Delivery of CRISPR/Cas9 by receptor-targeted nanoparticles as a corrective therapy for Cystic Fibrosis

Amy Walker

Great Ormond Street Institute of Child Health

University College London

A thesis submitted for the degree of Doctor of Philosophy

October 2019

Declaration

I, Amy Walker, declare that the work presented in this thesis is my own. Any contributions from collaborators, or where information has been derived from other sources, is explicitly referenced in the text.

Abstract

Cystic Fibrosis (CF) is the most common inherited genetic disorder, affecting around 1 in 2,500 babies born in the UK. Clinical manifestations are caused by mutations in the gene encoding the cystic fibrosis transmembrane regulator (CFTR), a membrane channel protein which regulates anion transport and mucociliary clearance. CRISPR/Cas9 is an attractive therapeutic option for CF as it could target the underlying cause of the disease, rather than treating symptoms. However, a major hurdle to overcome if CRISPR/Cas9 is to be deployed as a treatment is how to deliver the technology to the lung.

We hypothesise that receptor targeted nanocomplexes (RTNs), previously shown to deliver DNA and siRNA to the lung, can be used to package and deliver the gene editing components. Here, we have successfully optimised these nanocomplexes specifically for the delivery of Cas9 over its various platforms and extensively characterised the particles. Using an epithelial GFP reporter system, we were able to achieve higher transfection levels than commercially available reagents, for both Cas9 mRNA and Cas9 protein.

The 10th most common CF causing mutation, 3849+10kb C>T, generates a cryptic splice site, resulting in the formation of a pseudoexon containing a PTC, producing a truncated version of the protein. Using pairs of gRNAs, we were able to successfully remove the cryptic splice site using an NHEJ strategy, correcting aberrant splicing and, importantly, CFTR channel function as measured by Ussing Chamber.

A homology-independent targeted integration (HITI) strategy was used to investigate the potential of restoring CFTR function regardless of mutation type. As proof-of-concept, we delivered GFP into AAVS1 locus of HEK293 cells, achieving a targeted knock-in efficiency of 14%.

Finally, the use of RTNs to deliver CRISPR/Cas9 in vivo was explored successfully in an Ai9 mouse reporter model to restore tdTomato expression by paired gRNAs excision of a stop cassette, confirming RTNs can be used as a safe delivery method for repeated dosing of CRISPR/Cas9, and highlighting their translational potential.

Impact Statement

Cystic fibrosis (CF) is an autosomal recessive disease, which affects around 1 in 2,500 babies born in the UK. To date, several gene therapy approaches have been evaluated by delivery of CFTR cDNA by viral or non-viral vectors, however a clinically effective therapy has yet to emerge. Major challenges include delivery to the lung, limited levels of gene transfer, and persistence of transgene expression.

In recent years, there has been rapid advancement in the development of gene editing techniques, and this thesis explores the use of the CRISPR/Cas9 system for the treatment of CF. A number of advances towards a gene editing therapy have been described, with particular application to class V deep intronic splice mutations that make up 1.6% of the CF population. This editing strategy has broader application beyond CF, with the potential of treating 75 other genetic diseases caused by splice mutations. The HITI work described provides a proof-of-concept for developing a gene therapy strategy applicable to all CF patients, regardless of mutation site, and could significantly reduce the burden CF has on the healthcare system, particularly with the recent approval of high-cost, small modulator drugs in the UK. Finally, the optimisation of Receptor-Targeted Nanocomplexes for the delivery of the CRISPR/Cas9 platform provide huge potential for a highly efficient, non-immunogenic delivery method, not limited to the field of CF.

This work has been presented at both national and international conferences. Manuscripts of this work are in preparation and are expected to be submitted to journals in the near the future, where we believe will have an impact on the gene editing field.

Acknowledgements

This thesis is dedicated to my granny, who isn't here to see me complete my PhD, but is my whole reason for doing so.

I would like to express my sincere gratitude to Professor Steve Hart, for his supervision and encouragement throughout my PhD, as well as opening up so many opportunities to me. Thanks also to Prof. Chris O'Callaghan, Dr. Patrick Harrison and Prof. Debbie Baines for their guidance and expertise.

Thanks to Ahmad Aldossary and Ruhina Maeshima, particularly for assistance with *in vivo* experiments, and their general loveliness. To all the Hart group, particularly Preetha Purushothaman and Emily Young, who helped to make my time at ICH so enjoyable. A special thanks to Ileana Guerrini, who helped me find my feet as a PhD student, and taught me to ask the right scientific questions. A big thank you to Nick Harper, who steered me through my Masters and led me to this PhD.

Thanks also to collaborators, without whom this work wouldn't be possible. To Jayesh Kulkarni, Dominik Witzigmann, and Prof. Pieter Cullis, for not only their incredible knowledge, but for making my time in UBC so enjoyable. Thanks to Prof. Robin McNulty and David Pearce, for his endless help with mice lungs. To Max Woodall and Afroditi Avgerinou, not only for their collaborations but for their amazing friendships.

Thanks to the CF Trust and BRC for their funding, and to The Bogue Fellowship for giving me a once in a lifetime experience in Canada.

To all the PUW3 girls for providing friendship and fun throughout the three years, especially to Joy- I got very lucky.

Thanks to my lovely mum and Gavin for always believing in me, my dad for all of his support, and to my wonderful grandad. Thanks to Nick and Ruth for of all their praise and encouragement. Finally, thanks to Toby, who provided me with endless cupcakes, celebrated all the small wins, and kept me motivated throughout.

Publications

Boyd, C; Guo, S; Huang, S; Kerem, B; Oren, Y; **Walker, A** & Hart, S. New approaches to genetic therapies for Cystic Fibrosis. *Journal of Cystic Fibrosis.*

Hart, S; **Walker, A** & Harrison, P. New Genetic Approaches to Treating Cystic Fibrosis. *Advanced Textbook on Gene Transfer, Gene Therapy and Genetic Pharmacology,* Ch.170.

Hart, S. & **Walker, A.** New genetic approaches to treating diseases of the skin. *Harper's Textbook of Paediatric Dermatology,* Ch.38.3

Walker, A, *et. al.* Cas9-mediated repair of Cystic Fibrosis deep-intronic splice mutations using receptor-targeted nanocomplexes. Manuscript in preparation.

Walker, A, *et. al.* Cas9-mediated Homology- Independent Targeted Integration of donor templated as a tool for CF gene therapy. Manuscript in preparation.

Contents

Declaration 2

Abstract 3

Impact Statement 4

Acknowledgements 5

Publications 6

Contents 7

Abbreviations 13

1.1 Cystic Fibrosis 20

1.1.1 History of Cystic Fibrosis 20

1.1.2 The CFTR gene and protein 21

1.1.3 Cystic Fibrosis pathophysiology in the lungs 22

1.1.1.1 ASL in normal airways 23

1.1.1.2 The high-salt hypothesis 23

1.1.1.3 Low-volume hypothesis 23

1.1.1.4 Impaired bicarbonate secretion hypothesis 24

1.1.4 CFTR mutations 24

1.2 Additional channels that interact with CFTR 28

1.2.1 The epithelial sodium channel 28

1.2.2 Calcium-activated chloride channels 29

1.3 CFTR study model 31

1.3.1 *In vitro* study models 31

1.3.2 *In vivo* models 34

1.4 Current CF therapy 35

1.4.1 CF Modulator therapies 37

1.5 Gene Therapy 40

1.5.1 Viral vectors 41

1.5.1.1 Lentiviral vectors 41

1.5.1.2 Adeno-associated viral vectors 42
1.5.1.3 Bocavirus vectors 43
1.5.2 Non-viral vectors 43
1.5.2.1 Non-viral vectors for CF gene therapy 44
1.5.3 *In vitro* transcribed mRNA 45
1.5.4 Barriers to transfection in the CF lung 47
1.5.5 Cellular therapy 48
1.6 DNA repair *50*
1.6.1 Non-homologous end joining (NHEJ) 50
1.6.2 Homology Directed Repair (HDR) 53
1.7 Genome Editing *55*
1.7.1 Zinc-Finger Nucleases 55
1.7.2 Transcription activator-like effector nucleases 58
1.7.3 Clustered Regularly Interspaced Short Palindromic Repeats (CRISPR) 59
1.7.3.1 CRISPR/Cas variants 63
1.7.3.2 CRISPR/Cas9 applications 65
1.7.3.3 Base editing 66
1.7.3.4 Prime editing 67
1.7.3.5 Clinical applications 67
1.7.3.6 CRISPR/Cas9 off target effects 69
1.7.4 Gene editing for CF 71
1.2 Aims *74*
2.1 Materials and methods *76*
2.1.1 Equipment 76
2.1.2 Kits and reagents 76
2.1.3 Eukaryotic cells 78
2.1.4 Lipids 79
2.1.5 Liposome composition 79
2.1.6 Peptides 79
2.1.7 Plasmids 80
2.1.8 Primers 80

2.1.9 Taqman Assays (FAM-labelled probes) 80
2.1.10 Antibodies 80
2.1.11 Recipes 81
2.2 Methods *81*
2.2.1 Collagen coating of plastic ware 81
2.2.2 Cell Maintenance 81
2.2.3 ALI Culture 82
2.2.4 Transepithelial electrical resistance (TEER) measurements 82
2.2.5 Lentiviral transduction of GFP in NHBE BMI-1 cells 82
2.2.6 Cell cloning by serial dilution 82
2.2.7 Liposome formation 82
2.2.8 gRNA preparation 83
2.2.9 *In vitro* Cas9 mRNA and Cas9 RNP transfections 83
2.2.10 DNA extraction 83
2.2.11 RNA extraction 84
2.2.12 Polymerase chain reaction (PCR) 84
2.2.13 Agarose gel electrophoresis 85
2.2.14 PCR purification 85
2.2.15 T7 endonuclease I assay 85
2.2.16 Sanger sequencing 86
2.2.17 Determination of indel % using TIDE and ICE software 86
2.2.18 Analysis of CFTR protein expression by Western Blot 86
2.2.19 CFTR mRNA quantification by qRT-PCR 87
2.2.20 Flow cytometry for in vitro transfections 88
2.2.21 Biophysical characterisation of RTNs 88
2.2.22 Gel retardation assay 88
2.2.23 Cryogenic Transmission Electron Microscopy 88
2.2.24 Electrophysiology studies 89
2.2.25 Luciferase assay 89
2.2.26 Bradford protein assay 90
2.2.27 Flow cytometry for *in vivo* transfections 90

2.2.28 Histology 91
2.2.29 Immunohistochemistry protocol 91
2.2.30 Autostainer protocols 91
2.2.31 Data analysis 93
3.1 Introduction 95
3.2 Aims 97
3.3 Results 98
3.3.1 BMI-1 modified epithelial cells as a suitable model for CF 98
3.3.1.1 Proliferative potential of NHBE BMI-1 cells 98
3.3.1.2 NHBE BMI-1 cell morphology 99
3.3.1.3 CFTR profiling in NHBE cells 100
3.3.1.3.1 CFTR mRNA expression changes over time 100
3.3.1.3.2 CFTR protein expression in primary NHBEs 101
3.3.1.4 CFTR profiling in NHBE BMI-1 cells 102
3.3.1.4.1 CFTR mRNA expression in NHBE BMI-1 cells 102
3.3.2 Receptor-targeted nanocomplexes for nucleic acid delivery in NHBE BMI-1 cells 104
3.3.2.2.1 Determining optimal alkyl chain length for mRNA delivery 108
3.3.3 Development and validation of a cellular model for the optimisation of delivery of CRISPR/Cas9 110
3.3.3.2 Validation of NHBE-BMI1 GFP cellular model 112
3.3.3.3 Determining optimal hydrocarbon chain length for Cas9 mRNA delivery 115
3.3.3.4 Amount of Cas9 mRNA and gRNA affects editing efficiency 116
3.3.4 Delivery of Cas9 protein and gRNA as a pre-assembled RNP complex 118
3.3.4.1 Comparison of Cas9 mRNA vs. Cas9 RNP on editing efficiency 120
3.3.4.2 Comparison of lipid ratios for RNP delivery 122
3.3.4.3 Optimising peptide ratios for RNP delivery 124
3.3.4.4 Cell viability in RNP transfected cells 125
3.3.5 Biophysical characterisation of Receptor Targeted Nanocomplexes 126
3.3.5.1 Dynamic-Light Scattering 126

3.3.5.2 Transmission electron cryomicroscopy (CryoTEM) 127
3.3.5.3 Nucleic acid condensation and packaging properties of RTNs 129
3.4 Discussion *132*
4.1 Introduction *138*
4.2 Aims *141*
4.3 Results *142*
4.3.1 Targeted excision of a deep intronic splice mutation using an NHEJ-mediated editing strategy 142
4.3.1.1 Validation of gRNAs in HEK293T cells 142
4.3.1.2 Targeted excision using gRNA pairs in HEK293T cells 144
4.3.1.3 Validation of gRNAs in CFBE3849 cells 147
4.3.1.4 Targeted excision using gRNA pairs in CFBE3949 cells 150
4.3.1.5 Sequential transfections to increase targeted excision efficiency in CFBE3849 cells 152
4.3.1.6 Single cell cloning of corrected CFBE3849 cells 155
4.3.1.7 CFTR mRNA splicing analysis in CFBE3849 cells 156
4.3.1.8 CFTR mRNA splicing analysis in corrected CFBE3849 cells 159
4.3.1.9 Electrophysiology studies in an Ussing chamber 161
4.3.2 Investigating a NHEJ-mediated HITI editing strategy for targeted knock-in of GFP 163
4.3.2.1 Construction of HITI donor plasmid and edting plasmid 164
4.3.2.2 Transient transfection efficiency of RTNs containing donor and editing plasmids in HEK293T cells 165
4.3.2.3 Efficiency of HITI-mediated GFP integration in HEK293T cells 12 days post-transfection 167
4.3.2.4 Molecular analysis of HITI-mediated integration events via junction PCR167
4.3.2.5 Single cell clone molecular analysis of HITI-mediated integration events via junction PCR 169
4.3 Discussion *170*
5.1 Introduction *177*

5.2 Aims *179*

5.3 Results *180*

5.3.1 RTNs for delivery of Luciferase mRNA to the murine lung 180

5.3.2 RTNs delivery for Cre mRNA or Cas9 RNP in the Ai9 mouse 181

5.3.2.1 gRNAs targeting the STOP cassette in Ai9 mice 181

5.3.2.2 Delivery of Cas9 RNPs and Cre mRNA via RTNs 183

5.3.2.3 Weight of treated mice over time-course of experiment 183

5.3.2.4 Flow cytometry analysis to assess tdTomato+ve cells 184

5.3.2.5 Immunohistochemistry to assess tdTomato +ve cells 186

5.3.2.5.1 TdTomato embryo as a +ve control for IHC 186

5.3.1.1.1 IHC for detection of tdTomato +ve cells in Ai9 mice 188

5.3.2.6 Histology to assess in vivo toxicity of RTNs 189

5.3.2.7 Dynamic Light Scattering for biophysical characterisation of RTNs 190

5.4 Discussion *192*

6.1 Discussion *197*

Abbreviations

2- ΔΔCt	delta-delta Ct
µA	microampere
µm	micrometre
µg	microgram
µl	microlitre
γ	gamma
δ	delta
λ	lambda
ΔF508	CFTR mutation with deletion in phenylalanine at position 508
AAV	adeno-associated virus
AAVS1	adeno-associated virus integration site 1
ADP	adenosine diphosphate
ALI	air-liquid interface
Alt-NHEJ	alternative NHEJ
ANO1	anoctamin 1
ANOVA	analysis of variance
ASL	airway surface liquid
ATP	adenosine triphosphate
BEBM	bronchial epithelial basal media
BEGM	bronchial epithelial growth media
BFP	blue fluorescent protein
BMI-1	B-cell-specific Moloney murine leukaemia virus integration site 1
bp	base pair
C14	DTDTMA/DOPE
C16	DHDTMA /DOPE
C18	DOTMA/DOPE
Ca^{2+}	calcium ions
CaCCs	calcium-activated chloride channels
CAG	CMV enhancer fused to the chicken beta-actin promoter

cAMP	cyclic adenosine monophosphate
Cas	CRISPR associated system
CBF	ciliary beat frequency
CBh	Chicken Beta Actin Short)promoter.
cDNA	complementary DNA
CF	cystic Fibrosis
CFTR	cystic fibrosis transmembrane conductance regulator
CjCas9	*Campylobacter jejuni* Cas9
Cl^-	chloride ion
CMV	cytomegalovirus
Cpf1	CRISPR from *Prevotella* and *Francisella*
CRISPR	Clustered regularly interspaced short palindromic repeats
crRNA	CRISPR-RNA
Ct	cycle threshold
CtIP	C-terminal-binding protein interacting protein
dCas9	dead Cas9
DHDTMA	1,2-Di-((Z)-hexadec-11-enyloxy)-N,N,N trimethylammonium propane iodide
DMD	Duchenne muscular dystrophy
DMEM	Dulbecco's Modified Eagle Medium
DMSO	dimethyl sulfoxide
DNA	deoxyribonucleic acid
DNMT3A	DNA Methyltransferase 3 Alpha
dNTP	deoxyribonucleotide triphosphate
DOPE	dioleoyl L-α phosphatidyl ethanolamine
DOTMA	N-[1- (2,3-dioleyloxy)propyl]-N,N,N-trimethlylammonium chloride
DSB	double-strand break
DSBR	double-strand break repair
DTDTMA	1,2-Di-((Z)-tetradec-11-enyloxy)-N,N,N trimethylammonium propane chloride

E peptide	K16GACSERSMNFCG
E. coli	*Escherichia coli*
eGFP	enhanced green fluorescent protein
ENaC	epithelial sodium channel
EpCAM	epithelial cell adhesion molecule
ER	endoplasmic reticulum
ERAD	endoplasmic reticulum-associated degradation
FBS	foetal bovine serum
FEV1	forced expiratory volume in one second
G1	gap 1 phase
G2	gap 2 phase
GFP	Green fluorescent protein
GL67	genzyme lipid 67A/1,2-Dioleoyl-sn-Glycero-3-Phosphoethanolamine
gRNA	guide RNA
GSH	glutathione
hCFTR	human CFTR
HCO_3	bicarbonate
HDR	homology directed repair
HITI	homology-independent targeted integration
hMSCs	human mesenchymal stem cells
HN	hemagglutinin/neuraminidase protein
HR	Homologous recombination
ICAM-1	Intercellular Adhesion Molecule 1
iPSC	induced pluripotent stem cell
Isc	short-circuit current
IVT	in vitro transcribed
Kbp	kilo base pair
L	lipid
L2K	Lipofectamine 2000
LIG4	DNA *ligase* 4

LTR	long terminal repeat
M1	transmembrane domains 1
M2	transmembrane domains 2
MFI	mean fluorescence intensity
mg	milligram
ml	millilitre
mM	millimolar
MMEJ	microhomology-mediated end joining
MOI	multiplicity of infection
mRNA	messenger RNA
MSCs	mesenchymal stromal cells
MSD	membrane-spanning domains
Na+	sodium ion
NaCl	sodium chloride
NBD	nucleotide-binding domains
ng	nanogram
NGS	next-generation sequencing
NHBE	normal human bronchial epithelial
NHEJ	non-Homologous End Joining
NLS	nuclear localization signals
NS	not significant
P	probability value
PAM	Protospacer adjacent motif
PARP1	poly(ADP-ribose) polymerase 1
PBAEs	poly-beta amino esters polymer
PCD	primary ciliary dyskinesia
PCR	Polymerase Chain Reaction
PEG	polyethylene glycol
PKA	protein kinase A
pol θ	DNA polymerase theta
ppFEV1	percent predicted forced expiratory volume in one second

pre-crRNA	precursor-CRISPR-RNA
qRT-PCR	quantitative reverse transcription-PCR
RNA	ribonucleic acid
RNP	ribonucleoprotein
ROCK	rho-associated protein kinase
rpm	revolutions per minute
RTN	receptor-targeted nanocomplexes
RuvC	crossover junction endodeoxyribonuclease
RVD	repeat-variable di-residue
S	synthesis phase
SaCas9	*Staphylococcus aureus* Cas9
SAM	synergistic activation mediator protein complex
SD	standard deviation
SDSA	synthesis-dependent strand annealing
SIN	self-inactivating
siRNA	small-interfering RNA
SLC26A9	solute Carrier Family 26 Member 9
SpCas9	*Streptococcus pyogenes* Cas9
SPLUNC1	short palate lung and nasal epithelial clone 1
ssDNA	single strand DNA
ssODN	single-stranded oligonucleotide
TAE	tris-acetate- EDTA buffer
TALEN	transcription activator-like effector nuclease
TEER	transepithelial electrical resistance
TGF-β1	transforming growth factor beta 1
TIDE	tracking of Indels by Decomposition
TMEM16A	transmembrane member 16A
tracrRNA	trans-activating crRNA
UTR	untranslated region
V	volt
w/v	weight/volume

wt-CFTR	wild-type cystic fibrosis transmembrane conductance regulator
XLF	XRCC4-like factor
XRCC4	X-ray cross-complementing protein 4
Y peptide	K16GACYGLPHKFCG
ZFN	Zinc-finger nuclease

Chapter 1

Introduction

1.1 Cystic Fibrosis

Cystic fibrosis (CF) is an autosomal recessive disease categorised by pancreatic insufficiency, biliary and gastrointestinal disease and, importantly, respiratory disease. In the UK, there are 10,500 people currently living with CF, and it affects around 1 in 2,500 babies born [1]. With a carrier frequency of 1 in 25, it the most common inherited disease in the European population [2]. The estimated median life expectancy of a baby born in the UK with CF today is 47 years of age [1].

1.1.1 History of Cystic Fibrosis

"Woe to that child which when kissed on the forehead tastes salty. He is bewitched and soon must die". Uncovered from Medieval European folklore, this adage is thought to be one of the earliest references to Cystic Fibrosis [3]. A medical description of the disease followed in 1595 by Pieter Pauw, of Leiden, Netherlands, who conducted an autopsy of an 11-year old girl, and described the pancreas as swollen, with a hard, white viscous mass [4].

It wasn't until 1938 that the term 'cystic fibrosis of the pancreas' was coined by Dorothy Andersen, on the basis of features she observed in the pancreatic tissues obtained from the autopsies of infants and children. Andersen also noted that children often presented with infection and damage to the lung airways [5]. In 1945, Sidney Faber introduced the term *mucoviscidosis,* reasoning that cause of the disease was a generalised state of thickened mucus [6]. This term is still widely used outside English speaking countries. One year later, Andersen and Hodges presented the first compelling evidence that the disease was likely caused by a single gene, and inherited in a monogenic recessive fashion [7].

During a New York heatwave in the summer of 1948, it was observed that a disproportionate number of patients admitted to Columbia Hospital had cystic fibrosis, seemingly becoming dehydrated more readily than their healthy peers [8]. Based on this observation, Paul di Sant'Agnese and colleagues found that people with CF lose an excessive amount of salt in their sweat [9], corroborating the salty brows of 'bewitched' Medieval children. The finding of this unique, abnormal physiological function resulted in the development of the sweat chloride test, an accurate and accessible diagnostic parameter that remains in use today [4].

