APEX2 targeted to specific subcellular compartments, specifically the ER membrane (APEX2-ERM) and the cytosol (APEX2-NES), were generated. Stable APEX2 expression was confirmed through western blotting and RT-qPCR. Streptavidin-HRP dot blot assays verified the enzymatic activity of APEX2 on total RNA across HEK293T APEX2-ERM, HEK293T APEX2-NES, and A549 APEX2-ERM cell lines. Structured illumination microscopy (SIM) confirmed the localisation of APEX2's epitope tag with the ER membrane in HEK293T APEX2-ERM cells. However, overlap between the APEX2 epitope tag and cytoplasmic markers in the APEX2-NES HEK293T cell line, as well as between the APEX2 epitope tag and ER markers in the APEX2-ERM A549 cell line, was lower. SIM successfully demonstrated the presence of APEX2 enzymatic activity with its intended subcellular regions using streptavidin-conjugated Alexa Fluor dyes, confirming that APEX2's activity was generally restricted to the intended subcellular regions in the APEX2-ERM HEK293T and APEX2-NES HEK293T cell lines. However, in APEX2-ERM A549 cells SIM image analysis suggested that APEX2 labelling activity might extend beyond the ER membrane. Further functional assays are needed to clarify the extent and significance of this phenomenon.

An improved method was developed to refine the streptavidin pulldown technique essential for isolating APEX2-labelled RNA. This approach involved using a synthetically biotinylated spike-in RNA, both to enhance pulldown

efficiency and to provide a reliable normalisation control for RT-qPCR data analysis. APEX2-RT-qPCR was used to evaluate the selectivity of APEX2 labelling for both delivered NlucSec mRNA and cytosolic Nluc mRNA in APEX2-ERM HEK293T, APEX2-NES HEK293T and APEX2-ERM A549 cells. APEX2-RT-qPCR results from APEX2-ERM HEK293T cells demonstrated consistent labelling and successful pulldown of NlucSec mRNA, which is expected to be translated at the ER, establishing this cell line as the primary focus of the study. Further investigation involved comparing two SPCS in the delivered reporter mRNA. Stronger APEX2 labelling was observed for PMEL SPCS-driven mRNAs at the ER compared to those driven by IL6 SPCS. However, this did not consistently correlate with an increased translational output of PMEL SPCS-driven mRNAs.

Finally, several alternative approaches to increase APEX2 labelling and orthogonal methods for investigating mRNA localisation were tested. Efforts to enhance APEX2 labelling efficiency included synthesising IVT mRNA with varying proportions of S^6GTP, a synthetic modified nucleotide shown to enhance APEX labelling of nascently transcribed RNA (Li et al., 2022). However, despite using fresh reagents, poor reproducibility across experiments and inconsistent APEX2-mediated RNA labelling limited result interpretation. Additionally, incubation of APEX2-expressing cells with puromycin was tested as a method to enhance the accessibility of APEX2-mediated biotinyl radicals to proximal RNA, with preliminary data showing promise. A biochemical membrane fractionation method was also tested following the delivery of NlucSec mRNA in APEX2-ERM HEK293T cells. While this fractionation technique successfully enriched ER-localised endogenous mRNAs in the membrane fraction, it yielded less consistent data for the delivered IVT mRNA anticipated to localise to the ER.

The BBT (Between Body and Tail) labelling strategy for mRNA was investigated as a possible alternative to traditional CDS labelling and could offer new avenues for studying mRNA transfection dynamics. New plasmids were generated encoding eGFP and mCherry, flanked by therapeutically relevant UTRs. A pulse-chase experiment demonstrated that IVT mRNA encapsulated in Lipofectamine 2000 nanoparticles is internalised by cells within 20 minutes.

However, fluorescent protein expression remained at background levels until two hours after the pulse, highlighting a time lag between mRNA internalisation and protein synthesis. This underscores the limitations of reporter assays in capturing the early phases of mRNA transfection and emphasises the importance of developing tools such as APEX2, which have the potential to provide insights into the interactions between delivered RNA and cellular machinery prior to translation.

6.2. Methodological enhancements of streptavidin affinity purification for isolation of APEX2-labelled RNA and future optimisation work

The optimised streptavidin pulldown method developed in this study enhances the APEX methodology beyond existing published techniques, enabling for the first time the investigation of delivered mRNA *via* APEX2-RT-qPCR. This approach incorporates a biotinylated spike-in RNA to normalise post-streptavidin pulldown RT-qPCR data, representing a novel strategy compared to the existing method of normalising APEX2 data against that from APEX2-NES cells (Fazal et al., 2019; Padrón et al., 2019). By combining normalisation to a spike-in with the use of a no-biotin-aniline negative control, this methodology provides a robust alternative for addressing variability in pulldown efficiency and is expected to be broadly applicable to future APEX2 research.

The final optimised streptavidin pulldown protocol required just 500 ng of input APEX2-labelled RNA, marking a substantial reduction from the 25-100 µg typically used in the original protocols (Fazal et al., 2019; Padrón et al., 2019; Zhou et al., 2019). In APEX-seq, RNA is fragmented prior to pulldown, which might explain the need for larger input quantities. However, Zhou et al. (2019) reported that the same streptavidin pulldown technique worked effectively for analyses by RT-qPCR. Despite the variability observed in APEX2-RT-qPCR data, the reduction in input RNA is a noteworthy finding and should be considered by those aiming to optimise APEX2-RT-qPCR experiments. It is possible that the optimal input RNA level is dependent on APEX labelling efficiency, which may vary between different cell lines. Factors influencing this variability could

include APEX2 expression levels and the subcellular localisation of APEX, particularly whether it is localised to membrane-enclosed regions.

A challenge encountered in APEX2-RT-qPCR was the variability between experiments, which was not entirely alleviated by using a biotinylated spike-in RNA. This variability was likely attributed to several factors, including the low labelling efficiency of APEX2, some degree of non-specific interactions between streptavidin and non-biotinylated RNA, and difficulties in verifying the stability and purity of key APEX2 labelling reagents. Efforts were made to confirm APEX2-mediated RNA labelling by dot blot assay prior to conducting streptavidin pulldown experiments. One additional reason for the inconsistencies observed in some APEX2-RT-qPCR experiments could be linked to the heterogeneous enzymatic activity within each mixed cell population expressing APEX2, as observed through SIM imaging (Figure 3.8A, Figure 3.10A, Figure 3.12A and Figure 3.13). Interestingly, widespread expression of the V5 epitope tag was observed throughout the mixed cell population of APEX2-ERM HEK293T cells imaged by SIM. This suggests that the observed enzymatic heterogeneity is more likely due to variations in the uptake of APEX2 labelling reagents, rather than differences in APEX2 expression between cells. Further investigation is warranted to explore this hypothesis, including additional imaging to determine whether this heterogeneity is consistent across different biological replicates.

The RT step in APEX2-RT-qPCR may introduce bias towards non-labelled, non-oxidised RNA, which warrants further investigation. This bias could arise from the H_2O_2 induced oxidation of nucleobases (Liu et al., 2012) or from APEX2-mediated biotin labelling of nucleobases, potentially leading to misincorporation of nucleotides or disengagement of certain reverse transcriptases at adduct sites. (Potapov et al., 2018; Sexton et al., 2017). Although different reverse transcriptases were tested in this study without any observed effect (data not shown), the effective recovery of S^6G IVT mRNA exposed to Biotin-XX-MTSEA post-streptavidin pulldown RT-qPCR suggests that the GoScript reverse transcriptase can process guanine bases with biotin adducts. In proximity-dependent labelling studies, detecting false positives,

where the biological molecule is isolated despite not being biotinylated, is crucial (as reviewed by Kang & Rhee (2022)). When APEX2 is applied for spatial proteomics, this issue is mitigated by the detection of biotin-phenol adducts on small peptide fragments by mass spectrometry (Lee et al., 2016, 2017). However, current APEX RNA analysis relies on reverse transcription and, in the case of APEX-seq, subsequent sequencing of DNA. In contrast, direct RNA sequencing offers a promising alternative, with a growing range of algorithms capable of base-calling chemically modified nucleobases (as reviewed by Zhao et al. (2022)). However, no models have yet been developed to accurately distinguish APEX2-mediated biotinylated guanine from oxidised bases and other modifications. The potential for biotin itself to become oxidised (Lee et al., 2017) further complicates the development of such models.

6.3. APEX2 to investigate therapeutic mRNA delivery

As a proof of concept, this thesis focuses on using APEX2 to investigate the localisation of delivered IVT mRNA containing SPCS at the ER. SPCSs are a less explored feature of mRNA within the therapeutic field, and work is still required for more thorough application in therapeutic mRNA therapeutics. Recent research demonstrated that substituting the native SPCS of the SARS-CoV-2 receptor-binding domain (RBD) mRNA vaccine with IL-6 or tPA SPCS led to increased antigen expression and enhanced immune response *in vivo*, suggesting that SPCS engineering can improve mRNA vaccine efficacy (Zhang et al., 2024). Investigating the localisation of delivered mRNA to the ER via APEX2-RT-qPCR in an APEX2-ERM fusion cell line could offer valuable insights into SPCS engineering, particularly as the relationship between ER localisation and translational output remains to be fully understood.

To further investigate potential differences in mRNA trafficking, future experiments should incorporate APEX2 labelling at additional, earlier time points, such as 30 minutes or 1-hour post-transfection. This would allow assessment into whether any trafficking differences can be detected at these initial stages, particularly given that translational dynamics in the APEX2-ERM HEK293T systems varied between NlucSec mRNAs with different SPCSs.

Additionally, expanding the analysis to include a broader range of coding sequences and SPCSs would provide a more comprehensive understanding of how these features impact mRNA localisation and translation.

To facilitate the comparison of delivery and translation of mRNAs with varying CDSs, NanoBiT technology offers a promising complementary approach to APEX2-RT-qPCR for evaluating therapeutic mRNA. By adding a HiBiT tag to the 3' end of any coding region in the mRNA, this method enables rapid assessment of translational output in cells stably expressing the LgBiT component, as demonstrated by Ascanelli et al. (2024). When combined with APEX2-RT-qPCR, this approach could facilitate consistent measurement of both translational output and localisation across different mRNA sequences.

The investigation of transfection efficiency in therapeutic mRNA delivery in cells has typically relied on reporter mRNA assays, flow cytometry or ELISAs to detect therapeutic protein expression. However, these methods are indirect and do not provide insights into key factors that might influence total functional protein output. There is a growing need for new methodologies that can address these limitations and, in doing so, offer a more comprehensive understanding of mRNA trafficking and localisation post-endosomal escape.

APEX2-mediated RNA labelling could complement reporter assays by focusing on different stages of the mRNA delivery process, such as the period after endosomal escape but before engagement with the translational machinery. However, in this study, the APEX2-NES cell line appeared to label delivered mRNA with reduced specificity, consistent with observation by Fazal et al. (2019)(Figure 5.2). They attributed this to both the lower cytosolic concentration of RNA and the more reducing cytosolic environment compared to other subcellular compartments where APEX2 labelling has been more successful. This issue could similarly affect the detection of delivered mRNA that has escaped endosomes, potentially limiting APEX2's utility for addressing questions related to endosomal escape. Given the challenges identified with the APEX2-NES system, it is likely necessary to develop more targeted APEX2 systems for probing endosomal escape. One strategy may involve creating and

validating APEX2-fusion systems specifically designed to localise to the cytosolic-facing side of distinct endosomal compartments. This approach would enable proximity-based labelling of delivered mRNA that has escaped from these compartments. Alternatively, if proteins that interact with delivered mRNA following endosomal escape can be identified in future studies, these proteins could also serve as bait-proteins to generate novel APEX2-fusion systems.

Another important consideration is whether APEX2-mediated biotinyl radicals can penetrate nanoparticle delivery systems and label mRNA that has not yet disassembled from the delivery system. This could be tested *in vitro*, provided nanoparticle encapsulation efficiency remains at 100% in the presence of the APEX2 protein, $OH^{\cdot}$ radicals and biotinyl radicals.

If all these potential challenges are overcome, the next phase would involve delivering mRNA to APEX2 expressing cells using two types of nanoparticles: one known for efficient internalisation but poor endosomal escape, and another known for its high endosomal escape efficiency. This approach would enable a direct comparison between the nanoparticles' efficiency in mRNA transfection, from total cellular uptake to translational output, and localisation of delivered mRNA. By contrasting nanoparticles with varying endosomal escape efficiencies, it would provide insights into the role of endosomal release in successful mRNA delivery and cytosolic localisation.

6.4. Alternative approaches to proximity-dependent labelling to quantify subcellular localisation of delivered mRNA

The use of APEX2-RT-qPCR, as a proximity-dependent labelling tool, presents both significant opportunities and challenges. A key advantage of APEX2 labelling is its ability to capture a precise snapshot of mRNA localisation within a one-minute window, making it particularly useful for studying the dynamics of mRNA trafficking. However, APEX2 can only target one subcellular location at a time and requires careful validation of APEX2's localisation in each new cell line.

Investigating the trafficking of delivered mRNA to the ER could be advanced by subcellular fractionation, particularly when combined with polysomal fractionation to isolate actively translated mRNA (Reid & Nicchitta, 2012). This approach allows for the direct assessment of translation in membrane-bound versus cytosolic compartments, offering insights that complement APEX2 labelling studies. Although membrane fractionation in this study effectively enriched mRNAs predicted to be ER-associated, it was less sensitive than APEX2-RT-qPCR for detecting delivered mRNA at the ER. While APEX2 labelling combined with translation inhibitors, such as cycloheximide, has been employed to investigate translation-dependent localisation mechanisms, the integration of APEX2 labelling with polysomal fractionation remains largely unexplored. In contrast, Jan et al. (2014) previously examined the combination of a mutant Biotin Ligase (BirA*) fusion system targeted to the ER with polysomal fractionation to study mRNA translation at the ER. Directly implementing the methodology from Jan et al. (2014) or adapting it for use with an APEX2 system could offer a valuable alternative strategy for investigating ER-localised translation, thereby enhancing our understanding of mRNA delivery and localisation dynamics.

Cell lines expressing mutant base-editing enzymes, such as adenosine deaminase acting on RNA (ADAR), could provide an alternative approach to quantify the endosomal escape of delivered mRNA by RNA sequencing as a proxy for localisation. In the Targeted RNA Editing by ADAR (TRIBE) approach, ADAR's catalytic domain can be directed to specific RNAs by fusing the catalytic domain to a bait RNA-binding protein (McMahon et al., 2016). MS2-TRIBE, where ADAR's catalytic domain is fused to the MS2 coat protein (MCP) (Biswas et al., 2020), could be applied to mRNA delivery studies to quantify delivered mRNAs engineered with MS2 aptamers in its 3' UTR However, the application of TRIBE methods is limited by the costs associated with RNA sequencing. In addition, ADAR-based methods are typically assessed after a 24-hour incubation to allow detectable editing (Biswas et al., 2020; McMahon et al., 2016; Medina-Munoz et al., 2024), whereas APEX2 can yield results in only 1 minute (Fazal et al., 2019;

Padrón et al., 2019). RNA base-editing enzymes with more rapid editing kinetics would enhance the utility of this approach in mRNA drug delivery studies.

Single-molecule FISH (smFISH) is a powerful technique for visualising individual mRNA molecules by using multiple probes to enhance signal intensity (Haimovich & Gerst, 2018). It has the potential to assess endosomal escape but typically requires immunofluorescence (IF) to mark endosomal compartments and provide a reference point to understand whether delivered mRNA is in the cytosol. Paramasivam et al. (2022b) optimised and applied a dual smFISH:IF approach to investigate endocytic trafficking of delivered mRNA, though they noted difficulty detecting mRNA within intact endosomes, likely due to poor fixation of encapsulated mRNA in nanoparticles. While their study focused on endocytic trafficking rather than quantifying endosomal escape efficiency, this approach could potentially be adapted for endosomal escape studies.

Between Body and Tail (BBT) labelling involves the incorporation of N_3-dATPs in the poly(A) tail, proximal to the 3' UTR (Anhäuser et al., 2019). Compared to CDS-labelled mRNA, BBT labelling preserves translational efficiency, making it potentially valuable for imaging mRNA in cells. However, to extend BBT labelling beyond studies of mRNA internalisation, several critical aspects must be addressed. Firstly, it is essential to determine the extent to which the translational and degradation kinetics of BBT-labelled mRNA reflect those of non-labelled IVT mRNA, to understand the extent to which findings can be extrapolated to non-labelled mRNA. Additionally, it remains unclear to what extent the signal observed in live-cell imaging reflects intact, functional mRNA versus degraded fragments that retain fluorophores. A further challenge is determining whether the fluorophore dye itself influences mRNA subcellular localisation, as cyanine-based dyes, for instance, have been reported to accumulate near mitochondria due to charge interactions (Negwer et al., 2017; Rhee & Bao, 2010). Conducting smFISH experiments on BBT labelled mRNA in parallel to non-labelled mRNA could help clarify whether the observed localisation is authentic. Finally, it is important to investigate whether BBT

labelling currently achieves single-molecule detection and, if not, to optimise its sensitivity.

In conclusion, while no single method for quantifying mRNA delivery and localisation is without limitations, APEX2-RT-qPCR stands out as a valuable tool that can be employed orthogonally alongside other techniques. Its advantages include the ability to detect unlabelled mRNA and its rapid processing time, enabling a swift assessment of delivered mRNA dynamics. In this study, APEX2-RT-qPCR was utilised to simultaneously compare the translational output and ER localisation of delivered mRNAs with different SPCS in the context of two CDSs. The APEX2-RT-qPCR results indicated that higher localisation of mRNAs at the ER does not always correlate with increased translational output and further investigation is required to fully elucidate the underlying mechanisms. This underscores the importance of a multi-faceted approach to understanding mRNA sequence engineering, paving the way for more effective strategies in therapeutic applications.

Chapter 7 - References

Aditham, A., Shi, H., Guo, J., Zeng, H., Zhou, Y., Wade, S. D., Huang, J., Liu, J., & Wang, X. (2022). Chemically Modified mocRNAs for Highly Efficient Protein Expression in Mammalian Cells. *ACS Chemical Biology*, *17*(12), 3352–3366. https://doi.org/10.1021/acschembio.1c00569

Akinc, A., Maier, M. A., Manoharan, M., Fitzgerald, K., Jayaraman, M., Barros, S., Ansell, S., Du, X., Hope, M. J., Madden, T. D., Mui, B. L., Semple, S. C., Tam, Y. K., Ciufolini, M., Witzigmann, D., Kulkarni, J. A., Van Der Meel, R., & Cullis, P. R. (2019). The Onpattro story and the clinical translation of nanomedicines containing nucleic acid-based drugs. *Nature Nanotechnology*, *14*(12), 1084–1087. https://doi.org/10.1038/s41565-019-0591-y

Akopian, D., Shen, K., Zhang, X., & Shan, S. (2013). Signal Recognition Particle: An Essential Protein-Targeting Machine. *Annual Review of Biochemistry*, *82*(1), 693–721. https://doi.org/10.1146/annurev-biochem-072711-164732

Alberts, B., Johnson, A., Lewis, J., Morgan, D., Raff, M., Roberts, K., & Walter, P. (2015). *Molecular Biology of the Cell* (Sixth edition, Onkine-Ausgabe). Garland Science.

Anderson, B. R., Muramatsu, H., Nallagatla, S. R., Bevilacqua, P. C., Sansing, L. H., Weissman, D., & Karikó, K. (2010). Incorporation of pseudouridine into mRNA enhances translation by diminishing PKR activation. *Nucleic Acids Research*, *38*(17), 5884–5892. https://doi.org/10.1093/nar/gkq347

Anderson, S., Bankier, A. T., Barrell, B. G., De Bruijn, M. H. L., Coulson, A. R., Drouin, J., Eperon, I. C., Nierlich, D. P., Roe, B. A., Sanger, F., Schreier, P. H., Smith, A. J. H., Staden, R., & Young, I. G. (1981). Sequence and organization of the human mitochondrial genome. *Nature*, *290*(5806), 457–465. https://doi.org/10.1038/290457a0

Andries, O., Mc Cafferty, S., De Smedt, S. C., Weiss, R., Sanders, N. N., & Kitada, T. (2015). N1-methylpseudouridine-incorporated mRNA outperforms pseudouridine-incorporated mRNA by providing enhanced protein expression and reduced immunogenicity in mammalian cell lines and mice. *Journal of Controlled Release*, *217*, 337–344. https://doi.org/10.1016/j.jconrel.2015.08.051

Anhäuser, L., Hüwel, S., Zobel, T., & Rentmeister, A. (2019). Multiple covalent fluorescence labeling of eukaryotic mRNA at the poly(A) tail enhances translation and can be performed in living cells. *Nucleic Acids Research*, *47*(7). https://doi.org/10.1093/nar/gkz084

Ascanelli, C., Lawrence, E., Batho, C. A. P., & Wilson, C. H. (2024). *A flexible, high-throughput system for studying mRNA translation kinetics* in vitro *and* in cellulo *with HiBit technology*. https://doi.org/10.1101/2024.06.27.600987

Aviner, R. (2020). The science of puromycin: From studies of ribosome function to applications in biotechnology. *Computational and Structural Biotechnology Journal*, *18*, 1074–1083. https://doi.org/10.1016/j.csbj.2020.04.014

Azizi, H., Renner, T. M., Agbayani, G., Simard, B., Dudani, R., Harrison, B. A., Iqbal, U., Jia, Y., McCluskie, M. J., & Akache, B. (2024). Self-amplifying RNAs generated with the modified nucleotides 5-methylcytidine and 5-methyluridine mediate strong expression and immunogenicity *in vivo*. *NAR Molecular Medicine*, *1*(2). https://doi.org/10.1093/narmme/ugae004

Baer, B. W., & Kornberg, R. D. (1983). The protein responsible for the repeating structure of cytoplasmic poly(A)-ribonucleoprotein. *The Journal of Cell Biology*, *96*(3), 717–721. https://doi.org/10.1083/jcb.96.3.717

Bahiri-Elitzur, S., & Tuller, T. (2021). Codon-based indices for modeling gene expression and transcript evolution. *Computational and Structural Biotechnology Journal*, *19*, 2646–2663. https://doi.org/10.1016/j.csbj.2021.04.042

Baiersdörfer, M., Boros, G., Muramatsu, H., Mahiny, A., Vlatkovic, I., Sahin, U., & Karikó, K. (2019). A Facile Method for the Removal of dsRNA Contaminant from In Vitro-Transcribed mRNA. *Molecular Therapy - Nucleic Acids*, *15*, 26–35. https://doi.org/10.1016/j.omtn.2019.02.018

Barlowe, C. K., & Miller, E. A. (2013). Secretory Protein Biogenesis and Traffic in the Early Secretory Pathway. *Genetics*, *193*(2), 383–410. https://doi.org/10.1534/genetics.112.142810