By the mid 1980s, it had been discovered that epithelial tissue was malfunctioning in all organs affected by CF, and in-particular was relatively impermeable to chloride ions [10,

11], explaining the excessively salty sweat of patients. Finally, in 1989, a group of collaborators led by Lap Chee Tsui, Francis Collins and Jack Riordan announced that they had isolated the gene responsible for CF [12-14], and named the protein the Cystic Fibrosis Transmembrane Conductance Regulator (CFTR). Additionally, the researchers noted the prevalence of aberration in the DNA, a three-base deletion that results in the loss of phenylalanine at position 508, that appeared to account for upwards of 70% of CF cases [12, 13].

1.1.2 The CFTR gene and protein

The CFTR gene is located on chromosome 7 (long arm, region q31-q32). It is 250 kB in length and comprises 27 exons, encoding an mRNA of 6.5 kb and a protein of 1,480 amino acids [15]. The CFTR protein is an ABC (ATP-binding cassette) transporter class protein (sub-class ABCC7) [13] and functions as an ATP-gated anion channel regulated by cyclic AMP (cAMP)-dependent phosphorylation. Crucial for trans-epithelial chloride and bicarbonate transport, the channel is found in the secretory epithelia of many organs including the lung, pancreas, digestive and reproductive tracts and the skin [12].

The protein consists of two motifs, each containing a membrane spanning domain (MSD), that forms the translocation pathway, and a nucleotide-binding domain (NBD), that hydrolyses ATP [16]. In CFTR, unlike other ABC transporters, the MSD-NBD motifs are linked by a regulatory domain (R) [13] (Figure 1.1).

The CFTR protein is regulated by phosphorylation of the R domain by cyclic nucleotide-dependent protein kinases (e.g. protein kinase A, PKA). The R domain contains several consensus-site serines, and phosphorylation of different sites can have an additive effect on channel activity [17]. Once the channel is activated by phosphorylation, channel gating is driven by ATP binding to its cytosolic NBDs [18], which triggers association of the domains into a stable head to-tail-dimer that occlude two molecules of ATP in their interface [19]. This dimer formation initiates a burst of pore openings, while dimer dissociation is induced by ATP hydrolysis, leading to closing of the pore . There is substantial asymmetry between NBD1 and NBD2 and, where ATP is required to bind both domains, only NBD2 is capable of catalysing ATP hydrolysis [20].

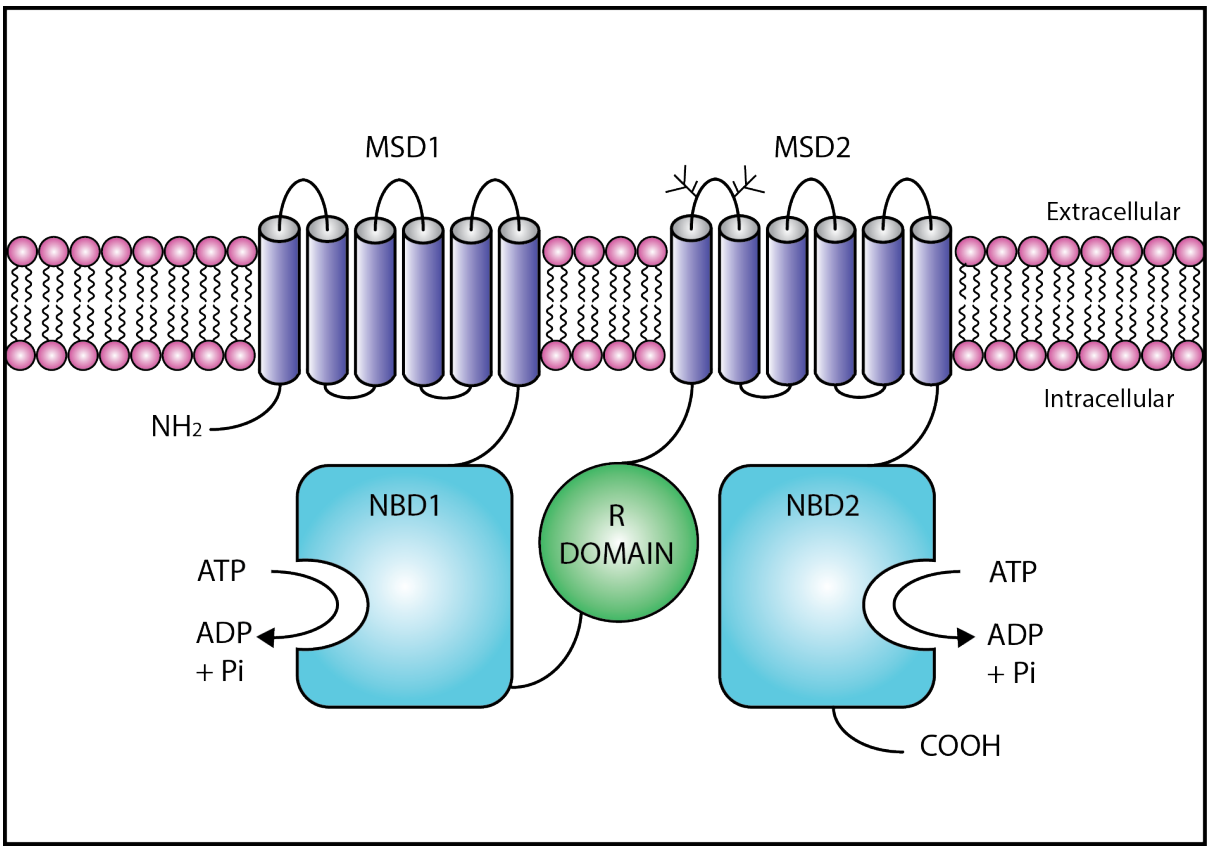


Figure 1.1. Schematic of the CFTR chloride channel. *The channel consists of two membrane spanning domains (MSD-1 and MSD-2), each of which has six alpha helix segments. there are two nucleotide binding domains (NBD-1 and NBD-2) and one regulatory (R) domain. CFTR chloride channel can be activated by phosphorylation of the R domain, followed by ATP binding to NBDs, followed by hydrolysis.*

1.1.3 Cystic Fibrosis pathophysiology in the lungs

Mucosal obstruction of exocrine glands is the principal cause of morbidity and mortality in patients with cystic fibrosis, particularly within the lung. The impaired conductance of chloride and bicarbonate ions across the airway epithelium results in abnormally viscous secretions and impaired mucociliary clearance in the airways. This presence of thick, neutrophil-dominated muco-purulent debris is the primary pathological hallmark of the disease [21]. The thickened mucus provides a perfect microenvironmental niche for bacteria, which permanently colonize the lung, ultimately leading to respiratory failure. Indeed, progressive lung disease is responsible for ~70% of CF deaths [22]. Pathogens such as *Pseudomonas aeruginosa*, *Staphylococcus aureus*, and *Haemophilus influenzae* become well established within the airways of patients with CF and are not effectively eradicated with

antibiotic treatment [23]. There are several conflicting hypotheses concerning the cause of the bacterial colonisation within the lungs, the high-salt hypothesis, the low-volume hypothesis, and the impaired bicarbonate secretion hypothesis, as discussed below.

1.1.1.1 ASL in normal airways

The airway surface liquid (ASL) is formed of two distinct layers. The periciliary layer (PCL) lies adjacent to the airway epithelial cells and surrounds the cilia. In healthy lungs, PCL has a depth of around 10 µm, providing a low viscosity fluid for ciliary beating [24]. The PCL also prevents the mucus layer from coming into contact with the epithelial cell surface, preventing mucus adhesion and the formation of plaques [25]. The mucus layer contains large gel-forming mucins which trap inhaled particles for removal from the lungs by mucociliary clearance (Figure 1.2). The layers combined work as innate immune system and form the first line of defence against bacterial pathogens, facilitating mucociliary clearance from the distal to proximal airways and toward the mouth [26].

1.1.1.2 The high-salt hypothesis

The high-salt hypothesis postulates that there is an increased salt concentration ([NaCl]) within the ASL in CF patients. This increased [NaCl] from ~50 mM to 100 mM at the apical side of the epithelium inhibits the actions of secreted antimicrobials such as human Beta defensin-1 (hBD-1), leaving the CF airways susceptible to infection and allowing bacteria to colonise there [27, 28].

1.1.1.3 Low-volume hypothesis

In contrast, the low-volume hypothesis proposes that normal and CF ASL is isotonic, but in CF has a low volume. In this theory, the lack of functioning CFTR enhances the sodium absorption from the ASL, leading to an intracellular flux of chloride ions (via a non-CFTR pathway) and a net movement of water to maintain isotonicity [29]. ASL volume depletion leads to collapsing of the cilia and cessation of mucociliary clearance. This prevents removal of bacteria from the lungs and allows for their permanent colonisation. Increasing evidence has emerged that the ASL height in CF is much lower than normal (3-4 µm vs. ~7 µm) supporting this hypothesis [30, 31]. Moreover, CF-like lung disease can be initiated in the pig by inducing

hyperabsorption alone, suggesting airway dehydration is the initiating factor in the development of CF-like lung disease [30].

1.1.1.4 Impaired bicarbonate secretion hypothesis

The final hypothesis focuses on impaired HCO_3^- ions rather than Cl^- ions. In healthy lungs, gel-forming mucins begin as large, highly condensed macromolecules compacted into intracellular granules, where high concentrations of Ca^{2+} and H^+ prevent the expansion of the mucins by shielding their repulsive forces of fixed negative charges [32]. Upon release by exocytosis, the mucins transform to immensely expanded polyanionic polymers that cover an protect the apical surfaces of epithelial tissue. This is achieved by removal of Ca^{2+} and H^+ ions exposing the anionic sites on the mucins. This hypothesis postulates that HCO_3^- is crucial to normal mucin expansion by forming complexes with these ions and that, in patients with cystic fibrosis, the lack of secreted extracellular HCO_3^- in impairs calcium removal, preventing normal mucin expansion, and promoting stasis of mucus in the ducts or on the luminal surfaces of affected organs [33].

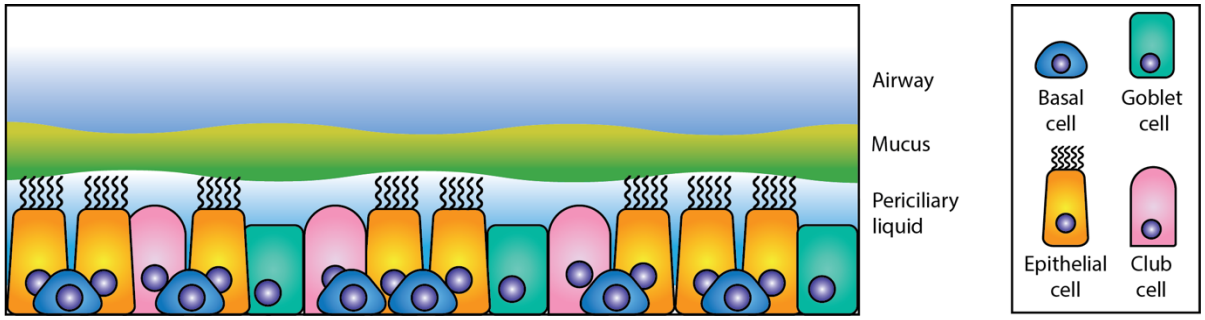


Figure 1.2 Schematic of the airway epithelium. *The airway surface liquid (ASL) is comprised of the periciliary liquid (PCL) and mucus. Epithelial cells, goblet cells, club cells and basal cells are the main constituents of the pseudostratified bronchi.*

1.1.4 CFTR mutations

Over 2,000 mutant variants of the CFTR gene are listed on the Cystic Fibrosis Mutation Database, a website initiated by the Cystic Fibrosis Genetic Analysis Consortium. Of these, 312 are thought to be disease causing [34]. Mutations in the CFTR gene have been grouped into six distinct functional classes according to the method by which they disrupt the synthesis, trafficking and function and stability of CFTR. Recently, it was suggested that a

seventh class should be included in the classification system, effectively splitting class I into two, where in class I, no CFTR protein is produces, while in class VII, no CFTR mRNA is successfully transcribed [35] (Figure 1.3).

Class I mutations have the most severe phenotype, and around 10% of CF patients carry at least one copy of a mutant variant that falls into this class [36]. Included are nonsense mutations; single point alterations that result in premature termination codons (PTCs) in the corresponding mRNA transcript, causing cessation of translation. Examples include Gly542X (where glycine (GGA) at position 542 is mutated to the translational stop codon TGA) (c.1624G>T) , and Trp1282X (c.3846G>A). Canonical splice mutations are also considered class I mutations, including 1717-1G→A (c.1585-1G>A), which destroys a splice site at the 3' end of intron 10 [37]. Additionally, lack of protein production can be caused by insertions or deletions leading to frame shifts. Originally, these mutations were considered class I mutations, however in the more recent classification would be considered class VII [35]. CFTRdele2,3(21 kb) deletes 21,080 bp spanning introns 1-3 of the CFTR gene, resulting in the loss of exons 2 and 3 in CFTR mRNA, thereby producing a premature termination signal within exon 4 [37]. In both of these mutations, there is no mRNA transcription and therefore lack of protein synthesis, and these mutations are considered 'unrescuable' by current modulator therapies.

In class II mutations, a full length CFTR transcript is produced, however missense and in frame deletions disrupt CFTR folding and trafficking to the apical cell membrane. The most common CF-causing mutation, Phe508del, belongs in class II, and around 88% of CF patients on the Cystic Fibrosis Foundation Patient Registry have at least one copy of this mutation [34]. The mutation is caused by the deletion of three nucleotides (CTT) spanning positions 507 and 508, resulting in the loss of a single codon for the amino acid phenylalanine (F) in the NBD1 domain. This causes misfolding of the protein, ER retention and targeting for ER-associated degradation (ERAD) [38]. Another class II mutation is N1303K (asparagine to lysine mutation at position 1303 of NBD2) (c.3909C>G). Like F508del, N1303K is a folding mutation, located at an equivalent but opposite position in NBD2 [39]. The mutation is particularly prevalent in Mediterranean populations, and is the second most common mutation in the southwestern region of France (8.7% of patients harbor at least one copy of this mutation) [40].

In addition, both F508del and N1303K are gating mutations. F508del sits at the interface between NBD1 and the fourth intracellular loop (ICL4), also known as the coupling

helix of TMD2. N1303K, on the other hand, resides in the interface between NBD2 and ICL2, and is thought to stabilize the Q-loop of NBD2 [41]. Both play important roles in CFTR gating because they relay signals from NBDs leading to conformational changes in the TMDs to open and close the gate [42]. Such mutations provide a good example that, while the class system is useful for understanding defects at a cellular level, and many mutations can be unambiguously categorised into certain classes, some mutations are much more complex and fall into several categories.

In class III mutations, full length CFTR reaches the apical cell membrane, however mutations impair the regulation of the CFTR channel, resulting in abnormal gating characterized by a reduced open probability of the channel. G551D, glycine to aspartate at position 551 (c.1652G>A), is a class III mutation. Located in the signature sequence of NBD1, it completely eliminates the ability of ATP to increase the opening rate of the channel [43]. Known as the Celtic mutation, it is the most common class III mutation and occurs in around 4% of CF patients [44].

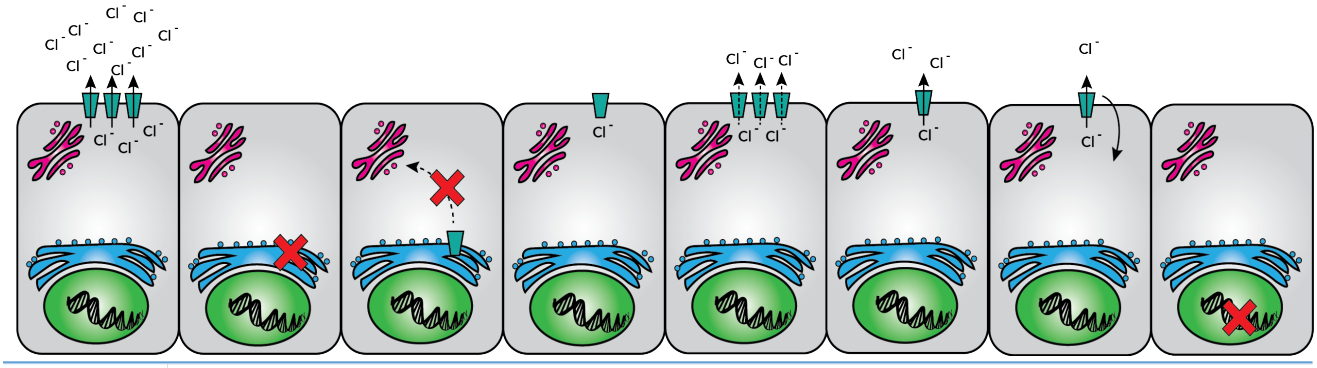
The remaining CF mutation classes are less severe than classes I- III and class VII. Class IV mutations are missense mutations, where substitution of an amino acid leads to changes in CFTR protein structure that forms the pore of the channel. A 'misshaped' pore restricts movement of Cl^- ions through the channel, and these mutations are referred to as conductance defects. R117H, an arginine to histidine substitution at position 117 (c.350G>A), is located in the first extracellular loop (ECL1) between the first and second transmembrane segments, and causes reduced conductance and open probability [45].

Class V are missense mutations that result in aberrant splicing. In such mutations, both alternative and normal transcripts of CFTR mRNA are produced, the latter in minor quantities, and as such there is some residual channel function at the cell surface. 3849+10 kb C > T (c.3178-2477C>T) is a C > T point mutation in intron 19 of the CFTR gene. This mutation generates a cryptic GT splice donor site, leading to the insertion of an 84bp pseudoexon between exons 19 and 20. This pseudoexon contains an in-frame TAA stop codon, which generates a truncated, non-functional protein [46]. Studies using RT-PCR showed considerable variation in the level of aberrantly spliced RNA transcribed from the mutation, with low levels of correct transcripts (0-28%) identified in the nasal epithelium from 10 patients [47]. Increased levels of aberrant mRNA correlated significantly with disease severity, as measured by FEV1 (forced expiratory volume in one second). Further studies have shown

that levels of aberrant splicing can vary among organs of the same patient, with these individuals having extremely variable disease expression ranging from minimal lung disease, male fertility and pancreatic sufficiency, to severe disease in all affected organs [48].

Finally, in class VI mutations, CFTR protein reaches the apical cell membrane, however there is rapid turnover of the channel at the cell surface, where the protein is unstable and quickly removed and degraded [49]. N287Y, where asparagine is replaced by tyrosine at position 287 (c.991A>T) is a class VI mutation located in the second intracellular loop [50].

Many CF patients are compound heterozygous for two CF mutations, with each allele having a different mutation. In such cases, they can be classified by which mutation has the dominant effect. For example, a patient heterozygous for G542X (class I) and F508del (class II) is considered to be placed in class I [51]. Class IV and V mutations are functionally dominant when occurring with class I, II or III mutations, leading to a less severe phenotype in patients harboring these mutations [52].



wt-CFTR	Class I	Class II	Class III	Class IV	Class V	Class VI	Class VII
CFTR defect	No protein	No traffic	Impaired gating	Decreased conductance	Less protein	Less stable	No mRNA
Mutation example	Gly542X, Trp1282X	Phe508del, Asn1303Lys,	Gly551Asp, Ser549Arg,	Arg117His, Arg334Trp,	3272-26A→G, 3849+10 kb C→T	c. 120del123, Gln1412X	dele2,3 (21 kb), Asn287Tyr
Corrective therapy	Rescue synthesis	Rescue traffic	Restore channel conductance	Restore channel conductance	Correct splicing	Promote stability	Unrescuable

Figure 1.3. ***CFTR mutation classes****. CFTR mutations can be grouped into seven functional classes according to their primary molecular defect. Adapted from [53].*

1.2 Additional channels that interact with CFTR

1.2.1 The epithelial sodium channel

The epithelial sodium channel (ENaC) is a heterotrimeric protein channel also located on the apical surface of epithelial cells. The channel is expressed in numerous tissues such as the kidney, male and female reproductive tracts, colon, skin, and lung. ENaC facilitates sodium and water reabsorption from the lumen across the apical membrane into the epithelial cell [54]. The channel is typically composed of 3 subunits, α, β, and γ, encoded by the genes SCNN1A, SCNN1B and SCNN1G respectively. In the brain and testes, an additional δ subunit has been reported (SCNN1D) which forms a functional channel with α, β subunits [55]. Each subunit consists of two transmembrane helices and an extracellular loop, with intracellular N and C termini (Figure 1.4). Mutations in the subunits of ENaC can result in pseudohypoaldosteronism (PHA), a rare autosomal recessive disorder causing a lack of responsiveness to aldosterone characterised by hypertension, hyperkalaemia, and metabolic acidosis [56].

In the lung, the ENaC channel is activated by proteolytic cleavage of the extracellular loops of α and γ- subunits. Both intracellular [57] or extracellular [58] cleavage leads to a conformation change in the channel by releasing inhibitory peptides from the ECLs [59]. This activates the channel by relieving Na^+ self-inhibition and increases the channel open probability [57].

ENaC and CFTR are the primary determinants of ASL hydration status, and therefore mucociliary clearance [60]. In normal airway epithelia, CFTR has an inhibitory effect over ENaC, and up-regulation and down-regulation of ENaC activity in response to CFTR is crucial in maintaining a relatively constant ASL volume [61]. In CF, loss of CFTR function causes an increase in ENaC activity, leading to hyperabsorption of Na^+. This dehydrates the PCL and abolishes mucus transport [25]. Mucus stasis results, providing conditions for intermittent, and then chronic bacterial infections.

Additional proof that heightened ENaC activity leads to a dehydrated ASL and reduction in mucociliary clearance comes from a mouse model that over-expresses ENaC [62]. In these transgenic mice, there is airway-specific overexpression ENaC using a cell-specific promoter, which leads to hyperabsorption of Na^+ ions, ASL dehydration, mucus obstruction

along with chronic inflammation, and early death in these animals [62]. This correlates with the 'low-volume hypothesis' of CF pathophysiology discussed above [29].

Several mechanisms by which CFTR regulates ENaC have been proposed, ranging from altered cellular trafficking of ENaC [63, 64], direct protein-protein interaction between CFTR and ENaC [65], or interaction via cytoskeletal proteins [66, 67]. It is of note that, while loss of CFTR function leads to increased ENaC activity and hyperabsorption of Na^+ the airways, this effect is tissue specific. In the sweat ducts, for example, the loss of Cl^- conductance subsequently reduces the absorption of Na^+ by ENaC, by means of electrochemical effects, with excess salt appearing in the sweat of CF patients [68].

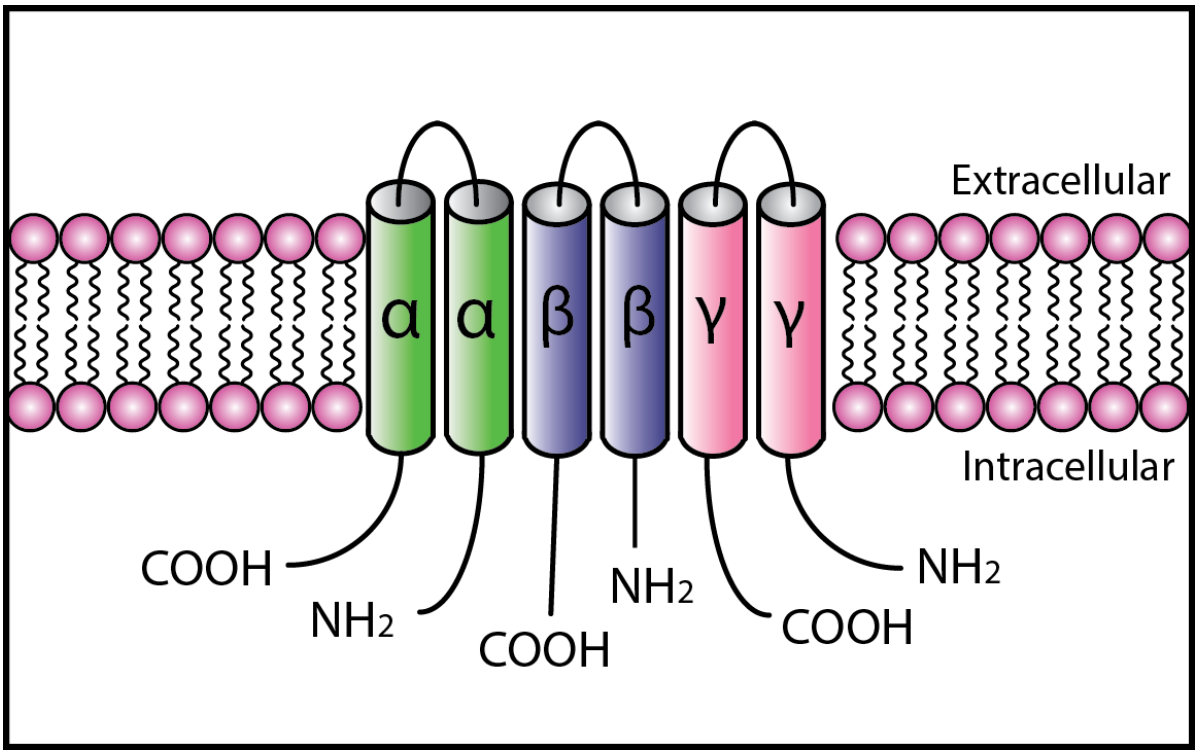


Figure 1.4. Schematic of the epithelial sodium channel (ENaC). *The channel consists of 3 subunits (α, β, and γ or δ), with each subunit formed of two transmembrane domains and an extracellular loop. In human airways, ENaC is activated by proteolytic cleavage of the extracellular loops of α and γ- subunits.*

1.2.2 Calcium-activated chloride channels

Calcium-activated chloride channels (CaCCs) are a class of chloride channels that are activated by intracellular calcium. They have been described in many mammalian cell types, including

excitable cell tissues such as the interstitial cells of Cajal (ICC) of the intestinal tract, where they have electrical pace-making activity triggering peristalsis [69], and in the smooth muscle cells of cardiac tissue, where the opening of CaCCs leads to depolarisation followed by opening of voltage-gated Ca2+ channels (VGCCs), and subsequent contraction [70]. CaCCs are also present in non-excitable tissue, such as the acinar cells of the salivary ducts, and the epithelial cells lining the G.I tract and lungs. In these non-excitable tissues, the channels play a significant role in regulating fluid secretion from cells [71].

Despite their physiological significance, the molecular identity of CaCCs remained unknown until 2008, when transmembrane protein 16A (TMEM16A), a plasma membrane protein with previously unknown function, was associated with calcium-dependent chloride current [72, 73]. Also referred to as anoctamin 1 (ANO1), the channel has eight transmembrane domains, with both N and C termini located in the intracellular face of the membrane (Figure 1.5). CaCC mediated cellular responses are instigated by the stimulation of $P2Y_2$ receptors by ligands such as ATP or UTP, increasing intracellular concentration of Ca^{2+}. The channel is then activated by binding of Ca^{2+} to two glutamic acid residues in the third intracellular loop [74, 75]. Recently, it has been suggested that ubiquitous Ca^{2+} sensor calmodulin (CaM) also regulates channel activity [76].

Belonging to a family that includes ten members in total, TMEM16A is particularly expressed in airway cells, although other members of the gene family, TMEM16F and TMEM16K, have also been detected at low levels in the $TMEM16A^{-/-}$ mouse. However, their residual function is insufficient in maintaining normal airway hydration, with the mice exhibiting neonatal mucus accumulation and a CF-like lung phenotype [77].

The precise mechanism of interaction between TMEM16a and CFTR has been investigated extensively. It has been suggested that physical interaction between TMEM16A and CFTR in the plasma membrane is mediated by PDZ adapter proteins, which play a key role in the organisation of the signalling complex at the cell membrane [78]. This is supported by co-immunoprecipitation of the two channels. Further, deletion of the PDZ-interacting motifs in both channels reduced membrane expression of the proteins [79]. Additionally, there has been growing evidence that CFTR can respond to Ca^{2+}. It was shown that stimulation of purinergic $P2Y_2$ receptors can activate CFTR in airway epithelial cells [80].

In mice, CaCCs are seemingly able to compensate for lack of CFTR protein in the CF mouse ($CFTR^{-/-}$), as shown by their lack of airway disease and no reduction in fluid secretion

[81]. However, this functional compensation does not seem to be present in humans with CF, where ANO1 activity was significantly decreased in CF versus healthy primary bronchial epithelial cells obtained from lung explants [82]. Nevertheless, there is evidence that activation of these alternative Cl^- channels could compensate lack of CFTR and restore ASL height [83, 84].

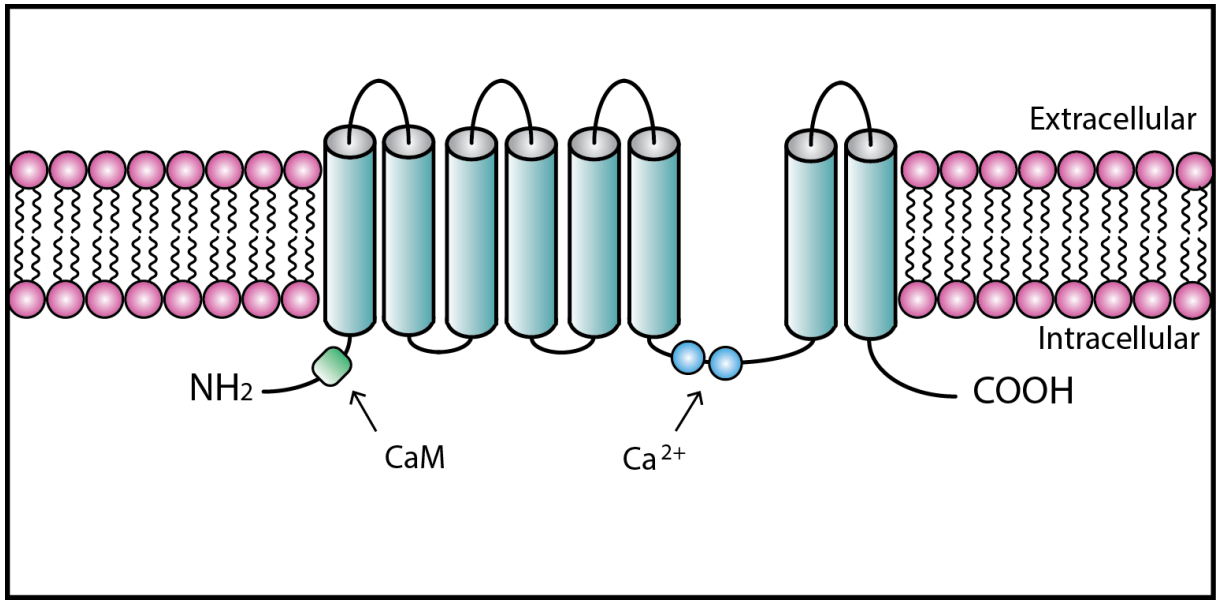


Figure 1.5 Schematic of TMEM16a (anoctamin1). *The channel consists of 8 transmembrane domains, with intracellular N and C termini. In human airways, TMEM16A is activated by binding of cytosolic Ca^{2+} at two glutamic residues in the third intracellular loop.*

1.3 CFTR study model

Disease models play a critical role in the understanding the pathophysiology of any disease, and in the development of therapies. *In vitro* models and animal models of CF have served to further our understanding of the mechanisms associated disease progression and pathophysiology, and also assist in the development of new therapies to treat patients.

1.3.1 *In vitro* study models

The development of immortalised cell lines, generally through introduction of a viral gene that partially deregulates the cell cycle, has been a significant benefit to the study of various human diseases, and immortalisation of cells from CF patients as well as non-CF individuals have been used successfully in CF research. Immortalised epithelial cells from a variety of

tissues and species have been used in CF research, however the airway epithelium has been the principal cell source for immortalised cells because of the significance of the lungs in CF morbidity and mortality (Table 1.1).

Whilst these cells lines are a useful tool in the study of CF, and can maintain some key properties of primary cell tissue such as polarization, tight junctions and ion transport, there are several limitations to the systems. Epithelial cells are generally quiescent, and immortalising these cells for proliferation necessitates the expression of one or more active oncogene [85]. The cells therefore do not present the complete phenotype of parental tissue and can be heterogeneous in long term culture. Further, they are intrinsically unstable with respect to karyotype and phenotype.

***Table 1.1**. Human airway epithelial cell lines used in CF research. Adapted from [85].*

Name/code	Cell type	Genotype	Immortalization	Tight junctions/ TEER (Ω cm²)
IB3-1	CF bronchial	F508del/W1282X	AdV12/SV40	no
AA	Bronchial cells	wt CFTR	hTERT /SV40 LT	~150
KK	Bronchial cells	F508del CFTR	hTERT /SV40 LT	~150
CFT43	Airway cells	F508del CFTR	SV40 LT	100–200
CFT1	Airway cells	F508del CFTR	E6/E7	400–500
CFBE45o-	Bronchial epithelial cells	F508del/R117H	pSVori-plasmid	unknown
56FHTE8o-	Tracheobronchial epithelial cells	WT	pSVori-plasmid	unknown
16HBE14o-	Bronchial epithelial cells	Endogenous WT CFTR	pSVori-plasmid	200–1000
CFBE41o-	Bronchial epithelial cells	F508del/F508del	pSVori-plasmid	100–500
CFTE29o-	Tracheobronchial epithelial cells	F508del/F508del	pSVori-plasmid	no
1HAEo-	Tracheobronchial epithelial cells	WT	pSVori-plasmid	100–1000
9HTEo-	Tracheobronchial epithelial cell	WT	pSVori-plasmid	no
HTE67tsa209o-	Tracheobronchial epithelial cell	WT	tsa209 plasmid	no
CFDEo-/6REP-CFTR	Tracheobronchial epithelial cell	WT-CFTR, RVS promoter	pSVori-plasmid	no

Primary epithelial cells are an obvious choice for an *in vitro* system. These cells were first grown in 1982 using defined culture medium and coating of plasticware with fibronectin [86]. Airway cells can be obtained by brushing of nasal or bronchial epithelium, which will contain basal progenitor cells in addition to the differentiated cells. On submerged culture, differentiated cells will rapidly die off while the basal cells will proliferate. This is followed by

maintenance at an air-liquid interface (ALI) culture which grows to replicate the pseudostratified, ciliated, columnar epithelium found in the airways. In this culture method, the basolateral surface of the cells is in contact with growth medium, whereas the apical side is exposed to air, mimicking the conditions within the human airways and driving differentiation of cells.

Primary epithelial cells grown on ALI culture replicate many of the features of a CF airway including production of thick, sticky mucus, a depleted PCL and failure to respond to cAMP-mediated activators of CFTR, such as forskolin [87, 88]. A key limitation of these basal cells is that they lose their differentiation capacity after two passages, necessitating frequent sampling from CF donors. It may not be possible to source the same donor, and differences in CF modifier genes, of which there are a growing number [89, 90], may affect experimental results. While immortalisation of cells generally results in loss of differentiation, recent advancements in the field have provided new ways of expanding cell proliferative potential and the ability to differentiate. One such method involves culturing basal epithelial cells with an irradiated fibroblast feeder layer and rho-kinase (ROCK) inhibitor, Y-27632 [91, 92]. The feeder layer provides extracellular secretions such as growth factors, while the ROCK inhibitor increases proliferative capacity [93]. A feeder free method involves the dual inhibition of SMAD signalling pathways using bone morphogenetic protein antagonist (DMH-1) and transforming growth factor β antagonist (A-83-01). This can enable robust expansion of primary basal epithelial cells, which remain proliferative after 17 passages [94]. Another approach to extend proliferative capacity is by co-transduction of telomerase reverse transcriptase (hTERT) and the anti-senescent, proto-oncogene B-cell-specific Moloney murine leukaemia virus integration site 1 (*BMI-1*) [95], or even *BMI-1* alone [96]. Through these methods, cells are able to be expanded for at least 20 population doublings, whilst maintaining genetic stability. Further, these *BMI-1* transduced basal cells differentiate on ALI culture with a pseudostratified morphology.

Another *in vitro* model that closely represents essential *in vivo* tissue architecture is the use of organoids as a long-term culture system. Organoids are 3D cellular clusters derived from stem cells that are capable of self-renewal and organisation, and retain functionality of the tissue of origin [97]. The culture technique is based on using culture media containing specific growth factors and an appropriate extracellular matrix such as Matrigel to mimic the *in vivo* environment. Organoid models have been successfully created for a variety of organs

including lung [98], kidneys [99] and brain [100]. Clevers' group were able to apply a CRISPR/Cas9 based gene editing approach to correct the CFTR locus in cultured intestinal stem cells from CF patients and demonstrate restored functionality of the corrected mutation in clonally-expanded organoids [101]. Further, the group previously demonstrated that GFP+ colon organoids could be successfully reintroduced into a superficially damaged mouse colon, where the transplanted cells adhere and readily integrate to the damaged area. Long term (6 month) engraftment was observed with transplantation of organoids derived from a single colon stem cell after *in vitro* expansion. Taken together, this work provides a potential *ex vivo* strategy for gene therapy in patients [102]. Tracheospheres have also been reported derived from airway basal cells that represent the epithelial and mucociliary features of the lung [103]. Recently, the development of nanospheroids has provided another model in which to measure CFTR activity. In contrast to intestinal organoids in which intestinal stem cells are obtained from a rectal biopsy, nasospheroids require a minimally invasive nasal biopsy. Nasospheroids form spontaneously in 2–5 days of culture, and can be maintained for at least 12 weeks. Conversely to intestinal organoids, the apical membranes face the outside and basolateral membranes face the inside [104].

1.3.2 *In vivo* models

Despite an arsenal of *in vitro* CF cell models due to advances in primary cell culturing and modelling, animal models are still needed in order to understand the pathophysiology of the disease, which can help elucidate disease mechanisms and help discover new therapies. Since the discovery of the CFTR gene in 1989, several research groups began to develop murine CF models, initially by making CFTR null ($^{-/-}$) mice [105-108], before creating models with specific CF mutations, such as homozygous F508del mice [109, 110] and G551D [111]. It soon became obvious that despite high similarity between human and murine CFTR sequence (78%), the pathophysiology of the disease in mice differed to humans, where lung disease is not evident in mice [112]. It has been shown that mice express alternative, CFTR independent chloride channels which are able to compensate for the lack of CFTR channel activity [113]. Mice overexpressing the β subunit of ENaC possess a lung phenotype resembling that of CF in humans, with bronchiectasis and poor airway clearance, and are often used as a surrogate model [114, 115].

Animals models that more closely resemble the pulmonary disease progression in humans are therefore needed. Porcine models provide a more robust representation of the human CF phenotype, with 92% CFTR sequence homology between humans and pigs [116]. Using AAV-mediated gene targeting and somatic cell nuclear transfer, Rogers and colleagues generated pigs heterozygous with either the F508del mutation or a null mutation. These heterozygous pigs were then employed to create animals homozygous for the F508del mutation [117]. Maintenance of these pigs requires immediate surgery to alleviate intestinal obstruction, another associated pathology of the condition, and the animals require significant veterinary care [118].

The ferret model has also been used extensively in CF research, particularly in the assessment of lung infection, due to the similarity in lung cell biology and anatomy to humans. Similar to the porcine model, AAV-mediated gene targeting created ferrets with heterozygous CFTR gene disruption, which were bred to create a homozygous model [119]. The ferret's relatively short gestation period, and smaller size when compared to the pig, make it a more cost-effective model to maintain.

Table 1.2 Phenotypic manifestations of CF in humans and selected animal models. *PI; pancreatic insufficiency, MI; meconium ileus. *phenotype may vary between mouse models. Adapted from [120].*

	Spontaneous lung infection	Pancreatic disease	Intestinal disease	Liver and gallbladder disease	Reproduction
Human	Yes	PI	MI	Biliary cirrhosis	Severe vas deferens defect
Mouse*	No	No	Intestinal obstruction, often fatal	No	Reduced fertility in females
Pig	Yes	PI	100% MI	Biliary cirrhosis	Severe vas deferens defect
Ferret	Yes	PI	75% MI	Liver disease	Severe vas deferens defect

1.4 Current CF therapy

Patients with CF today are living much longer, healthier lives than their counterparts 20 years ago, largely due to well-designed treatment plans and an improved understanding of the disease as a whole. Combined with better management and understanding of the disease,

over half of CF patients are now adults, and the life expectancy of a baby born with CF in the UK today estimated to be 47 years [1]. Treatment of the disease requires a multidisciplinary approach to care where doctors, nurses, physiotherapists, dietitians and social workers, who work as a team to tackle this multi-faceted disease. Generally, care is provided in in specialist CF care centres and this has improved the clinical outcome compared with patients in the general community [121].

The primary objectives of CF treatment include maintaining lung function as close to the normal range as possible, preventing and controlling lung infections while managing the associated complications and administering nutritional therapy for normal growth. The latter is largely controlled in most cases through the use of pancreatic enzyme supplements and fat soluble vitamins which the intestines are unable to absorb.

Exercise and healthy diet are strongly encouraged in CF patients. Physiotherapy remains a mainstay treatment of CF, and helps to loosen and remove mucus from the lungs. Inhalation of hypertonic saline can also aid with mucocilliary clearance by temporarily increasing hydration of the airway surface [122]. In addition, several mucolytic agents are available, which further aid the clearance of mucus from the lungs. The heightened inflammatory processes in CF lead to the accumulation and subsequent degradation of neutrophils within the lungs, leading to large amounts of extracellular DNA, which increase the viscosity of sputum. DNases such as Dornase Alfa can cleave this extracellular DNA and air lung clearance [123]. Impaired bicarbonate secretion leads to insufficient mucin unfolding and mucus attachment in the lung [33]. OligoG, a guluronate rich alginate derived from the seaweed *Laminaria hyperborean,* is able to chelate calcium, allowing for detachment of CF mucus [124]. AlgiPharma recently completed a Phase IIB clinical trial of the product (NCT03822455).

Throughout their lives, virtually all CF sufferers develop repeated viral and bacterial infections within the lung, with common pathogens including *Staphylococcus aureus* and *Pseudomonas aeruginosa* [125]. Antibiotics are crucial therefore in the treatment of both acute exacerbations and chronic infections in CF lung disease. These may be administered orally at home but can often require a hospital stay with intravenous administration. CF individuals must spend several hours each day taking their medication, however, in recent years there has been dramatic development of new drugs termed modulators that are revolutionising the treatment of CF, as discussed below.

1.4.1 CF Modulator therapies

Over the last few years there has been rapid progress in the treatment of CF with the emergence of CFTR modulator therapies [126]. Modulators are small molecules designed to target the underlying defect in CF by improving production, intracellular processing, and/or function of the defective CFTR protein channel. They include drugs that act as 'potentiators', which modulate the gating of CFTR at the cell membrane and increase the open probability of the channel. The first potentiator drug, ivacaftor (also known by the trade name Kalydeco®), was approved in 2012. The drug is effective for Class III mutations such as G511D, where the CFTR protein is successfully transported to the apical membrane but for which the gating properties are defective. In clinical trials, the treated patients showed improved lung function, fewer exacerbations and indicated a better quality of life. In lung function tests, absolute FEV1 increased by almost 10% [127]. Ivacaftor is now FDA approved for 37 different CF causing mutations, including splice mutations and residual function mutations, along with gating mutations, enabling effective treatment of 5-6% of CF patients [128]. Further, it is now approved for use in CF patients as young as six months of age [129]. Proteostasis Therapeutics drug PTI-808 (NCT03500263) is another potentiator which is currently in phase II clinical trials in combination with other modulators. Vertex Therapeutics also have an altered from of ivacaftor in phase II trials (NCT03911713).

'Correctors' are another class of modulator, which enhance the folding and trafficking of the CFTR protein to the plasma membrane. Lumacaftor is a corrector molecule which is able to rescue the CFTR protein from the endoplasmic reticulum through enhancing interactions between the NBD1, MSD1, and MSD2 domains [130]. It was thought it might benefit patients with the most common F508del mutation. However, this mutated protein displays multifaceted defects including aberrant glycosylation, misfolding leading to retention in the endoplasmic reticulum and poor trafficking to the plasma membrane, as well as decreased stability leading to rapid protein turnover, and finally defective gating [131-134]. This renders single correctors ineffective [135] but combinations of potentiators and orrectors, as in Orkambi®, a combination of ivacaftor and lumacaftor, is more effective in maintaining lung function [136]. Orkambi® is approved in the USA and EU for use in treating patients homozygous for the F508del mutation. The benefits of Orkambi® are moderate in

comparison to those achieved with lumacaftor for patients with class III mutations, with an with an improvement in absolute FEV1 of 3% [137]; meaning that while the drugs are approved for use in several European countries, they are not currently funded by healthcare systems. However, as of 24th October 2019, Orkambi® has received NHS funding for use in the UK [138]. Symdeko® is a new combination of drugs approved for patients in the USA that have at least one copy of F508del. It is comprised of two modulators; tezacaftor, a potentiator, and the corrector lumacaftor, and showed an increase in FEV1 of 4% [139]. Further, Symdeko® showed reduced side effects when compared to Orkambi®, which range from ocular and hepatic issues to dyspnoea, which can occur during initiation of treatment [140]. A triple combination therapy, also from Vertex pharmaceuticals, involves the use of two correctors with different mechanisms of action, elexacaftor and tezacaftor, along with potentiator ivacaftor, and has completed two phase III clinical trials, with improvements in FEV1 of up to 13% in patients heterozygous for F508del and one minimal function mutation (NCT03525548) (NCT03525444) [141]. The triple combination has recently received FDA approval for use in patients aged 12+ with at least one copy of F508del mutation under the name Trikafta® [142]. Abbvie are also recruiting a phase II trial with their corrector ABBV-2222 and potentiator, ABBV-3067 (NCT03969888). Flatley Discovery Labs have their own corrector, FDL169, entering phase II clinical trials (NCT03756922).

Another class of modulators, called 'amplifiers', enhance translation of CFTR mRNA, increasing the amount of CFTR protein within the cell for other modulators, such as ivacaftor and lumacaftor, to work on. PTI-428 from Proteostasis Therapeutics is in Phase 2 clinical trials for use in combination with ivacaftor (NCT03258424), or with other Proteostasis modulators (NCT03500263).