Bartel, D. P. (2004). MicroRNAs. *Cell*, *116*(2), 281–297. https://doi.org/10.1016/S0092-8674(04)00045-5

Beach, M. A., Teo, S. L. Y., Chen, M. Z., Smith, S. A., Pouton, C. W., Johnston, A. P. R., & Such, G. K. (2022). Quantifying the Endosomal Escape of pH-Responsive Nanoparticles Using the Split Luciferase Endosomal Escape Quantification Assay. *ACS Applied Materials & Interfaces*, *14*(3), 3653–3661. https://doi.org/10.1021/acsami.1c18359

Beissert, T., Perkovic, M., Vogel, A., Erbar, S., Walzer, K. C., Hempel, T., Brill, S., Haefner, E., Becker, R., Türeci, Ö., & Sahin, U. (2020). A Trans-amplifying RNA Vaccine Strategy for Induction of Potent Protective Immunity. *Molecular Therapy*, *28*(1), 119–128. https://doi.org/10.1016/j.ymthe.2019.09.009

Ben-Akiva, E., Karlsson, J., Hemmati, S., Yu, H., Tzeng, S. Y., Pardoll, D. M., & Green, J. J. (2023). Biodegradable lipophilic polymeric mRNA nanoparticles for ligand-free targeting of splenic dendritic cells for cancer vaccination. *Proceedings of the National Academy of Sciences*, *120*(26), e2301606120. https://doi.org/10.1073/pnas.2301606120

Bentley, D. L. (2014). Coupling mRNA processing with transcription in time and space. *Nature Reviews Genetics*, *15*(3), 163–175. https://doi.org/10.1038/nrg3662

Bernard, M.-C., Bazin, E., Petiot, N., Lemdani, K., Commandeur, S., Verdelet, C., Margot, S., Perkov, V., Ripoll, M., Garinot, M., Ruiz, S., Boudet, F., Rokbi, B., & Haensler, J. (2023). The impact of nucleoside base modification in mRNA vaccine is influenced by the chemistry of its lipid nanoparticle delivery system. *Molecular Therapy - Nucleic Acids*, *32*, 794–806. https://doi.org/10.1016/j.omtn.2023.05.004

Bernstein, P., Peltz, S. W., & Ross, J. (1989). The Poly(A)-Poly(A)-Binding Protein Complex Is a Major Determinant of mRNA Stability In Vitro. *Molecular and Cellular Biology*, *9*(2), 659–670. https://doi.org/10.1128/mcb.9.2.659-670.1989

Bettinger, T. (2001). Peptide-mediated RNA delivery: A novel approach for enhanced transfection of primary and post-mitotic cells. *Nucleic Acids Research*, *29*(18), 3882–3891. https://doi.org/10.1093/nar/29.18.3882

Bevan, M. J. (2006). Cross-priming. *Nature Immunology*, *7*(4), 363–365. https://doi.org/10.1038/ni0406-363

Bicknell, A. A., Reid, D. W., Licata, M. C., Jones, A. K., Cheng, Y. M., Li, M., Hsiao, C. J., Pepin, C. S., Metkar, M., Levdansky, Y., Fritz, B. R., Andrianova, E. A., Jain, R., Valkov, E., Köhrer, C., & Moore, M. J. (2024). Attenuating ribosome load improves protein output from mRNA by limiting translation-dependent mRNA decay. *Cell Reports*, *43*(4), 114098. https://doi.org/10.1016/j.celrep.2024.114098

Biswas, J., Rahman, R., Gupta, V., Rosbash, M., & Singer, R. H. (2020). MS2-TRIBE Evaluates Both Protein-RNA Interactions and Nuclear Organization of Transcription by RNA Editing. *iScience*, *23*(7), 101318. https://doi.org/10.1016/j.isci.2020.101318

Blanchet, S., & Ranjan, N. (2022). Translation Phases in Eukaryotes. In K.-D. Entian (Ed.), *Ribosome Biogenesis* (Vol. 2533, pp. 217–228). Springer US. https://doi.org/10.1007/978-1-0716-2501-9_13

Boman, J., Marušič, T., Seravalli, T. V., Skok, J., Pettersson, F., Nemec, K. Š., Widmark, H., & Sekirnik, R. (2024). Quality by design approach to improve quality and decrease cost of in vitro transcription of mRNA using design of experiments. *Biotechnology and Bioengineering*. https://doi.org/10.1002/bit.28806

Bonin, M., Oberstrass, J., Lukacs, N., Ewert, K., Oesterschulze, E., Kassing, R., & Nellen, W. (2000). Determination of preferential binding sites for anti-dsRNA antibodies on double-stranded RNA by scanning force microscopy. *RNA (New York, N.Y.)*, *6*(4), 563–570. https://doi.org/10.1017/s1355838200992318

Boo, S. H., & Kim, Y. K. (2020). The emerging role of RNA modifications in the regulation of mRNA stability. *Experimental & Molecular Medicine*, *52*(3), 400–408. https://doi.org/10.1038/s12276-020-0407-z

Boye, E., & Grallert, B. (2020). eIF2α phosphorylation and the regulation of translation. *Current Genetics*, *66*(2), 293–297. https://doi.org/10.1007/s00294-019-01026-1

Branon, T. C., Bosch, J. A., Sanchez, A. D., Udeshi, N. D., Svinkina, T., Carr, S. A., Feldman, J. L., Perrimon, N., & Ting, A. Y. (2018). Efficient proximity labeling in living cells and organisms with TurboID. *Nature Biotechnology*, *36*(9), 880–887. https://doi.org/10.1038/nbt.4201

Brito Querido, J., Díaz-López, I., & Ramakrishnan, V. (2024). The molecular basis of translation initiation and its regulation in eukaryotes. *Nature Reviews Molecular Cell Biology*, *25*(3), 168–186. https://doi.org/10.1038/s41580-023-00624-9

Bucci, C., Parton, R. G., Mather, I. H., Stunnenberg, H., Simons, K., Hoflack, B., & Zerial, M. (1992). The small GTPase rab5 functions as a regulatory factor in the early endocytic pathway. *Cell*, *70*(5), 715–728. https://doi.org/10.1016/0092-8674(92)90306-W

Buccitelli, C., & Selbach, M. (2020). mRNAs, proteins and the emerging principles of gene expression control. *Nature Reviews Genetics*, *21*(10), 630–644. https://doi.org/10.1038/s41576-020-0258-4

Buschmann, M. D., Carrasco, M. J., Alishetty, S., Paige, M., Alameh, M. G., & Weissman, D. (2021). Nanomaterial Delivery Systems for mRNA Vaccines. *Vaccines*, *9*(1), 65. https://doi.org/10.3390/vaccines9010065

Castillo-Hair, S., Fedak, S., Wang, B., Linder, J., Havens, K., Certo, M., & Seelig, G. (2024). Optimizing 5'UTRs for mRNA-delivered gene editing using deep learning. *Nature Communications*, *15*(1), 5284. https://doi.org/10.1038/s41467-024-49508-2

Castillo-Hair, S. M., & Seelig, G. (2022). Machine Learning for Designing Next-Generation mRNA Therapeutics. *Accounts of Chemical Research*, *55*(1), 24–34. https://doi.org/10.1021/acs.accounts.1c00621

Chatterjee, S., Kon, E., Sharma, P., & Peer, D. (2024). Endosomal escape: A bottleneck for LNP-mediated therapeutics. *Proceedings of the National Academy of Sciences*, *121*(11), e2307800120. https://doi.org/10.1073/pnas.2307800120

Chen, B.-M., Cheng, T.-L., & Roffler, S. R. (2021). Polyethylene Glycol Immunogenicity: Theoretical, Clinical, and Practical Aspects of Anti-Polyethylene Glycol Antibodies. *ACS Nano*, *15*(9), 14022–14048. https://doi.org/10.1021/acsnano.1c05922

Chen, C., & Sarnow, P. (1995). Initiation of Protein Synthesis by the Eukaryotic Translational Apparatus on Circular RNAs. *Science*, *268*(5209), 415–417. https://doi.org/10.1126/science.7536344

Chen, C.-C., Butz, E. S., Chao, Y.-K., Grishchuk, Y., Becker, L., Heller, S., Slaugenhaupt, S. A., Biel, M., Wahl-Schott, C., & Grimm, C. (2017). Small Molecules for Early Endosome-Specific Patch Clamping. *Cell Chemical Biology*, *24*(7), 907-916.e4. https://doi.org/10.1016/j.chembiol.2017.05.025

Chen, X., Zhong, S., Hou, Y., Cao, R., Wang, W., Li, D., Dai, Q., Kim, D., & Xi, P. (2023). Superresolution structured illumination microscopy reconstruction algorithms: A review. *Light: Science & Applications*, *12*(1), 172. https://doi.org/10.1038/s41377-023-01204-4

Chen, Y. G., Chen, R., Ahmad, S., Verma, R., Kasturi, S. P., Amaya, L., Broughton, J. P., Kim, J., Cadena, C., Pulendran, B., Hur, S., & Chang, H. Y. (2019). N6-Methyladenosine Modification Controls Circular RNA Immunity. *Molecular Cell*, *76*(1), 96-109.e9. https://doi.org/10.1016/j.molcel.2019.07.016

Cheng, Q., Farbiak, L., Vaidya, A., Guerrero, E., Lee, E. E., Rose, E. K., Wang, X., Robinson, J., Lee, S. M., Wei, T., Miller, W. E., Alvarez Benedicto, E., Lian, X., Wang, R. C., & Siegwart, D. J. (2023). In situ production and secretion of proteins endow therapeutic benefit against psoriasiform dermatitis and melanoma. *Proceedings of the National Academy of Sciences of the United States of America*, *120*(52), e2313009120. https://doi.org/10.1073/pnas.2313009120

Cheng, Q., Wei, T., Farbiak, L., Johnson, L. T., Dilliard, S. A., & Siegwart, D. J. (2020). Selective organ targeting (SORT) nanoparticles for tissue-specific mRNA delivery and CRISPR-Cas gene editing. *Nature Nanotechnology*, *15*(4), 313–320. https://doi.org/10.1038/s41565-020-0669-6

Child, J. R., Chen, Q., Reid, D. W., Jagannathan, S., & Nicchitta, C. V. (2021). Recruitment of endoplasmic reticulum-targeted and cytosolic mRNAs into membrane-associated stress granules. *RNA (New York, N.Y.)*, *27*(10), 1241–1256. https://doi.org/10.1261/rna.078858.121

Choi-Rhee, E., Schulman, H., & Cronan, J. E. (2004). Promiscuous protein biotinylation by *Escherichia coli* biotin protein ligase. *Protein Science*, *13*(11), 3043–3050. https://doi.org/10.1110/ps.04911804

Colgan, D. F., & Manley, J. L. (1997). Mechanism and regulation of mRNA polyadenylation. *Genes & Development*, *11*(21), 2755–2766. https://doi.org/10.1101/gad.11.21.2755

Comes, J. D. G., Pijlman, G. P., & Hick, T. A. H. (2023). Rise of the RNA machines – self-amplification in mRNA vaccine design. *Trends in Biotechnology*, *41*(11), 1417–1429. https://doi.org/10.1016/j.tibtech.2023.05.007

Corre, G., Seye, A., Frin, S., Ferrand, M., Winkler, K., Luc, C., Dorange, F., Rocca, C. J., & Galy, A. (2022). Lentiviral standards to determine the sensitivity of assays that quantify lentiviral vector copy numbers and genomic insertion sites in cells. *Gene Therapy*, *29*(9), 536–543. https://doi.org/10.1038/s41434-022-00315-8

Cullis, P. R., & Hope, M. J. (2017). Lipid Nanoparticle Systems for Enabling Gene Therapies. *Molecular Therapy: The Journal of the American Society of Gene Therapy*, *25*(7), 1467–1475. https://doi.org/10.1016/j.ymthe.2017.03.013

Custer, T. C., & Walter, N. G. (2017). In vitro labeling strategies for in cellulo fluorescence microscopy of single ribonucleoprotein machines. *Protein Science: A Publication of the Protein Society*, *26*(7), 1363–1379. https://doi.org/10.1002/pro.3108

Dastgerdi, N. K., Gumus, N., Bayraktutan, H., Jackson, D., Polra, K., McKay, P. F., Atyabi, F., Dinarvand, R., Shattock, R. J., Martinez-Pomares, L., Gurnani, P., & Alexander, C. (2024). Charge neutralized poly(β-amino ester) polyplex nanoparticles for delivery of self-amplifying RNA. *Nanoscale Advances*, *6*(5), 1409–1422. https://doi.org/10.1039/D3NA00794D

Davis, D. R. (1995). Stabilization of RNA stacking by pseudouridine. *Nucleic Acids Research*, *23*(24), 5020–5026. https://doi.org/10.1093/nar/23.24.5020

de Sousa Abreu, R., Penalva, L. O., Marcotte, E. M., & Vogel, C. (2009). Global signatures of protein and mRNA expression levels. *Molecular bioSystems*, *5*(12), 1512–1526. https://doi.org/10.1039/b908315d

Deyhimfar, R., Izady, M., Shoghi, M., Kazazi, M. H., Ghazvini, Z. F., Nazari, H., Fekrirad, Z., & Arefian, E. (2024). The clinical impact of mRNA therapeutics in the treatment of cancers, infections, genetic disorders, and autoimmune diseases. *Heliyon*, *10*(5), e26971. https://doi.org/10.1016/j.heliyon.2024.e26971

Dilliard, S. A., Cheng, Q., & Siegwart, D. J. (2021). On the mechanism of tissue-specific mRNA delivery by selective organ targeting nanoparticles. *Proceedings of the National Academy of Sciences of the United States of America*, *118*(52), e2109256118. https://doi.org/10.1073/pnas.2109256118

Dimitriadis, G. J. (1978). Translation of rabbit globin mRNA introduced by liposomes into mouse lymphocytes. *Nature*, *274*(5674), 923–924. https://doi.org/10.1038/274923a0

Dolgin, E. (2021). The tangled history of mRNA vaccines. *Nature*, *597*(7876), 318–324. https://doi.org/10.1038/d41586-021-02483-w

Dousis, A., Ravichandran, K., Hobert, E. M., Moore, M. J., & Rabideau, A. E. (2023). An engineered T7 RNA polymerase that produces mRNA free of immunostimulatory byproducts. *Nature Biotechnology*, *41*(4), 560–568. https://doi.org/10.1038/s41587-022-01525-6

Dowdy, S. F. (2017). Overcoming cellular barriers for RNA therapeutics. In *Nature Biotechnology*. https://doi.org/10.1038/nbt.3802

Dowdy, S. F., Setten, R. L., Cui, X.-S., & Jadhav, S. G. (2022). Delivery of RNA Therapeutics: The Great Endosomal Escape! *Nucleic Acids Therapeutics*. https://pubmed.ncbi.nlm.nih.gov/35612432/

Downie Ruiz Velasco, A., Parsons, A. L., Heatley, M. C., Martin, A. R. G., Smart, A. D., Shah, N., & Jopling, C. L. (2024). MicroRNA biogenesis is broadly disrupted by inhibition of the splicing factor SF3B1. *Nucleic Acids Research*, *52*(15), 9210–9229. https://doi.org/10.1093/nar/gkae505

Duffy, E. E., Rutenberg-Schoenberg, M., Stark, C. D., Kitchen, R. R., Gerstein, M. B., & Simon, M. D. (2015). Tracking Distinct RNA Populations Using Efficient and Reversible Covalent Chemistry. *Molecular Cell*, *59*(5), 858–866. https://doi.org/10.1016/j.molcel.2015.07.023

Dunn, K. W., Kamocka, M. M., & McDonald, J. H. (2011). A practical guide to evaluating colocalization in biological microscopy. *American Journal of Physiology. Cell Physiology*, *300*(4), C723-742. https://doi.org/10.1152/ajpcell.00462.2010

Eltoukhy, A. A., Chen, D., Alabi, C. A., Langer, R., & Anderson, D. G. (2013). Degradable Terpolymers with Alkyl Side Chains Demonstrate Enhanced Gene Delivery Potency and Nanoparticle Stability. *Advanced Materials*, *25*(10), 1487–1493. https://doi.org/10.1002/adma.201204346

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., Coler, R. N., Diamond, M. S., Crowe, J. E., Reed, S. G., Thackray, L. B., Carnahan, R. H., & Van Hoeven, N. (2020). Intramuscular Delivery of Replicon RNA Encoding ZIKV-117 Human Monoclonal Antibody Protects against Zika Virus Infection. *Molecular Therapy - Methods & Clinical Development*, *18*, 402–414. https://doi.org/10.1016/j.omtm.2020.06.011

Fazal, F. M., Han, S., Parker, K. R., Kaewsapsak, P., Xu, J., Boettiger, A. N., Chang, H. Y., & Ting, A. Y. (2019a). Atlas of Subcellular RNA Localization Revealed by APEX-Seq. *Cell*. https://doi.org/10.1016/j.cell.2019.05.027

Fazal, F. M., Han, S., Parker, K. R., Kaewsapsak, P., Xu, J., Boettiger, A. N., Chang, H. Y., & Ting, A. Y. (2019b). Atlas of Subcellular RNA Localization Revealed by APEX-Seq. *Cell*. https://doi.org/10.1016/j.cell.2019.05.027

Finnegan, M., Linley, E., Denyer, S. P., McDonnell, G., Simons, C., & Maillard, J.-Y. (2010). Mode of action of hydrogen peroxide and other oxidizing agents: Differences between liquid and gas forms. *Journal of Antimicrobial Chemotherapy*, *65*(10), 2108–2115. https://doi.org/10.1093/jac/dkq308

Forterre, A. V., Wang, J.-H., Delcayre, A., Kim, K., Green, C., Pegram, M. D., Jeffrey, S. S., & Matin, A. C. (2020). Extracellular Vesicle-Mediated In Vitro Transcribed mRNA Delivery for Treatment of HER2+ Breast Cancer Xenografts in Mice by Prodrug CB1954 without General Toxicity. *Molecular Cancer Therapeutics*, *19*(3), 858–867. https://doi.org/10.1158/1535-7163.MCT-19-0928

Fuchs, A.-L., Neu, A., & Sprangers, R. (2016). A general method for rapid and cost-efficient large-scale production of 5' capped RNA. *RNA*, *22*(9), 1454–1466. https://doi.org/10.1261/rna.056614.116

Furuichi, Y. (2015). Discovery of m^7G-cap in eukaryotic mRNAs. *Proceedings of the Japan Academy, Series B*, *91*(8), 394–409. https://doi.org/10.2183/pjab.91.394

Gallie, D. R. (1991). The cap and poly(A) tail function synergistically to regulate mRNA translational efficiency. *Genes & Development*, *5*(11), 2108–2116. https://doi.org/10.1101/gad.5.11.2108

Galloway, A., & Cowling, V. H. (2019). mRNA cap regulation in mammalian cell function and fate. *Biochimica et Biophysica Acta (BBA) - Gene Regulatory Mechanisms*, *1862*(3), 270–279. https://doi.org/10.1016/j.bbagrm.2018.09.011

Gholamalipour, Y., Karunanayake Mudiyanselage, A., & Martin, C. T. (2018). 3’ end additions by T7 RNA polymerase are RNA self-templated, distributive and diverse in character-RNA-Seq analyses. *Nucleic Acids Research*, *46*(18), 9253–9263. https://doi.org/10.1093/nar/gky796

Gonatopoulos-Pournatzis, T., & Cowling, V. H. (2014). Cap-binding complex (CBC). *Biochemical Journal*, *457*(2), 231–242. https://doi.org/10.1042/BJ20131214

Goswami, R., Chatzikleanthous, D., Lou, G., Giusti, F., Bonci, A., Taccone, M., Brazzoli, M., Gallorini, S., Ferlenghi, I., Berti, F., O’Hagan, D. T., Pergola, C., Baudner, B. C., & Adamo, R. (2019). Mannosylation of LNP Results in Improved Potency for Self-Amplifying RNA (SAM) Vaccines. *ACS Infectious Diseases*, *5*(9), 1546–1558. https://doi.org/10.1021/acsinfecdis.9b00084

Gote, V., Bolla, P. K., Kommineni, N., Butreddy, A., Nukala, P. K., Palakurthi, S. S., & Khan, W. (2023). A Comprehensive Review of mRNA Vaccines. *International Journal of Molecular Sciences*, *24*(3), 2700. https://doi.org/10.3390/ijms24032700

Green, T. M., Alpaugh, M. L., Barsky, S. H., Rappa, G., & Lorico, A. (2015). Breast Cancer-Derived Extracellular Vesicles: Characterization and Contribution to the Metastatic Phenotype. *BioMed Research International*, *2015*, 1–13. https://doi.org/10.1155/2015/634865

Grier, A. E., Burleigh, S., Sahni, J., Clough, C. A., Cardot, V., Choe, D. C., Krutein, M. C., Rawlings, D. J., Jensen, M. C., Scharenberg, A. M., & Jacoby, K. (2016). pEVL: A Linear Plasmid for Generating mRNA IVT Templates With Extended Encoded Poly(A) Sequences. *Molecular Therapy - Nucleic Acids*, *5*, e306. https://doi.org/10.1038/mtna.2016.21

Haimovich, G., & Gerst, J. (2018). Single-molecule Fluorescence in situ Hybridization (smFISH) for RNA Detection in Adherent Animal Cells. *Bio-Protocol*, *8*(21), 1–17. https://doi.org/10.21769/bioprotoc.3070

Hammell, C. M., Gross, S., Zenklusen, D., Heath, C. V., Stutz, F., Moore, C., & Cole, C. N. (2002). Coupling of Termination, 3' Processing, and mRNA Export. *Molecular and Cellular Biology*, *22*(18), 6441–6457. https://doi.org/10.1128/MCB.22.18.6441-6457.2002

Hansen, T. B., Jensen, T. I., Clausen, B. H., Bramsen, J. B., Finsen, B., Damgaard, C. K., & Kjems, J. (2013). Natural RNA circles function as efficient microRNA sponges. *Nature*, *495*(7441), 384–388. https://doi.org/10.1038/nature11993

Hanson, G., & Coller, J. (2018). Codon optimality, bias and usage in translation and mRNA decay. *Nature Reviews Molecular Cell Biology*, *19*(1), 20–30. https://doi.org/10.1038/nrm.2017.91

Hassett, K. J., Benenato, K. E., Jacquinet, E., Lee, A., Woods, A., Yuzhakov, O., Himansu, S., Deterling, J., Geilich, B. M., Ketova, T., Mihai, C., Lynn, A., McFadyen, I., Moore, M. J., Senn, J. J., Stanton, M. G., Almarsson, Ö., Ciaramella, G., & Brito, L. A. (2019). Optimization of Lipid Nanoparticles for Intramuscular Administration of mRNA Vaccines. *Molecular Therapy. Nucleic Acids*, *15*, 1–11. https://doi.org/10.1016/j.omtn.2019.01.013