Table 1.3. Currently licensed CF modulator therapies

Drug	Modulation	CF mutation
ivacaftor (Kalydeco®)	potentiator	Class III, + 28 specified mutations
lumacaftor/ivacaftor (Orkambi®)	corrector + potentiator	Two copies of F508del
tezacaftor/ivacaftor (Symdeko®)	corrector + potentiator	Two copies of F508del, + 26 specified mutations
elexacaftor/tezacaftor/ ivacaftor (Trikafta™)	2x corrector + potentiator	At least one copy of F508del

While the different modulator drugs are useful for treating CFTR mutations that produce a defective protein, individual with Class I mutations do not produce a full-length protein due to point mutations or indels, creating premature termination codons (PTC), while those in Class V have splice mutations or promoter mutations that result in lower mRNA levels which results in lower protein production. So, in spite of the recent advancements in drug therapies, there is still a great need for new treatments that are amenable to a larger cohort of CF individuals. Read-through drugs such as Ataluren, which was identified by chemical library screening, insert an amino acid at internal premature termination codons. The drug is conditionally approved by the European Medicines Agency for patients aged 5 years and above with Duchenne's muscular dystrophy caused by a nonsense mutation [143]. The drug entered clinical trials for CF Class I mutations although recent Phase 3 trials (NCT02456103) failed to show efficacy (0.6% change in predicted FEV1) and so the study was terminated [126]. Additionally, ELX-02 is a synthetic aminoglycoside optimized as a translational read-through drug (TRID). A phase 2 study is planned to test the safety and tolerability of ELX-02 for patients with at least one copy of the class I mutation G542X (NCT03309605).

Investigations into therapies exploiting alternative the chloride channels TMEM16A [144] and SLC26A9 [145] are in progress. Another strategy is to reduce sodium uptake by the epithelial sodium channel (ENaC) [146, 147]. These strategies are designed to correct the ion imbalance, improving hydration of the airway and so reducing mucus plugging and enhancing mucociliary clearance.

1.5 Gene Therapy

Gene therapy involves the therapeutic delivery of nucleic acid into an individual's cells in order to treat a specific genetic disease. CF is a monogenic, recessively inherited pulmonary disorder and so it has long been regarded as particularly amenable to gene replacement therapy [148]. The lung is directly accessible from the external environment, providing an attractive route of delivery by inhalation after nebulisation of the therapeutic formulation. A further favourable feature is that it has been shown that as little as 5% expression of normal levels of CFTR corrects the ion transport defect *in vitro* [149], however that level increases to 25% for correction of mucus transport in a human epithelial model [150]. Epithelial studies in ALI cultures showed that epithelia containing 20% of wild type cells mixed with CF cells generated approximately 70% of the transepithelial chloride ion current of a fully wild-type epithelium [151].

Gene therapy approaches have been evaluated over the last 25 years by delivery of CFTR cDNA by viral or non-viral vectors, yet, despite the advantageous features, no clinically effective therapy has emerged to date [152]. It has become clear that the lung is far more resistant to gene transfer than initially anticipated resulting in the low levels of transfection observed. The difficulties of pulmonary gene transfer should not have been so surprising as the lung has evolved and adapted to provide a barrier to viral and bacterial infections that is equally effective in eliminating viral vectors and non-viral nanoparticles.

Genetic therapies, however, remain attractive as they tackle the basic molecular defect of the disease rather than treating symptoms, with the potential for long-term therapy for all patients regardless of mutation class. Moreover, in recent years there have been several outstanding examples of clinical efficacy of *in vivo* gene therapy for other genetic diseases including haemophilia, restoring Factor IX levels sufficient to reduce bleeding episodes [153, 154], and retinopathies, with the first FDA-approved gene therapy product, Luxturna by Spark Therapeutics available for treatment of inherited retinal dystrophy caused by biallelic mutations in RPE65 [155].

In recent years, improved understanding of the biology of the lung and its barriers, such as the periciliary brush [156], mucus biology [157] and epithelial biology is enabling the rational development of more efficient viral vectors as well as improved nanoparticles for non-viral gene transfer. Further, nucleic acid technologies including minicircle DNA, siRNA and

mRNA, have also improved so that higher and more sustained levels of correction and CFTR expression may be anticipated. Viral and non-viral delivery methods for CF gene therapy are discussed below.

1.5.1 Viral vectors

Viral vectors can be used to deliver genetic material into cells, and have been used in gene therapy for over two decades. With regards to CF, the earliest clinical study involving viruses was in 1994, when CFTR cDNA was administrated by adenovirus, where E1 region was deleted to render the virus replication incompetent [158, 159]. In the early 2000s, several clinical trials took place, including the successful treatment of several X-linked severe combined immunodeficiency disease (X-SCID) patients, marking a major milestone in the field of gene therapy. This *ex* vivo trial used pseudotyped gammaretroviral vector to deliver the γc transgene to $CD34^+$ cells harvested from the bone marrow of patients, before infusion back to the patient [160]. Viral integration, however, also presents risks of insertional oncogenesis should the transgene integrate in the location of a proto-oncogene. Such was the case in the X-SCID trial, where integration at the *LMO2* locus caused T-cell leukaemia in two of the five patients, discovered two years after the trial took place [161]. Elsewhere, another patient with ornithine-cytosine transferase (OCT) deficiency disease died as a result of a severe immune reaction of administration adenovirus vector on a clinical trial [162]. These events led to the suspension of numerous clinical trials involving hundreds of patients, and halted progress within the field of gene therapy. However, advances in the development vector systems, and increasing clinical experience has once more advanced the use of viral vectors in the treatment of human disease.

1.5.1.1 Lentiviral vectors

Lentiviral (LV) vectors integrate into the genome, thus providing the benefits of long term expression, potentially for the lifetime of the epithelial cells, which in the human airway epithelium could be periods of weeks to months. LV vectors, unlike gamma retroviral vectors, can transduce non-mitotic cells, such as the airway epithelium. The latest generation of LV vectors, with self-inactivating (SIN) long terminal repeats that reduce their ability to activate oncogenes in the genome, have so far proven much safer. Pseudotyping of viruses, by using

foreign viral envelope proteins to alter host tropism, or increase the stability of the particles, has also been extensively investigated [163, 164].

LV vectors have, so far, not been tested in clinical trials for CF but they are under development. For example, a vesicular stomatitis virus glycoprotein G (VSV-G) pseudotyped lentiviral vector, which is administered after dosing with a mild detergent, lysophosphatidylcholine (LPC), to disrupt tight junctions, has demonstrated promising levels of gene transfer in marmosets, a non-human primate model [165]. However, the use of LPC remains controversial, and the clinical translation of disrupting the tight junctions of CF patients raises safety concerns. Given the chronic lung infections CF patients harbour, opening tight junctions may leave patients susceptible to systemic infiltration of pathogens and subsequent life-threatening sepsis [166]

An aerosolised feline immunodeficiency virus (FIV) lentivirus pseudotyped with a GP64 envelope protein nebulised to CF new-born pigs was reported to provide partial correction of the CF defect [167]. More recently, a simian immunodeficiency virus (SIV) derived lentiviral vector, pseudotyped with Sendai virus fusion protein (F) and Hemagglutinin/Neuraminidase (HN) envelope proteins was reported to demonstrate stable transduction of murine airway epithelium for a year [168] and are in preclinical development [169]. The UK CF Gene Therapy Consortium have announced a partnership with Boehringer Ingelheim and Oxford BioMedica to translate the viral vector into clinical trials [170].

1.5.1.2 Adeno-associated viral vectors

AAV vectors have a very low chromosomal integration frequency but, achieve long-term, episomal transgene expression through the formation of stable concatameric structures of the vector genome [171]. AAVs have become the most widely used vector for *in vivo* gene delivery, including the first EU-approved gene therapy product, Glybera, for lipoprotein lipase deficiency a metabolic disease.

CF gene therapy has been attempted with AAV, particularly with serotypes that have a greater affinity for airways, such as rAAV2 [172, 173], but transgene expression was low or undetectable, even after repeated dosing. Current AAV vector development for CF is focused on improving lung epithelial transduction efficiency by viral capsid selection and promoter development [174], as well as naturally occurring alternative capsid variants [175].

The limitations of AAV vectors include their packaging capacity of 4.7 kb, which necessitates using a truncated version of CFTR, and their high degree of immunogenicity which limits the efficacy of repeated administration in the lung. Exposure to AAV is widespread with more than 90% of the population having pre-existing immunity to AAV2 [176]. CF patients harbour neutralising antibodies to several strains of AAV, so it is imperative this issue is considered when developing AAV–based therapies for CF [177].

1.5.1.3 Bocavirus vectors

Recently, a new approach described the use of human Bocavirus-Type-1 (HBoV1) derived vectors for the transduction of human airway epithelial cells grown on ALI culture [178]. Like AAV, HBoV1 are human parvoviruses, and display high tropism for the apical membrane of human airway epithelial cells, with a larger capsid than AAV. This enables the packaging of full length CFTR in an rAAV2 genome along with a full length, strong CMV-β-actin promoter into the capsid of HBoV1 to produce a chimeric rAAV2/HBoV1 vector comprising genetic material from both the rAAV2 and HBoV1 viruses [178]; AAV2 itself can only accommodate a truncated CFTR with shortened, synthetic promoters of lower potency. The chimeric vector have been successfully used for the transduction of ferret airways [179] and viral production optimised to achieve higher yield going forward to preclinical studies [180].

1.5.2 Non-viral vectors

Non-viral vectors such as liposome and polymers provide a method of targeted delivery and controlled release of therapeutic agents such as drugs and nucleic acids to cells. Lipid based vectors, particularly cationic lipids, have long been investigated as a non-viral vector for gene delivery due to their favourable pharmacokinetic profile and ease of production. Liposomes are a spherical lipid bilayer, mostly comprised of phospholipids with a hydrophilic head group and hydrophobic tail. The positively-charged head groups of cationic lipids can facilitate spontaneous electrostatic binding with negatively charged phosphate groups on DNA molecules, forming entropically favourable nanoparticles [181]. The net charge of the complexes are determined by the ratio between the cationic lipids and nucleic acids, which, in turn, affects the size, stability, and structure of the particles [182]. The nanocomplexes can interact with the cell surface, which leads to internalization, typically by endocytosis.

The first application of cationic lipids as gene vectors was in 1987, where the cationic lipid DOTMA (1,2-Di-((Z)-octadec-9-enyloxy)-N,N,N trimethyl ammonium propane chloride) was found to deliver a DNA plasmid containing chloramphenicol acetyltransferase, used as a reporter system for tissue expression, with greater efficiency than existing methods, including dextran mediated delivery or calcium phosphate precipitated DNA [183]. There are now many commercially available, widely used lipids designed and synthesised specifically for transfection, with different structural aspects such as head group and hydrocarbon tail length. A neutral, 'helper' lipid such as DOPE (Dioleoyl L-α phosphatidyl ethanolamine) is now often formulated with cationic lipids because of its membrane stabilising properties, which are believed to aid endosomal escape [184]. Cholesterol, an important component of biological cell membranes, can be added to increase circulation time of lipid nanoparticles by imparting stability to the complexes [185, 186]. Meanwhile, PEGylation of liposomal particles, (with poly(ethylene glycol)), can prevent aggregation of nanocomplexes, with particles remaining small and discrete both outside and within cells, whilst conveying a stealth coating on the particles, enhancing serum stability [187, 188].

Non-viral delivery methods have several advantages over viral vectors. Firstly, the systems are not likely to elicit an immune response, which is important when repeated doses need to be administered, as in the case of CF. In addition, non-viral vectors have the capacity to deliver molecules irrespective of size. Combined with their cost-effectiveness and availability to be produced on a large scale [189], these systems have increased in popularity for gene therapy in recent years.

1.5.2.1 Non-viral vectors for CF gene therapy

Gene therapy with non-viral vectors, particularly liposomes [190-196], but also compacted, PEGylated polylysine nanoparticles [197], have been investigated extensively in clinical trials for CF. The first non-viral trial in the UK took place in 1994, where DC-chol:DOPE (3ß-[N-(N',N'-dimethylaminoethane)-carbamoyl]cholesterol/DOPE) was delivered to CF and non-CF patients to confirm the safety of the formulations for inhaled delivery, and to test whether the particles themselves altered the biophysical properties of the airways [198], with the view of using these particles for CFTR cDNA delivery in the future. Indeed, these nanoparticles were used by Hyde and colleagues in 2000 to establish safety and tolerance of three repeat administrations of CFTR cDNA to the nasal epithelium of CF patients. There was no evidence

of toxicity, inflammation or an immune response toward the nanoparticles, and six of ten patients treated showed evidence of CFTR gene transfer [199].

In 2001, the UK Cystic Fibrosis Gene Therapy Consortium was formed by three university groups with the aim to advance CF research with the goal of bringing CF gene therapy to the clinic. Early pre-clinical work involved the development of a codon-optimised, CpG-free plasmid, which allowed for sustained transgene expression with the absence of inflammation [200]. The liposomal vector GL67A, comprised of the cationic lipid GL67 (Genzyme lipid 67A/1,2-Dioleoyl-sn-Glycero-3-Phosphoethanolamine), DOPE and DMPE-PEG5000 (1,2-dimyristoyl-sn-glycero-3-phosphoethanolamine-N-methoxy(polyethylene glycol)-5000), which had been used in mice [201] and in healthy humans [202], was selected as a vector for future use in clinical trials.

The first clinical trial (phase I/IIa) (NCT00789867) using GL67A to deliver the plasmid pGM169, which expresses CFTR under the control of the human cytomegalovirus enhancer/elongation factor 1α sequence [200] and a modified EF1a promoter [203], aimed to evaluate the safety outcome of a single dose with three volumes of nebulization (5, 10 and 20 ml). Thirty five CF patients, received a nebulised dose of pGM169/GL67A formulation, and 5 ml was selected as the optimal dose for the next phase of the trial [204]. The phase IIb trial (NCT01621867) commenced in 2012 on 140 patients, receiving either 5 ml of pGM169/GL67A or a saline placebo every 28 days in a 12-month study. The findings from the study showed a significant, but modest improvement in FEV1 (3.7%) compared to placebo, associated with a stabilisation of lung function, while the safety of repeated delivery was validated [205]. The authors concluded that further improvements to the current formulation were needed before the therapy would be suitable for clinical care.

1.5.3 *In vitro* transcribed mRNA

In vitro transcribed messenger RNA (mRNA) as a template for CFTR protein production is an approach to CF therapy that could be effective in all CF patients regardless of mutation type [206, 207]. The structure of mRNA from 5’ to 3’ comprises the 5’cap, 5’ untranslated region (UTR), coding region, 3’ UTR and poly adenine (polyA) tail, and the structure of each can be readily optimised. Chemical modifications, particularly chemically-modified bases such as pseudouridine, methylcytosine and others, have resulted in mRNA with improved stability,

translational efficiency and reduced immunogenicity [208]. Equally important have been the developments in the untranslated regions of the mRNA template, including the 5'-capping procedure using anti-reverse capping reagents (ARCA), capping analogues that reduce the rate of decapping [209], the structure of the 5' and 3' untranslated regions (UTRs) and poly A region [210]. The coding region itself may be optimised for higher levels of expression by codon optimisation in the template DNA; this ensures use of codons for which there is a higher concentration of tRNA with the relevant anti-codon although [211].

CFTR mRNA therapies have several important advantages over plasmid-based CF gene therapies. Translation of mRNA in the cytoplasm means that, unlike plasmids, the nuclear envelope is not a barrier to transfection and so expression levels are potentially much higher. From a safety perspective, mRNA cannot integrate into the genome and so the risk of insertional mutagenesis is nullified, in contrast to plasmid DNA where there is a low but significant risk of insertional oncogenesis or germline transmission. Production of mRNA is relatively simple and inexpensive without the need for bacterial fermentation processes required for plasmid production (apart from providing the plasmid template for *in vitro* transcription) and so reduces the risk of bacterial endotoxin contamination.

Although chemical modifications to mRNA have reduced inflammatory toxicity as well as improved stability, mRNA expression is more transient than plasmid-based expression, necessitating administration at more frequent intervals than plasmid-based expression vectors although this may not be a limiting factor depending on protein stability. For example, delivery of surfactant protein B (SpB) mRNA to murine lungs in a mouse model of SpB deficiency led to restoration of 71% of normal levels of SpB from twice weekly administration to the airways, which was sufficient to achieve complete survival of the mice for the duration of the experiments [212]. CFTR levels of approximately 25% were reported to be sufficient for restoration of mucociliary clearance [150] and wild type CFTR protein, once in the apical epithelial membrane, has a half-life of at least 24 hours [133] with others reporting more than 48 hours [213]. Thus, it is likely that a single dose of CFTR mRNA would lead to CFTR channel activity of at least several days implying a repeated dosing requirement of approximately weekly to maintain membrane CFTR protein levels and ion transport activity.

While naked mRNA delivery by intradermal injection for vaccination purposes has been reported [214, 215], most applications of mRNA, particularly for lung delivery by nebulisation, will demand a nanoparticle formulation since even chemically modified mRNA

remain susceptible to nuclease degradation and would be unlikely to penetrate the mucociliary barriers to epithelial transfection discussed above. Nanoparticle-mediated delivery *in vivo* has been described for pulmonary delivery [206, 216, 217] but further developments are required for optimal lung delivery. The first CFTR mRNA clinical trial, drug MRT5005 by Translate Bio, is currently in phase I clinical trials (NCT03375047) for patients regardless of mutation type.

1.5.4 Barriers to transfection in the CF lung

The lung has evolved and adapted to provide a barrier to viral and bacterial infections that is equally effective in eliminating viral vectors and non-viral nanoparticles. Challenges to delivery include nebulisation and deposition in the correct region of the lung followed by mucus penetration through to the apical membrane of the airway epithelium, then endocytic uptake. The favoured method of pulmonary delivery of gene therapy formulations is by inhalation as this is minimally invasive and allows pain-free access. Several studies have demonstrated the stability of nanoparticles to nebulisation [218-220]. Nanoparticle formulations can be labelled with 99m-technetium for scintigraphic imaging of lung biodistribution which allows assessment of the region of deposition in the lung. Optimised aerosol conditions combined with intubation was shown to direct deposition to the central and lower airways in a normal pig, where CFTR is most abundantly expressed [219].

Once deposited in the airways by nebulisation, nanoparticles must then penetrate through the airway surface liquid (ASL) which comprises mucus and the periciliary liquid layer (PCL), the watery layer that bathes the cilia [221]. Mucus, a gel-like layer containing mucin glycoproteins, entraps inhaled particles, including viruses, bacteria, pollutants and potentially nanoparticles, which are then removed rapidly in a distal to proximal direction by ciliary beating. CF mucus may present more of a physical barrier as it is thicker and stickier than non-CF mucus. Impaired CFTR-mediated bicarbonate secretion lowers the ASL pH and leads to retention of calcium ions by the mucins, maintaining them in a compacted form rather than the open-mesh network of normal mucus [222, 223]. As aforementioned, mucus viscosity is further contributed by chromosomal DNA released by dead inflammatory cells such as neutrophils [224]. Approaches to overcome the mucus barrier include treating patients in early childhood before the onset of mucus thickening, the use of mucolytics such as N-acetyl

cysteine to mobilise the mucus [225], magnetofection to pull paramagnetic nanoparticles through the mucus layer by a magnetic field [226], while attention to the chemical properties of the nanoparticles themselves may enable them to better penetrate mucus. For example, a dense coating of PEG on the surface of the DNA nanoparticles was shown to reduce charge-mediated interactions with mucus components, improving penetration of mucus [227]. Once through the mucus, the periciliary brush presents a further barrier which is almost impenetrable to particles larger than 40 nm [156], thus, the production of smaller particles would likely be beneficial to transfection efficiency.

There are many diverse cell types in the epithelium including ciliated epithelial cells, goblet and club cells, as well as macrophages and neutrophils. Targeted transfection of the epithelial cells that express CFTR is required to maximise efficacy of treatment. CFTR expression appears to be most evident in submucosal glands [228] but the narrow opening to the gland largely excludes gene therapy vectors [229]. Thus, gene therapy strategies are usually designed to transfect the surface ciliated epithelium. *In vivo* studies demonstrated that liposomal transfection of the nasal epithelium in CF mice, which lacks submucosal glands, was sufficient to correct the electrophysiological defect [230, 231] while correction of human CF epithelial cells *in vitro* by retroviral transduction restored chloride transport across the polarized epithelium *in vitro*.

A further challenge is that epithelial cells are terminally differentiated with a half-life of approximately 6 months in mice [232] although epithelial turnover may be shorter in the CF lung. Transfecting terminally-differentiated cells such as the airway epithelium, presents two problems, the first of which is that plasmid-based, CFTR gene transfer is inefficient due to the highly selective nature of uptake through pores in the nuclear envelope limiting nuclear uptake of plasmid [233, 234]. This barrier is likely to be one of the main reasons why clinical gene therapy studies with liposomal nanoparticles have demonstrated low levels of transgene expression. The second issue is that transfected cells have a limited life-span of a few weeks and so, periodically, the therapy will have to be re-administered.

1.5.5 Cellular therapy

Cellular therapy could be an alternative therapeutic approach for CF, either by correcting the patient's own stem cells *ex vivo* and re-administering them, or by the use of allogeneic stem cells from a healthy donor. Gene editing of the basal cell progenitor cell population which

resides deep within the lung tissue would provide a longer-term therapy, but targeting of these cells would require preconditioning to create a niche for engraftment of engineered basal cells with agents such as polidocanol [235], by irradiation [236] or sulphur dioxide treatment [237], all of which would be problematic clinically in CF patients.

Proof of concept was provided by Coraux and colleagues, who analysed the capacity of human adult basal stem cells to restore a well-differentiated and functional airway epithelium [238]. Adult basal cells were obtained by cell sorting adult airway tissue, before seeding onto denuded rat tracheae, which were then grafted into nude mice for 35 days, or onto ALI culture membranes. They found the cells were capable of reconstituting a fully differentiated epithelium, both *in vivo* in xenografts and *in vitro* in ALI cultures, concluding that these basal cells are at least surface transit-amplifying epithelial cells. Similarly, Reynolds *et al.* showed that by using a mixture of human tracheobronchial epithelial tissue stem cells (TSCs) and progenitor cells, they were able to repopulate an injured airway epithelium in mice. Two weeks after transplant, the TSCs/progenitor cells had differentiated into the three major epithelial cell types; basal, secretory and ciliated [239].

Bone marrow provides a source of various stem cells, including hematopoietic stem cells and mesenchymal stromal cells (MSCs), and have been used for cellular therapy in the combat of various blood and bone disorders for several decades [240, 241]. To study the benefit of stem cells from bone marrow in CF, researchers optimised the delivery of WT bone marrow cells, and found that they contributed to apical CFTR expression in $CFTR^{-/-}$ mice. This reduced *Pseudomonas aeruginosa* lung infection and increased survival of $CFTR^{-/-}$ recipients [241]. A safety study (phase I) of 15 CF patients aiming to test the safety of delivery of allogeneic human mesenchymal stem cells (MSCs) (NCT02866721) is expected to be completed in December 2019. The study will focus firstly on safety, and report any alteration of disease state, including infection and inflammation biomarkers, and FEV1%.

Manipulation of cells *ex vivo* has several advantages over an *in vivo* approach. In an adult CF patient, thick sticky mucus provides a barrier to reaching the basal cells, which is circumvented in an *ex vivo* approach. Further, there is potential to sort cells before re-transplantation, improving the safety and efficacy of the therapy.

1.6 DNA repair

The maintenance of genome stability and accurate transmission of genetic material to daughter cells are of utmost importance. The genome of our cells is continuously damaged due to normal processes within cells, including by-products of metabolism such as reactive-oxygen species (ROS), or due to exogenous factors such as exposure to radiation, UV light, or chemotherapeutics [242]. Damage presents itself in several different ways, including base lesions, inter and intra-strand crosslinks, and both singles and double-stranded breaks (DSB) [243]. In DSBs, the phosphate backbones of each complementary DNA strand are broken simultaneously, causing one of the most cytotoxic forms of lesion [242]. Studies estimating the frequency of spontaneous DSBs suggest that they occur as often as one DSB per 10^8 bp per cell cycle [244, 245]. These types of damage are all hallmarks of cancer and can drive tumorigenesis. Failure to repair even a single DSB can result in cell death [246]. There are two mechanistically distinct repair pathways which have evolved to repair these DSBs: non-homologous end joining (NHEJ) and homology-directed repair, as discussed below.

1.6.1 Non-homologous end joining (NHEJ)

The NHEJ cellular repair pathway involves modification and ligation of the broken DNA ends with very little or no homology, often creating small insertions or deletions (indels) [242]. NHEJ can occur in any stage of the cell cycle, and is particularly prevalent in the G_1 phase, where lack of cyclin-dependent kinase (CDK1) activity prevents homologous recombination by inhibiting ssDNA resection [247]. The NHEJ repair pathway can be divided into two subclasses: canonical non-homologous end joining (c-NHEJ) and alternative NHEJ (alt-NHEJ).

The c-NHEJ pathway is the predominant repair pathway in mammalian cells, and is initiated by the binding of protein Ku, a heterodimer of two related proteins, Ku70 and Ku80, to the break site, which protects and stabilises the DNA ends [248]. Ku localises to the site of DSBs within seconds of their creation [249], and binds to the sugar backbone of the DNA rather than the bases, explaining its sequence independent binding [248]. Once bound, Ku serves as a scaffold to recruit other proteins required for DSB repair. Artemis, which has endonuclease activity on 5' and 3' overhangs, is phosphorylated by the catalytic subunit of DNA–protein kinase catalytic subunit (DNA-PKcs), and blunts the ends of the DSBs, preparing them for ligation [250]. DNA polymerases μ and λ (Pol μ and Pol λ) can also interact with Ku to

resynthesise missing nucleotides and extend the DNA ends, another processing mechanism to prepare the ends for ligation [251]. DNA ligase IV (ligIV) is also recruited to the scaffold, and is stabilised by X-ray cross-complementing protein 4 (XRCC4) [252]. XRCC4-like factor (XLF) then stimulates the ability of the XRCC4-ligIV complex to ligate and join even incompatible ends [253].

In contrast, alt-NHEJ is seemingly independent of the above factors, and is considered a 'back-up' mechanism of DNA repair when other methods fail, with much slower kinetics [254]. Alt-NHEJ frequently often uses short stretches of complementary nucleotides, situated either at the ends of the DNA break, and the pathway is also referred to as microhomology-mediated end joining (MMEJ). Poly(ADP-ribose) polymerase 1 (PARP1) acts in place of Ku, identifying and binding to DNA ends and recruiting subsequent proteins [255]. Mre11, Rad50 and Nbs1 proteins, which form the MRN complex then bind to the DNA, along with Ctp1. Together, they function to promote strand resection, preparing the DSBs for ligation [256, 257]. Error-prone polymerase θ (Pol θ), is also involved in end processing [258]. In contrast to c-NHEJ, it is ligase III (ligIII) that is responsible for ligating the ends of the DSB and therefore DNA repair [259].

The NHEJ pathways are error-prone on two counts; firstly, there is no built-in mechanism to ensure the restoration of the original sequence at the site of DSB, leading to sequence alterations, random nucleotide exchange, and indels at the junction. Secondly, there is no inherent mechanism to ensure restoration of the original DNA molecule as a whole, and so the system can, in principle, join any DNA ends irrespective of molecular origin, which can lead to chromosomal translocations [260].

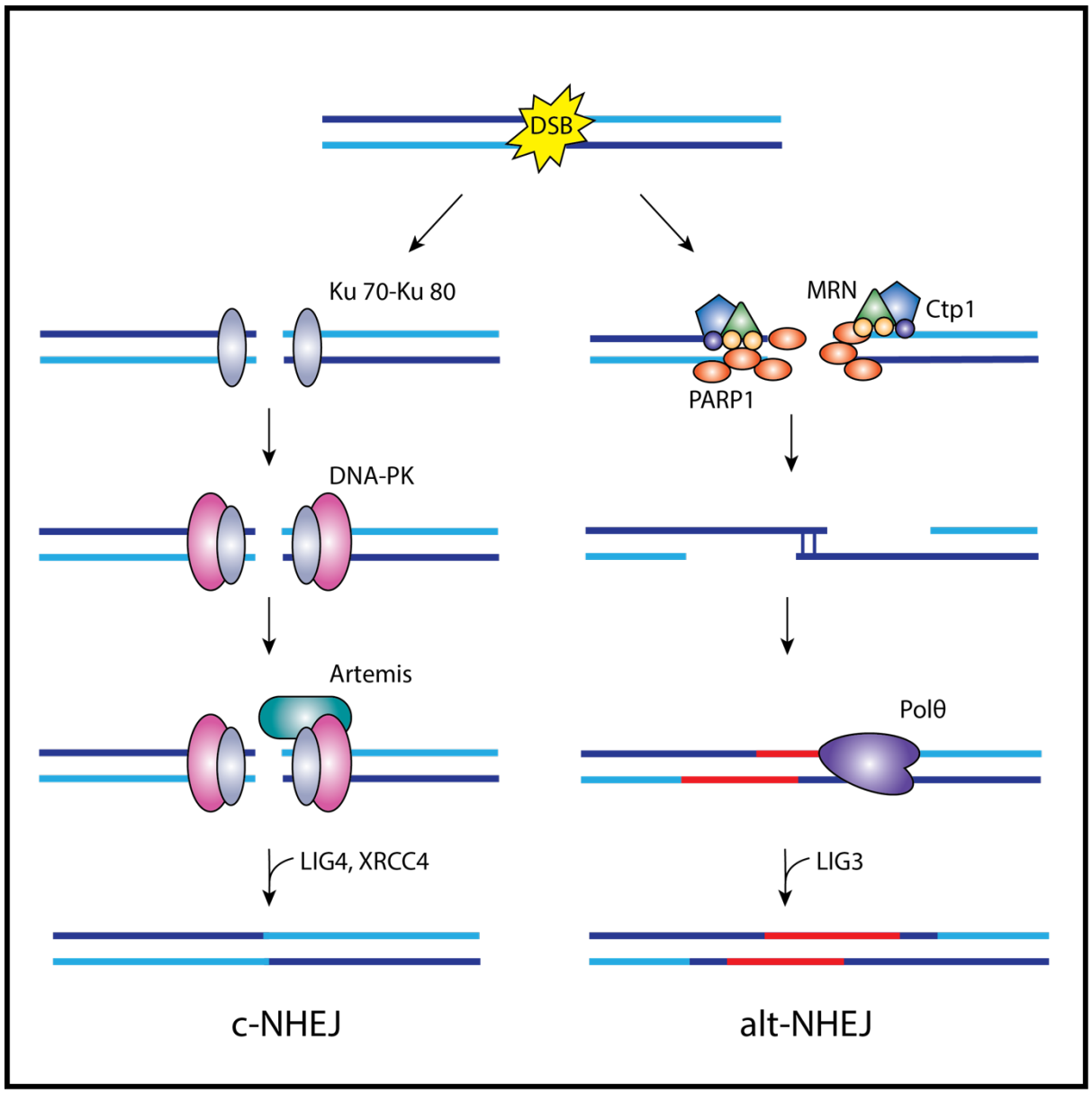


Figure.1.6. c-NHEJ and alt-NHEJ repair pathway. *c-NHEJ is the predominant repair pathway in mammalian cells. It is initiated by the dimerisation of Ku (Ku70- Ku80). Ku serves as a scaffold at the DNA ends, recruiting DNA-PKcs to the break site, which phosphorylate the Artemis protein. DNA ligase IV is recruited to the scaffold along with XRCC4 to ligate and re-join the ends. In the alt-NHEJ, PARP1 acts in place of Ku, identifying and binding to DNA ends to recruit subsequent proteins Mre11, Rad50 and Nbs1, forming the MRN complex, along with Cpt1 and Pol ϑ, promoting strand resection. DNA ligase is responsible for ligation and re-joining of DNA ends.*

1.6.2 Homology Directed Repair (HDR)

HDR is another mechanism in cells to repair DNA DSBs, the most common form of HDR being homologous recombination (HR). In contrast to NHEJ, HR uses an undamaged DNA template on the sister chromatid or homologous chromosome to repair the DSB, leading to the reconstitution of the original sequence [261]. As such, it is considered an error-free mechanism for removal of DSBs during the mitotic DNA replication phase (S and G2 phase) in the presence of a DNA template [262]. HR is also active as the fundamental repair pathway in germ cells, where programmed DSBs occur during meiosis [263].

HR can be conceptually divided into three stages (pre-synapsis, synapsis, post-synapsis). There are three subclasses for this kind of recombination: classical DSB repair (DSBR), synthesis-dependent strand-annealing (SDSA), and break-induced repair (BIR); all three subclasses are initiated by the same mechanisms during pre-synapsis and synapsis stages [264]. Firstly, the DNA ends are processed to form a 3' overhang extended region of ssDNA, mediated by binding of the MRN complex, and stabilised by replication protein A (RPA) [265]. These ssDNA tails complex with Rad51, forming a filament and performing homology search and DNA strand invasion [266], generating a displacement loop (D-loop structure where the invading strand primes repair synthesis) [263].

In post-synapsis, the 3 subclasses of HR are distinguished from each other. In classical DSBR, there is formation of a double Holliday junction (dHJs), four-stranded branch structures that have symmetrical sequences. To complete repair of the DSB, the dHJs have to be resolved by nucleases called HJ resolvases, which cleave the dHJs [263]. This can produce crossover recombinants that harbour a reciprocal exchange of the arms of the recombining chromosomes [266].

In SDSA, no Holliday junction is formed. Instead, the extended D-loop is reversed, leading to annealing of the newly synthesized strand with the resected strand of the other 3' over-hang strand [267]. Unlike DSBR, this pathway fundamentally avoids crossovers, which reduces the potential for genomic rearrangements [264].

The last subclass is less well defined. BIR occurs when only one D-loop is present and therefore only one invasive end can be used for repair. This single strand searches for and invades a homologous sequence, initiating the synthesis of both leading and lagging strand, and forming only one Holliday junction [268]. While this pathway can repair the DSB, and

restore the integrity of the of the chromosome, the second end is never engaged and therefore the genetic information distal to the DSB is lost [263].

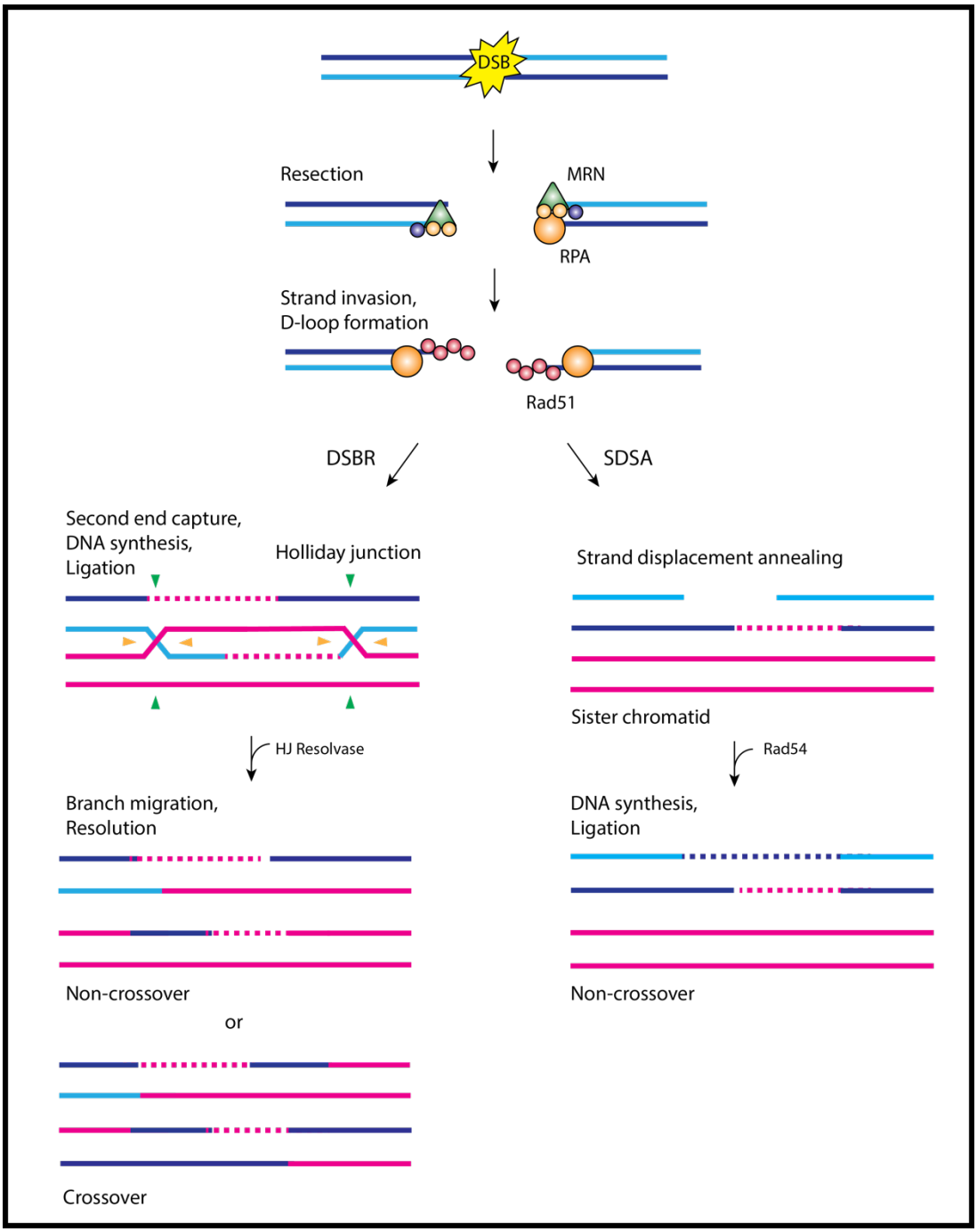


Figure 1.7. DSBR and SDSA HDR repair pathways. Firstly, the DNA ends are processed to form a 3' region of ssDNA, mediated by binding of the MRN complex, and stabilised by RPA. These ssDNA tails complex with Rad51, mediating the formation of a filament to invade the undamaged template and a displacement loop. In DSBR, there is formation of a double Holliday junction , resolved by HJ resolvases, leading to crossover or non-crossover events. In

the SDSA pathway, recombination occurs due to invasion by one 3'-ssDNA strand without crossover between chromatids.

1.7 Genome Editing

Genome editing, or genetic engineering, is a type of genetic modification in which DNA is inserted, deleted, altered or replaced at a particular location in the genome of a living organism. Homologous recombination can typically be observed in only 10^{-6} to 10^{-9} of treated mammalian cells [269]; this can be stimulated two to three orders of magnitude by a DSB at the target locus [270]. However, the challenge of using the cellular repair machinery lies in creating a precise DSB in the genome.

Restriction endonucleases, capable of cleaving DNA into fragments at or near specific recognition sites, transformed molecular biology in the seventies by enabling cloning. However, these are not useful for delivering a targeted chromosomal DSB, since they recognize short DNA sites (usually 4 to 8 bp sites), which occur much too frequently within most genomes to be applicable in targeted gene editing. Progress was made by the discovery of the function of meganuclease I-SceI in yeast *Sacharomyces cerevisiae.* Meganuclease I-SceI is able to target a specific 18 bp sequence (5'-TAGGGATAACAGGGTAAT-3') and achieve a DSB with four-base, 3' overhangs [271]. Researchers identified a chromosomal neomycin resistant gene interrupted by an I-SceI recognition site. They were able to use this as a reporter to monitor gene correction upon delivery of an I-SceI expression vector and DNA repair template, and achieved an HR gene correction rate of $4x10^{-4}$ in NIH3T3 cells [272]. While this technique has been used numerous studies and helped elucidate the cellular mechanisms of the DNA repair pathways, the rarity of the targeted sequence occurring naturally in the genome limits the application of I-SceI for therapeutic genome editing.

Recently, the development of three core technologies has revolutionised the field of genome editing, as discussed below, and compared in Table 1.3.

1.7.1 Zinc-Finger Nucleases

To be useful for targeted genome engineering, an endonuclease must exhibit a unique combination of qualities: precise recognition of long target sequences (long enough for unique occurrence in a eukaryotic genome), combined with sufficient adaptability for

retargeting desired sequences. Zinc-Finger Endonucleases (ZNFs), which are hybrids of the DNA cleavage domain from a bacterial protein and zinc-finger (ZF) DNA binding domain, meet these requirements.

Repetitive zinc-finger binding domains were first discovered in transcription factor IIIA from *Xenopus oocytes,* which were later termed ZF motifs [273]. Each ZF motif consists of ~30 amino acids folded into a ββα structure, which is stabilized by chelation of a zinc ion by the conserved Cys_2His_2 residues. The crystal structure of a set of three fingers bound to DNA showed that each finger contacts primarily 3 bp. The ZF motifs bind DNA by inserting the α-helix into the major groove of the DNA double helix [274].

Restriction enzymes have a dual function; DNA recognition and DNA cleavage; type IIS restriction enzymes, such as *Fok*I*,* have two discrete domains for each of these functions [275]. It was therefore proposed that *Fok*I would be an ideal candidate for DSB induction in gene editing, as the DNA recognition sequence could be manipulated to change sequence specificity to recognise longer sequences, without altering the DNA cleavage domain and maintaining cleavage activity. Single amino acid substitutions could uncouple the DNA-binding and nuclease activities of *Fok*I [276], and the possibility of achieving targeted DSB cleavage in the genome was realised by fusing the *Fok*I cleavage domain to two ZF motifs [277], and the term zinc-finger nucleases was coined. It was shown that dimerization of the cleavage domain is necessary [278], and two recognition sites for the ZFNs must be in close proximity and in inverted repeat orientation, with a separation of 5-7 bp shown to cleave most efficiently. Importantly, this method of targeted DSB formation was shown to promote HR *Xenopus laevis* oocytes [279, 280].

In order to identify ZF domains that recognise each of the 64 possible DNA triplets within our genome, phage display techniques were initially used [281-283]. A library of known of ZF, recognising most of the triplicates of 3 bp DNA, was made commercially available by Sangamo Biosciences, and subsequent ZFs could be assembled to create a full array, in a technique known as modular assembly. However this was a laborious method, and it was discovered that in some cases the strict modularity observed in ZFs was disrupted by binding to bases outside their proximal DNA triplet, leading to failures in recognising the desired DNA target [284].

To overcome these issues, a technique called oligomerized pool engineering (OPEN) was developed. Based on a pool of pre-constructed 3-finger arrays that recognise 9 bp target

sequences, the array with highest sequence affinity is identified in a screening method in *Escherichia coli* using a bacterial two-hybrid (B2H) system. The B2H reporter system measures expression of the *lacZ* reporter gene as a result of transcriptional activation between the ZF array and its target site upstream of the promoter [285]. Despite the success rate, this method is still laborious in terms of screening, taking around 8 weeks.

Most recently, context-dependent assembly (CoDA) provides a platform for engineering ZFNs using only standard cloning techniques or custom DNA synthesis [286]. CoDa is a publicly available platform of reagents and software. The CoDA approach assembles 3-finger arrays using N- and C- terminal fingers which have been identified previously in other arrays containing a common middle finger, and are engineered to function when positioned next to one of 18 fixed middle finger units [286].Unlike modular assembly, CoDA treats the 3-finger array as a whole, therefore accounting for context-depended effects between adjacent fingers, as previously mentioned.

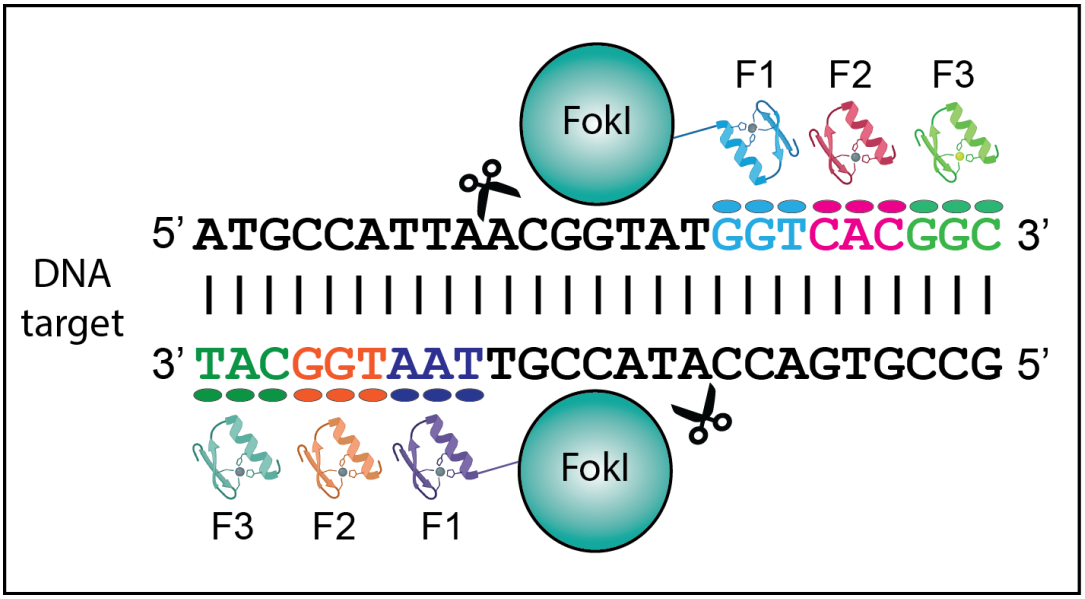


Figure 1.8. Schematic of two ZNF subunits bound to DNA. *Each subunit encompasses three zinc-fingers arranged in a tandem array, each of which binds to three nucleotides. The C-terminus of the ZF array is fused to the catalytic domain of the FokI endonuclease. After dimerization, the nuclease is activated and creates a DSB at the DNA location specified by the binding site associated with the ZF array.*

1.7.2 Transcription activator-like effector nucleases

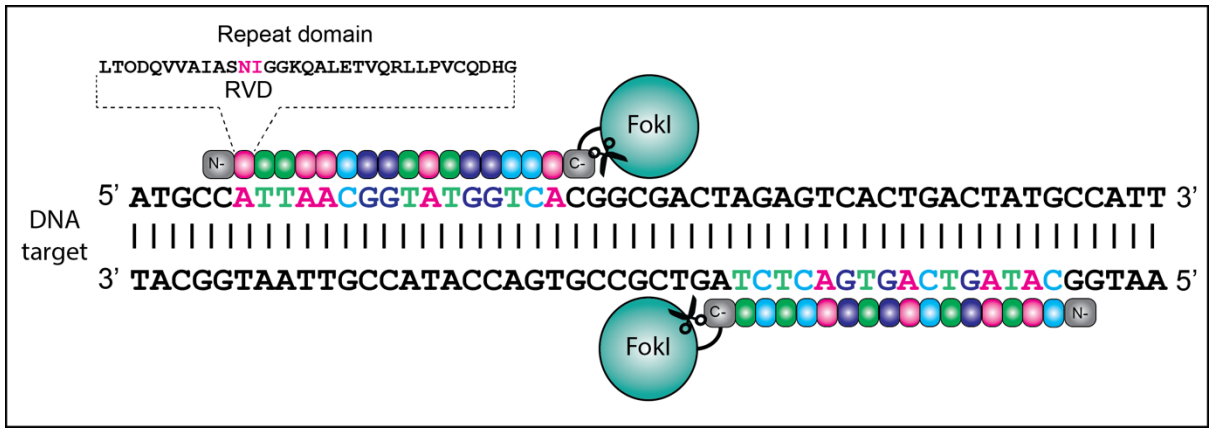
Similar to ZFNs, transcription activator-like effector nucleases (TALENs) TALENs use the catalytic domain from *FokI* endonuclease for DNA cleavage, which is instead fused to a DNA binding domain derived from transcription activator-like effectors (TALEs) [287]. TALEs from plant pathogen *Xanthomonas* are important virulent factors that act as transcriptional activators in plant cells, where a central domain of tandem repeats mediates DNA binding [288, 289]. Each repeat is a sequence of 33-35 amino acids which specifically target a single nucleotide. Nucleotide preference is determined by two highly variable adjacent amino acids at positions 12 and 13, referred to as the repeat-variable di-residue (RVD) [290]. The most common RVDs are HD (Histidine, Aspartic Acid), NG (Asparagine, Glycine), NI (Asparagine, Isoleucine) and NN (Asparagine, Asparagine) which recognize C, T, A, and G nucleotide respectively [290, 291].