Hegde, R. S., & Keenan, R. J. (2022). The mechanisms of integral membrane protein biogenesis. *Nature Reviews Molecular Cell Biology*, *23*(2), 107–124. https://doi.org/10.1038/s41580-021-00413-2

Henderson, J. M., Ujita, A., Hill, E., Yousif-Rosales, S., Smith, C., Ko, N., McReynolds, T., Cabral, C. R., Escamilla-Powers, J. R., & Houston, M. E. (2021). Cap 1 Messenger RNA Synthesis with Co-transcriptional CleanCap ® Analog by In Vitro Transcription. *Current Protocols*, *1*(2), e39. https://doi.org/10.1002/cpz1.39

Hồ, N. T., Hughes, S. G., Ta, V. T., Phan, L. T., Đỗ, Q., Nguyễn, T. V., Phạm, A. T. V., Thị Ngọc Đặng, M., Nguyễn, L. V., Trịnh, Q. V., Phạm, H. N., Chử, M. V., Nguyễn, T. T., Lương, Q. C., Tường Lê, V. T., Nguyễn, T. V., Trần, L.-T.-L., Thi Van Luu, A., Nguyen, A. N., … Nguyen, X.-H. (2024). Safety, immunogenicity and efficacy of the self-amplifying mRNA ARCT-154 COVID-19 vaccine: Pooled phase 1, 2, 3a and 3b randomized, controlled trials. *Nature Communications*, *15*(1), 4081. https://doi.org/10.1038/s41467-024-47905-1

Holtkamp, S., Kreiter, S., Selmi, A., Simon, P., Koslowski, M., Huber, C., Türeci, Ö., & Sahin, U. (2006). Modification of antigen-encoding RNA increases stability, translational efficacy, and T-cell stimulatory capacity of dendritic cells. *Blood*, *108*(13), 4009–4017. https://doi.org/10.1182/blood-2006-04-015024

Hou, X., Zaks, T., Langer, R., & Dong, Y. (2021). Lipid nanoparticles for mRNA delivery. In *Nature Reviews Materials*. https://doi.org/10.1038/s41578-021-00358-0

Hu, B., Zhong, L., Weng, Y., Peng, L., Huang, Y., Zhao, Y., & Liang, X.-J. (2020). Therapeutic siRNA: State of the art. *Signal Transduction and Targeted Therapy*, *5*(1), 101. https://doi.org/10.1038/s41392-020-0207-x

Hu, C., Bai, Y., Liu, J., Wang, Y., He, Q., Zhang, X., Cheng, F., Xu, M., Mao, Q., & Liang, Z. (2024). Research progress on the quality control of mRNA vaccines. *Expert Review of Vaccines*, *23*(1), 570–583. https://doi.org/10.1080/14760584.2024.2354251

Huang, H., Weng, H., & Chen, J. (2020). The Biogenesis and Precise Control of RNA m6A Methylation. *Trends in Genetics: TIG*, *36*(1), 44–52. https://doi.org/10.1016/j.tig.2019.10.011

Huayamares, S. G., Lokugamage, M. P., Rab, R., Da Silva Sanchez, A. J., Kim, H., Radmand, A., Loughrey, D., Lian, L., Hou, Y., Achyut, B. R., Ehrhardt, A., Hong, J. S., Sago, C. D., Paunovska, K., Echeverri, E. S., Vanover, D., Santangelo, P. J., Sorscher, E. J., & Dahlman, J. E. (2023). High-throughput screens identify a lipid nanoparticle that preferentially delivers mRNA to human tumors in vivo. *Journal of Controlled Release*, *357*, 394–403. https://doi.org/10.1016/j.jconrel.2023.04.005

Hung, V., Lam, S. S., Udeshi, N. D., Svinkina, T., Guzman, G., Mootha, V. K., Carr, S. A., & Ting, A. Y. (2017). Proteomic mapping of cytosol-facing outer mitochondrial and ER membranes in living human cells by proximity biotinylation. *eLife*, *6*. https://doi.org/10.7554/eLife.24463

Hunter, M. R., Cui, L., Porebski, B. T., Pereira, S., Sonzini, S., Odunze, U., Iyer, P., Engkvist, O., Lloyd, R. L., Peel, S., Sabirsh, A., Ross-Thriepland, D., Jones, A. T., & Desai, A. S. (2023). Understanding Intracellular Biology to Improve mRNA Delivery by Lipid Nanoparticles. *Small Methods*, *7*(9), 2201695. https://doi.org/10.1002/smtd.202201695

Hwang, J., & Espenshade, P. J. (2016). Proximity-dependent biotin labelling in yeast using the engineered ascorbate peroxidase APEX2. *The Biochemical Journal*, *473*(16), 2463–2469. https://doi.org/10.1042/BCJ20160106

Ibba, M. L., Ciccone, G., Esposito, C. L., Catuogno, S., & Giangrande, P. H. (2021). Advances in mRNA non-viral delivery approaches. *Advanced Drug Delivery Reviews*, *177*, 113930. https://doi.org/10.1016/j.addr.2021.113930

Ishigaki, Y., Li, X., Serin, G., & Maquat, L. E. (2001). Evidence for a pioneer round of mRNA translation: mRNAs subject to nonsense-mediated decay in mammalian cells are bound by CBP80 and CBP20. *Cell*, *106*(5), 607–617. https://doi.org/10.1016/s0092-8674(01)00475-5

Izaurralde, E., Lewis, J., McGuigan, C., Jankowska, M., Darzynkiewicz, E., & Mattaj, I. W. (1994). A nuclear cap binding protein complex involved in pre-mRNA splicing. *Cell*, *78*(4), 657–668. https://doi.org/10.1016/0092-8674(94)90530-4

Jan, C. H., Williams, C. C., & Weissman, J. S. (2014). Principles of ER cotranslational translocation revealed by proximity-specific ribosome profiling. *Science*, *346*(6210), 1257521. https://doi.org/10.1126/science.1257521

Jayaraman, M., Ansell, S. M., Mui, B. L., Tam, Y. K., Chen, J., Du, X., Butler, D., Eltepu, L., Matsuda, S., Narayanannair, J. K., Rajeev, K. G., Hafez, I. M., Akinc, A., Maier, M. A., Tracy, M. A., Cullis, P. R., Madden, T. D., Manoharan, M., & Hope, M. J. (2012). Maximizing the potency of siRNA lipid nanoparticles for hepatic gene silencing in vivo. *Angewandte Chemie (International Ed. in English)*, *51*(34), 8529–8533. https://doi.org/10.1002/anie.201203263

Jia, X., He, X., Huang, C., Li, J., Dong, Z., & Liu, K. (2024). Protein translation: Biological processes and therapeutic strategies for human diseases. *Signal Transduction and Targeted Therapy*, *9*(1), 44. https://doi.org/10.1038/s41392-024-01749-9

Jiang, A. Y., Witten, J., Raji, I. O., Eweje, F., MacIsaac, C., Meng, S., Oladimeji, F. A., Hu, Y., Manan, R. S., Langer, R., & Anderson, D. G. (2024). Combinatorial development of nebulized mRNA delivery formulations for the lungs. *Nature Nanotechnology*, *19*(3), 364–375. https://doi.org/10.1038/s41565-023-01548-3

Jirikowski, G. F., Sanna, P. P., Maciejewski-Lenoir, D., & Bloom, F. E. (1992). Reversal of Diabetes Insipidus in Brattleboro Rats: Intrahypothalamic Injection of Vasopressin mRNA. *Science*, *255*(5047), 996–998. https://doi.org/10.1126/science.1546298

Joazeiro, C. A. P. (2019). Mechanisms and functions of ribosome-associated protein quality control. *Nature Reviews. Molecular Cell Biology*, *20*(6), 368–383. https://doi.org/10.1038/s41580-019-0118-2

Johansson, J. M., Du Rietz, H., Hedlund, H., Eriksson, H. C., Oude Blenke, E., Pote, A., Harun, S., Nordenfelt, P., Lindfors, L., & Wittrup, A. (2024). *Cellular and biophysical barriers to lipid nanoparticle mediated delivery of RNA to the cytosol*. https://doi.org/10.1101/2024.05.31.596627

Joshi, B. S., de Beer, M. A., Giepmans, B. N. G., & Zuhorn, I. S. (2020). Endocytosis of Extracellular Vesicles and Release of Their Cargo from Endosomes. *ACS Nano*, *14*(4), 4444–4455. https://doi.org/10.1021/acsnano.9b10033

Ju, Y., Lee, W. S., Pilkington, E. H., Kelly, H. G., Li, S., Selva, K. J., Wragg, K. M., Subbarao, K., Nguyen, T. H. O., Rowntree, L. C., Allen, L. F., Bond, K., Williamson, D. A., Truong, N. P., Plebanski, M., Kedzierska, K., Mahanty, S., Chung, A. W., Caruso, F., … Kent, S. J. (2022). Anti-PEG Antibodies Boosted in Humans by SARS-CoV-2 Lipid Nanoparticle mRNA Vaccine. *ACS Nano*, *16*(8), 11769–11780. https://doi.org/10.1021/acsnano.2c04543

Kaczmarek, J. C., Patel, A. K., Kauffman, K. J., Fenton, O. S., Webber, M. J., Heartlein, M. W., DeRosa, F., & Anderson, D. G. (2016). Polymer-Lipid Nanoparticles for Systemic Delivery of mRNA to the Lungs. *Angewandte Chemie (International Ed. in English)*, *55*(44), 13808–13812. https://doi.org/10.1002/anie.201608450

Kaewsapsak, P., Shechner, D. M., Mallard, W., Rinn, J. L., & Ting, A. Y. (2017). Live-cell mapping of organelle-associated RNAs via proximity biotinylation combined with protein-RNA crosslinking. *eLife*, *6*. https://doi.org/10.7554/eLife.29224

Kang, M.-G., & Rhee, H.-W. (2022). Molecular Spatiomics by Proximity Labeling. *Accounts of Chemical Research*, *55*(10), 1411–1422. https://doi.org/10.1021/acs.accounts.2c00061

Karikó, K., Buckstein, M., Ni, H., & Weissman, D. (2005). Suppression of RNA Recognition by Toll-like Receptors: The Impact of Nucleoside Modification and the Evolutionary Origin of RNA. *Immunity*, *23*(2), 165–175. https://doi.org/10.1016/j.immuni.2005.06.008

Karikó, K., Muramatsu, H., Ludwig, J., & 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 Research*, *39*(21), e142. https://doi.org/10.1093/nar/gkr695

Karikó, K., Muramatsu, H., Welsh, F. A., Ludwig, J., Kato, H., Akira, S., & Weissman, D. (2008). Incorporation of pseudouridine into mRNA yields superior nonimmunogenic vector with increased translational capacity and biological stability. *Molecular Therapy*, *16*(11), 1833–1840. https://doi.org/10.1038/mt.2008.200

Karlsson, J., Rhodes, K. R., Green, J. J., & Tzeng, S. Y. (2020). Poly(beta-amino ester)s as gene delivery vehicles: Challenges and opportunities. *Expert Opinion on Drug Delivery*, *17*(10), 1395–1410. https://doi.org/10.1080/17425247.2020.1796628

Kasiewicz, L. N., Biswas, S., Beach, A., Ren, H., Dutta, C., Mazzola, A. M., Rohde, E., Chadwick, A., Cheng, C., Garcia, S. P., Iyer, S., Matsumoto, Y., Khera, A. V., Musunuru, K., Kathiresan, S., Malyala, P., Rajeev, K. G., & Bellinger, A. M. (2023). GalNAc-Lipid nanoparticles enable non-LDLR dependent hepatic delivery of a CRISPR base editing therapy. *Nature Communications*, *14*(1), 2776. https://doi.org/10.1038/s41467-023-37465-1

Katahira, J. (2012). mRNA export and the TREX complex. *Biochimica et Biophysica Acta (BBA) - Gene Regulatory Mechanisms*, *1819*(6), 507–513. https://doi.org/10.1016/j.bbagrm.2011.12.001

Katahira, J. (2015). Nuclear Export of Messenger RNA. *Genes*, *6*(2), 163–184. https://doi.org/10.3390/genes6020163

Kauffman, K. J., Oberli, M. A., Dorkin, J. R., Hurtado, J. E., Kaczmarek, J. C., Bhadani, S., Wyckoff, J., Langer, R., Jaklenec, A., & Anderson, D. G. (2018). Rapid, Single-Cell Analysis and Discovery of Vectored mRNA Transfection In Vivo with a loxP-Flanked tdTomato Reporter Mouse. *Molecular Therapy. Nucleic Acids*, *10*, 55–63. https://doi.org/10.1016/j.omtn.2017.11.005

Keenan, R. J., Freymann, D. M., Stroud, R. M., & Walter, P. (2001). The Signal Recognition Particle. *Annual Review of Biochemistry*, *70*(1), 755–775. https://doi.org/10.1146/annurev.biochem.70.1.755

Kijima, T., Eguchi, T., Neckers, L., & Prince, T. L. (2018). Monitoring of the Heat Shock Response with a Real-Time Luciferase Reporter. *Methods in Molecular Biology (Clifton, N.J.)*, *1709*, 35–45. https://doi.org/10.1007/978-1-4939-7477-1_3

Kirschman, J. L., Bhosle, S., Vanover, D., Blanchard, E. L., Loomis, K. H., Zurla, C., Murray, K., Lam, B. C., & Santangelo, P. J. (2017). Characterizing exogenous mRNA delivery, trafficking, cytoplasmic release and RNA–protein correlations at the level of single cells. *Nucleic Acids Research*, *45*(12), e113–e113. https://doi.org/10.1093/nar/gkx290

Koeberl, D., Schulze, A., Sondheimer, N., Lipshutz, G. S., Geberhiwot, T., Li, L., Saini, R., Luo, J., Sikirica, V., Jin, L., Liang, M., Leuchars, M., & Grunewald, S. (2024). Interim analyses of a first-in-human phase 1/2 mRNA trial for propionic acidaemia. *Nature*, *628*(8009), 872–877. https://doi.org/10.1038/s41586-024-07266-7

Komar, A. A., Lesnik, T., & Reiss, C. (1999). Synonymous codon substitutions affect ribosome traffic and protein folding during in vitro translation. *FEBS Letters*, *462*(3), 387–391. https://doi.org/10.1016/s0014-5793(99)01566-5

Konarska, M. M., Padgett, R. A., & Sharp, P. A. (1984). Recognition of cap structure in splicing in vitro of mRNA precursors. *Cell*, *38*(3), 731–736. https://doi.org/10.1016/0092-8674(84)90268-X

Kore, A. R., Charles, I., Yang, L., & Kuersten, S. (2009). Synthesis and Activity of Modified Cytidine 5'-Monophosphate Probes for T4 RNA Ligase 1. *Nucleosides, Nucleotides and Nucleic Acids*, *28*(4), 292–302. https://doi.org/10.1080/15257770902946181

Kornberg, R. D. (1974). Chromatin Structure: A Repeating Unit of Histones and DNA: Chromatin structure is based on a repeating unit of eight histone molecules and about 200 DNA base pairs. *Science*, *184*(4139), 868–871. https://doi.org/10.1126/science.184.4139.868

Kornberg, R. D., & Lorch, Y. (1999). Twenty-Five Years of the Nucleosome, Fundamental Particle of the Eukaryote Chromosome. *Cell*, *98*(3), 285–294. https://doi.org/10.1016/S0092-8674(00)81958-3

Kozak, M. (1986). Point mutations define a sequence flanking the AUG initiator codon that modulates translation by eukaryotic ribosomes. *Cell*, *44*(2), 283–292. https://doi.org/10.1016/0092-8674(86)90762-2

Kozak, M. (1994). Features in the 5' non-coding sequences of rabbit α and β-globin mRNAs that affect translational efficiency. *Journal of Molecular Biology*, *235*(1), 95–110. https://doi.org/10.1016/S0022-2836(05)80019-1

Kubitz, L., Bitsch, S., Zhao, X., Schmitt, K., Deweid, L., Roehrig, A., Barazzone, E. C., Valerius, O., Kolmar, H., & Béthune, J. (2022). Engineering of ultraID, a compact and hyperactive enzyme for proximity-dependent biotinylation in living cells. *Communications Biology*, *5*(1), 657. https://doi.org/10.1038/s42003-022-03604-5

Kühn, U., Nemeth, A., Meyer, S., & Wahle, E. (2003). The RNA Binding Domains of the Nuclear poly(A)-binding Protein. *Journal of Biological Chemistry*, *278*(19), 16916–16925. https://doi.org/10.1074/jbc.M209886200

Kulkarni, J. A., Darjuan, M. M., Mercer, J. E., Chen, S., Van Der Meel, R., Thewalt, J. L., Tam, Y. Y. C., & Cullis, P. R. (2018). On the Formation and Morphology of Lipid Nanoparticles Containing Ionizable Cationic Lipids and siRNA. *ACS Nano*, *12*(5), 4787–4795. https://doi.org/10.1021/acsnano.8b01516

Kwon, B. S., Chintamaneni, C., Kozak, C. A., Copeland, N. G., Gilbert, D. J., Jenkins, N., Barton, D., Francke, U., Kobayashi, Y., & Kim, K. K. (1991). A melanocyte-specific gene, Pmel 17, maps near the silver coat color locus on mouse chromosome 10 and is in a

syntenic region on human chromosome 12. *Proceedings of the National Academy of Sciences*, *88*(20), 9228–9232. https://doi.org/10.1073/pnas.88.20.9228

Lam, S. S., Martell, J. D., Kamer, K. J., Deerinck, T. J., Ellisman, M. H., Mootha, V. K., & Ting, A. Y. (2014). Directed evolution of APEX2 for electron microscopy and proximity labeling. *Nature Methods*. https://doi.org/10.1038/nmeth.3179

Le Hir, H., Izaurralde, E., Maquat, L. E., & Moore, M. J. (2000). The spliceosome deposits multiple proteins 20-24 nucleotides upstream of mRNA exon-exon junctions. *The EMBO Journal*, *19*(24), 6860–6869. https://doi.org/10.1093/emboj/19.24.6860

Lee, S.-Y., Kang, M.-G., Shin, S., Kwak, C., Kwon, T., Seo, J. K., Kim, J.-S., & Rhee, H.-W. (2017). Architecture Mapping of the Inner Mitochondrial Membrane Proteome by Chemical Tools in Live Cells. *Journal of the American Chemical Society*, *139*(10), 3651–3662. https://doi.org/10.1021/jacs.6b10418

Lee, S.-Y., Lee, H., Lee, H.-K., Lee, S.-W., Ha, S. C., Kwon, T., Seo, J. K., Lee, C., & Rhee, H.-W. (2016). Proximity-Directed Labeling Reveals a New Rapamycin-Induced Heterodimer of FKBP25 and FRB in Live Cells. *ACS Central Science*, *2*(8), 506–516. https://doi.org/10.1021/acscentsci.6b00137

Lee, S.-Y., Seo, J. K., & Rhee, H.-W. (2019). Direct Identification of Biotinylated Proteins from Proximity Labeling (Spot-BioID). *Methods in Molecular Biology (Clifton, N.J.)*, *2008*, 97–105. https://doi.org/10.1007/978-1-4939-9537-0_8

Lee, Y., Choe, J., Park, O. H., & Kim, Y. K. (2020). Molecular Mechanisms Driving mRNA Degradation by m6A Modification. *Trends in Genetics*, *36*(3), 177–188. https://doi.org/10.1016/j.tig.2019.12.007

Lee, Y., Jeong, M., Park, J., Jung, H., & Lee, H. (2023). Immunogenicity of lipid nanoparticles and its impact on the efficacy of mRNA vaccines and therapeutics. *Experimental & Molecular Medicine*, *55*(10), 2085–2096. https://doi.org/10.1038/s12276-023-01086-x

Lei, J., Qi, S., Yu, X., Gao, X., Yang, K., Zhang, X., Cheng, M., Bai, B., Feng, Y., Lu, M., Wang, Y., Li, H., & Yu, G. (2024). Development of Mannosylated Lipid Nanoparticles for mRNA Cancer Vaccine with High Antigen Presentation Efficiency and Immunomodulatory Capability. *Angewandte Chemie International Edition*, *63*(13), e202318515. https://doi.org/10.1002/anie.202318515

Lenk, R., Kleindienst, W., Szabó, G. T., Baiersdörfer, M., Boros, G., Keller, J. M., Mahiny, A. J., & Vlatkovic, I. (2024). Understanding the impact of in vitro transcription byproducts and contaminants. *Frontiers in Molecular Biosciences*, *11*, 1426129. https://doi.org/10.3389/fmolb.2024.1426129

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., Tiu, G. C., Wellington-Oguri, R., Fujii, K., Sharma, E., Watkins, A. M., Nicol, J. J., Romano, J., Tunguz, B., Diaz, F., … Das, R. (2022). Combinatorial optimization of mRNA structure, stability, and translation for RNA-based therapeutics. *Nature Communications*, *13*(1), 1536. https://doi.org/10.1038/s41467-022-28776-w

Leppek, K., Das, R., & Barna, M. (2018). Functional 5' UTR mRNA structures in eukaryotic translation regulation and how to find them. *Nature Reviews Molecular Cell Biology*, *19*(3), 158–174. https://doi.org/10.1038/nrm.2017.103

Li, J., He, Y., Wang, W., Wu, C., Hong, C., & Hammond, P. T. (2017). Polyamine-Mediated Stoichiometric Assembly of Ribonucleoproteins for Enhanced mRNA Delivery. *Angewandte Chemie*, *129*(44), 13897–13900. https://doi.org/10.1002/ange.201707466

Li, J., Wang, W., He, Y., Li, Y., Yan, E. Z., Zhang, K., Irvine, D. J., & Hammond, P. T. (2017). Structurally Programmed Assembly of Translation Initiation Nanoplex for Superior mRNA Delivery. *ACS Nano*, *11*(3), 2531–2544. https://doi.org/10.1021/acsnano.6b08447

Li, R., Zou, Z., Wang, W., & Zou, P. (2022). Metabolic incorporation of electron-rich ribonucleosides enhances APEX-seq for profiling spatially restricted nascent transcriptome. *Cell Chemical Biology*, *29*(7), 1218-1231.e8. https://doi.org/10.1016/j.chembiol.2022.02.005

Li, Y., & Breaker, R. R. (1999). Kinetics of RNA Degradation by Specific Base Catalysis of Transesterification Involving the 2‘-Hydroxyl Group. *Journal of the American Chemical Society*, *121*(23), 5364–5372. https://doi.org/10.1021/ja990592p

Liaci, A. M., Steigenberger, B., Telles De Souza, P. C., Tamara, S., Gröllers-Mulderij, M., Ogrissek, P., Marrink, S. J., Scheltema, R. A., & Förster, F. (2021). Structure of the human signal peptidase complex reveals the determinants for signal peptide cleavage. *Molecular Cell*, *81*(19), 3934-3948.e11. https://doi.org/10.1016/j.molcel.2021.07.031