Reber and colleagues first reported the development of TALENs capable of mediating efficient endogenous DNA modification in mammalian cells. They developed truncation variants of the TALE proteins that enable efficient cleavage *in vitro* when linked to the catalytic domain of *FokI.* The TALENs were able to mediate DSB cleavage in human genes *NFT3* and *CCR5,* which were subsequently repaired by either NHEJ or HDR [287]. The nucleases have subsequently been utilized in a large number of applications, from creating targeted alterations in model organisms from zebrafish [292] to pigs [293], to introduction of specific insertions in human somatic and pluripotent stem cells using HR donor templates [294].

The crystal structure of TALE-DNA binding domains bound to their cognate sites show that each TALE repeat comprises two helices connected by a RVD-containing loop, which contacts the major groove of DNA. The first RVD residue stabilises the loop via contact with the protein backbone, whereas the second residue makes base-specific contact [295, 296].

The cloning and synthesis of DNA encoding engineered TALE repeat arrays can prove to be difficult due to the necessity of assembling multiple, nearly-identical, repeat sequences, variable only at the RVD. Several platforms have been designed to overcome these technical difficulties, including 'Golden Gate' multi-fragment ligation strategy [297], and Fast Ligation-based Automatable Solid-phase High-throughput (FLASH) [298]. Reagent kits for these cloning assembly techniques and others are available from the plasmid distribution service Addgene for academic use.

TALENs have several advantages over ZFNs, namely their ease of design. Like ZFNs, TALENs can be assembled in a modular fashion, however, there is no risk of interaction between neighbouring motifs [284]. The genes encoding TALENs are about three times larger than ZFNs; despite the TALE motifs being similar in size to ZFs, they recognize a single base, while ZFs recognize 3- to 4-bp sequences. While this improves the specificity of the platform, it may affect delivery efficiency of plasmids encoding the systems.



Repeat domain	Target base	Repeat-variable di-residue (RVD)
	G	NN (asparagine, asparagine)
	C	HD (histidine, aspartic acid)
	T	NG (asparagine, glycine)
	A	NI (asparagine, isoleucine)

Figure 1.9 Schematic of customised TALENs bound to DNA. *The DNA binding domain is composed of repeat domains, consisting of 34 amino acid residues, and recognises one nucleotide. Nucleotide preference is determined by two highly variable adjacent amino acids at positions 12 and 13, referred to as the repeat-variable di-residue (RVD). The most common RVDs, NN, HD, NG and NI recognise G, C, T and A nucleotides respectively. Adapted from [299].*

1.7.3 Clustered Regularly Interspaced Short Palindromic Repeats (CRISPR)

Although both ZFN and TALENs are effective, ease of use means that they have been largely supplanted in *vitro* by the recent developments of the CRISPR/Cas9 gene editing system. The term Clustered Regularly Interspaced Short Palindromic Repeats (CRISPR) and their CRISPR-

associated (Cas) nucleases was first mentioned in 2002 [300], however, CRISPRs were first detected in 1987 in *Escherichia coli*, when five identical cluster repeat sequences were discovered whilst analysing the gene responsible for isozyme conversion of alkaline phosphatase (*iap* gene) [301]. In 1993, Francisco Mojica and colleagues discovered similar CRISPR sequences in *Haloferax mediterranei,* and subsequently in a number of bacterial and archaeal genomes [302], and by 2005 Mojica hypothesised that these highly conserved sequences were part of the bacterial immune system [303].

Meanwhile, DNA repair proteins in hyperthermophilic archaea were identified as being exclusively associated with CRISPR, and therefore termed *cas* (CRISPR-associated) genes [304], and it was hypothesised that CRISPR and Cas proteins work together as an acquired immune system, protecting prokaryotic cells against invading viruses [305]. One such *cas* protein (now known as Cas9) was discovered in *Streptococcus thermophilus*, and predicted to have nuclease activity [305]. Additionally, the presence of spacer sequences, which have homology to viral genes, were discovered at the CRISPR locus, and a common sequence at one end of the spacer region, the protospacer adjacent motif (PAM), was recognised to be required for target recognition [306].

In 2007, it was experimentally proven that after viral challenge, bacteria integrated new spacers derived from phage genomic sequences [307], confirming that the CRISPR systems are indeed an adaptive immune system. These phage-derived spacer sequences are transcribed into small RNAs, termed CRISPR RNAs (crRNAs) [308], which guide the Cas9 to the target DNA of the attacking phage [309]. It was demonstrated that Cas9 creates a double-stranded break (DSB) 3 base pairs upstream of the PAM sequence. Further, they confirmed that the Cas9 protein is the only one required for DNA cleavage, a distinguishing feature of class II CRISPR systems [310]. This is in contrast to class I systems, where several *cas* proteins come together as a multi-protein effector complex to cleave DNA [311].

Emmanuelle Charpentier, in 2011, discovered that a second small RNA exists, which they called trans-activating CRISPR RNA (tracrRNA). These tracrRNA direct the maturation of crRNAs, forming a duplex and guiding Cas9 to its target [312]. In collaboration with Jennifer Doudna, the groups showed that the crRNA and tracrRNA could be fused together to create a single, synthetic guide. Further, by changing the sequence of the crRNA, the Cas9 endonuclease could be re-programmed to cleave any double-stranded DNA sequence of

interest, highlighting the potential to exploit the system for RNA-programmable genome editing [313] [314] (Figure 1.10).

Shortly after, the CRISPR system was successfully used for genome editing in eukaryotic cells. Zhang and colleagues engineered a plasmid encoding a codon-optimized *S. pyogenes* Cas9 (SpCas9), along with nuclear localization signals (NLSs) to ensure nuclear compartmentalization in mammalian cells. Using a crRNA:tracrRNA duplex targeting various loci in both human and mouse cell lines, they were able to achieve targeted cleavage of mammalian chromosomes. Because DSBs in mammalian DNA are partially repaired by the indel-forming non-homologous end joining (NHEJ) pathway, the group used the SURVEYOR® assay to detect endogenous target cleavage [315, 316]. Further, by providing a homology repair template, they were able to introduce a DNA sequence into the genomic locus via homology-directed repair (HDR) [316].

The crystal structure of spCas9 complexed with sgRNA and its target DNA at 2.5 A˚ resolution was reported in 2014 single-wavelength anomalous dispersion (SAD) [317]. The protein has two distinct lobes, the alpha-helical recognition (REC) lobe and the nuclease (NUC) lobe. The REC lobe contains a long helix denoted a bridge helix, and RECI and RECII domains. The NUC lobe consists of the RuvC domain, a histidine–asparagine–histidine (HNH) domain and PAM-interacting domain (PI) domains. While the recognition lobe is essential for binding the sgRNA and DNA, it is the HNH nuclease domain that cleaves the DNA strand complimentary to the 20-nucleotide guide sequence, while the RuvC cleaves the non-complimentary strand [317].

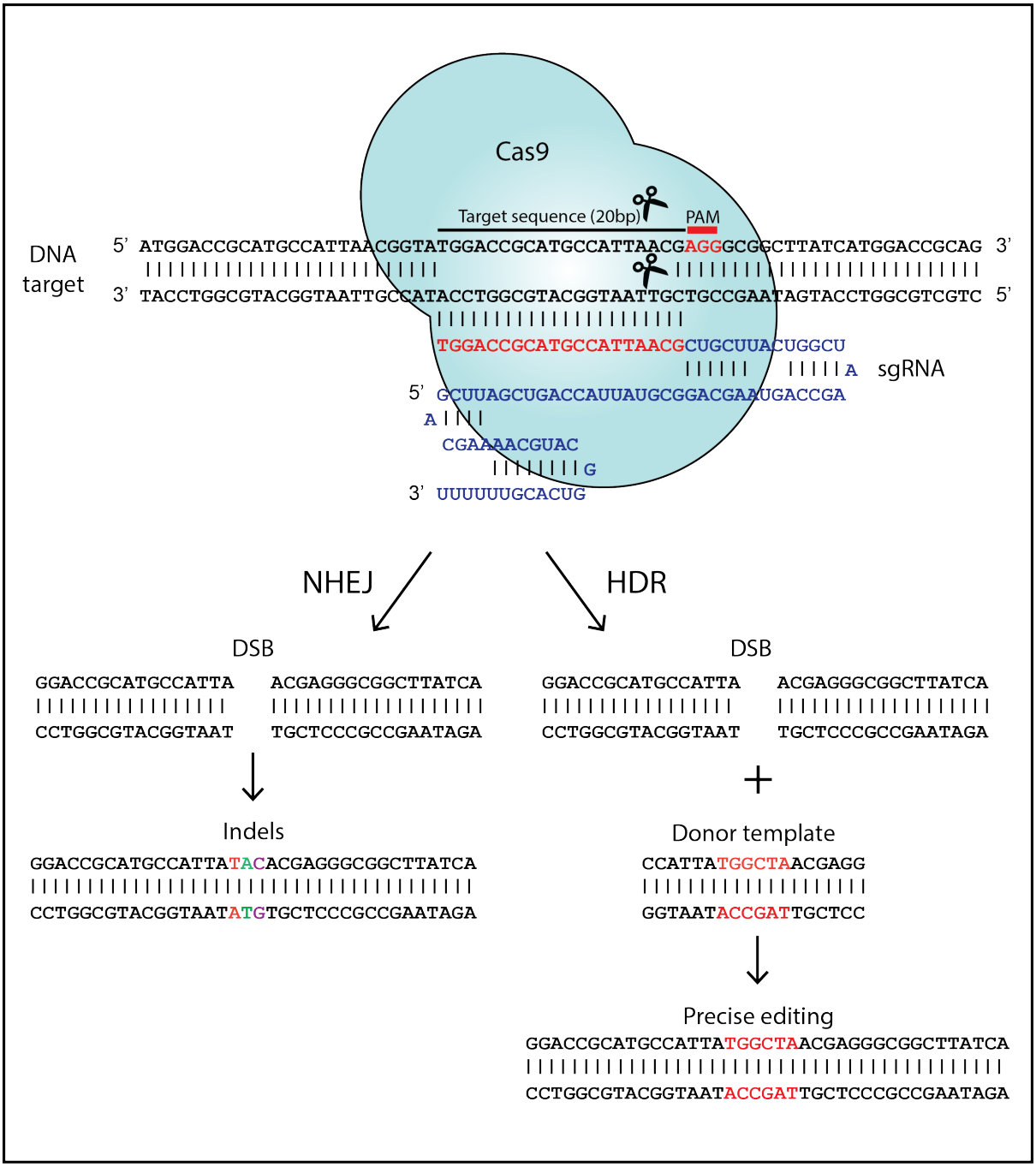


Figure 1.10. Schematic of the RNA-guided Cas9 nuclease. *The Cas9 is targeted to genomic DNA by a sgRNA consisting of a 20 nucleotide guide sequence (red) and a scaffold (blue). The guide sequence pairs with the DNA target directly upstream of a 5'-NGG protospacer adjacent motif (PAM) and Cas9 mediates a DSB 3 bp upstream of the PAM. DSB repair via one of two cellular repair pathways; in NHEJ or HDR.*

***Table 1.4. Comparison of programmable nucleases for use in genome editing**. Adapted from [318].*

	Meganuclease	ZFN	TALEN	Cas9
Target sequence (bp)	14–40	9–18	14–20	~23
Number of target sites	limited	many	many	many
Enzyme engineering	difficult	difficult	easy	very easy
Size (kb)	~1	~1	~3	~3.5–4.5
Target recognition	protein-DNA	protein-DNA	protein-DNA	RNA-DNA base pairing and protein-DNA

1.7.3.1 CRISPR/Cas variants

As discussed, there are two classes of CRISPR systems, with class 2 CRISPR systems most comprehensively characterised. Class 2 systems, with single-protein effector modules, are further divided into types II, IV, V, and VI. The most widely used system, with it's simple 'NGG' PAM sequence requirement, is the type II Cas9 from *Streptococcus pyogenes* [319]. Indeed, the 'NGG' sequence is abundant throughout the human genome, with a 'GG' dinucleotide occurring approximately once every 42 bases [320]. However, targeting an 'AT' rich area, or where DSBs are required at a specific position can be challenging, and so directed evolution processes have been utilised to generate Cas9 variants with altered PAM specificities [321, 322], such as xCas9, which can recognise 5'- 'NG', 'GAA' and 'GAT' PAM sequences, the broadest PAM compatibility reported, and greater specificity than spCas9 [322]. Additionally, the spCas9 variant VQR targets 'NGAN' and 'NGCG' PAMs, while EQR variant has specificity for 'NGAG' [321].

The discovery and development of Cas9 orthologues from different species has further expanded the range of recognition sites that can be targeted within the genome. Additionally, several of these orthologues are smaller in size than spCas9, with a 4 kb coding sequence, preventing efficient packaging into AAV. SaCas9, from *Staphylococcus aureus,* recognises the PAM sequence 'NNGRRT' and is 3.16 kb in size [323], while CjCas9 from

Campylobacter jejuni recognises 'NNNNRYAC', where 'Y' is a pyrimidine base, is 2.95 kb in size, and has been successfully used *in vivo* in a mouse model for Age-Related Macular Degeneration (AMD) [324]. Despite their smaller size, these Cas9 proteins have a relatively limited targeting scope due to the greater complexity of their PAM sequence as compared to spCas9.

Another prominent CRISPR/Cas variant is the Class 2 type V Cpf1 (CRISPR from *Prevotella* and *Francisella*) [325], later renamed Cas12a [326]. The PAM site was initially described as 'TTTN', but it was later discovered that the first nucleotide could be A, C or G, but not T [327]. Unlike Cas9 variants, Cas12a is a single RNA-guided endonuclease, which doesn't require the presence of a tracrRNA, and generates a staggered DSB with 5'-overhang. Furthermore, Cas12a cleaves DNA at the distal end of its PAM sequence, meaning that indels will be located further from the target site, preserving the PAM for subsequent rounds of Cas12a cleavage [325].

In addition to these developments, efforts have also been made to re-engineer the well characterised spCas9 protein, including enhancing target specificity. Joung and colleagues describe SpCas9-HF1, a high-fidelity variant designed to reduce non-specific DNA contacts by the introduction of four amino acid residue substitutions (N497A, R661A, Q695A, and Q926A) [328]. Similarly, Zhang *et al.* engineered an aspartate-to-alanine substitution (D10A) in the RuvC I domain of SpCas9 to convert the nuclease into a DNA nickase, which are capable of introducing a single-strand cut with the same specificity as a regular Cas9 nuclease [316]. By using pairs of nickases, with gRNAs targeting opposite DNA strands, overhangs are produced, providing greater control over precise gene integration [315]. Further, since these approaches require two separate gRNAs to be in a certain proximal distance, the probability of off-target modification is significantly reduced [316].

A further modification to spCas9 involves inactivating the nuclease domains by silent mutations D10A and H840A on the RuvC and HNH domains respectively, producing catalytically inactive dCas9 that will successfully bind to the target region however will not induce a DSB [329]. This system can be modified by fusing heterologous domains to change methylation patterns, or to include transcriptional activators to increase expression of genes of interest [330, 331].

1.7.3.2 CRISPR/Cas9 applications

Since its inception as a genome editing tool in 2012, more than 15,250 research articles have been published about it, according to a search on PubMed. The CRISPR/Cas9 system can be used to facilitate a wide variety of targeted genome engineering applications, including the generation of cellular disease models by means of knocking out specific genes. The CRISPR/Cas system can be easily introduced into target cells using transient transfection of plasmids carrying Cas9 and an appropriate gRNA, targeted to the coding region of a gene of interest, and introducing indels which can cause premature termination codons (PTCs) or frameshift mutations to knock out the gene. Through single cell cloning, a population of cells containing a knockout can be obtained. Moreover, the multiplexing capabilities of Cas9 can offer a means to study complex, polygenic human diseases such as diabetes or autism.

Further, CRISPR/Cas9 has enabled efficient and targeted genome modification in many species that have been intractable using traditional genetic manipulation techniques. The system can be directly injected into fertilised zygotes to achieve heritable gene modification at one or multiple alleles. Numerous transgenic animal models have been created using this method, from a mouse model of Duchenne's muscular dystrophy [332], to transgenic cattle with increased resistance to tuberculosis [333].

Thanks to its high editing efficiency, CRISPR/Cas9 has been use for unbiased, genome wide functional screens in order to identify genes that play important roles in a phenotype of interest, and lentiviral delivery of gRNAs directed against all genes can be used to perturb thousands of genomic elements in parallel [334]. Zhang and colleagues designed a library of 64,751 unique gRNA sequences targeting 18,080 genes (genome-scale CRISRP/Cas9 knockout- GeCKO library) to enable negative and positive selection screening in human cells. Using the GeCKO library, they were able to identify genes integral for cell viability in cancer and pluripotent stem cells [335].

As mentioned, delivery of Cas9 and gRNA without the use of a repair template can be used to knockout specific genes. Following Cas9 cleavage, eukaryotes predominantly repair DSBs through the error-prone NHEJ pathway, which can lead to the accumulation of indels. This pathway can be harnessed to knockout specific genes via delivery of Cas9 and gRNA without the use of a repair template. However, by simultaneously targeting two sites within a gene using separate gRNAs, a deletion can be generated between the DSBs. Ousterout *et al.* used this strategy to restore the expression of the dystrophin gene in cells carrying

mutations that cause Duchenne muscular dystrophy (DMD). By designing pairs of gRNAs targeting the mutational hotspot at exons 45-55, they were able to generate a single large deletion, restoring the dystrophin reading frame and leading to the expression of a truncated, but functional, dystrophin protein. Following enrichment, these cells were engrafted into immunodeficient mice and human dystrophin was found to be expressed [336].

By providing a donor template, it is also possible to insert new elements into a specific location via the HDR pathway following Cas9 mediated DNA cleavage. The donor template requires sequences of homology, termed homology arms, flanking the cutting site. By designing an HDR template with mutations in the PAM sequence or gRNA seed sequence, is it possible to prevent further rounds of Cas9 cutting. Various types of HDR repair templates have been used, including circular and linear double-stranded DNA molecules, and single-stranged oligodeoxynucleotides (ssODNS). When large sequence changes are required, such as the insertion of fluorescent proteins or expression cassettes, researchers generally use dsDNA plasmid templates. For smaller modifications, ssODNs may be appropriate, and can require as little as 30-50 bp homology arms, and asymmetric arms can further increase HDR efficiency [337].

As mentioned, DSB repair by HDR is less efficient than by the canonical NHEJ pathway, with HDR restricted to late S and G2 phases in the cell cycle, where sister chromatids are available to serve as repair templates [262]. As such, several methods have been implemented in order to increase efficiency, including precisely timed CRISPR/cas9 delivery combined with established cell cycle synchronization techniques [338]. Other researchers have used chemical interruption of the NHEJ pathway, such as targeting DNA ligase IV, a key enzyme in the NHEJ pathway, using an Scr7 inhibitor, however there are reported issues regarding specificity [339]. Meanwhile, Gordon and colleagues have shown that by covalently tethering an ssODN repair template to the Cas9-gRNA ribonucleoprotein complex (RNP), they can co-localise the donor template to the DSB repair machinery, and demonstrated a 30-fold enhancement in HDR efficiency [340]. Through these methods and others, much higher rates of precise genome editing can now be achieved using the HDR pathway.

1.7.3.3 Base editing

The most common genetic variants associated with human diseases are point mutations [341], and so the ability to edit single nucleotide bases is of great importance in the

development of corrective therapeutics. This can be achieved via co-delivery of Cas9 and a homologous repair template containing the edited nucleotide of choice [342]. However, HDR competes with NHEJ during DSB resolution, and so indels are generally a more abundant outcome. Such strategies therefore remain inefficient, particularly in post-mitotic cells. By fusion of the catalytically inactive dCas9, or Cas9 nickase, to cytosine deaminase, researchers have achieved targeted cytosine to uracil nucleotide conversion within a specific window located within the spacer sequence, where uracil becomes thymine through DNA replication [343, 344]. A uracil glycosylase inhibitor can be used to prevent the DNA repair pathways from restoring the edited bases, and permanent nucleotide conversion of up to 75% can be achieved in mammalian cells [343]. Liu's lab have now created a further base editor that can deaminate adenine to yield inosine, which is treated as guanine by polymerases. Using this base editor, they achieved an adenine to guanine transition efficiency rate of up to 50% in mammalian cells [345]. Together, these base editors enable direct, programmable introduction of all four transition mutations without the need of DNA cleavage, and have been utilised in zebrafish, mice and human embryos [346-348].

1.7.3.4 Prime editing

Prime editing involves the use of a catalytically impaired Cas9 fused to an engineered reverse transcriptase, along with a programmed prime editing guide RNA (pegRNA) that both specifies the target site and encodes the desired edit, enabling the copying of genetic information directly from the extension on the pegRNA into the target genomic locus [349]. Like base editors, prime editing offers much lower off-targeting than HDR due to lack of any DSBs within the genome. The technology can enable targeted insertions, deletions and all 12 possible classes of point mutation without the need on an HDR template, advancing the study and correction of the vast majority of pathogenic alleles [349].

1.7.3.5 Clinical applications

The technology has started to enter its first clinical trials, with several trails currently recruiting patients (Table 1.4). Most involve an *ex vivo* approach, removing cells from a patient and using CRISPR to modify the cells in such a way as to fight disease, as in the β-thalassemia trails, where stem cells are edited to increase the oxygen-carrying capacity of red

blood cells (ClinicalTrials.gov Identifier: NCT03655678), and in cancer treatments, where a patient's T cells are genetically engineered to produce chimeric antigen receptors (CAR-T cells), allowing them to recognise cancers more effectively (NCT04035434). The cells are then transplanted back into the patient's body. There are several advantages to an *ex vivo* approach: the delivery of the CRISPR tools to the cells is much simpler, there is the potential to sort cells before re-transplantation, improving the safety and efficacy of the therapy, and it minimises the risk of editing non-target cell types. The first *in vivo* trial, involving direct delivery of the CRISPR/Cas9 tool to the retina of patients suffering from Leber congenital amaurosis 10, a disorder caused by a mutation in the CEP290 gene and resulting in blindness, is currently recruiting patients (NCT03872479) [350].

Table 1.5. Current clinical trials involving CRISPR/Cas9. *Trails are actively recruiting patients, or have completed recruitment, as listed on clinicaltrails.gov [351]*

Study Title	Disease	Status	Location	I.D. number
A Safety and Efficacy Study of TALEN and CRISPR/Cas9 in the Treatment of HPV-related Cervical Intraepithelial Neoplasia I	HPV-related Malignant Neoplasm	Recruiting	Guangdong, China	NCT03057912
Safety of Transplantation of CRISPR CCR5 Modified CD34+ Cells in HIV-infected Subjects With Haematological Malignancies	HIV-1-infection	Recruiting	Beijing, China	NCT03164135
NY-ESO-1-redirected CRISPR (TCRendo and PD1) Edited T Cells (NYCE T Cells)	Multiple Myeloma, Synovial Sarcoma, Melanoma	Recruiting	Pennsylvania, USA	NCT03399448
Study of CRISPR-Cas9 Mediated PD-1 and TCR Gene-knocked Out Mesothelin-directed CAR-T Cells in Patients With Mesothelin Positive Multiple Solid Tumours	Solid tumour (adult)	Recruiting	Beijing, China	NCT03545815
CRISPR (HPK1) Edited CD19-specific CAR-T Cells (XYF19 CAR-T Cells) for CD19+ Leukaemia or Lymphoma.	Acute Lymphocytic Leukaemia (relapsed)	Recruiting	Shannxi, China	NCT04037566
A Safety and Efficacy Study Evaluating CTX001 in Subjects With Transfusion-Dependent β-Thalassemia	Beta-Thalassemia	Recruiting	Toronto, Canada	NCT03655678
A Safety and Efficacy Study Evaluating CTX110 in Subjects With Relapsed or Refractory B-Cell Malignancies	Non-Hodgkin and B cell lymphoma	Recruiting	Tennessee, USA	NCT04035434
A Safety and Efficacy Study Evaluating CTX001 in Subjects With Severe Sickle Cell Disease	Sickle cell Disease	Recruiting	New York, USA	NCT03745287
Study of PD-1 Gene-knocked Out Mesothelin-directed CAR-T Cells With the Conditioning of PC in Mesothelin Positive Multiple Solid Tumours	Solid tumour, adult	Recruiting	Beijing, China	NCT03747965
A Study Evaluating UCART019 in Patients With Relapsed or Refractory CD19+ Leukaemia and Lymphoma	B-cell leukaemia and lymphoma	Recruiting	Beijing, China	NCT03166878
PD-1 Knockout Engineered T Cells for Advanced Esophageal Cancer	Esophageal Cancer	Completed	Zhejiang, China	NCT03081715
PD-1 Knockout EBV-CTLs for Advanced Stage Epstein-Barr Virus (EBV) Associated Malignancies	Esophageal Cancer	Recruiting	Zhejiang, China	NCT03044743
Single Ascending Dose Study in Participants With LCA10	Leber Congenital Amaurosis 10	Recruiting	Florida, USA	NCT03872479

1.7.3.6 CRISPR/Cas9 off target effects

The simple target-site recognition code of CRISPR–Cas9-based systems accounts for its robustness and wide applicability. However, this simplicity has the potential to accommodate mismatches within the genome, and thus Cas9 mediated DNA cleavage at off-target sites. Early studies showed that Cas9 could bind to alternate sites which harbour up to five mismatched bases compared to the gRNA sequence, highlighting potential fidelity issues

when using the system [352], whilst more recent studies have shown in mice and human embryos have shown much lower levels of off-target cutting [353-355].

Design of gRNAs is of upmost importance in reducing off-target effects, and several online programmes can aid in designing sequences to minimise the potential for off-target cleavage whilst maintaining high, on-target specificity. It has been shown that the seeding region, that is, the 12 nucleotides directly before the PAM sequence, are particularly important in minimising off-target effects [329] whilst a shorter gRNA sequence of 17-18 bp may also decrease off-targeting without significantly effecting editing efficiency [356].

As well as gRNA optimisation, it has been shown that decreasing Cas9 concentration can significantly improve the on- to off- target ratio, however, this is at the expense of on-target efficiency [357]. Additionally, the Cas9 variants described in section 1.7.3.1, including Cas9 nickase system, can also reduce off target effects; the requirement of two separate gRNAs significantly reducing the probability of off-target modification [316]. Direct delivery of Cas9 protein plus gRNA as a ribonucleoprotein complex (RNP) has also been shown to reduce the frequency of off-target mutations [358, 359].

Several ways of assessing off-targets effects exist. Bioinformatic analysis to identify putative off-target sites, followed by design of PCR primers to amplify these sequences from genomic DNA and blasting against WT genome sequence is one such method that can be easily performed. Other groups have utilised whole-genome sequencing to profile off-target mutations, such as Digenome-seq, (*in vitro* Cas9-digested whole-genome sequencing). This digest yields sequence reads with the same 5' ends at cleavage sites that can be computationally identified [360]. In addition, GUIDE-seq (genome-wide, unbiased identification of DSBs enabled by sequencing) relies on capture of double-stranded oligodeoxynucleotides into DSBs [361]. These techniques have limitations such as abundant false positives and an inability to operate during *in vivo* editing. DISCOVER-Seq (discovery of *in situ* Cas off-targets and verification by sequencing) is a recently described tool which tracks the recruitment of Mre11, one of the first proteins recruited to a DSB cut site, and can uncover the molecular nature of Cas activity in cells with single-base resolution [358].

1.7.4 Gene editing for CF

As a monogenic disease, CF is an ideal candidate for gene therapy. The delivery of CFTR cDNA has been studied extensively over the last 25 years, but the most significant clinical benefit reported so far was stabilisation of lung function in patients partaking in the most recent clinical trial [205]. In recent years, there has been a significant rise in the use of precise gene editing techniques in CF research, which may be more effective than cDNA addition. Promising results from other fields, including the first in-human trial on a patient with Hunter's Syndrome (NCT03041324) just six years after *in vivo* data was first published from a mouse model [362], has encouraged this.

The above approach involved the use of a 'superexon'; partial cDNA encompassing a particular set of exons. This strategy has been adapted for use in CF, where researchers used ZFNs to create a DSB at exon 11 of CFBE41o- cell line, and designed a donor template comprising wtCFTR exons 11-27. This was able to restore CFTR mRNA expressed from the endogenous CFTR promoter, and restored ion transport as measured by Ussing chamber [363]. Also using ZFNs was a study for HDR correction of the F508del mutation in CFTR29- cells, however correction rate was low (<1%), not surprising due to a low DSB efficiency of 7.8% [364], while no functional validation was conducted. In another study, skin fibroblasts were obtained from compound heterozygous CF patients (I507del/F508del) patients, from which induced pluripotent stem cells (iPSCs) were obtained. Using ZFNs to target the CFTR allele, the researchers found a sensitive homology-dependent preference for one CFTR allele versus the other, highlighting just how precise gene editing systems are. The corrected iPSC cells were differentiated *in vitro*, and restored expression of mature CFTR protein and chloride channel function was observed [365].

TALENs have also been used to correct the F508del mutation in an iPSC cell line. Using electroporation, TALENs and a single-stranded oligo DNA nucleotide (ssODN) donor template were delivered into the cells and, after generating single cell colonies by limiting dilution, researchers identified 5 clones with mono-allelic correction. One of the clones, which expressed endogenous core pluripotency markers, could be differentiated *in vitro* into derivatives of all three germ layers [366]. Recently, TALENs were used to insert a CFTR minigene into the AAVS1 locus of IB3-1 cells, delivered by helper-dependent adenoviral (HD-Ad) vectors, with an insertion efficiency of 5%. In this study, the group showed the rapid

clearance of TALENs within the transduced cells, as measured by mRNA and protein levels, limiting the chance of host immune response [367].

CRISPR/Cas9 systems have also been used to investigate corrective strategies for CF. Schwank and colleagues used CRISPR/Cas9 to correct intestinal stem cells obtained from two different paediatric patients homozygous for F508del mutation via HR. Upon correction, fully functional CFTR was expressed, as measured in clonally expanded organoids, with a normal response to forskolin-induced swelling [101]. Recently, a 20% correction rate in an iPSC line homozygous for the F508del mutation, without the use of viral vector, drug selection or reporter enrichment (VDR free), which restored CFTR function in iPSC-derived lung organoids. The group used electroporation to deliver gRNA plus Cas9 protein in a pre-formed ribonucleoprotein complex (RNP) [368]. Porteus and colleagues recently achieved even higher editing efficiency in primary upper airway basal stem cells obtained from F508del homozygous patients. Also using electroporation to deliver Cas9 + gRNA RNP complex, followed by immediate transduction of donor template by AAV6 vector, the group were able to achieve precise editing in 43% of alleles in a feeder-free, selection-free model [369].

Splice mutations 3272–26A>G and 3849+10kbC>T create a new splice donor site and acceptor site respectively, both leading to the inclusion of in-frame stop codons and consequent production of truncated, non-functional CFTR. Maule and colleagues used an alternative Cas9 variant, *Acidaminococcus sp*. BV3L6 (AsCas12a) and a single gRNA to correct these genetic defects [370]. Allelic discrimination was obtained by designing the gRNA target sequence on the mutations, with an absence of off-target effects. Complete functional recovery of the CFTR channel was verified in intestinal organoids and airway epithelial cells from patients harbouring these mutations.

In addition to using the technology for correction of CF mutations, researchers have generated CF animal models using CRISPR/Cas9, most recently a CF sheep. Owing to the similarity between sheep and human lung anatomy and development, the sheep model may be particularly relevant in CF research. Newborn $CFTR^{-/-}$ lambs developed severe disease, consistent with CF pathology in humans, although lungs at birth appeared grossly normal [371].

Although these studies present encouraging proof-of-concept data, their translational potential is limited because most these strategies rely on HDR mechanisms. HDR pathways are only functional in mitotic cells yet, as discussed above, it is most likely that for a CF

therapy, cells targeted by aerosol mediated systems will be deposited on the non-dividing, terminally-differentiated, surface epithelium. Thus, strategies that exploit non-HDR pathways are gaining more attention for *in vivo* gene editing including PITCh, (Precise Integration into Target Chromosomes), which exploits micro-homology mediated end joining [372] and HITI (Homology Independent Targeted Integration), which exploits the NHEJ pathway and preferentially integrates a donor template in the correct orientation. This strategy was used to successfully restore a large deletion in the *Mertk* gene in a rat model of Retinitis pigmentosa [373].

1.2 Aims

CRISPR/Cas9 could provide a simple, versatile and precise method of genetic manipulation for the treatment of Cystic Fibrosis. However, a lack of suitable delivery systems remains a major hurdle in the development of a gene editing therapy *in vivo*.

The aims of this thesis are:

1. To develop a suitable cellular model that closely mimics the CF lung, in which to assess delivery and editing efficiency.

2. Optimise receptor-targeted nanocomplexes for CRISPR/Cas9 delivery in mRNA and RNP format.

3. Design an NHEJ-mediated correction strategy for the 3949 + 10 kb C>T mutation in *BMI-1* modified bronchial epithelial cells, and assess functional restoration of CFTR.

4. Evaluate the homology-independent targeted integration (HITI) strategy as alternative approach to HDR for transgene knock-in in HEK293T cells.

5. Assess the use of RTNs for delivery of CRISPR/Cas9 to the airways of mice.

Chapter 2
Materials and Methods

2.1 Materials and methods

2.1.1 Equipment

Name	Supplier
Ussing chambers	World Precision Instruments
Bio-Rad 96 CFX	Bio-Rad
Canon Rebel XS DS126191 Digital Camera	Canon
BD LSRII	BD Biosciences
FACS Calibur	BD Biosciences
MoFlo XDP Cell Sorter	Beckman Coulter
Zeiss Axioplan Scopes	Zeiss
Trans-Blot Turbo Transfer System	Bio-Rad
NanoDrop ND-1000 Spectrophotometer	Thermo Fisher Scientific
Olympus IX70 microscope	Olympus
Thermocycler (Mastercycler ep)	Eppendorf
Gel Doc XR System	Bio-Rad
Ohmmeter (EVOM)	World Precision Instruments
Zetasizer Nano ZS	Malvern
NanoZoomer	Hamamatsu
Mark IV Vitrobot	Thermo Fisher Scientific

2.1.2 Kits and reagents

Name	Supplier
DMEM+GlutaMAX	Thermo Fisher Scientific
GeneRuler 1 kb Plus DNA Ladder	Thermo Fisher Scientific

6x DNA Loading Dye	Thermo Fisher Scientific
Q5 Hot Start High-Fidelity DNA Polymerase	New England Biolabs
Lipofectamine 2000	Thermo Fisher Scientific
L-glutamine (100x)	Thermo Fisher Scientific
NuPAGE 4x loading dye buffer	Thermo Fisher Scientific
NP40 Cell Extraction Buffer	Thermo Fisher Scientific
dNTPs	New England Biolabs
Eagle's Minimal Essential Medium (MEM)	Sigma-Aldrich
Bolt Sample Reducing Agent 10X	Thermo Fisher Scientific
0.05% Trypsin-EDTA	Thermo Fisher Scientific
Foetal Bovine Serum (FBS)	Sigma-Aldrich
B-Mercaptoethanol	Sigma-Aldrich
GeneJET Gel Extraction	Thermo Fisher Scientific
GeneJET PCR Purification	Thermo Fisher Scientific
GeneJET Genomic DNA Purification	Thermo Fisher Scientific
PureCol, bovine collagen	Nutacon
Pierce ECL Western Blotting Substrate	Thermo Fisher Scientific
SuperSignal West Femto Max. Sensitivity Substrate	Thermo Fisher Scientific
Penicillin/Streptomycin (100x)	Thermo Fisher Scientific
OptiMEM	Thermo Fisher Scientific
BEBM Basal Medium	Lonza
SingleQuot Kit Suppl. & Growth Factors	Lonza

SOC Media	New England Biolabs
T4 DNA Ligase and Buffer	Thermo Fisher Scientific
4-20% pre-cast polyacrylamide gels	Bio-Rad
10x Tris Buffered Saline (TBS)	Bio-Rad
Trans-Blot Turbo Mini PVDF Transfer Packs	Bio-Rad
CloneJET PCR Cloning Kit	Thermo Fisher Scientific
Non-Essential Amino Acids (NEAA) (100x)	Thermo Fisher Scientific
Precision Plus Protein Dual Color standards	Bio-Rad
NEBuilder HiFi DNA Assembly	New England Biolabs
HiScribe T7 High Yield RNA Synthesis Kit	New England Biolabs
T7 Endonuclease I	New England Biolabs
CleanCap Cas9 mRNA (modified)	TriLink BioTechnologies
CleanCap Cre mRNA (5moU)	TriLink BioTechnologies
Alt-R S.p. Cas9 Nuclease 3NLS	Integrated DNA Technologies
SensiFAST Probe Hi-ROX one-step kit	Bioline
SYBR Safe DNA Gel Stain	Thermo Fisher Scientific
Qiagen RNeasy mini kit	Qiagen
Pierce Protein assay BCA kit	Thermo Fisher Scientific

2.1.3 Eukaryotic cells

Cells	**Supplier**	**Growth media**
HEK293T	ATCC	DMEM+ GlutaMAX

NHBE	Lonza	BEGM or ALI media
CFBE 3849	UNC	BEGM or ALI media

2.1.4 Lipids

Lipid	Supplier
1,2-Di-((Z)-tetradec-11-enyloxy)-N,N,N trimethylammonium propane chloride	Avanti Polar Lipids
1,2-Di-((Z)-hexadec-11-enyloxy)-N,N,N trimethylammonium propane iodide	Avanti Polar Lipids
N-[1- (2,3-dioleyloxy)propyl]-N,N,N-trimethlylammonium chloride	Avanti Polar Lipids
Dioleoyl L-α phosphatidyl ethanolamine	Avanti Polar Lipids

2.1.5 Liposome composition

Liposome name	Cationic Lipid	Helper Lipid
DOTMA:DOPE (C18)	DOTMA (50%)	DOPE (50%)
DHDTMA:DOPE (C16)	DHDTMA (50%)	DOPE (50%)
DTDTMA:DOPE (C14)	DTDTMA (50%)	DOPE (50%)

2.1.6 Peptides

Name	Composition	Supplier
Y	K16GACYGLPHKFCG	China Peptides Co.
E	K16GACSERSMNFCG	China Peptides Co.

2.1.7 Plasmids

Name	Transgene and resistance gene
pMC_A78_SA_GFP_T7	GFP and kanamycin
64323_AAVS1_A78	BFP and ampicillin
pRRLSIN_cPPT_PGKGFP_WPRE	eGFP and kanamycin
pLVX.Puro.BMI1	puromycin

2.1.8 Primers

Primer	Sequence (5' > 3')	Amplified product size (bp)
CFTR intron 22 F	TCCTCTCAAAATGCCTACTGG	993
CFTR intron 22 R	GGTTGGGAAAGACTGGATGA	
HsAAVS1 F	AGTCTTCTTCCTCCAACCCG	630
HsAAVS1 R	CAAACTGCTTCTCCTCTTGG	
HsAAVS1 F	AGTCTTCTTCCTCCAACCCG	250
TRANS 5E-2 R	CTCACAATTCCACACAACA	

2.1.9 Taqman Assays (FAM-labelled probes)

Name	Catalogue #	Supplier
Human ACTB	Hs01060665_g1	Thermo Fisher Scientific
Human CFTR	Hs00357011_m1	Thermo Fisher Scientific

2.1.10 Antibodies

Name	Supplier	Catalogue #	Dilution
Mouse anti-CFTR 596	UNC-Chapel Hill	A4	1:2,000
Mouse anti- Na+/K+-ATPase	Santa Cruz	sc-48345	1:20,000
Goat-anti-mouse IgG-HRP	Santa Cruz	sc-2005	1:10,000
APC Rat Anti-Mouse CD45	BD Biosciences	561018	

BV421 Rat Anti-Mouse CD326	BD Biosciences	563214	
Rabbit anti-RFP	Rockland	600-401-379	1:100

2.1.11 Recipes

Solution	Composition
TBST	50 mM Tris/HCl, pH 7.5, 150 mM NaCl, 0.1% Tween 20™
Citrate buffer	1.14 g Sodium Citrate dihydrate, 500 ml H2O, pH 6.0
Tris-acetate- EDTA buffer(TAE)	4.84 g/L Tris base, 0.114% (v/v) glacial acetic acid and 0.2% (v/v) 0.5M EDTA at pH 8.0
ALI media	250 ml of BEBM, 250 ml of DMEM, 1x SingleQuot Kit Suppl. & Growth Factors (without gentamicin) and 1nM retinoic acid

2.2 Methods

2.2.1 Collagen coating of plastic ware

Collagen coating: PureCol® was prepared as a 1% (0.3mg/ml) solution in phosphate buffered saline (PBS). A volume sufficient to completely cover the surface was added to plates/flasks/transwells. Plasticware was then incubated for 1 h at room temperature, before one was in PBS to remove any acidity from the collagen coat. The plasticware can then be air dried and stored in a sealed bag.

2.2.2 Cell Maintenance

NHBE BMI-1 cells were seeded at a density of 3 x 10^5 in a collagen coated 75 cm^2 flask with 15 ml BEGM media. Media was changed every 48 h. After 7 days, when flask is 90-100% confluent, all medium was removed and cells washed with 1 ml Trypsin/EDTA. This was replaced with a further 1 ml of Trypsin/EDTA and incubated at room temperature, agitating the flask occasionally until cells begin to detatch. 5 ml Opti-MEM was then added to the flask and any remaining cells were washed off by pipetting the medium over the surface of the flask. The cells suspension is then transferred to a 15 ml tube and cells pelleted at 1200 rpm for 10 m. Supernatant was removed and cells re-suspended in 2 ml BEGM for counting.

2.2.3 ALI Culture

Cells were seeded in collagen coated Snap-well plates at a density of 0.5 x 10^6 per 1.2cm^2 in 250 μl BEGM. 1 ml per well of BEGM was added to the basolateral side. Twenty four h later, media was aspirated from both the apical and basolateral side of the membrane, and the basolateral side replaced with 1 ml ALI media. Media was replaced every 48 h.

2.2.4 Transepithelial electrical resistance (TEER) measurements

The TEER was measured using an Ohmmeter (EVOM) after replacing the apical media with 250 μl of PBS. The PBS was removed immediately after measuring the TEER, re-exposing the cells to air.

2.2.5 Lentiviral transduction of GFP in NHBE BMI-1 cells

NHBE BMI-1 cells were seeded at a density of 1 x 10^5 per well of a 6 well plate. The next day, cells were transduced with a pLenti PGK GFP Puro plasmid (Addgene) (multiplicity of infection (MOI) of 1) in 1 ml BEGM. 24 h later, each well was topped up with 1 ml BEGM. 48 h later, media was replaced with 2 ml BEGM. When confluent, cells were transferred to a collagen-coated 75 cm^2 flask. 1 week later, cells were sorted for GFP expression on the MoFlo XDP Cell Sorter, and GFP positive cells expanded in 75 cm^2 flasks.

2.2.6 Cell cloning by serial dilution

Cells were detached and pelleted as described above, before re-suspending in BEGM at a concentration of 1 x 10^6 cells per ml. Cells were serially diluted to a final concentration of 1 x 10^3 cells per ml. Cells were seeded at a density of one cell per well in x2 collagen-coated 96-well plates in 200 μl media (1:1 ratio of BEGM media: conditioned media (BEGM harvested from NHBE BMI-1 cells at 70% confluency). Plates were periodically inspected between days 7 and 10 to assess if colonies are establishing. Successful single cell colonies were chosen for expansion.

2.2.7 Liposome formation

Liposomes were formed using a NanoAssemblr (Precision Nanosystems, Canada), a microfluidic based platform. The cationic lipid (C14 (DTDTMA), C16 (DHDTMA) or C18 (DOTMA)) was mixed together with DOPE in ethanol at a molar ratio of 1:1, and injected into the cartridge at a flow rate of 12 ml min^{-1}. The newly formed liposomes were then dialysed

overnight in SnakeSkin™ Dialysis Tubing (10K MWCO, 22 mm) in 2 L sterile MilliQ water at room temperature with stirring, to remove residual ethanol. Liposomes were then sonicated in a water bath for 20 m to reduce the size, producing small, unilamellar vesicles. Liposomes were diluted to 1 mg/ ml for storage at 4°C.

2.2.8 gRNA preparation

Guides were ordered as Alt-R® CRISPR-Cas9 crRNA and tracrRNA (IDT Technologies, Belgium). crRNA and tracrRNA were assembled at an equimolar concentration of 30 μM and incubated at 95°C for 5 m to anneal.