Liang, Y., Liu, N., Yang, L., Tang, J., Wang, Y., & Mei, M. (2021). A Brief Review of circRNA Biogenesis, Detection, and Function. *Current Genomics*, *22*(7), 485–495. https://doi.org/10.2174/1389202922666210331130722

Lima, S. A., Chipman, L. B., Nicholson, A. L., Chen, Y.-H., Yee, B. A., Yeo, G. W., Coller, J., & Pasquinelli, A. E. (2017). Short poly(A) tails are a conserved feature of highly expressed genes. *Nature Structural & Molecular Biology*, *24*(12), 1057–1063. https://doi.org/10.1038/nsmb.3499

Lin, L., Su, K., Cheng, Q., & Liu, S. (2023). Targeting materials and strategies for RNA delivery. *Theranostics*, *13*(13), 4667–4693. https://doi.org/10.7150/thno.87316

Lindsay, K. E., Bhosle, S. M., Zurla, C., Beyersdorf, J., Rogers, K. A., Vanover, D., Xiao, P., Araínga, M., Shirreff, L. M., Pitard, B., Baumhof, P., Villinger, F., & Santangelo, P. J. (2019). Visualization of early events in mRNA vaccine delivery in non-human primates via PET-CT and near-infrared imaging. *Nature Biomedical Engineering*, *3*(5), 371–380. https://doi.org/10.1038/s41551-019-0378-3

Liu, M., Gong, X., Alluri, R. K., Wu, J., Sablo, T., & Li, Z. (2012). Characterization of RNA damage under oxidative stress in Escherichia coli. *Biological Chemistry*, *393*(3), 123–132. https://doi.org/10.1515/hsz-2011-0247

Liu, X., Zhang, Y., Zhou, S., Dain, L., Mei, L., & Zhu, G. (2022). Circular RNA: An emerging frontier in RNA therapeutic targets, RNA therapeutics, and mRNA vaccines. *Journal of Controlled Release: Official Journal of the Controlled Release Society*, *348*, 84–94. https://doi.org/10.1016/j.jconrel.2022.05.043

Liu, Y., Yang, Q., & Zhao, F. (2021). Synonymous but Not Silent: The Codon Usage Code for Gene Expression and Protein Folding. *Annual Review of Biochemistry*, *90*, 375–401. https://doi.org/10.1146/annurev-biochem-071320-112701

Loan Young, T., Chang Wang, K., James Varley, A., & Li, B. (2023). Clinical delivery of circular RNA: Lessons learned from RNA drug development. *Advanced Drug Delivery Reviews*, *197*, 114826. https://doi.org/10.1016/j.addr.2023.114826

Lorenz, R., Wolfinger, M. T., Tanzer, A., & Hofacker, I. L. (2016). Predicting RNA secondary structures from sequence and probing data. *Methods*, *103*, 86–98. https://doi.org/10.1016/j.ymeth.2016.04.004

Los, G. V., Encell, L. P., McDougall, M. G., Hartzell, D. D., Karassina, N., Zimprich, C., Wood, M. G., Learish, R., Ohana, R. F., Urh, M., Simpson, D., Mendez, J., Zimmerman, K., Otto, P., Vidugiris, G., Zhu, J., Darzins, A., Klaubert, D. H., Bulleit, R. F., & Wood, K. V. (2008). HaloTag: A Novel Protein Labeling Technology for Cell Imaging and Protein Analysis. *ACS Chemical Biology*, *3*(6), 373–382. https://doi.org/10.1021/cb800025k

Louloupi, A., Ntini, E., Conrad, T., & Ørom, U. A. V. (2018). Transient N-6-Methyladenosine Transcriptome Sequencing Reveals a Regulatory Role of m6A in Splicing Efficiency. *Cell Reports*, *23*(12), 3429–3437. https://doi.org/10.1016/j.celrep.2018.05.077

Lu, B., Javidi-Parsijani, P., Makani, V., Mehraein-Ghomi, F., Sarhan, W. M., Sun, D., Yoo, K. W., Atala, Z. P., Lyu, P., & Atala, A. (2019). Delivering SaCas9 mRNA by lentivirus-like bionanoparticles for transient expression and efficient genome editing. *Nucleic Acids Research*, *47*(8), e44–e44. https://doi.org/10.1093/nar/gkz093

Lu, J., Lu, G., Tan, S., Xia, J., Xiong, H., Yu, X., Qi, Q., Yu, X., Li, L., Yu, H., Xia, N., Zhang, T., Xu, Y., & Lin, J. (2020). A COVID-19 mRNA vaccine encoding SARS-CoV-2 virus-like particles induces a strong antiviral-like immune response in mice. *Cell Research*, *30*(10), 936–939. https://doi.org/10.1038/s41422-020-00392-7

Lu, Y., Huang, W., Li, M., & Zheng, A. (2023). Exosome-Based Carrier for RNA Delivery: Progress and Challenges. *Pharmaceutics*, *15*(2), 598. https://doi.org/10.3390/pharmaceutics15020598

Makino, S., Kawamata, T., Iwasaki, S., & Ohsumi, Y. (2021). Selectivity of mRNA degradation by autophagy in yeast. *Nature Communications*, *12*(1), 2316. https://doi.org/10.1038/s41467-021-22574-6

Mandel, C. R., Bai, Y., & Tong, L. (2008). Protein factors in pre-mRNA 3'-end processing. *Cellular and Molecular Life Sciences*, *65*(7–8), 1099–1122. https://doi.org/10.1007/s00018-007-7474-3

Maracci, C., Motta, S., Romagnoli, A., Costantino, M., Perego, P., & Di Marino, D. (2022). The mTOR/4E-BP1/eIF4E Signalling Pathway as a Source of CancerDrug Targets. *Current Medicinal Chemistry*, *29*(20), 3501–3529. https://doi.org/10.2174/0929867329666220224112042

Maresh, G. A., Marken, J. S., Neubauer, M., Aruffo, A., Hellström, I., Hellström, K. E., & Marquardt, H. (1994). Cloning and Expression of the Gene for the Melanoma-

Associated ME20 Antigen. *DNA and Cell Biology*, *13*(2), 87–95. https://doi.org/10.1089/dna.1994.13.87

Marotta, C. A., Wilson, J. T., Forget, B. G., & Weissman, S. M. (1977). Human beta-globin messenger RNA. III. Nucleotide sequences derived from complementary DNA. *The Journal of Biological Chemistry*, *252*(14), 5040–5053.

Martell, J. D., Deerinck, T. J., Lam, S. S., Ellisman, M. H., & Ting, A. Y. (2017). Electron microscopy using the genetically encoded APEX2 tag in cultured mammalian cells. *Nature Protocols*, *12*(9), 1792–1816. https://doi.org/10.1038/nprot.2017.065

Martell, J. D., Deerinck, T. J., Sancak, Y., Poulos, T. L., Mootha, V. K., Sosinsky, G. E., Ellisman, M. H., & Ting, A. Y. (2012). Engineered ascorbate peroxidase as a genetically encoded reporter for electron microscopy. *Nature Biotechnology*, *30*(11), 1143–1148. https://doi.org/10.1038/nbt.2375

Martinon, F., Krishnan, S., Lenzen, G., Magné, R., Gomard, E., Guillet, J., Lévy, J., & Meulien, P. (1993). Induction of virus-specific cytotoxic T lymphocytes *in vivo* by liposome-entrapped mRNA. *European Journal of Immunology*, *23*(7), 1719–1722. https://doi.org/10.1002/eji.1830230749

Mathews, D. H., Sabina, J., Zuker, M., & Turner, D. H. (1999). Expanded sequence dependence of thermodynamic parameters improves prediction of RNA secondary structure. *Journal of Molecular Biology*, *288*(5), 911–940. https://doi.org/10.1006/jmbi.1999.2700

Mauer, J., Luo, X., Blanjoie, A., Jiao, X., Grozhik, A. V., Patil, D. P., Linder, B., Pickering, B. F., Vasseur, J.-J., Chen, Q., Gross, S. S., Elemento, O., Debart, F., Kiledjian, M., & Jaffrey, S. R. (2017). Reversible methylation of m6Am in the 5' cap controls mRNA stability. *Nature*, *541*(7637), 371–375. https://doi.org/10.1038/nature21022

Mauger, D. M., Cabral, B. J., Presnyak, V., Su, S. V., Reid, D. W., Goodman, B., Link, K., Khatwani, N., Reynders, J., Moore, M. J., & McFadyen, I. J. (2019). mRNA structure regulates protein expression through changes in functional half-life. *Proceedings of the National Academy of Sciences of the United States of America*, *116*(48), 24075–24083. https://doi.org/10.1073/pnas.1908052116

Mazia, D., Schatten, G., & Sale, W. (1975). Adhesion of cells to surfaces coated with polylysine. Applications to electron microscopy. *The Journal of Cell Biology*, *66*(1), 198–200. https://doi.org/10.1083/jcb.66.1.198

McMahon, A. C., Rahman, R., Jin, H., Shen, J. L., Fieldsend, A., Luo, W., & Rosbash, M. (2016). TRIBE: Hijacking an RNA-Editing Enzyme to Identify Cell-Specific Targets of RNA-Binding Proteins. *Cell*, *165*(3), 742–753. https://doi.org/10.1016/j.cell.2016.03.007

Medina-Munoz, H. C., Kofman, E., Jagannatha, P., Boyle, E. A., Yu, T., Jones, K. L., Mueller, J. R., Lykins, G. D., Doudna, A. T., Park, S. S., Blue, S. M., Ranzau, B. L., Kohli, R. M., Komor, A. C., & Yeo, G. W. (2024). Expanded palette of RNA base editors for comprehensive RBP-RNA interactome studies. *Nature Communications*, *15*(1), 875. https://doi.org/10.1038/s41467-024-45009-4

Mellman, I., & Nelson, W. J. (2008). Coordinated protein sorting, targeting and distribution in polarized cells. *Nature Reviews Molecular Cell Biology*, *9*(11), 833–845. https://doi.org/10.1038/nrm2525

Mo, C., Li, X., Wu, Q., Fan, Y., Liu, D., Zhu, Y., Yang, Y., Liao, X., Zhou, Z., Zhou, L., Li, Q., Zhang, Q., Liu, W., Zhou, R., & Tian, X. (2023). SARS-CoV-2 mRNA vaccine requires signal peptide to induce antibody responses. *Vaccine*, *41*(46), 6863–6869. https://doi.org/10.1016/j.vaccine.2023.09.059

Mockey, M., Gonçalves, C., Dupuy, F. P., Lemoine, F. M., Pichon, C., & Midoux, P. (2006). mRNA transfection of dendritic cells: Synergistic effect of ARCA mRNA capping with Poly(A) chains in cis and in trans for a high protein expression level. *Biochemical and Biophysical Research Communications*, *340*(4), 1062–1068. https://doi.org/10.1016/j.bbrc.2005.12.105

Mu, X., Greenwald, E., Ahmad, S., & Hur, S. (2018). An origin of the immunogenicity of in vitro transcribed RNA. *Nucleic Acids Research*, *46*(10), 5239–5249. https://doi.org/10.1093/nar/gky177

Mu, X., & Hur, S. (2021). Immunogenicity of *In Vitro* -Transcribed RNA. *Accounts of Chemical Research*, *54*(21), 4012–4023. https://doi.org/10.1021/acs.accounts.1c00521

Mulroney, T. E., Pöyry, T., Yam-Puc, J. C., Rust, M., Harvey, R. F., Kalmar, L., Horner, E., Booth, L., Ferreira, A. P., Stoneley, M., Sawarkar, R., Mentzer, A. J., Lilley, K. S., Smales, C. M., Von Der Haar, T., Turtle, L., Dunachie, S., Klenerman, P., Thaventhiran, J. E. D., & Willis, A. E. (2024). N1-methylpseudouridylation of mRNA causes +1 ribosomal frameshifting. *Nature*, *625*(7993), 189–194. https://doi.org/10.1038/s41586-023-06800-3

Munson, M. J., O'Driscoll, G., Silva, A. M., Lázaro-Ibáñez, E., Gallud, A., Wilson, J. T., Collén, A., Esbjörner, E. K., & Sabirsh, A. (2021). A high-throughput Galectin-9 imaging assay for quantifying nanoparticle uptake, endosomal escape and functional RNA delivery. *Communications Biology*, *4*(1), 211. https://doi.org/10.1038/s42003-021-01728-8

Nakamura, K., Sawada, K., Kinose, Y., Yoshimura, A., Toda, A., Nakatsuka, E., Hashimoto, K., Mabuchi, S., Morishige, K., Kurachi, H., Lengyel, E., & Kimura, T. (2017). Exosomes Promote Ovarian Cancer Cell Invasion through Transfer of CD44 to Peritoneal Mesothelial Cells. *Molecular Cancer Research*, *15*(1), 78–92. https://doi.org/10.1158/1541-7786.MCR-16-0191

Negwer, I., Hirsch, M., Kaloyanova, S., Brown, T., Peneva, K., Butt, H., Koynov, K., & Helm, M. (2017). Modulation of Mitochondriotropic Properties of Cyanine Dyes by in Organello Copper-Free Click Reaction. *ChemBioChem*, *18*(18), 1814–1818. https://doi.org/10.1002/cbic.201700286

Nickless, A., Bailis, J. M., & You, Z. (2017). Control of gene expression through the nonsense-mediated RNA decay pathway. *Cell & Bioscience*, *7*(1), 26. https://doi.org/10.1186/s13578-017-0153-7

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., Bukur, V., Kuhn, A. N., Türeci, Ö., & Sahin, U. (2019). Improving mRNA-Based Therapeutic Gene Delivery by Expression-Augmenting 3' UTRs Identified by Cellular Library Screening. *Molecular Therapy: The Journal of the American Society of Gene Therapy*, *27*(4), 824–836. https://doi.org/10.1016/j.ymthe.2018.12.011

Ostro, M. J., Giacomoni, D., Lavelle, D., Paxton, W., & Dray, S. (1978). Evidence for translation of rabbit globin mRNA after liposomemediated insertion into a human cell line. *Nature*, *274*(5674), 921–923. https://doi.org/10.1038/274921a0

Owensiii, D., & Peppas, N. (2006). Opsonization, biodistribution, and pharmacokinetics of polymeric nanoparticles. *International Journal of Pharmaceutics*, *307*(1), 93–102. https://doi.org/10.1016/j.ijpharm.2005.10.010

Padrón, A., Iwasaki, S., & Ingolia, N. T. (2019). Proximity RNA Labeling by APEX-Seq Reveals the Organization of Translation Initiation Complexes and Repressive RNA Granules. *Molecular Cell*, *75*(4), 875-887.e5. https://doi.org/10.1016/j.molcel.2019.07.030

Paramasivam, P., Franke, C., Stöter, M., Höijer, A., Bartesaghi, S., Sabirsh, A., Lindfors, L., Arteta, M. Y., Dahlén, A., Bak, A., Andersson, S., Kalaidzidis, Y., Bickle, M., & Zerial, M. (2022). Endosomal escape of delivered mRNA from endosomal recycling tubules visualized at the nanoscale. *Journal of Cell Biology*, *221*(2). https://doi.org/10.1083/jcb.202110137

Paramasivam, P., Stöter, M., Corradi, E., Dalla Costa, I., Höijer, A., Bartesaghi, S., Sabirsh, A., Lindfors, L., Yanez Arteta, M., Nordberg, P., Andersson, S., Baudet, M.-L., Bickle, M., & Zerial, M. (2022). Quantitative intracellular retention of delivered RNAs through optimized cell fixation and immunostaining. *RNA*, *28*(3), 433–446. https://doi.org/10.1261/rna.078895.121

Parhiz, H., Atochina-Vasserman, E. N., & Weissman, D. (2024). mRNA-based therapeutics: Looking beyond COVID-19 vaccines. *The Lancet*, *403*(10432), 1192–1204. https://doi.org/10.1016/S0140-6736(23)02444-3

Passmore, L. A., & Coller, J. (2022). Roles of mRNA poly(A) tails in regulation of eukaryotic gene expression. *Nature Reviews Molecular Cell Biology*, *23*(2), 93–106. https://doi.org/10.1038/s41580-021-00417-y

Paunovska, K., Sago, C. D., Monaco, C. M., Hudson, W. H., Castro, M. G., Rudoltz, T. G., Kalathoor, S., Vanover, D. A., Santangelo, P. J., Ahmed, R., Bryksin, A. V., & Dahlman, J. E. (2018). A Direct Comparison of in Vitro and in Vivo Nucleic Acid Delivery Mediated by Hundreds of Nanoparticles Reveals a Weak Correlation. *Nano Letters*, *18*(3), 2148–2157. https://doi.org/10.1021/acs.nanolett.8b00432

Pei, D., & Buyanova, M. (2019). Overcoming Endosomal Entrapment in Drug Delivery. *Bioconjugate Chemistry*, *30*(2), 273–283. https://doi.org/10.1021/acs.bioconjchem.8b00778

Peraro, L., Deprey, K. L., Moser, M. K., Zou, Z., Ball, H. L., Levine, B., & Kritzer, J. A. (2018). Cell Penetration Profiling Using the Chloroalkane Penetration Assay. *Journal of*

the American Chemical Society, *140*(36), 11360–11369. https://doi.org/10.1021/jacs.8b06144

Perenkov, A. D., Sergeeva, A. D., Vedunova, M. V., & Krysko, D. V. (2023). In Vitro Transcribed RNA-Based Platform Vaccines: Past, Present, and Future. *Vaccines*, *11*(10), 1600. https://doi.org/10.3390/vaccines11101600

Pike, J. A., Styles, I. B., Rappoport, J. Z., & Heath, J. K. (2017). Quantifying receptor trafficking and colocalization with confocal microscopy. *Methods (San Diego, Calif.)*, *115*, 42–54. https://doi.org/10.1016/j.ymeth.2017.01.005

Popovitz, J., Sharma, R., Hoshyar, R., Soo Kim, B., Murthy, N., & Lee, K. (2023). Gene editing therapeutics based on mRNA delivery. *Advanced Drug Delivery Reviews*, *200*, 115026. https://doi.org/10.1016/j.addr.2023.115026

Potapov, V., Fu, X., Dai, N., Corrêa, I. R., Tanner, N. A., & Ong, J. L. (2018). Base modifications affecting RNA polymerase and reverse transcriptase fidelity. *Nucleic Acids Research*, *46*(11), 5753–5763. https://doi.org/10.1093/nar/gky341

Prel, A., Caval, V., Gayon, R., Ravassard, P., Duthoit, C., Payen, E., Maouche-Chretien, L., Creneguy, A., Nguyen, T. H., Martin, N., Piver, E., Sevrain, R., Lamouroux, L., Leboulch, P., Deschaseaux, F., Bouillé, P., Sensébé, L., & Pagès, J.-C. (2015). Highly efficient in vitro and in vivo delivery of functional RNAs using new versatile MS2-chimeric retrovirus-like particles. *Molecular Therapy - Methods & Clinical Development*, *2*, 15039. https://doi.org/10.1038/mtm.2015.39

Proudfoot, N. J. (1977). Complete 3' noncoding region sequences of rabbit and human β-globin messenger RNAs. *Cell*, *10*(4), 559–570. https://doi.org/10.1016/0092-8674(77)90089-7

Qin, Y., Ou, L., Zha, L., Zeng, Y., & Li, L. (2023). Delivery of nucleic acids using nanomaterials. *Molecular Biomedicine*, *4*(1), 48. https://doi.org/10.1186/s43556-023-00160-0

Qu, L., Yi, Z., Shen, Y., Lin, L., Chen, F., Xu, Y., Wu, Z., Tang, H., Zhang, X., Tian, F., Wang, C., Xiao, X., Dong, X., Guo, L., Lu, S., Yang, C., Tang, C., Yang, Y., Yu, W., … Wei, W. (2022). Circular RNA vaccines against SARS-CoV-2 and emerging variants. *Cell*, *185*(10), 1728-1744.e16. https://doi.org/10.1016/j.cell.2022.03.044

Reid, D. W., & Nicchitta, C. V. (2012). Primary role for endoplasmic reticulum-bound ribosomes in cellular translation identified by ribosome profiling. *The Journal of Biological Chemistry*, *287*(8), 5518–5527. https://doi.org/10.1074/jbc.M111.312280

Reid, D. W., & Nicchitta, C. V. (2015). Diversity and selectivity in mRNA translation on the endoplasmic reticulum. *Nature Reviews Molecular Cell Biology*, *16*(4), 221–231. https://doi.org/10.1038/nrm3958

Reshetnikov, V., Terenin, I., Shepelkova, G., Yeremeev, V., Kolmykov, S., Nagornykh, M., Kolosova, E., Sokolova, T., Zaborova, O., Kukushkin, I., Kazakova, A., Kunyk, D., Kirshina, A., Vasileva, O., Seregina, K., Pateev, I., Kolpakov, F., & Ivanov, R. (2024). Untranslated Region Sequences and the Efficacy of mRNA Vaccines against Tuberculosis. *International Journal of Molecular Sciences*, *25*(2), 888. https://doi.org/10.3390/ijms25020888

Rhee, H.-W., Zou, P., Udeshi, N. D., Martell, J. D., Mootha, V. K., Carr, S. A., & Ting, A. Y. (2013). Proteomic Mapping of Mitochondria in Living Cells via Spatially Restricted Enzymatic Tagging. *Science*, *339*(6125), 1328–1331. https://doi.org/10.1126/science.1230593

Rhee, W. J., & Bao, G. (2010). Slow non-specific accumulation of 2'-deoxy and 2'-O-methyl oligonucleotide probes at mitochondria in live cells. *Nucleic Acids Research*, *38*(9), e109. https://doi.org/10.1093/nar/gkq050

Roberts, A. P. E., Lewis, A. P., & Jopling, C. L. (2011). miR-122 activates hepatitis C virus translation by a specialized mechanism requiring particular RNA components. *Nucleic Acids Research*, *39*(17), 7716–7729. https://doi.org/10.1093/nar/gkr426

Roberts, T. C., Langer, R., & Wood, M. J. A. (2020). Advances in oligonucleotide drug delivery. *Nature Reviews Drug Discovery*, *19*(10), 673–694. https://doi.org/10.1038/s41573-020-0075-7

Roeder, R. G. (1996). The role of general initiation factors in transcription by RNA polymerase II. *Trends in Biochemical Sciences*, *21*(9), 327–335.

Rosenfeld, M. G., Lunyak, V. V., & Glass, C. K. (2006). Sensors and signals: A coactivator/corepressor/epigenetic code for integrating signal-dependent programs of transcriptional response. *Genes & Development*, *20*(11), 1405–1428. https://doi.org/10.1101/gad.1424806

Roux, K. J., Kim, D. I., Raida, M., & Burke, B. (2012). A promiscuous biotin ligase fusion protein identifies proximal and interacting proteins in mammalian cells. *Journal of Cell Biology*, *196*(6), 801–810. https://doi.org/10.1083/jcb.201112098

Sago, C. D., Lokugamage, M. P., Paunovska, K., Vanover, D. A., Monaco, C. M., Shah, N. N., Gamboa Castro, M., Anderson, S. E., Rudoltz, T. G., Lando, G. N., Munnilal Tiwari, P., Kirschman, J. L., Willett, N., Jang, Y. C., Santangelo, P. J., Bryksin, A. V., & Dahlman, J. E. (2018). High-throughput in vivo screen of functional mRNA delivery identifies nanoparticles for endothelial cell gene editing. *Proceedings of the National Academy of Sciences*, *115*(42). https://doi.org/10.1073/pnas.1811276115

Sahin, U., Karikó, K., & Türeci, Ö. (2014). mRNA-based therapeutics—Developing a new class of drugs. *Nature Reviews Drug Discovery*, *13*(10), 759–780. https://doi.org/10.1038/nrd4278

Sample, P. J., Wang, B., Reid, D. W., Presnyak, V., McFadyen, I. J., Morris, D. R., & Seelig, G. (2019). Human 5' UTR design and variant effect prediction from a massively parallel translation assay. *Nature Biotechnology*, *37*(7), 803–809. https://doi.org/10.1038/s41587-019-0164-5

Sanchez, A. J. D. S., Loughrey, D., Echeverri, E. S., Huayamares, S. G., Radmand, A., Paunovska, K., Hatit, M., Tiegreen, K. E., Santangelo, P. J., & Dahlman, J. E. (2024). Substituting Poly(ethylene glycol) Lipids with Poly(2-ethyl-2-oxazoline) Lipids Improves Lipid Nanoparticle Repeat Dosing. *Advanced Healthcare Materials*, *13*(17), 2304033. https://doi.org/10.1002/adhm.202304033

Sano, T., & Cantor, C. R. (1995). Intersubunit contacts made by tryptophan 120 with biotin are essential for both strong biotin binding and biotin-induced tighter subunit

association of streptavidin. *Proceedings of the National Academy of Sciences*, *92*(8), 3180–3184. https://doi.org/10.1073/pnas.92.8.3180

Sato, H., & Maquat, L. E. (2009). Remodeling of the pioneer translation initiation complex involves translation and the karyopherin importin β. *Genes & Development*, *23*(21), 2537–2550. https://doi.org/10.1101/gad.1817109

Sayers, E. J., Peel, S. E., Schantz, A., England, R. M., Beano, M., Bates, S. M., Desai, A. S., Puri, S., Ashford, M. B., & Jones, A. T. (2019). Endocytic Profiling of Cancer Cell Models Reveals Critical Factors Influencing LNP-Mediated mRNA Delivery and Protein Expression. *Molecular Therapy: The Journal of the American Society of Gene Therapy*, *27*(11), 1950–1962. https://doi.org/10.1016/j.ymthe.2019.07.018

Schenborn, E. T., & Mierendorf, R. C. (1985). A novel transcription property of SP6 and T7 RNA polymerases: Dependence on template structure. *Nucleic Acids Research*, *13*(17), 6223–6236. https://doi.org/10.1093/nar/13.17.6223

Scherr, M., Battmer, K., Blömer, U., Ganser, A., & Grez, M. (2001). Quantitative Determination of Lentiviral Vector Particle Numbers by Real-Time PCR. *BioTechniques*, *31*(3), 520–526. https://doi.org/10.2144/01313st05

Schmidt, C., & Schnierle, B. S. (2023). Self-Amplifying RNA Vaccine Candidates: Alternative Platforms for mRNA Vaccine Development. *Pathogens (Basel, Switzerland)*, *12*(1), 138. https://doi.org/10.3390/pathogens12010138

Schoenberg, D. R., & Maquat, L. E. (2012). Regulation of cytoplasmic mRNA decay. *Nature Reviews Genetics*, *13*(4), 246–259. https://doi.org/10.1038/nrg3160

Schonborn, J., Oberstraβ, J., Breyel, E., Tittgen, J., Schumacher, J., & Lukacs, N. (1991). Monoclonal antibodies to double-stranded RNA as probes of RNA structure in crude nucleic acid extracts. *Nucleic Acids Research*, *19*(11), 2993–3000. https://doi.org/10.1093/nar/19.11.2993

Semple, S. C., Klimuk, S. K., Harasym, T. O., Dos Santos, N., Ansell, S. M., Wong, K. F., Maurer, N., Stark, H., Cullis, P. R., Hope, M. J., & Scherrer, P. (2001). Efficient encapsulation of antisense oligonucleotides in lipid vesicles using ionizable aminolipids: Formation of novel small multilamellar vesicle structures. *Biochimica et Biophysica Acta (BBA) - Biomembranes*, *1510*(1–2), 152–166. https://doi.org/10.1016/S0005-2736(00)00343-6

Seo, J. J., Jung, S.-J., Yang, J., Choi, D.-E., & Kim, V. N. (2023). Functional viromic screens uncover regulatory RNA elements. *Cell*, *186*(15), 3291-3306.e21. https://doi.org/10.1016/j.cell.2023.06.007

Sergiev, P. V., Lavrik, I. N., Wlasoff, V. A., Dokudovskaya, S. S., Dontsova, O. A., Bogdanov, A. A., & Brimacombe, R. (1997). The path of mRNA through the bacterial ribosome: A site-directed crosslinking study using new photoreactive derivatives of guanosine and uridine. *RNA (New York, N.Y.)*, *3*(5), 464–475.

Sexton, A. N., Wang, P. Y., Rutenberg-Schoenberg, M., & Simon, M. D. (2017). Interpreting Reverse Transcriptase Termination and Mutation Events for Greater Insight into the Chemical Probing of RNA. *Biochemistry*, *56*(35), 4713–4721. https://doi.org/10.1021/acs.biochem.7b00323

Shatkin, A. (1976). Capping of eucaryotic mRNAs. *Cell*, *9*(4), 645–653. https://doi.org/10.1016/0092-8674(76)90128-8

Shatkin, A. J., & Manley, J. L. (2000). The ends of the affair: Capping and polyadenylation. *Nature Structural Biology*, *7*(10), 838–842. https://doi.org/10.1038/79583

Shin, J., Douglas, C. J., Zhang, S., Seath, C. P., & Bao, H. (2024). Targeting Recycling Endosomes to Potentiate mRNA Lipid Nanoparticles. *Nano Letters*, *24*(17), 5104–5109. https://doi.org/10.1021/acs.nanolett.3c04415

Shin, S., Lee, S.-Y., Kang, M.-G., Jang, D.-G., Kim, J., Rhee, H.-W., & Kim, J.-S. (2024). Super-resolution proximity labeling with enhanced direct identification of biotinylation sites. *Communications Biology*, *7*(1), 554. https://doi.org/10.1038/s42003-024-06112-w

Silva-Pilipich, N., Beloki, U., Salaberry, L., & Smerdou, C. (2024). Self-Amplifying RNA: A Second Revolution of mRNA Vaccines against COVID-19. *Vaccines*, *12*(3), 318. https://doi.org/10.3390/vaccines12030318

Smith, B. L., Gallie, D. R., Le, H., & Hansma, P. K. (1997). Visualization of Poly(A)-Binding Protein Complex Formation with Poly(A) RNA Using Atomic Force Microscopy. *Journal of Structural Biology*, *119*(2), 109–117. https://doi.org/10.1006/jsbi.1997.3864

Sonawane, N. D., Szoka, F. C., & Verkman, A. S. (2003). Chloride Accumulation and Swelling in Endosomes Enhances DNA Transfer by Polyamine-DNA Polyplexes. *Journal of Biological Chemistry*, *278*(45), 44826–44831. https://doi.org/10.1074/jbc.M308643200

Staring, J., Raaben, M., & Brummelkamp, T. R. (2018). Viral escape from endosomes and host detection at a glance. *Journal of Cell Science*. https://doi.org/10.1242/jcs.216259

Steensel, B. V., & Henikoff, S. (2000). Identification of in vivo DNA targets of chromatin proteins using tethered Dam methyltransferase. *Nature Biotechnology*, *18*(4), 424–428. https://doi.org/10.1038/74487

Stepinski, J., Waddell, C., Stolarski, R., Darzynkiewicz, E., & Rhoads, R. E. (2001). Synthesis and properties of mRNAs containing the novel 'anti-reverse' cap analogs 7-methyl(3'-O-methyl)GpppG and 7-methyl (3'-deoxy)GpppG. *RNA (New York, N.Y.)*, *7*(10), 1486–1495.

Sunshine, J. C., Peng, D. Y., & Green, J. J. (2012). Uptake and Transfection with Polymeric Nanoparticles Are Dependent on Polymer End-Group Structure, but Largely Independent of Nanoparticle Physical and Chemical Properties. *Molecular Pharmaceutics*, *9*(11), 3375–3383. https://doi.org/10.1021/mp3004176

Svitkin, Y. V., Cheng, Y. M., Chakraborty, T., Presnyak, V., John, M., & Sonenberg, N. (2017). N1-methyl-pseudouridine in mRNA enhances translation through eIF2α-dependent and independent mechanisms by increasing ribosome density. *Nucleic Acids Research*, *45*(10), 6023–6036. https://doi.org/10.1093/nar/gkx135

Tan, B., Peng, S., Yatim, S. M. J. M., Gunaratne, J., Hunziker, W., & Ludwig, A. (2020). An Optimized Protocol for Proximity Biotinylation in Confluent Epithelial Cell Cultures Using the Peroxidase APEX2. *STAR Protocols*, *1*(2), 100074. https://doi.org/10.1016/j.xpro.2020.100074

Tan, E., Chin, C. S. H., Lim, Z. F. S., & Ng, S. K. (2021). HEK293 Cell Line as a Platform to Produce Recombinant Proteins and Viral Vectors. *Frontiers in Bioengineering and Biotechnology*, *9*, 796991. https://doi.org/10.3389/fbioe.2021.796991

Tang, X., Zhang, Y., & Han, X. (2023). Ionizable Lipid Nanoparticles for mRNA Delivery. *Advanced NanoBiomed Research*, *3*(8), 2300006. https://doi.org/10.1002/anbr.202300006

Tatu, U., & Helenius, A. (1997). Interactions between Newly Synthesized Glycoproteins, Calnexin and a Network of Resident Chaperones in the Endoplasmic Reticulum. *The Journal of Cell Biology*, *136*(3), 555–565. https://doi.org/10.1083/jcb.136.3.555

Taverner, N. V., Smith, J. C., & Wardle, F. C. (2004). Identifying transcriptional targets. *Genome Biology*, *5*(3), 210. https://doi.org/10.1186/gb-2004-5-3-210

Tenchov, R., Bird, R., Curtze, A. E., & Zhou, Q. (2021). Lipid Nanoparticles–From Liposomes to mRNA Vaccine Delivery, a Landscape of Research Diversity and Advancement. *ACS Nano*, *15*(11), 16982–17015. https://doi.org/10.1021/acsnano.1c04996

Theos, A. C., Berson, J. F., Theos, S. C., Herman, K. E., Harper, D. C., Tenza, D., Sviderskaya, E. V., Lamoreux, M. L., Bennett, D. C., Raposo, G., & Marks, M. S. (2006). Dual Loss of ER Export and Endocytic Signals with Altered Melanosome Morphology in the *silver* Mutation of Pmel17. *Molecular Biology of the Cell*, *17*(8), 3598–3612. https://doi.org/10.1091/mbc.e06-01-0081

Thompson, J. E., Venegas, F. D., & Raines, R. T. (1994). Energetics of Catalysis by Ribonucleases: Fate of the 2',3'-Cyclic Phosphodiester Intermediate. *Biochemistry*, *33*(23), 7408–7414. https://doi.org/10.1021/bi00189a047

Tian, Y., Deng, Z., & Yang, P. (2022). mRNA vaccines: A novel weapon to control infectious diseases. *Frontiers in Microbiology*, *13*, 1008684. https://doi.org/10.3389/fmicb.2022.1008684

Trepotec, Z., Geiger, J., Plank, C., Aneja, M. K., & Rudolph, C. (2019). Segmented poly(A) tails significantly reduce recombination of plasmid DNA without affecting mRNA translation efficiency or half-life. *RNA*, *25*(4), 507–518. https://doi.org/10.1261/rna.069286.118

Tuller, T., Carmi, A., Vestsigian, K., Navon, S., Dorfan, Y., Zaborske, J., Pan, T., Dahan, O., Furman, I., & Pilpel, Y. (2010). An Evolutionarily Conserved Mechanism for Controlling the Efficiency of Protein Translation. *Cell*, *141*(2), 344–354. https://doi.org/10.1016/j.cell.2010.03.031

Usman, W. M., Pham, T. C., Kwok, Y. Y., Vu, L. T., Ma, V., Peng, B., Chan, Y. S., Wei, L., Chin, S. M., Azad, A., He, A. B.-L., Leung, A. Y. H., Yang, M., Shyh-Chang, N., Cho, W. C., Shi, J., & Le, M. T. N. (2018). Efficient RNA drug delivery using red blood cell

extracellular vesicles. *Nature Communications*, *9*(1), 2359. https://doi.org/10.1038/s41467-018-04791-8

van Stipdonk, M. J. B., Badia-Martinez, D., Sluijter, M., Offringa, R., van Hall, T., & Achour, A. (2009). Design of Agonistic Altered Peptides for the Robust Induction of CTL Directed towards H-2Db in Complex with the Melanoma-Associated Epitope gp100. *Cancer Research*, *69*(19), 7784–7792. https://doi.org/10.1158/0008-5472.CAN-09-1724

Verbeke, R., Hogan, M. J., Loré, K., & Pardi, N. (2022). Innate immune mechanisms of mRNA vaccines. *Immunity*, *55*(11), 1993–2005. https://doi.org/10.1016/j.immuni.2022.10.014

Vermeulen, L. M. P., De Smedt, S. C., Remaut, K., & Braeckmans, K. (2018). The proton sponge hypothesis: Fable or fact? *European Journal of Pharmaceutics and Biopharmaceutics*, *129*, 184–190. https://doi.org/10.1016/j.ejpb.2018.05.034

Victor, M. P., Acharya, D., Begum, T., & Ghosh, T. C. (2019). The optimization of mRNA expression level by its intrinsic properties—Insights from codon usage pattern and structural stability of mRNA. *Genomics*, *111*(6), Article 6. https://doi.org/10.1016/j.ygeno.2018.08.009

Vogel, A. B., Kanevsky, I., Che, Y., Swanson, K. A., Muik, A., Vormehr, M., Kranz, L. M., Walzer, K. C., Hein, S., Güler, A., Loschko, J., Maddur, M. S., Ota-Setlik, A., Tompkins, K., Cole, J., Lui, B. G., Ziegenhals, T., Plaschke, A., Eisel, D., … Sahin, U. (2021). BNT162b vaccines protect rhesus macaques from SARS-CoV-2. *Nature*, *592*(7853), 283–289. https://doi.org/10.1038/s41586-021-03275-y

von der Haar, T., Mulroney, T. E., Hedayioglu, F., Kurusamy, S., Rust, M., Lilley, K. S., Thaventhiran, J. E., Willis, A. E., & Smales, C. M. (2023). Translation of in vitro-transcribed RNA therapeutics. *Frontiers in Molecular Biosciences*, *10*. https://doi.org/10.3389/fmolb.2023.1128067

Wachter, A. (2014). Gene regulation by structured mRNA elements. *Trends in Genetics*, *30*(5), 172–181. https://doi.org/10.1016/j.tig.2014.03.001

Wang, J., Yang, S., & Zhang, Y. (2020). One-electron oxidation and redox potential of nucleobases and deoxyribonucleosides computed by QM/MM simulations. *Chemical Physics Letters*, *739*, 136948. https://doi.org/10.1016/j.cplett.2019.136948

Warminski, M., Trepkowska, E., Smietanski, M., Sikorski, P. J., Baranowski, M. R., Bednarczyk, M., Kedzierska, H., Majewski, B., Mamot, A., Papiernik, D., Popielec, A., Serwa, R. A., Shimanski, B. A., Sklepkiewicz, P., Sklucka, M., Sokolowska, O., Spiewla, T., Toczydlowska-Socha, D., Warminska, Z., … Kowalska, J. (2024). Trinucleotide mRNA Cap Analogue *N* 6-Benzylated at the Site of Posttranscriptional $^{m6}A_m$ Mark Facilitates mRNA Purification and Confers Superior Translational Properties In Vitro and In Vivo. *Journal of the American Chemical Society*, *146*(12), 8149–8163. https://doi.org/10.1021/jacs.3c12629

Wayment-Steele, H. K., Kim, D. S., Choe, C. A., Nicol, J. J., Wellington-Oguri, R., Watkins, A. M., Parra Sperberg, R. A., Huang, P.-S., Participants, E., & Das, R. (2021). Theoretical basis for stabilizing messenger RNA through secondary structure design. *Nucleic Acids Research*, *49*(18), 10604–10617. https://doi.org/10.1093/nar/gkab764

Weissman, D. (2015). mRNA transcript therapy. *Expert Review of Vaccines*, *14*(2), 265–281. https://doi.org/10.1586/14760584.2015.973859

Wells, S. E., Hillner, P. E., Vale, R. D., & Sachs, A. B. (1998). Circularization of mRNA by Eukaryotic Translation Initiation Factors. *Molecular Cell*, *2*(1), 135–140. https://doi.org/10.1016/S1097-2765(00)80122-7

Weng, Q., Wang, Y., Xie, Y., Yu, X., Zhang, S., Ge, J., Li, Z., Ye, G., & Guo, J. (2022). Extracellular vesicles-associated tRNA-derived fragments (tRFs): Biogenesis, biological functions, and their role as potential biomarkers in human diseases. *Journal of Molecular Medicine (Berlin, Germany)*, *100*(5), 679–695. https://doi.org/10.1007/s00109-022-02189-0

Westerich, K. J., Chandrasekaran, K. S., Gross-Thebing, T., Kueck, N., Raz, E., & Rentmeister, A. (2020). Bioorthogonal mRNA labeling at the poly(A) tail for imaging localization and dynamics in live zebrafish embryos. *Chemical Science*, *11*(11). https://doi.org/10.1039/c9sc05981d

Winz, M.-L., Samanta, A., Benzinger, D., & Jäschke, A. (2012). Site-specific terminal and internal labeling of RNA by poly(A) polymerase tailing and copper-catalyzed or copper-free strain-promoted click chemistry. *Nucleic Acids Research*, *40*(10), e78–e78. https://doi.org/10.1093/nar/gks062

Wittrup, A., Ai, A., Liu, X., Hamar, P., Trifonova, R., Charisse, K., Manoharan, M., Kirchhausen, T., & Lieberman, J. (2015). Visualizing lipid-formulated siRNA release from endosomes and target gene knockdown. *Nature Biotechnology*. https://doi.org/10.1038/nbt.3298

Wolff, J. A., Malone, R. W., Williams, P., Chong, W., Acsadi, G., Jani, A., & Felgner, P. L. (1990). Direct Gene Transfer into Mouse Muscle in Vivo. *Science*, *247*(4949), 1465–1468. https://doi.org/10.1126/science.1690918

Woods, N.-B. (2003). Lentiviral vector transduction of NOD/SCID repopulating cells results in multiple vector integrations per transduced cell: Risk of insertional mutagenesis. *Blood*, *101*(4), 1284–1289. https://doi.org/10.1182/blood-2002-07-2238

Wu, J. J., & Palazzo, A. F. (2020). Visualization of Endoplasmic Reticulum-Associated mRNA in Mammalian Cells. In M. Heinlein (Ed.), *RNA Tagging* (Vol. 2166, pp. 35–49). Springer US. https://doi.org/10.1007/978-1-0716-0712-1_3

Wu, M. Z., Asahara, H., Tzertzinis, G., & Roy, B. (2020). Synthesis of low immunogenicity RNA with high-temperature in vitro transcription. *RNA*, *26*(3), 345–360. https://doi.org/10.1261/rna.073858.119

Xia, X. (2021). Detailed Dissection and Critical Evaluation of the Pfizer/BioNTech and Moderna mRNA Vaccines. *Vaccines*, *9*(7), 734. https://doi.org/10.3390/vaccines9070734

Xie, Y., & Ren, Y. (2019). Mechanisms of nuclear mRNA export: A structural perspective. *Traffic*, *20*(11), 829–840. https://doi.org/10.1111/tra.12691

Yamashita, T., Takahashi, Y., & Takakura, Y. (2018). Possibility of Exosome-Based Therapeutics and Challenges in Production of Exosomes Eligible for Therapeutic

Application. *Biological and Pharmaceutical Bulletin*, *41*(6), 835–842. https://doi.org/10.1248/bpb.b18-00133

Yanez Arteta, M., Kjellman, T., Bartesaghi, S., Wallin, S., Wu, X., Kvist, A. J., Dabkowska, A., Székely, N., Radulescu, A., Bergenholtz, J., & Lindfors, L. (2018). Successful reprogramming of cellular protein production through mRNA delivered by functionalized lipid nanoparticles. *Proceedings of the National Academy of Sciences of the United States of America*, *115*(15), E3351–E3360. https://doi.org/10.1073/pnas.1720542115

Yang, J.-R., Chen, X., & Zhang, J. (2014). Codon-by-Codon Modulation of Translational Speed and Accuracy Via mRNA Folding. *PLoS Biology*, *12*(7), e1001910. https://doi.org/10.1371/journal.pbio.1001910

Yang, M., Lu, Y., Piao, W., & Jin, H. (2022). The Translational Regulation in mTOR Pathway. *Biomolecules*, *12*(6), 802. https://doi.org/10.3390/biom12060802

Yang, Y., Fan, X., Mao, M., Song, X., Wu, P., Zhang, Y., Jin, Y., Yang, Y., Chen, L.-L., Wong, C. C., Xiao, X., & Wang, Z. (2017). Extensive translation of circular RNAs driven by N6-methyladenosine. *Cell Res.*, *27*(3), 626-641. https://doi.org/10.1038/cr.2017.31

Yang, Y., & Wang, Z. (2019). IRES-mediated cap-independent translation, a path leading to hidden proteome. *Journal of Molecular Cell Biology*, *11*(10), 911–919. https://doi.org/10.1093/jmcb/mjz091

Yang, Z., Shi, J., Xie, J., Wang, Y., Sun, J., Liu, T., Zhao, Y., Zhao, X., Wang, X., Ma, Y., Malkoc, V., Chiang, C., Deng, W., Chen, Y., Fu, Y., Kwak, K. J., Fan, Y., Kang, C., Yin, C., … Lee, L. J. (2019). Large-scale generation of functional mRNA-encapsulating exosomes via cellular nanoporation. *Nature Biomedical Engineering*, *4*(1), 69–83. https://doi.org/10.1038/s41551-019-0485-1

Young, A. P., Jackson, D. J., & Wyeth, R. C. (2020). A technical review and guide to RNA fluorescence in situ hybridization. *PeerJ*, *8*, e8806. https://doi.org/10.7717/peerj.8806

Yousefi Adlsadabad, S., Hanrahan, J. W., & Kakkar, A. (2024). mRNA Delivery: Challenges and Advances through Polymeric Soft Nanoparticles. *International Journal of Molecular Sciences*, *25*(3), 1739. https://doi.org/10.3390/ijms25031739

Yu, C.-H., Dang, Y., Zhou, Z., Wu, C., Zhao, F., Sachs, M. S., & Liu, Y. (2015). Codon Usage Influences the Local Rate of Translation Elongation to Regulate Co-translational Protein Folding. *Molecular Cell*, *59*(5), 744–754. https://doi.org/10.1016/j.molcel.2015.07.018

Zhang, H., Zhang, L., Lin, A., Xu, C., Li, Z., Liu, K., Liu, B., Ma, X., Zhao, F., Jiang, H., Chen, C., Shen, H., Li, H., Mathews, D. H., Zhang, Y., & Huang, L. (2023). Algorithm for optimized mRNA design improves stability and immunogenicity. *Nature*, *621*(7978), 396–403. https://doi.org/10.1038/s41586-023-06127-z

Zhang, T., Yin, H., Li, Y., Yang, H., Ge, K., Zhang, J., Yuan, Q., Dai, X., Naeem, A., Weng, Y., Huang, Y., & Liang, X.-J. (2024). Optimized lipid nanoparticles (LNPs) for organ-selective nucleic acids delivery in vivo. *iScience*, *27*(6), 109804. https://doi.org/10.1016/j.isci.2024.109804

Zhang, Y., Zhai, S., Huang, H., Qin, S., Sun, M., Chen, Y., Lan, X., Li, G., Huang, Z., Wang, D., Luo, Y., Xiao, W., Li, H., He, X., Chen, M., Peng, X., & Song, X. (2024). Efficient signal sequence of mRNA vaccines enhances the antigen expression to expand the immune protection against viral infection. *Journal of Nanobiotechnology*, *22*(1), 295. https://doi.org/10.1186/s12951-024-02488-3

Zhang, Z., Fu, Y., Ju, X., Zhang, F., Zhang, P., & He, M. (2024). Advances in Engineering Circular RNA Vaccines. *Pathogens*, *13*(8), 692. https://doi.org/10.3390/pathogens13080692

Zhao, J., Hyman, L., & Moore, C. (1999). Formation of mRNA 3' Ends in Eukaryotes: Mechanism, Regulation, and Interrelationships with Other Steps in mRNA Synthesis. *Microbiology and Molecular Biology Reviews*, *63*(2), 405–445. https://doi.org/10.1128/MMBR.63.2.405-445.1999

Zhao, X., Zhang, Y., Hang, D., Meng, J., & Wei, Z. (2022). Detecting RNA modification using direct RNA sequencing: A systematic review. *Computational and Structural Biotechnology Journal*, *20*, 5740–5749. https://doi.org/10.1016/j.csbj.2022.10.023

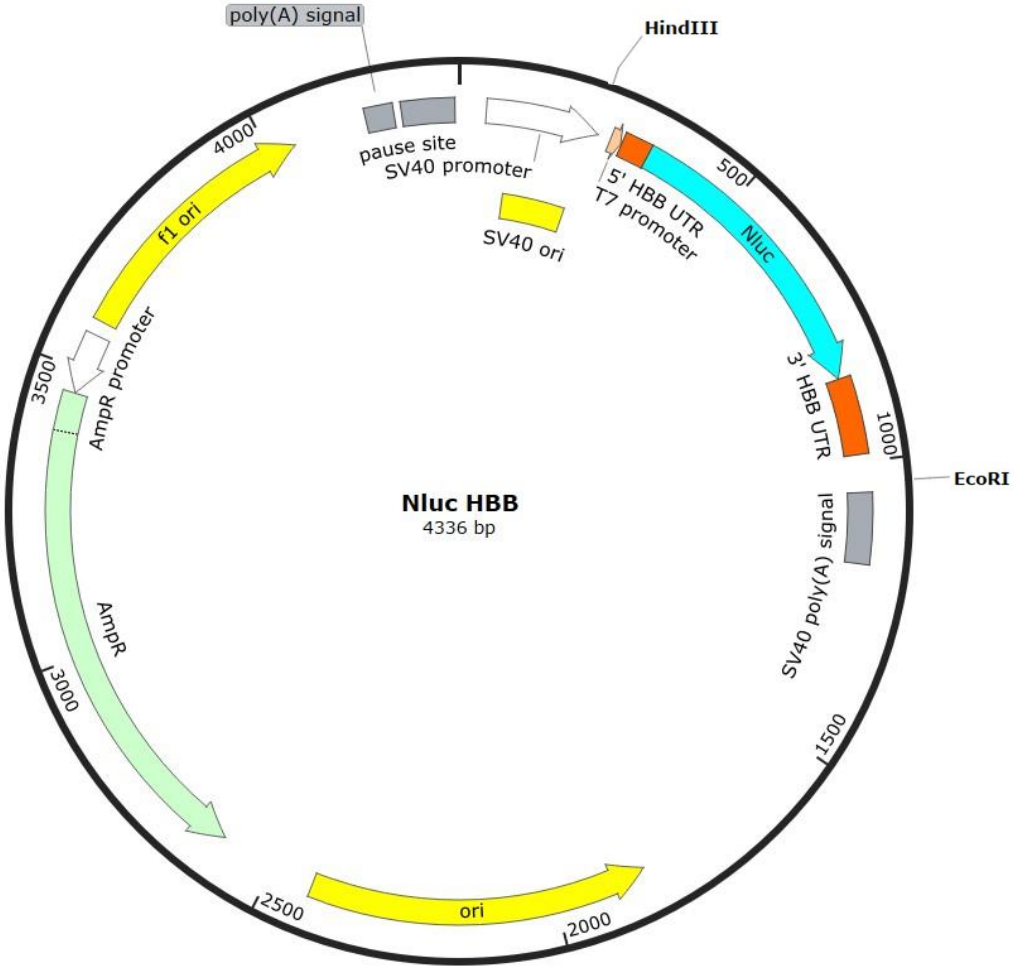
Zhou, X., Liao, W.-J., Liao, J.-M., Liao, P., & Lu, H. (2015). Ribosomal proteins: Functions beyond the ribosome. *Journal of Molecular Cell Biology*, *7*(2), 92–104. https://doi.org/10.1093/jmcb/mjv014

Zhou, Y., Wang, G., Wang, P., Li, Z., Yue, T., Wang, J., & Zou, P. (2019). Expanding APEX2 Substrates for Proximity-Dependent Labeling of Nucleic Acids and Proteins in Living Cells. *Angewandte Chemie - International Edition*. https://doi.org/10.1002/anie.201905949

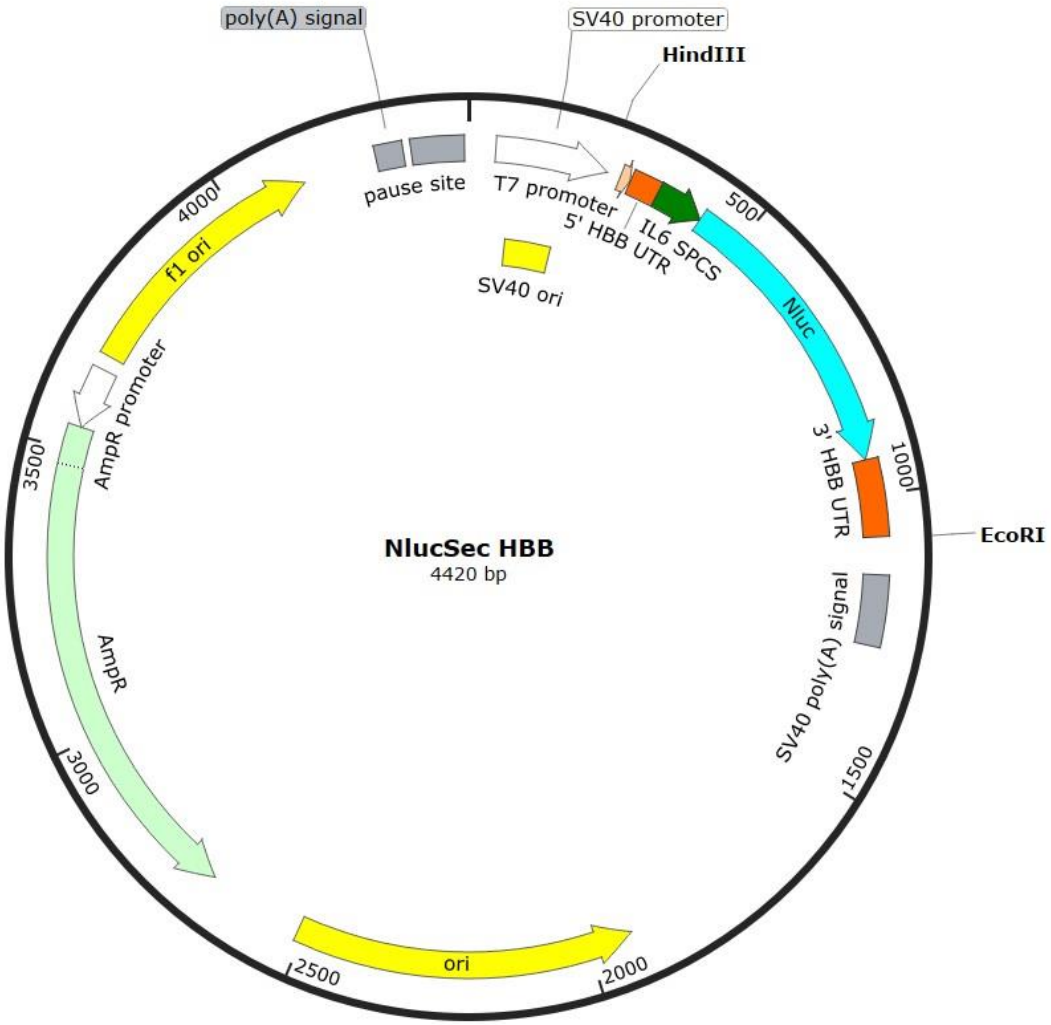
Ziegenhals, T., Frieling, R., Wolf, P., Göbel, K., Koch, S., Lohmann, M., Baiersdörfer, M., Fesser, S., Sahin, U., & Kuhn, A. N. (2023). Formation of dsRNA by-products during in vitro transcription can be reduced by using low steady-state levels of UTP. *Frontiers in Molecular Biosciences*, *10*, 1291045. https://doi.org/10.3389/fmolb.2023.1291045

Zuker, M. (1989). On finding all suboptimal foldings of an RNA molecule. *Science (New York, N.Y.)*, *244*(4900), 48–52. https://doi.org/10.1126/science.2468181

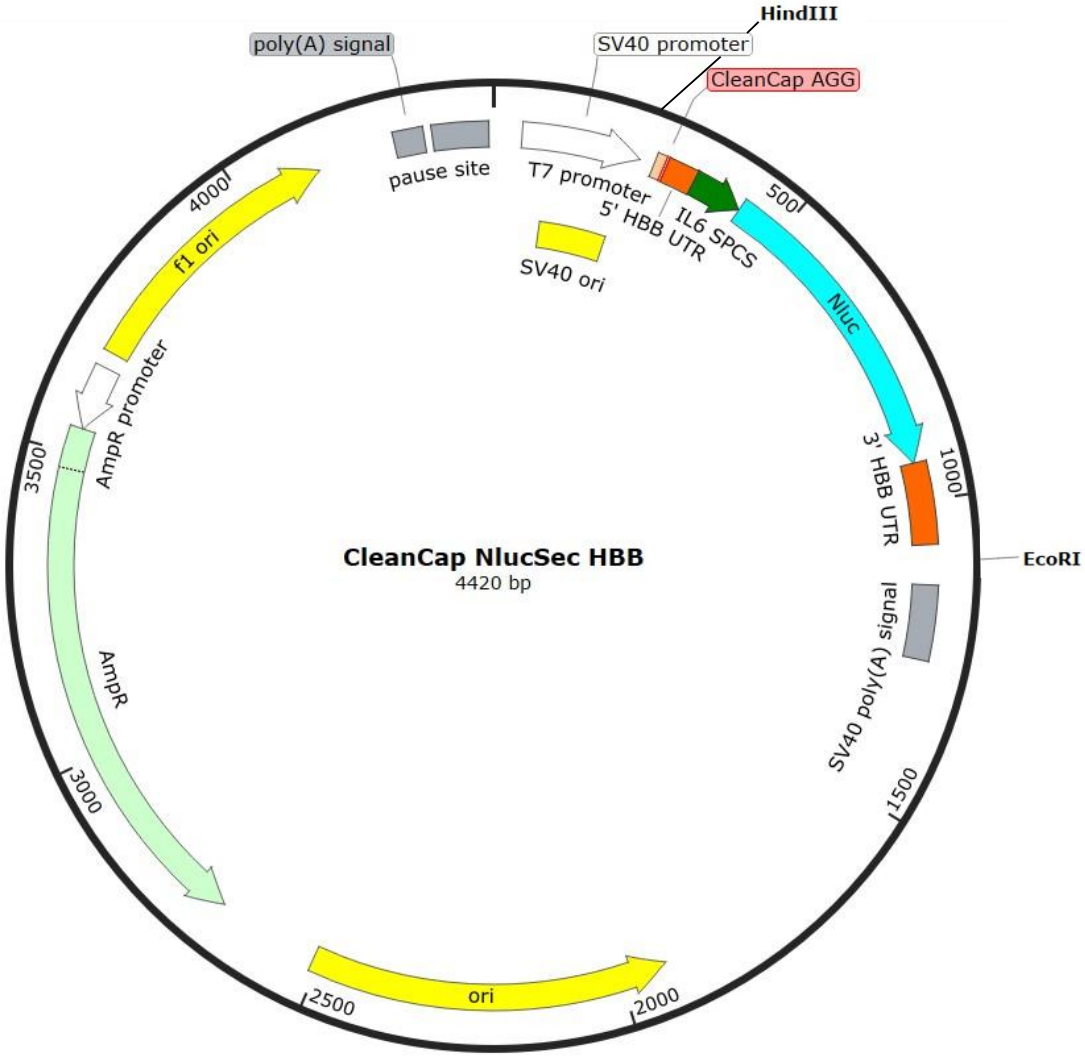
Chapter 8 - Supplementary Information



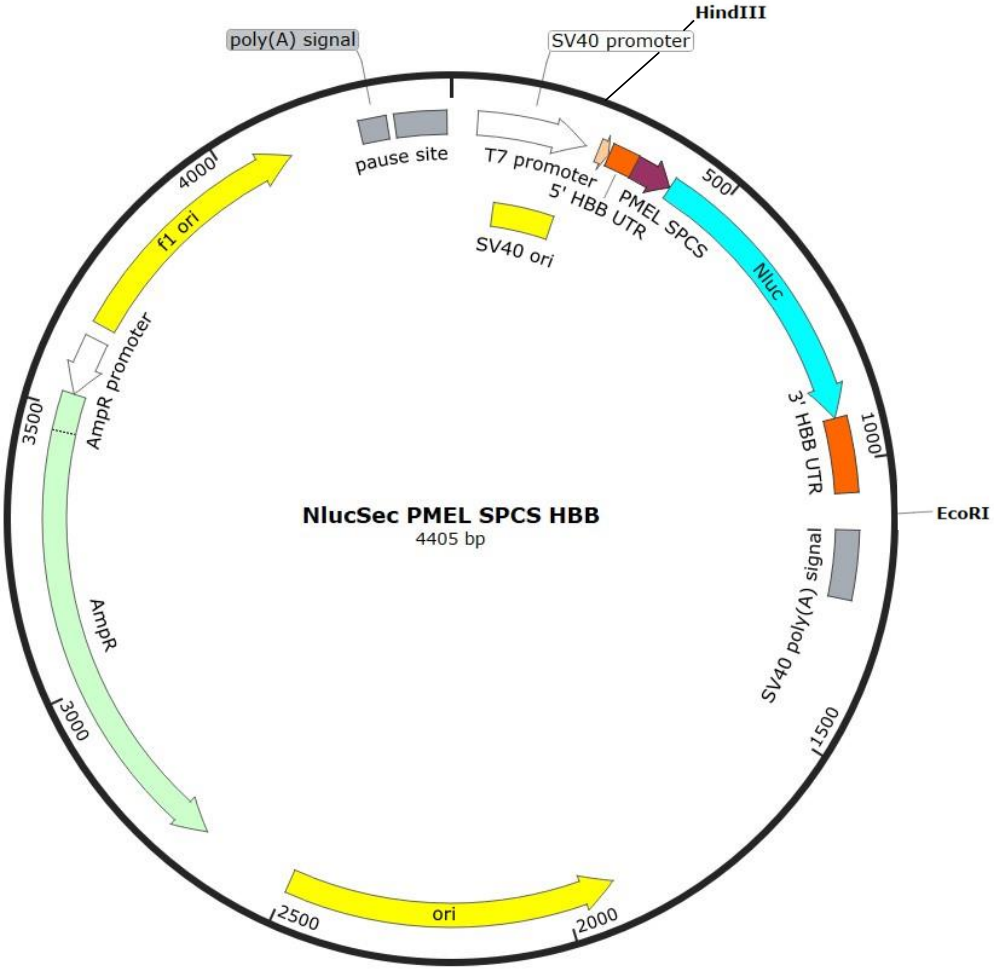
Supplementary Figure 1. Plasmid map of Nluc HBB. Encodes mRNA referred to in the main text as Nluc mRNA. Shows the location of the T7 promoter (Apricot), Nanoluciferase coding region (Cyan), human haemoglobin β subunit (HBB) UTRs (Orange), and restriction sites for HindIII and EcoRI.



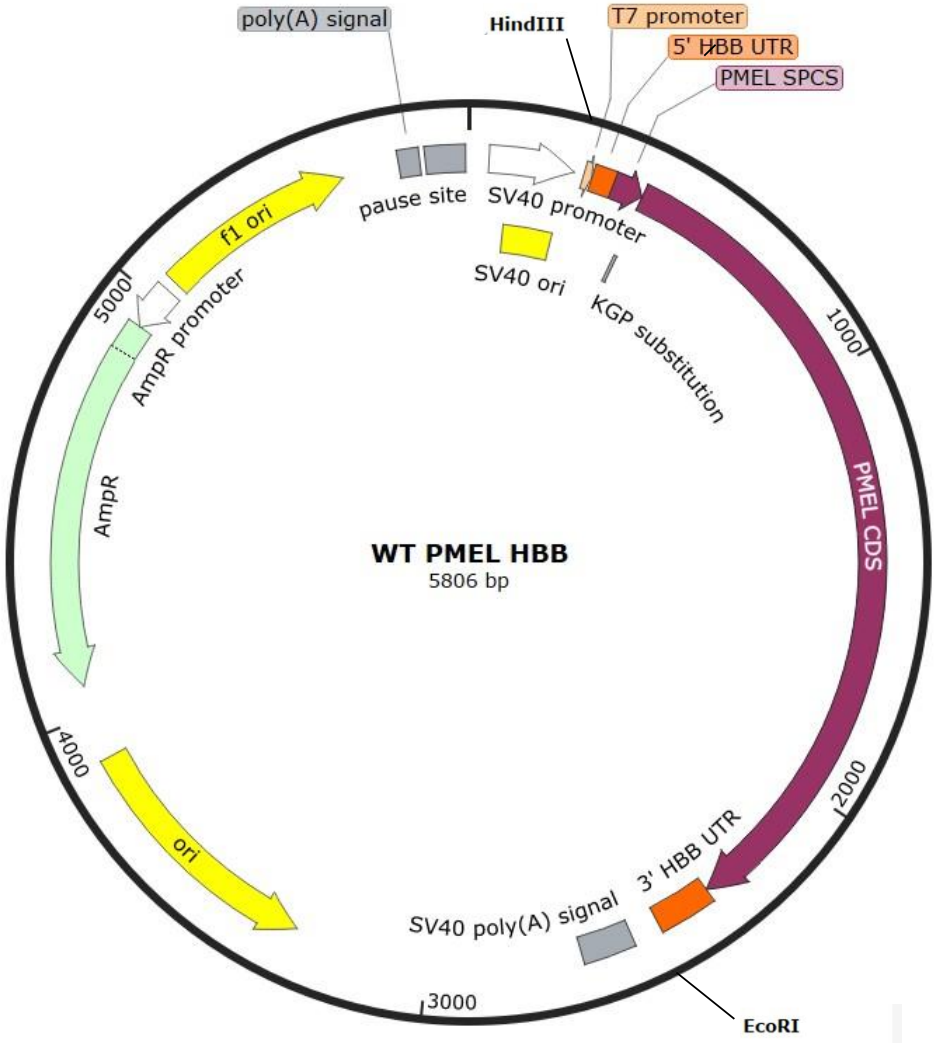
Supplementary Figure 2. Plasmid map of NlucSec HBB. Encodes mRNA referred to in the main text as NlucSec mRNA. Shows the location of the T7 promoter (Apricot), Interleukin-6 SPCS (Green), Nanoluciferase coding region (Cyan), HBB UTRs (Orange), and restriction sites for HindIII and EcoRI.



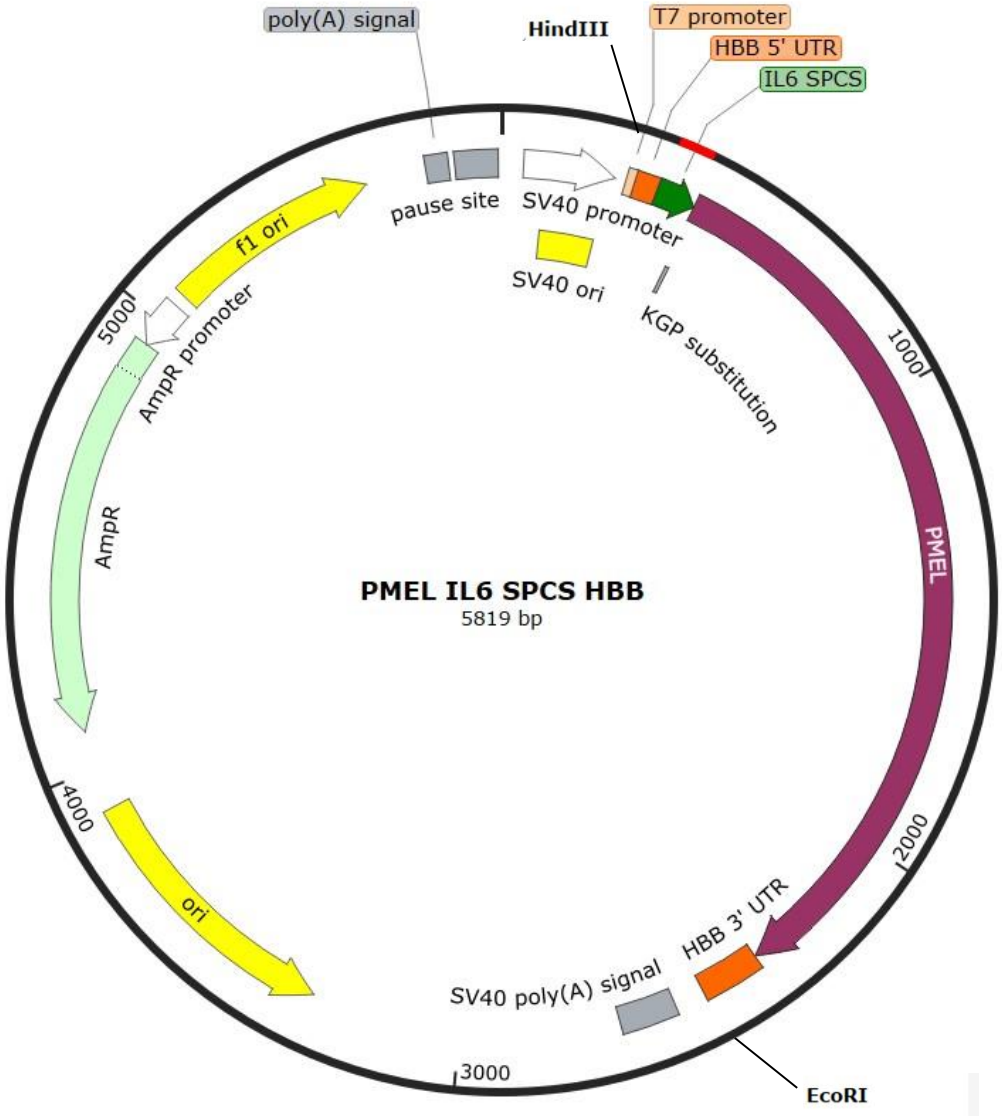
Supplementary Figure 3. Plasmid map of CleanCap NlucSec HBB. Encodes mRNA referred to in the main text as CleanCapped NlucSec mRNA. Shows the location of the T7 promoter (Apricot), CleanCap AGG sequence downstream of T7 promoter (Red), Interleukin-6 SPCS (Green), Nanoluciferase coding region (Cyan), HBB UTRs (Orange), and restriction sites for HindIII and EcoRI.



Supplementary Figure 4. Plasmid map of Nluc PMEL SPCS HBB. Encodes mRNA referred to in the main text as NlucSec PMEL SPCS mRNA. Shows the location of the T7 promoter (Apricot), PMEL SPCS (Purple), Nanoluciferase coding region (Cyan), HBB UTRs (Orange), and restriction sites for HindIII and EcoRI.

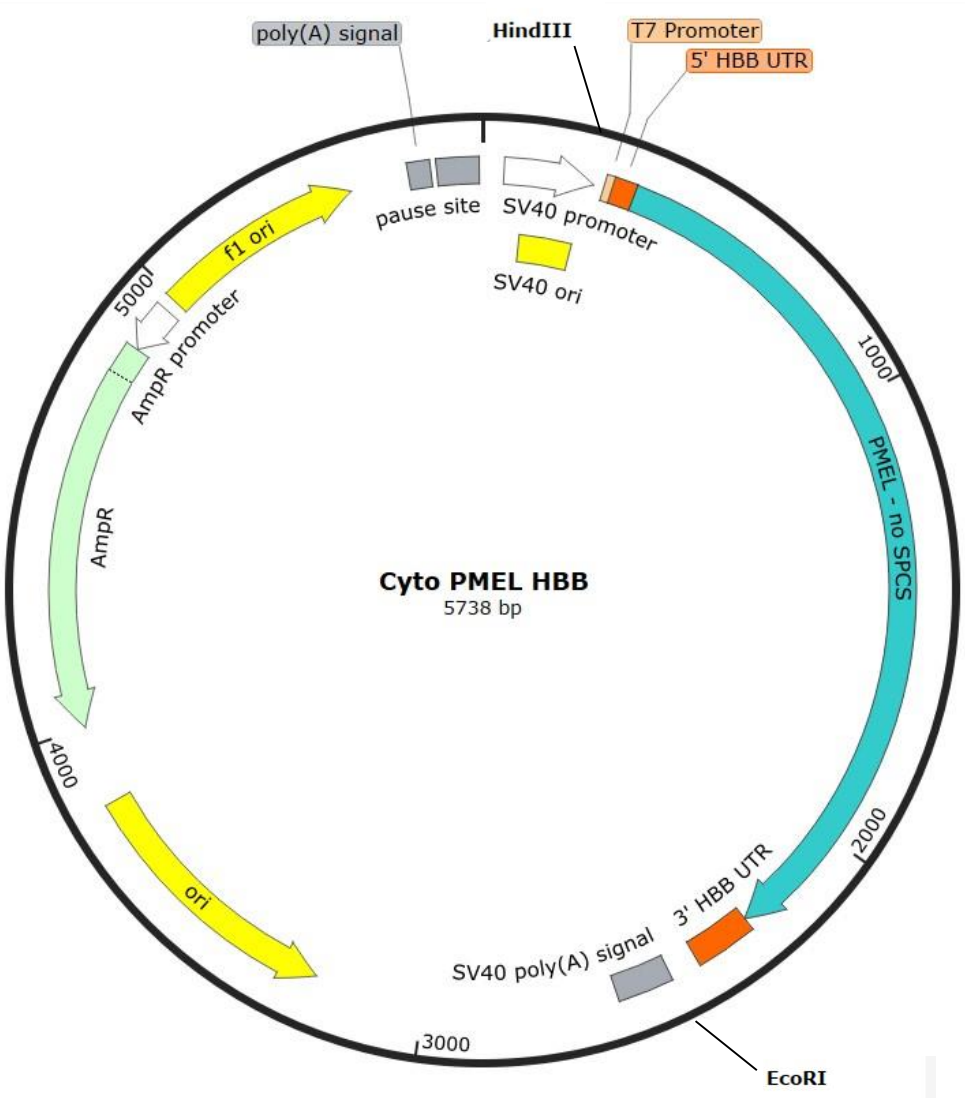


Supplementary Figure 5. Plasmid map of Human WT PMEL HBB. Encodes mRNA referred to in the main text as WT PMEL mRNA. Shows the location of the T7 promoter (Apricot), PMEL SPCS (Purple), PMEL coding region (Purple), HBB UTRs (Orange), and restriction sites for HindIII and EcoRI.

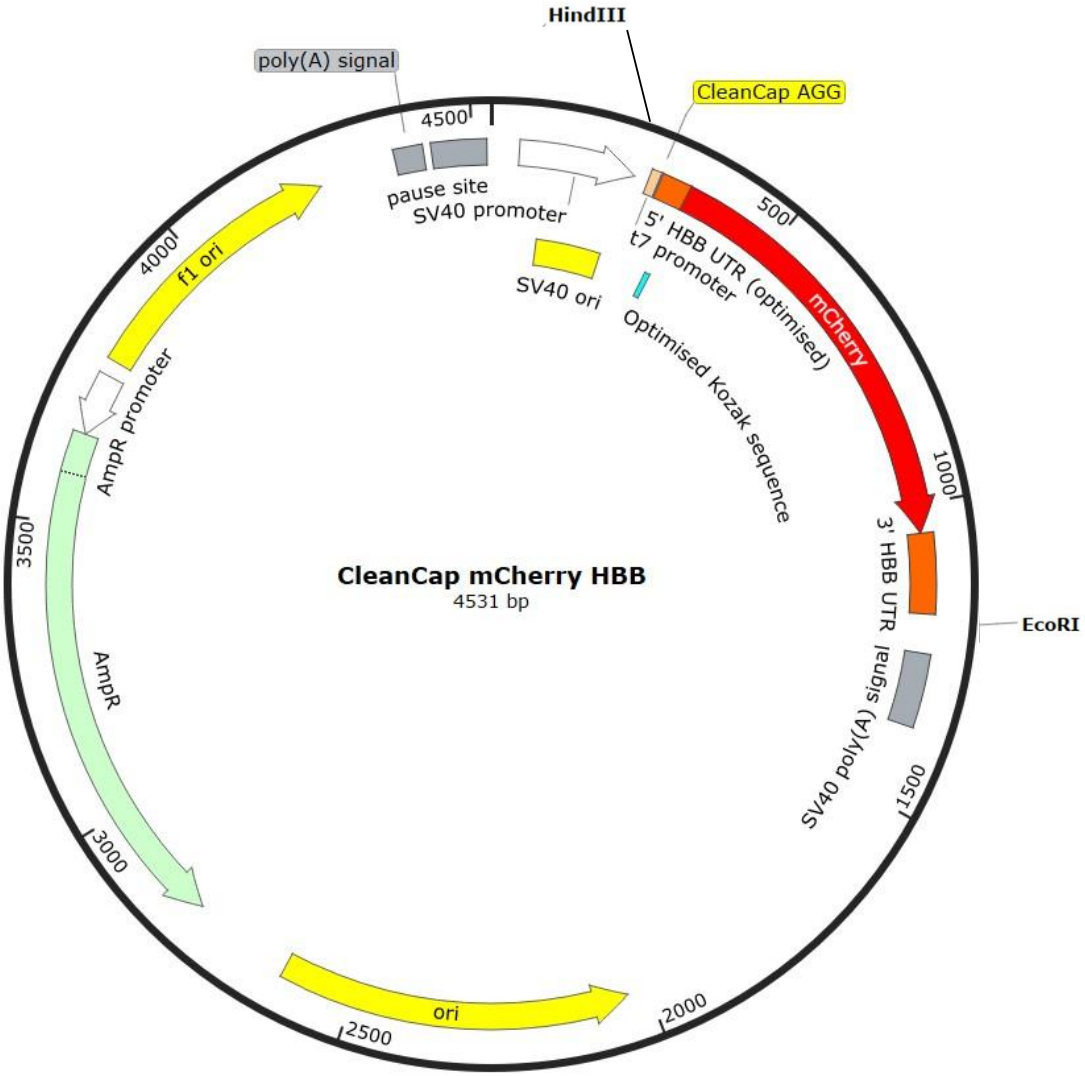


Supplementary Figure 6. Plasmid map of Human PMEL IL6 SPCS HBB. Encodes mRNA referred to in the main text as PMELIL6 SPCS mRNA. Shows the location of the T7 promoter (Apricot), IL6 SPCS (Green), PMEL coding region (Purple), HBB UTRs (Orange), and restriction sites for HindIII and EcoRI.

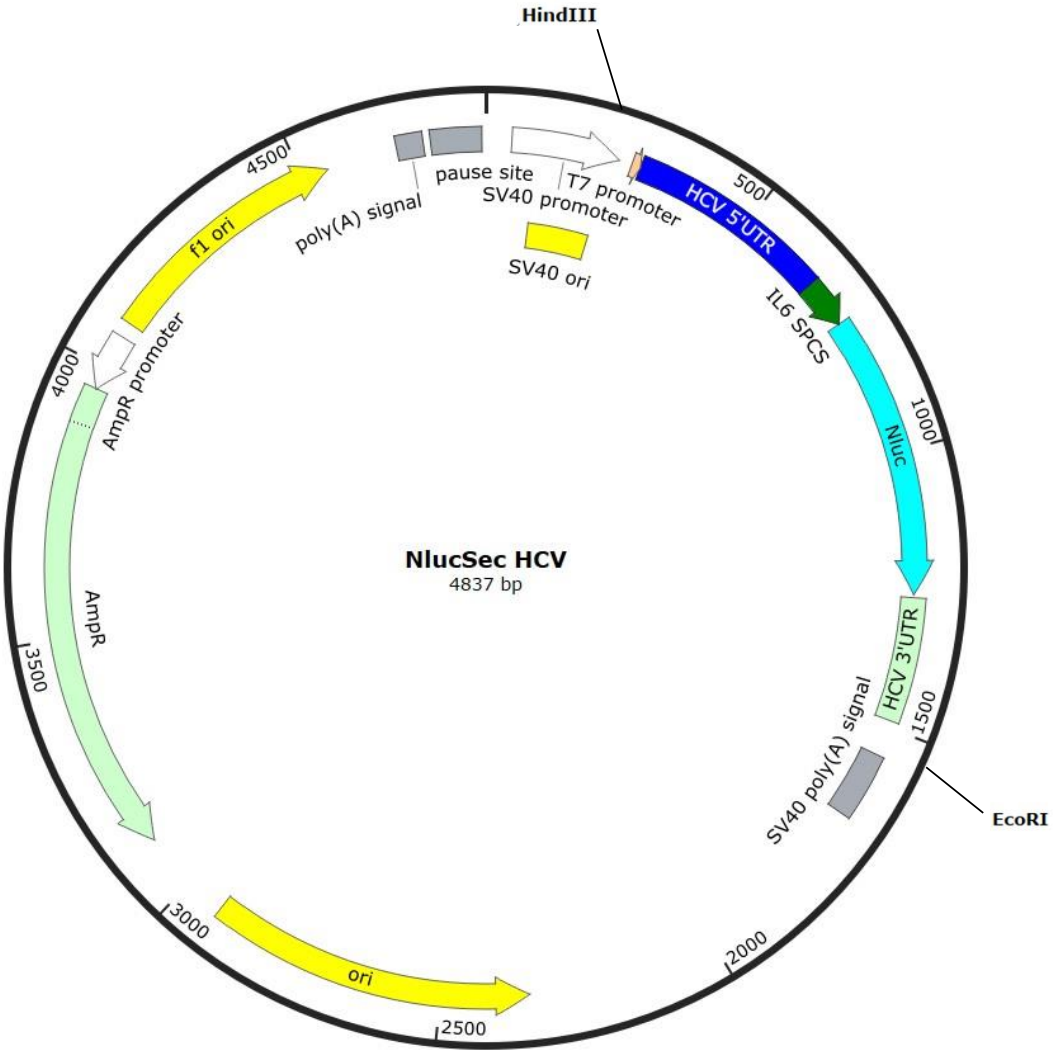
.



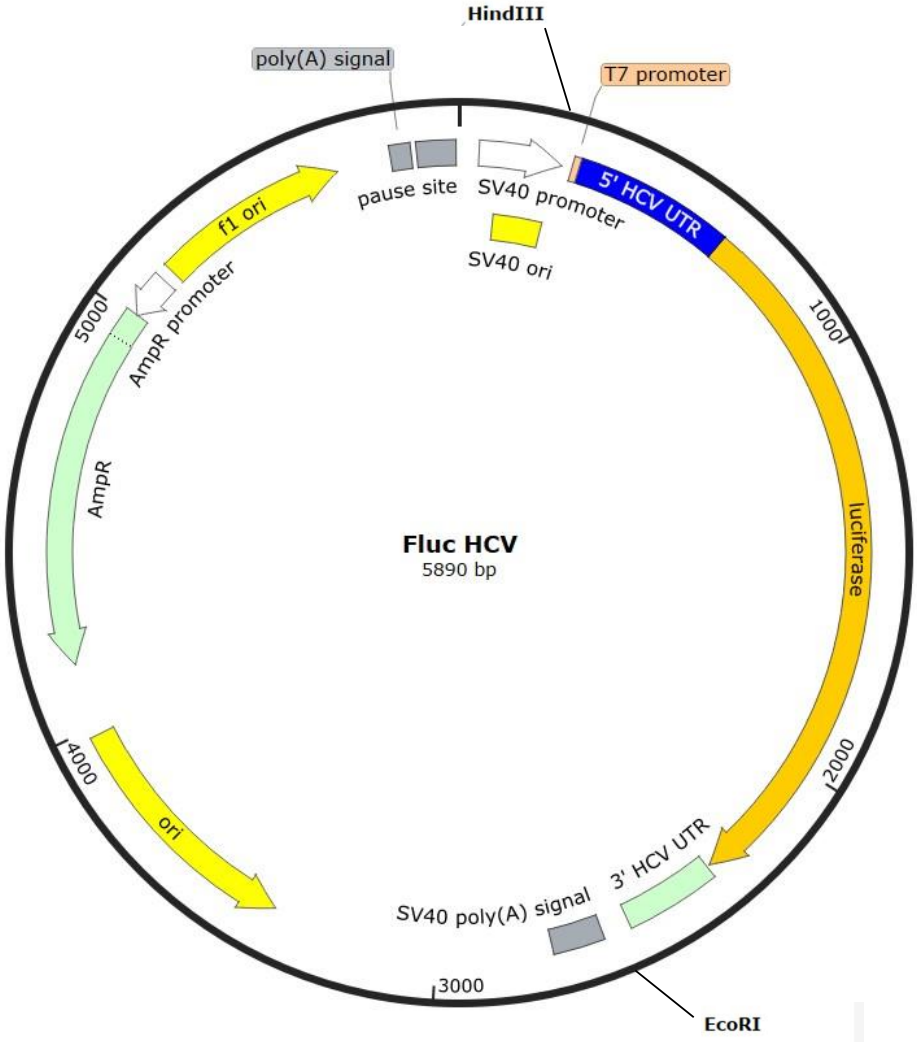
Supplementary Figure 7. Plasmid map of Cyto PMEL HBB. Encodes mRNA referred to in the main text as Cyto PMEL mRNA. Shows the location of the T7 promoter (Apricot), PMEL coding region excluding the native SPCS (Teal), HBB UTRs (Orange), and restriction sites for HindIII and EcoRI.



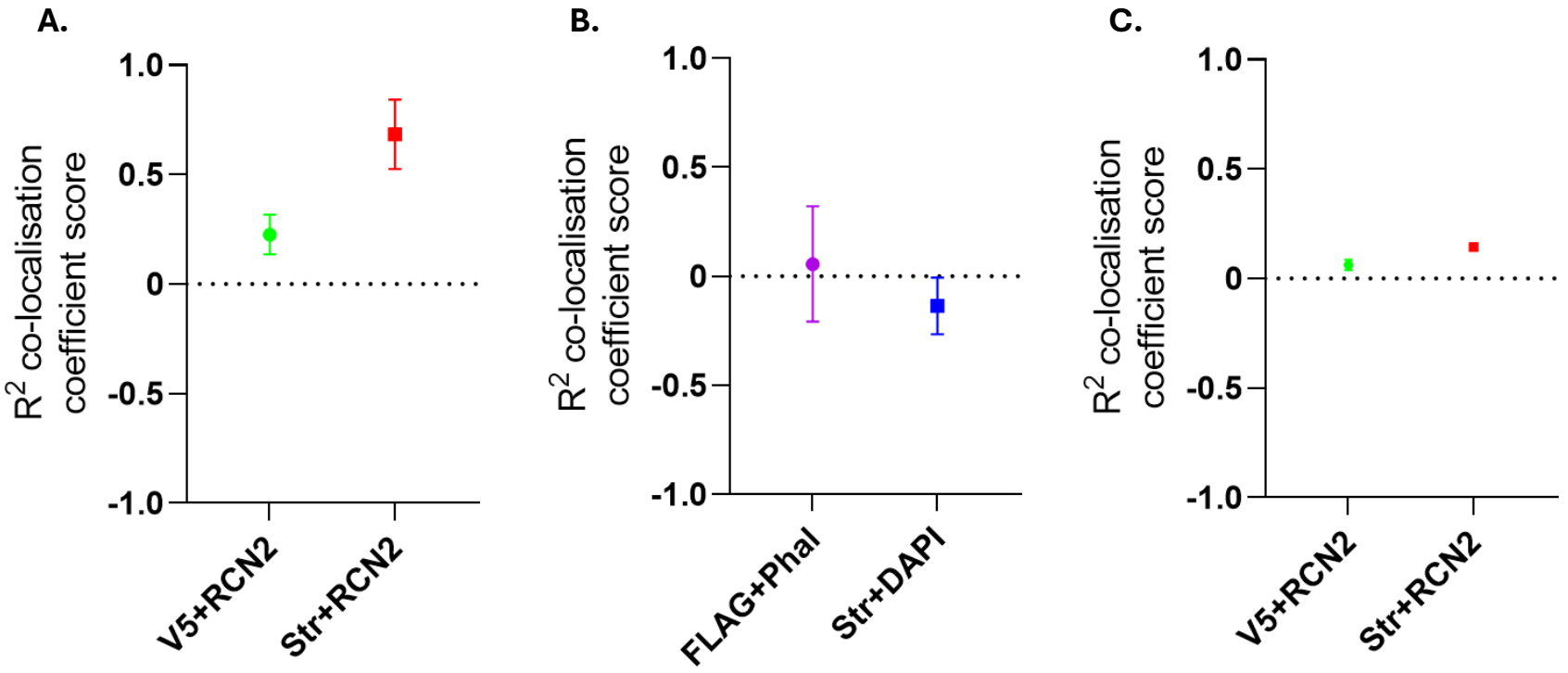
Supplementary Figure 8. Plasmid map of CleanCap mCherry HBB. Encodes mRNA referred to in the main text as mCherry mRNA. Shows the location of the T7 promoter (Apricot), CleanCap AGG sequence downstream of T7 promoter (Yellow), mCherry coding region (Red), HBB UTRs (Orange), and restriction sites for HindIII and EcoRI.



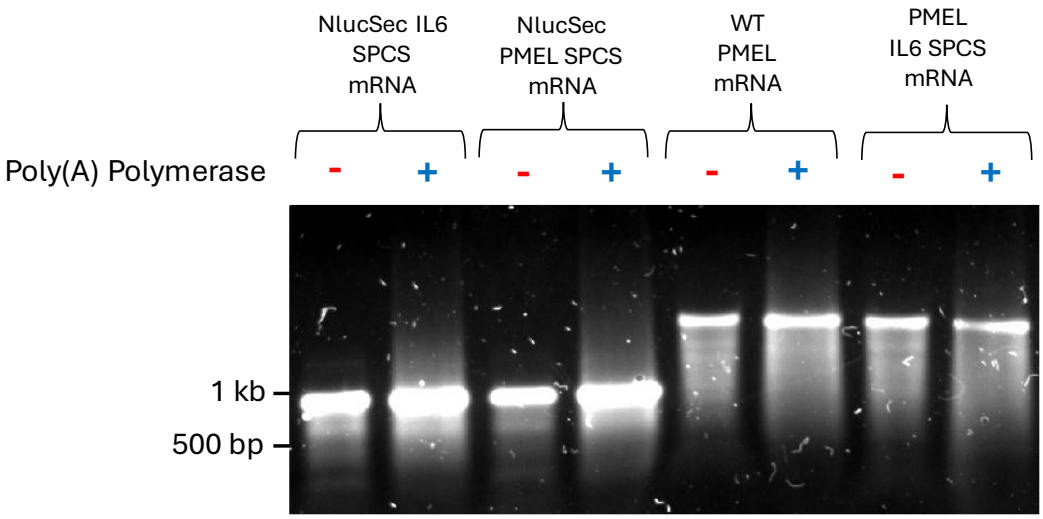
Supplementary Figure 9. Plasmid map of NlucSec HCV. Encodes mRNA referred to in the main text as NlucSec HCV mRNA. Shows the location of the T7 promoter (Apricot), Hepatitis C Virus 5' UTR (Blue), Interleukin-6 SPCS (Green), Nanoluciferase coding region (Cyan), Hepatitis C Virus 3' UTR (Green), and restriction sites for HindIII and EcoRI.



Supplementary Figure 10. Plasmid map of Fluc HCV. Encodes mRNA referred to in the main text as Fluc RNA. Shows the location of the T7 promoter (Apricot), Hepatitis C Virus 5' UTR (Blue), Firefly luciferase coding region (Yellow), Hepatitis C Virus 3' UTR (Green), and restriction sites for HindIII and EcoRI.



Supplementary Figure 11. Pearson's (R^2) colocalisation coefficient scores from SIM experiments in Chapter 3. Analysis was performed on manually delineated objects from images taken at the mid-section of the nucleus. (A) APEX2-ERM HEK293T cells: V5+RCN2 indicates APEX2-ERM's V5 epitope colocalisation with RCN2 as an ER-membrane marker, and Str+RCN2 indicates colocalisation of APEX2 enzymatic activity with the ER membrane. (B) APEX2-NES HEK293T cells: FLAG+Phal indicates APEX2-NES's FLAG epitope colocalisation with phalloidin as a cytosolic marker, and Str+DAPI indicates colocalisation of APEX2-NES enzymatic activity with DAPI as a nuclear stain. (C) APEX2-ERM A549 cells: V5+RCN2 indicates APEX2-ERM's V5 epitope colocalisation with RCN2 as an ER-membrane marker, and Str+RCN2 indicates colocalisation of APEX2 enzymatic activity with the ER membrane.



Supplementary Figure 12. Non-denaturing agarose gel showing *in vitro* transcribed (IVT) RNA before and after enzymatic poly(A) tailing by *E. coli* poly(A) polymerase. The "-" lanes represent IVT RNA before enzymatic polyadenylation, while the "+" lanes represent IVT mRNA after enzymatic polyadenylation. The smearing observed above the main bands shows the increased molecular weight of polyadenylated RNA, confirming successful poly(A) tailing.

Supplementary Table 1. GeneBlocks supplied by IDT utilised in this study. Yellow text highlights HindIII restriction site and cyan text highlights EcoRI restriction sites.

Name	Sequence	Cloning method
NLucHBB	GATCGTCAAAGCTTACTTGGTCAGCGGCCGCTAATACGACTCACTATAGGGACATTTGCTTCTGACACAACTGTGTTCACTAGCAACCTCAAA CAGACACCATGGTCTTCACACTCGAAGATTTCGTTGGGGACTGGCGACAGACAGCCGGCTACAACCTGGACCAAGTCCTTGAACAGGGAG GTGTGTCCAGTTTGTTTCAGAATCTCGGGGTGTCCGTAACTCCGATCCAAAGGATTGTCCTGAGCGGTGAAAATGGGCTGAAGATCGACATC CATGTCATCATCCCGTATGAAGGTCTGAGCGGCGACCAAATGGGCCAGATCGAAAAAATTTTTAAGGTGGTGTACCCTGTGGATGATCATCA CTTTAAGGTGATCCTGCACTATGGCACACTGGTAATCGACGGGGTTACGCCGAACATGATCGACTATTTCGGACGGCCGTATGAAGGCATC GCCGTGTTCGACGGCAAAAAGATCACTGTAACAGGGACCCTGTGGAACGGCAACAAAATTATCGACGAGCGCCTGATCAACCCCGACGG CTCCCTGCTGTTCCGAGTAACCATCAACGGAGTGACCGGCTGGCGGCTGTGCGAACGCATTCTGGCGTAAGCTCGCTTTCTTGCTGTCCAA TTTCTATTAAAGGTTCCTTTGTTCCCTAAGTCCAACTACTAAACTGGGGGATATTATGAAGGGCCTTGAGCATCTGGATTCTGCCTAATAAAAAA CATTTATTTTCATTGCAAGAATTCGTCAGATC	PCR amplification for classical restriction style cloning
NLucSec **HBB**	GATCGTCAAAGCTTACTTGGTCAGCGGCCGCTAATACGACTCACTATAGGGACATTTGCTTCTGACACAACTGTGTTCACTAGCAACCTCAAA CAGACACCATGGACTCCTTCTCCACAAGCGCCTTCGGTCCAGTTGCCTTCTCCCTGGGCCTGCTCCTGGTGTTGCCTGCTGCCTTCCCTGC CCCAGTCTTCACACTCGAAGATTTCGTTGGGGACTGGCGACAGACAGCCGGCTACAACCTGGACCAAGTCCTTGAACAGGGAGGTGTGTC CAGTTTGTTTCAGAATCTCGGGGTGTCCGTAACTCCGATCCAAAGGATTGTCCTGAGCGGTGAAAATGGGCTGAAGATCGACATCCATGTCA TCATCCCGTATGAAGGTCTGAGCGGCGACCAAATGGGCCAGATCGAAAAAATTTTTAAGGTGGTGTACCCTGTGGATGATCATCACTTTAAGG TGATCCTGCACTATGGCACACTGGTAATCGACGGGGTTACGCCGAACATGATCGACTATTTCGGACGGCCGTATGAAGGCATCGCCGTGTT CGACGGCAAAAAGATCACTGTAACAGGGACCCTGTGGAACGGCAACAAAATTATCGACGAGCGCCTGATCAACCCCGACGGCTCCCTGC TGTTCCGAGTAACCATCAACGGAGTGACCGGCTGGCGGCTGTGCGAACGCATTCTGGCGTAAGCTCGCTTTCTTGCTGTCCAATTTCTATTA AAGGTTCCTTTGTTCCCTAAGTCCAACTACTAAACTGGGGGATATTATGAAGGGCCTTGAGCATCTGGATTCTGCCTAATAAAAAACATTTATTTT CATTGCAAGAATTCGTCAGATC	PCR amplification for classical restriction style cloning
PMEL_5'	ATGGATCTGGTGCTAAAAAGATGCCTTCTTCATTTGGCTGTGATAGGTGCTTTGCTGGCTGTGGGGGCTACAAAAGGCCCCAGAAACCAGGA CTGGCTTGGTGTCTCAAGGCAACTCAGAACCAAAGCCTGGAACAGGCAGCTGTATCCAGAGTGGACAGAAGCCCAGAGACTTGACTGCTG GAGAGGTGGTCAAGTGTCCCTCAAGGTCAGTAATGATGGGCCTACACTGATTGGTGCAAATGCCTCCTTCTCTATTGCCTTGAACTTCCCTGG AAGCCAAAAGGTATTGCCAGATGGGCAGGTTATCTGGGTCAACAATACCATCATCAATGGGAGCCAGGTGTGGGGAGGACAGCCAGTGTAT CCCCAGGAAACTGACGATGCCTGCATCTTCCCTGATGGTGGACCTTGCCCATCTGGCTCTTGGTCTCAGAAGAGAAGCTTTGTTTATGTCTG GAAGACCTGGGGCCAATACTGGCAAGTTCTAGGGGGCCCAGTGTCTGGGCTGAGCATTGGGACAGGCAGGGCAATGCTGGGCACACACA CCATGGAAGTGACTGTCTACCATCGCCGGGGATCCCGGA	NEBuilder HiFi Assembly Cloning
PMEL_middle	GCTATGTGCCTCTTGCTCATTCCAGCTCAGCCTTCACCATTACTGACCAGGTGCCTTTCTCCGTGAGCGTGTCCCAGTTGCGGGCCTTGGAT GGAGGGAACAAGCACTTCCTGAGAAATCAGCCTCTGACCTTTGCCCTCCAGCTCCATGACCCCAGTGGCTATCTGGCTGAAGCTGACCTCT CCTACACCTGGGACTTTGGAGACAGTAGTGGAACCCTGATCTCTCGGGCACTTGTGGTCACTCATACTTACCTGGAGCCTGGCCCAGTCAC TGCCCAGGTGGTCCTGCAGGCTGCCATTCCTCTCACCTCCTGTGGCTCCTCCCCAGTTCCAGGCACCACAGATGGGCACAGGCCAACTG CAGAGGCCCCTAACACCACAGCTGGCCAAGTGCCTACTACAGAAGTTGTGGGTACTACACCTGGTCAGGCGCCAACTGCAGAGCCCTCT	NEBuilder HiFi Assembly Cloning

	GGAACCACATCTGTGCAGGTGCCAACCACTGAAGTCATAAGCACTGCACCTGTGCAGATGCCAACTGCAGAGAGCACAGGTATGACACCT GAGAAGGTGCCAGTTTCAGAGGTCATGGGTACCACACTGGCAGAGATGTCAACTCCAGAGGCT	
PMEL_3'	ACAGGTATGACACCTGCAGAGGTATCAATTGTGGTGCTTTCTGGAACCACAGCTGCACAGGTAACAACTACAGAGTGGGTGGAGACCACAG CTAGAGAGCTACCTATCCCTGAGCCTGAAGGTCCAGATGCCAGCTCAATCATGTCTACGGAAAGTATTACAGGTTCCCTGGGCCCCCTGCT GGATGGTACAGCCACCTTAAGGCTGGTGAAGAGACAAGTCCCCCTGGATTGTGTTCTGTATCGATATGGTTCCTTTTCCGTCACCCTGGACAT TGTCCAGGGTATTGAAAGTGCCGAGATCCTGCAGGCTGTGCCGTCCGGTGAGGGGGATGCATTTGAGCTGACTGTGTCCTGCCAAGGCGG GCTGCCCAAGGAAGCCTGCATGGAGATCTCATCGCCAGGGTGCCAGCCCCCTGCCCAGCGGCTGTGCCAGCCTGTGCTACCCAGCCCA GCCTGCCAGCTGGTTCTGCACCAGATACTGAAGGGTGGCTCGGGGACATACTGCCTCAATGTGTCTCTGGCTGATACCAACAGCCTGGCA GTGGTCAGCACCCAGCTTATCATGCCTGGTCAAGAAGCAGGCCTTGGGCAGGTTCCGCTGATCGTGGGCATCTTGCTGGTGTTGATGGCTG TGGTCCTTGCATCTCTGATATATAGGCGCAGACTTATGAAGCAAGACTTCTCCGTACCCCAGTTGCCACATAGCAGCAGTCACTGGCTGCGT CTACCCCGCATCTTCTGCTCTTGTCCCATTGGTGAGAACAGCCCCCTCCTCAGTGGGCAGCAGGTCTGA	NEBuilder HiFi Assembly Cloning
Cytosolic PMEL_5'	ATGGGTAAAGGCCCCAGAAACCAGGACTGGCTTGGTGTCTCAAGGCAACTCAGAACCAAAGCCTGGAACAGGCAGCTGTATCCAGAGTGG ACAGAAGCCCAGAGACTTGACTGCTGGAGAGGTGGTCAAGTGTCCCTCAAGGTCAGTAATGATGGGCCTACACTGATTGGTGCAAATGCCT CCTTCTCTATTGCCTTGAACTTCCCTGGAAGCCAAAAGGTATTGCCAGATGGGCAGGTTATCTGGGTCAACAATACCATCATCAATGGGAGC CAGGTGTGGGGAGGACAGCCAGTGTATCCCCAGGAAACTGACGATGCCTGCATCTTCCCTGATGGTGGACCTTGCCCATCTGGCTCTTGG TCTCAGAAGAGAAGCTTTGTTTATGTCTGGAAGACCTGGGGCCAATACTGGCAAGTTCTAGGGGGCCCAGTGTCTGGGCTGAGCATTGGGA CAGGCAGGGCAATGCTGGGCACACACACCATGGAAGTGACTGTCTACCATCGCCGGGGATCCCGGA	NEBuilder HiFi Assembly Cloning

Supplementary Table 2. Primers utilised for NEBuilder HiFi Assembly cloning. Details include primer name, sequence, annealing conditions and the plasmids each fragment was created to generate. Purple text highlights Fragment overlap with next fragment/vector, Grey text highlights nucleotides that have been altered in the primers, analogous to site-directed mutagenesis.

Primer name	Primer sequence 5'-3'	Primer conditions	Used to generate
Nluc_alt_frag_F	CAGACACCATGGATCTGGTGCT	60 °C	Nluc altSPCSHBB
Nluc_alt_frag_R	TGTGAAGACTGTAGCCCCCACAGCCAG		
Amp_F	GCCATTGCTACAGGCATCGTGGTGTCACGC	56 °C	
Nluc_alt_Vec_R	CAGATCCATGGTGTCTGTTTGAGGTTG		
Nluc_alt_Vec_F	GGGCTACAGTCTTCACACTCGAAGATTTCGT	56 °C	
Amp_R	CTGTAGCAATGGCAACAACGTTGCGC		
PMEL_5'end_F	ACAGACACCATGGATCTGGTGCTAAAAAGATGCC	64 °C	PMEL_HBB
PMEL_5'end_R	CACATAGCTCCGGGATCCCCGGC		
PMEL_mid_F	ATCCCGGAGCTATGTGCCTCTTGCTCATTCC	65 °C	
PMEL_mid_R	CATACCTGTAGCCTCTGGAGTTGACATCTCTG		
PMEL_3'end_F	CCAGAGGCTACAGGTATGACACCTGCAGAGG	65 °C	
PMEL_3'end_R	AAGCGAGCTCAGACCTGCTGCCCACTGA		
Amp_F	GCCATTGCTACAGGCATCGTGGTGTCACGC	57 °C	
PMEL_Vec_R	CAGATCCATGGTGTCTGTTTGAGGT		
PMEL_Vec_F	GGTCTGAGCTCGCTTTCTTGC	55 °C	
Amp_R	CTGTAGCAATGGCAACAACGTTGCGC		
OPMEL_IL6_Frag_F	CAGACACCATGGACTCCTTCTCC	60 °C	PMEL_alt_SPCSHBB
OPMEL_IL6_Frag_R	GGGCCTTTTGGGGCAGGGAAGGCA		
Amp_F	GCCATTGCTACAGGCATCGTGGTGTCACGC	58 °C	
OPMEL_IL6_Vec_R	GGAGTCCATGGTGTCTGTTTGAGGTTGC		
OPMEL_IL6_Vec_F	TGCCCCAAAAGGCCCCAGAAACCAGG	58 °C	
Amp_R	CTGTAGCAATGGCAACAACGTTGCGC		
Cyto_PMEL_5'end_F	ACAGACACCATGGGTAAAGGCCCCAGAAAC	58 °C	Cyto_PMEL_hbb
Cyto_PMEL_5'end_R	CACATAGCTCCGGGATCCCCGGC		
Cyto_PMEL_mid_F	ATCCCGGAGCTATGTGCCTCTTGCTCATTCC	65 °C	
Cyto_PMEL_mid_R	CATACCTGTAGCCTCTGGAGTTGACATCTCTG		
Cyto_PMEL_3'end_F	CCAGAGGCTACAGGTATGACACCTGCAGAGG	65 °C	
Cyto_PMEL_3'end_R	AAGCGAGCTCAGACCTGCTGCCCACTGA		
Amp_F	GCCATTGCTACAGGCATCGTGGTGTCACGC	55 °C	
Cyto_PMEL_Vec_R	TTTACCCATGGTGTCTGTTTGAGGTTG		
Cyto_PMEL_Vec_F	GGTCTGAGCTCGCTTTCTTGC	55 °C	
Amp_R	CTGTAGCAATGGCAACAACGTTGCGC		
PMEL_mut_Frag_F	TGTGTCTCTGGCTGATACCAACAGCCTGGCAG	66 °C	PMEL_mutant_HBB
PMEL_mut_Frag_R	GCTGCCCACTGGCGGCGGGGCTGTTCTC	66 °C	
Amp_F	GCCATTGCTACAGGCATCGTGGTGTCACGC	60 °C	
PMEL_mut_Vec_R	TGGTATCAGCCAGAGACACATTGAGGCAGTATG		
PMEL_mut_Vec_F	CCCCGCCGCCAGTGGGCAGCAGGTCTGAG	59 °C	
Amp_R	CTGTAGCAATGGCAACAACGTTGCGC		
CleanCap_Frag_F	CTATCGATAGGTACCGAGCTCTTACGCG	64 °C	

CleanCap_Frag_R	CACAGTTGTGTCAGAAGCAAATGTCCTTATAGTGAGTCGT		CleanCap_NlucSec_HBB
Amp_F	GCCATTGCTACAGGCATCGTGGTGTCACGC	58 °C	
CleanCap_Vec_R	GCTCGGTACCTATCGATAGAGAAATGTTC		
CleanCap_Vec_F	GCTTCTGACACAACTGTGTTCACT	58 °C	
Amp_R	CTGTAGCAATGGCAACAACGTTGCGC		
CC_Opt_Frag_F	CTATCGATAGGTACCGAGCTCTTACGCG	64 °C	CleanCap_NlucSec_Opt_HBB
CC_Opt_Frag_R	TCCATGGTGGCGGCTTGAGGTTGCT		
Amp_F	GCCATTGCTACAGGCATCGTGGTGTCACGC	58 °C	
CC Opt_Vec_R	GTACCTATCGATAGAGAAATGTTCTGGCAC		
CC_Opt_Vec_F	AGCCGCCACCATGGACTCCTTCTCCACAAG	56 °C	
Amp_R	CTGTAGCAATGGCAACAACGTTGCGC		
mCher_Frag_F	AGCCGCCACCATGGTGAGCAAGGGCGAG	56 °C	CleanCap_mCher_Opt_HBB
mCher_Frag_R	GAAAGCGAGCTTACTTGTACAGCTCGTCCATG		
Amp_F	GCCATTGCTACAGGCATCGTGGTGTCACGC	60 °C	
mCher_Vec_R	TGCTCACCATGGTGGCGGCTTGAGGTTG		
mCher_Vec_F	GTACAAGTAAGCTCGCTTTCTTGCTGTC	56 °C	
Amp_R	CTGTAGCAATGGCAACAACGTTGCGC		
eGFP_frag_F	AGCCGCCACCATGGTGAGCAAGGGCGAG	57 °C	CleanCap_eGFP_Opt_HBB
eGFP_frag_R	GAAAGCGAGCTTACTTGTACAGCTCGTCCATG		
Amp_F	GCCATTGCTACAGGCATCGTGGTGTCACGC	57 °C	
eGFP_Vec_R	TGCTCACCATGGTGGCGGCTTGAGGTTG		
eGFP_Vec_F	GTACAAGTAAGCTCGCTTTCTTGCTGTCC	58 °C	
Amp_R	CTGTAGCAATGGCAACAACGTTGCGC		

Supplementary Table 3. PCR primers utilised to generate *in vitro* transcription templates. T7 promoter sequence is highlighted in green.

Name	Sequence	Annealing temperature
Renilla_IVT_F	TAATACGACTCACTATAGGGCTAGAGCTTGGC	59 °C
Renilla_IVT_R	TCGAGGTCGACGGTATCGATAAGCTC	
eGFP_IVT_F	TAATACGACTCACTATAGGGATGGTGAGCAAGGGCGAGG	60 °C
eGFP_IVT_R	GCGGCCGCTTTACTTGTACAGC	
HBB_IVT_F	GCCGCTAATACGACTCACTAT	58 °C
HBB_IVT_R	TTGCAATGAAAATAAATGTTTTTTATTAGGCAGAATCCAGA	

Supplementary Table 4. Primers utilised for sequencing plasmids.

Primer	Sequence	Plasmid
SV40_F	TCTGCGATCTGCATCTCAATTA	NlucHbb, NlucSecHbb, PMELHbb, CytoPMELHbb, PMELIL6SPCSHbb, PMELmutantHbb, CCNlucSecHbb, CCNlucHbbOpt, CCmCherHbbOpt, CCeGFPHbbOpt
PMEL_mid_R	CATACCTGTAGCCTCTGGAGTTGACATCTCTG	PMELHbb, CytoPMELHbb, PMELIL6SPCSHbb
PMEL_end_F	CCAGAGGCTACAGGTATGACACCTGCAGAGG	PMELHbb, CytoPMELHbb, PMELIL6SPCSHbb

Supplementary Table 5. RT-qPCR primers used to quantify mRNA levels.

Primer Name	Sequence 5'-3'	Predicted subcellular location of endogenous transcript	Source
APEX F	TTCGGAACCATCAAGCACCC	N/A	AS
APEX R	CCAGGGTGGAATGGAACCTTAG		
18S F	CAGCCACCCGAGATTGAGCA	Both	Jopling lab
18S R	TAGTAGCGACGGGCGGTGTG		
TMX1 F	ACGGACGAGAACTGGAGAGA	ER	(Fazal et al., 2019)
TMX1 R	ATTTTGACAAGCAGGGCACC		
SSR2 F	GTTTGGGATGCCAACGATGAG	ER	(Fazal et al., 2019)
SSR2 R	CTCCACGGCGTATCTGTTCA		
FAU1 F	TCCTAAGGTGGCCAAACAGG	Cytosolic	(Fazal et al., 2019)
FAU1 R	GTGGGCACAACGTTGACAAA		
MTCO2 F	AACCAAACCACTTTCACCGC	Mitochondrial	(Fazal et al., 2019)
MTCO2 R	CGATGGGCATGAAACTGTGG		
SFT2D2 F	CCATCTTCCTCATGGGACCAG	ER	(Fazal et al., 2019)
SFT2D2 R	GCAGAACACAGGGTAAGTGC		
PFDN4 F	CGGCTGCAGAAGATGTCAATG	Cytosolic	AS. Inspired by Padrón et al. (2019)
PFDN4 R	TCTTCTAGGTTTTGGAGTTGTTTCTT		

GRP94	GGATCAAGGACGGATGATGAA	ER	AS. Inspired by Padrón et al. (2019)
GRP94	CAGAGAGAGGAAGAAGCTATTCAG		
Nluc_F	CCGTATGAAGGTCTGAGCGG	N/A	Jopling lab
Nluc_R	TCTTTTTGCCGTCGAACACG		
Nluc_F_alternative	CACACTCGAAGATTTCGTTGG	N/A	AS
Nluc_R_alternative	GAAACAAACTGGACACACCTC		
Fluc_F	CTAAGGAAGTCGGGGAAGCG	N/A	Jopling lab
Fluc_R	ATCCCCCTCGGGTGTAATCA		
PMEL_F	TTAAGGCTGGTGAAGAGACAAG	ER	AS
PMEL_R	AGGATCTCGGCACTTTCAATAC		
PMEL_F_alt	CTCCAGAGGCTACAGGTATGA	ER	AS
PMEL_R_alt	TCTCCACCCACTCTGTAGTT		
eGFP_F	GAACCGCATCGAGCTGAA	N/A	AS
eGFP_R	TGCTTGTCGGCCATGATATAG		
Renilla_F	TAACGCGGCCTCTTCTTATTT	N/A	AS
Renilla_R	GATTTGCCTGATTTGCCCATAC		