2.2.9 *In vitro* Cas9 mRNA and Cas9 RNP transfections

Bronchial epithelial cells were seeded in 24 well plates at a density of 0.6 x 10^5 cells per well in a total volume of 1 ml BEGM media 24 h before transfection. Nanocomplexes were prepared in 100 ul Opti-MEM per well at a weight ratio of 1:3:4 mRNA: lipid: peptide, or ribonucleoprotein complex (RNP): lipid: peptide, unless otherwise stated. For DNA transfections, the ratio was 1:1:4 DNA: lipid: peptide. Components were incubated for 30 m at room temperature, allowing the complexes to self-assemble. Nanocomplexes were then diluted in 1 ml Opti-MEM per well and mixed by vigorous pipetting. BEGM media was removed from cells and diluted nanocomplexes added. Cells were incubated at 37°C for 4 h, after which Opti-MEM was replaced with BEGM. Cells were maintained for 8 days before assessment by Flow Cytometry.

2.2.10 DNA extraction

Total DNA was extracted using a GeneJET Genomic DNA Purification Kit following the manufacturer's guidelines, based on the reversible binding of DNA to a silica-gel membrane followed by elution in low-salt conditions. Cells were harvested then pelleted at 13,000 rpm for 5 minutes. The cells were resuspended in 200 µl of Lysis Solution and 20 µl of Proteinase K Solution, followed by 10 minutes incubation at 56°C. After that, 20 µl of RNase A Solution was added and the cell lysates were incubated for 10 minutes at room temperature, followed by the addition of 400 µl of 50% ethanol and transfer to the binding column. Two washes were required for the column, first with 500 µl of washing buffer I then with 500 µl of washing buffer II, followed by 2 minutes spinning at 13, 000 rpm to

ensure that no ethanol was carried over during the DNA elution step. Finally, the genomic DNA was eluted into a new tube with 50 µl of nuclease-free water. The DNA quantity was measured by OD_{260} using NanoDrop ND-1000. DNA samples with 260/280 of 1.8-2.0 were sufficiently pure for use in further experiments. The genomic DNA obtained was stored at 4°C or -20^{o}C for long term storage.

2.2.11 RNA extraction

Total RNA was extracted using an RNeasy mini kit following the manufacturer's guidelines which are based on a silica-based membrane column selective for RNA binding of samples followed by elution in low-salt conditions. Cells were harvested then pelleted at 13,000 rpm for 5 minutes. The cells were resuspended in 350 µl RLT buffer supplemented with 1% β-mercaptoethanol. 350 µl of 70% Ethanol was then added into cell lysates and were transferred to a binding column. Three washes were required for the column - first with 700µl RW1 washing buffer, then twice with 500 µl of RPE washing buffer, followed by 2 minutes spinning at 13,000 rpm to ensure that no ethanol was carried over during the RNA elution step. Finally, total RNA was eluted by adding 30 µl RNase-free water to the column membrane, then centrifuging at 13,000 rpm for 1 minute. Elution step was repeated with the same elute to achieve greatest RNA concentration. RNA quality and concentration was measured by OD_{260} using NanoDrop ND-1000. RNA samples with 260/280 of 1.9-2.1 were sufficiently pure for use in further experiments. The RNA obtained was stored at -80^{o}C for long term storage.

2.2.12 Polymerase chain reaction (PCR)

PCR reactions were performed with 50 µl reaction volume using 100 ng of DNA template, 10 µl of 5X Q5 Reaction Buffer, 2.5 µl of 10µM primers (forward and reverse), 1 µl of 10mM dNTPs, 0.5 µl Q5 Hot Start High-Fidelity DNA Polymerase, made up to 50 µl with nuclease-free water. The cycling conditions were as follows: an initial denaturation and a polymerase activation step at 98°C for 30 s, followed by 35 cycles of denaturation (10 s, 98°C), annealing (30 s, 3°C above the melting temperature (Tm) of the lower Tm primer) and extension (30 seconds/kb, 72°C). A final extension cycle (5 min, 72°C) completed the PCR reaction. The PCR reaction was then analysed on an agarose gel and visualised using SYBR Safe staining.

2.2.13 Agarose gel electrophoresis

Agarose gel for DNA electrophoresis was prepared in 1X TAE with the appropriate amount of agarose depending on the size of the DNA product to be visualised, typically 1-2%. Addition of 1 µl of SYBR Safe DNA Gel Stain per 10 ml of 1X TAE allows for DNA product visualisation. The agarose was dissolved by heating in a microwave oven, then poured into an appropriately-sized casting tray. To run samples on the agarose gel, the DNA was mixed with 6X DNA loading dye at a final concentration of 1X, then separated by running in 1X TAE buffer at 90 V for 60 min. The DNA bands were visualised and documented using the UVIdoc system.

2.2.14 PCR purification

The PCR fragment was purified when it was necessary using a GeneJET PCR Purification Kit following the manufacturer's instructions. The PCR reaction was mixed with same volume of binding buffer, then transferred into a binding column, followed by washing with 700 µl of wash buffer. Finally, the PCR product was eluted into a new tube with 50 µl of nuclease-free water. The DNA quantity was measured by OD_{260} using NanoDrop ND-1000. DNA samples with 260/280 of 1.8-2.0 were sufficiently pure for use in further experiments. The DNA obtained was stored at 4°C or -20^{o}C for long term storage.

2.2.15 T7 endonuclease I assay

200 ng of purified PCR product was mixed with 2 µl of NEB buffer 2, and nuclease-free water up to 19 µl. This mixture was heated at 95°C for 5 min for denaturation, then left at room temperature for 20 min for re-annealing. 1 µl of T7 endonuclease I enzyme was added to the mixture, before incubation at 37°C for 20 min. Immediately, the reaction was loaded on 2 % agarose gel which was visualised using SYBR Safe staining.

2.2.16 Sanger sequencing

50ng of DNA sample, along with 5 µM of the appropriate forward primer was sent for sequencing to the GENEWIZ company. The sequence results were examined using the SnapGene Viewer version 4.1 software.

2.2.17Determination of indel % using TIDE and ICE software

To determine the indel percentage using the Tracking of Indels provided by Decomposition software (TIDE) or Inference of CRISPR Edits (ICE), a purified PCR product from the transfected and un-transfected cell pool were sent first for Sanger sequencing. A chromatogram sequence was then uploaded to http://tide.nki.nl/ or https://ice.synthego.com/#/ for analysis.

2.2.18 Analysis of CFTR protein expression by Western Blot

The culture medium was removed from the 12-well support plate and Snapwell wiped and washed briefly in PBS to remove excess media. 150 µl of PBS per well was added to the top of the Snapwell and cells were gently scraped with cell scraper. Cell suspension was transferred to microcentrifuge tube. An additional 150 µl was added to collect remaining cells, and the cell suspensions combined. The sample was centrifuged at 15,000 g for 2.5 min. PBS was removed and 60 µl ice cold lysis buffer was added to fully disrupt pellet. The sample then rested on ice for 20 min, before centrifugation at 15,000 x g for 20 min in a 4°C pre-cooled centrifuge. Supernatant was then transferred to a new tube.

Total protein concentration was measured using Peirce™ BCA Protein Assay Kit. Eight µg of each sample was loaded into wells of 4-20% pre-cast polyacrylamide gels (BioRad) and run for 90 min at 125 V, before transfer using Trans-Blot Turbo Transfer System at 25 V for 30 min. Membrane was then washed in TBST (50 mM Tris/HCl, pH 7.5, 150 mM NaCl, 0.1% Tween 20™) before blocking in TBSMT (TBST containing 5% Marvel™) for 1 h at room temperature. Membrane was then incubated with the anti-CFTR UNC-596 Primary Antibody (1/2000) in blocking solution overnight at 4°C. Membrane was washed in TBST before incubation with anti-mouse IgG-HRP labelled Secondary Antibody (1/10,000) in blocking buffer at room temperature for 1 h. Membrane was incubated with SuperSignal West Femto Substrate and image acquired using GelDox. Membrane was washed before incubation with anti-Na/K-

ATPase primary antibody (1/20,000) in blocking buffer at room temperature, followed by washing and incubation anti-mouse IgG-HRP labelled secondary antibody (1/10,000) in blocking buffer at room temperature for 1 h. Membrane was washed before signal development using Clarity Western ECL Substrate, and an image was acquired using GelDox.

2.2.19 CFTR mRNA quantification by qRT-PCR

Culture medium was removed from the 12-well support plate and Snapwell wiped and washed briefly in PBS to remove excess media. 150 µl of PBS per well was added to the top of the Snapwell and cells were gently scraped with cell scraper. Cell suspension was transferred to microcentrifuge tube. An additional 150 µl was added to collect remaining cells, and the cell suspensions combined. Cells were pelleted at 2,500 rpm for 2.5 min. PBS was removed and the pellet re-suspended in 350 µl IP lysis buffer and 1% β-mercaptoethanol. RNA was extracted using the RNeasy Mini Kit (Qiagen), according to the manufacturer's instructions. Concentration was measured using a Nanodrop 2000 (Thermo Fisher Scientific). RNA quality was verified on a TAE agarose gel as follows. 60% formamide was added to 500 ng RNA along with 6X loading dye. Mixture was incubated at 65°C for 10 min before resting on ice for 5 min. Samples were ran on a 1.5% TAE agarose gel (60 V, 1 h). Gel was incubated in 100 ml TAE + 15 µl SYBR™ Safe gel stain for 30 min, before destaining in TAE for 15 min. Image was acquired using a GelDox. RNA was reversed transcribed into cDNA using the SuperScript IV Reverse Transcriptase (Invitrogen), according to the manufacturer's protocol. Reverse transcription reaction was used immediately for PCR amplification.

Table 2.1: PCR reaction mix

Reagent	Volume
20X Primer & Probe	1 µl
2X Taqman Mastermix	10 µl
cDNA	2 µl

Table 2.1 qRT- PCR cycling conditions

Step	Temp.	Time	Cycle

Initial denaturation	95 °C	15 min	X1
Denature	95 °C	15 s	
Anneal/extend	60 °C	1 min	X45

2.2.20 Flow cytometry for in vitro transfections

Cells grown on 24 well plates were detached using 100 µl Trypsin-EDTA, and re-suspended with 200µl of the respective culture media. Cells were then acquired with a BD LSRII or FACS Calibur flow cytometer. Data analysis was performed using the FlowJo software v.10 software.

2.2.21 Biophysical characterisation of RTNs

The size and charge (ζ potential) of the nanocomplexes were measured using a Malvern Zetasizer nano ZS. The nanocomplex was first prepared on water with a minimum of 1-2 µg of nucleic acid (DNA or mRNA) or RNP, then incubated for 30 min at room temperature. The sample was then diluted in final volume of 1 ml and transferred to a cuvette for measurements. Size was measured first before ζ potential.

2.2.22 Gel retardation assay

Complexes were prepared in water and either left untreated or treated with 20 U/mL heparan sulphate for 1 h at room temperature. 250 ng of plasmid DNA complexed into particles (or plasmid DNA alone as a control) were loaded onto a 1% agarose gel, made in TAE buffer and stained with 1 µg/ml ethidium bromide, and electrophoresed at a voltage of 100 V for 1 h with TAE as the running buffer.

2.2.23 Cryogenic Transmission Electron Microscopy

RTNs were concentrated to a final concentration of 20 mg/mL of total liposome. 2–4 µl of RTN suspension was added to glow-discharged copper grids and plunge-frozen using a FEI Mark IV Vitrobot (FEI, Hillsboro, OR) to generate vitreous ice. Grids were stored in liquid nitrogen until

imaged. Grids were moved into a Gatan 70° cryo-tilt transfer system pre-equilibrated to −180 °C before insertion into the microscope. An FEI LaB6 G2 TEM (FEI, Hillsboro, OR) operating at 200 kV under low-dose conditions was used to image all samples. A bottom-mount FEI Eagle 4K CCD camera was used to capture all images. All samples were imaged at a 55,000× magnification with a nominal under-focus of 1–2 µm to enhance contrast. Sample preparation and imaging were performed at the UBC Bioimaging Facility (Vancouver, BC).

2.2.24 Electrophysiology studies

Primary epithelium cells grown on Air-liquid interface (ALI) were bathed in Ussing chambers in Krebs-Henseleit (K-H) solution consisting of NaCl (117 mM), $CaCl_2$ (2.5 mM), KCl (4.7 mM), $MgSO_4$ (1.2 mM), $NaHCO_3$ (25 mM), KH_2PO_4 (1.2 mM) and D-glucose (11 mM), where pH was 7.4 and temperature was 37°C. The buffer was bubbled at 21% O_2 and 5% CO_2 gas. The cell monolayers were maintained in open-circuit conditions until the transepithelial potential difference (V_t) and resistance (R_t) were stable. Following that, the monolayer cells were short-circuited, applied through DVC-4000 voltage/current clamp to bring V_t to 0 mV. The short circuit current I_{sc} was measured and recorded using a PowerLab computer interface as µA/cm^2. Every 30 s the settings returned to open circuit for 3 s so that the spontaneous V_t could be measured and resistance could be calculated. Drugs circulated in bathing buffer were as follows: amiloride (10 µM) at the apical side to block the sodium channel activity, followed by CFTR activator drugs added both basally and apically; forskolin (25 µM) and 3-isobutyl-1-methylxanthine (IBMX) (100 µM) Finally, to block CFTR-dependent I_{sc}, ,10 µM of CFTRinh-172 was added to the apical side.

2.2.25 Luciferase assay

Twenty four h post transfection, mice were euthanized with 100 µl of pentobarbital sodium (Euthatal) administered intraperitoneally. The chest cavity was exposed post-mortem and the heart perfused with 5 ml of PBS, after which the lungs were dissected and stored at -80°C for a minimum of 24 h. Lungs were weighed, placed in 2 ml tubes containing 1.4 mm ceramic beads and 4 µl of Reporter Gene Assay Lysis Buffer added per milligram of tissue. The mixture was then frozen at -80^{o}C for 20 min, left to thaw at room temperature for 10 minutes and

then placed on ice for a further 10 min. Tissue homogenisation was then performed using a Percellys 24 tissue homogeniser at a 3D motion speed of 5,700 rpm for 30 s, performed twice with a 30 s delay between the two steps. Samples were then centrifuged at 14,196 x g for 10 min at 4°C to pellet the ceramic beads and tissue debris. The supernatant was transferred into autoclaved microfuge tubes and assayed for luciferase activity using the Luciferase Assay System and luminescence detected with a FLUOstar Optima plate reader. Results were standardised for protein content using the Bradford protein assay and expressed as RLU/mg of protein.

2.2.26 Bradford protein assay

20 µl of sample (lysate + luciferase substrate) was mixed with 180 µl of 1x Bradford protein assay reagent in a clear 96 well plate, and incubated at least for 5 min at RT. The absorbance at 595 mm was measured by the FLUOstar OPTIMA plate reader. Total protein concentration per well was calculated using a protein standard curve obtained from several known concentrations of bovine serum albumin (BSA).

2.2.27 Flow cytometry for *in vivo* transfections

Transfected mice were euthanized using pentobarbital administered intraperitoneally. The chest cavity was exposed by midline sternotomy, then the heart was perfused with 5 ml of PBS. The lungs were then dissected and kept in PBS, then transferred to a falcon tube for digestion in 2 mg/ml of Collagenase D, then incubated for 30 min at 37°C with automated rotation. The cells were passed through a strainer (40 µm) to ensure single cell suspension. The cells were washed twice with PBS then counted, and $2x10^6$ cells per sample were stained with DAPI for viability, CD45 (APC) and CD326 (BV421) 30 min at room temperature then acquired with a BD LSRII flow cytometer and the data were then analysed using the FlowJo software v.10.

2.2.28 Histology

24 hours following transfection mice were euthanized with 100 µl of pentobarbital sodium (Euthatal) administered intraperitoneally. The chest cavity was exposed by midline sternotomy post-mortem and the lungs inflation-fixed by intratracheal instillation of 4% (w/v) paraformaldehyde in PBS at a pressure of 20 cm 4% paraformaldehyde. The lungs were then dissected and immersed in fixative overnight at 4°C and then immersed in a 15% (w/v) sucrose solution in PBS overnight at 4°C. The tissue was then dehydrated in alcohol and embedded in paraffin wax. 5 µm thick sections were cut, dried onto microscope slides, de-waxed, rehydrated, stained with haematoxylin, counter-stained with eosin, dehydrated and mounted. Images were then taken using a NanoZoomer automated digital slide scanner and processed with NDP.view software (v2.2.1; Hamamatsu Photonics).

2.2.29 Immunohistochemistry protocol

Samples were dewaxed and transferred into H_20, before washing in TBS for 5 min. Antigen retrieval required microwaving in 10 mM citrate buffer 10 min at 80% power followed 10 min at 60 % power, before allowing to cool for a further 10 min. Samples were washed in TBS for x2 5 min, before incubation in 3% hydrogen peroxide in ddH_20 to block endogenous peroxidase. Slides were then blocked for 20 min with ImmPRESS ready-to-use 2.5% normal horse serum blocking solution. Slides were drained, before overnight incubation at 4°C with anti-RFP (rabbit) at a 1:100 dilution, with a total volume of 200 µl per slide. The following morning, slides were washed in TBS for 2 x 5 min, before a 45 min incubation with secondary antibody- ImmPRESS peroxidase anti-rabbit Ig reagent (Vector Labs, MP-7452). Slides were washed in TBS for 2x 5 min, followed by incubation with ImmPACT NovaRed substrate until colour develops as seen under a microscope. Slides remained in ddH_2O before counterstaining, dehydration, and clearance in xylene. Coverslips were applied and mages were then taken using a NanoZoomer automated digital slide scanner and processed with NDP.view software (v2.2.1; Hamamatsu Photonics).

2.2.30 Autostainer protocols

2.2.30.1 Dewax

Step	Time
Dry	30 s
Xylene	3 min
Xylene	3 min
100% EtOH	2 min
100% EtOH	2 min
70% EtOH	2 min
30% EtOH	2 min

2.2.30.2 Counterstain

Step	Time
ddH2O	30 s
Haematoxylin	60 s
Tap water	20 s
1 % HCL, 70% EtOH	15 s
Tap water	2 min
ddH2O	30 s
70% EtOH	2 min
100% EtOH	2 min
100% EtOH	2 min
Xylene	3 min
Xylene	3 min

2.2.30.3 H&E staining

Step	Time
Dry	30 s
Xylene	3 min
Xylene	3 min
100% EtOH	2 min
100% EtOH	2 min

70% EtOH	2 min
30% EtOH	2 min
Distilled water	1 min
Haematoxylin	2 min 30 s
Tap water	20 s
1 % HCL, 70% EtOH	15 s
Tap water	2 min
Eosin	2 min
Tap water	2 min
70% EtOH	30 s
100% EtOH	2 min
100% EtOH	2 min
Xylene	2 min
Xylene	2 min

2.2.31 Data analysis

Data was analysed using GraphPad version 7.0, and Microsoft Excel 2016, and were then expressed as Mean ± Standard Deviation (SD). The significant differences between 2 groups was calculated using the Student's T. Test. When more than 2 groups were being compared, the one-way analysis of variance (ANOVA) test was used, with multiple comparisons. P values of less than 0.05 were marked with *, p values of less than 0.01 were marked with **, p values of less than 0.001 were marked with ***, and p values of less than 0.0001 were marked with ****.

Chapter 3
Results

Optimisation and characterisation of Receptor-Targeted Nanocomplexes for delivery of CRISPR/Cas9

3.1 Introduction

The CRISPR/Cas9 genome editing technology is widely applied as a powerful tool in scientific research, and is a potential therapeutic for genome regulation in a myriad of genetic diseases. However, delivery of the CRISPR components into cells and target organs remains a challenge, which must be overcome if the system is to be translated to a therapeutic purpose in humans. A delivery method must be both safe and efficient, in order to realise the potential of this robust and precise technology.

A suitable model is required to test the delivery and efficiency of any gene editing strategy, and to assess the functional consequences of editing. For CF, primary epithelial cells are the obvious choice for an *in vitro* system. These cells can be obtained by brushing of nasal or bronchial epithelium and, when maintained on ALI culture, the basal cells obtained from these brushings will differentiate into a pseudostratified, ciliated, columnar epithelium, typical of a lung airway. A key limitation of basal cells is that they lose their differentiation capacity after two passages, necessitating frequent sampling from CF donors. Genetic variability among individuals with CF requires that samples are obtained from the same donor. While immortalisation of cells generally results in loss of differentiation, recent advancements in the field have provided new ways of expanding cell proliferative potential and the ability to differentiate by transduction with telomerase hTERT and the anti-senescent, proto-oncogene *BMI-1* [95], or even *BMI-1* alone, as our lab has previously shown [96]. It is essential that these modified cells maintain the key properties of primary bronchial epithelial cells. For example, it has been suggested that CFTR expression decreases over time [374]; if our gene editing strategies are to correct CFTR, we need to ensure CFTR is expressed. It is therefore necessary to characterise these modified cells to ensure they maintain these basic properties.

As mentioned, the prerequisite for effective genome editing is efficient delivery in cells, tissues, and organs. To date, the delivery of Cas9 and sgRNA remains challenging as a result of the large size of Cas9 (4.1 kb). Our lab has previously used receptor-targeted nanocomplexes (RTNs) comprised of cationic lipid and targeting peptide moieties to deliver nucleic acids, in the form of plasmid, mRNA and siRNA, to a variety of cell types and tissues, including the mouse lung by means of aerosolisation [375-380]. These RTNs have several advantages over their viral counterparts, namely an unrestricted carrying capacity and

decreased likelihood of eliciting an immune response. Here, we repurpose and optimise RTNs for CRISPR/Cas9 delivery, and biophysically characterise these particles.

3.2 Aims

In this chapter, we aim to:

1. Develop a suitable cellular model that closely mimics the CF lung, in which to assess delivery and editing efficiency.

2. Optimise receptor-targeted nanocomplexes for CRISPR/Cas9 delivery in mRNA and RNP format.

3. Biophysically characterise RTNs and assess packaging efficiency.

3.3 Results

3.3.1 BMI-1 modified epithelial cells as a suitable model for CF

BMI-1 is an anti-senescence protein that is known to function in the self-renewal and maintenance of the multipotency of stem cells [381, 382]. Transduction of *BMI-1* has been previously shown to allow cells to be expanded over at least 20 population doublings [96], and can therefore be used to overcome the main limitation of use of primary bronchial epithelial cells. Full length human cDNA (10 kb) of the *BMI*-1 gene was cloned into a plasmid, and lentivirus produced as previously described [383] by another researcher within our lab. Primary NHBE cells (P2) were then transduced with virus at different MOIs- 1, 4 & 16, where MOI is the ratio of plaque forming units (pfu) of virus per cell.

3.3.1.1 Proliferative potential of NHBE BMI-1 cells

In order to assess the proliferative potential of the *BMI-1* cells, population doubling capacity was calculated for each of the MOIs. While passage number refers to the number of times the cells in culture have been sub-cultured, without consideration of inoculation density, analysis of population doubling refers to the total number of times the cells in a population have doubled since their primary isolation *in vitro*. The formula to calculate population doublings is as follows: n= In(A/B), where A is the cumulative cell number and B is the cell number used as inoculum to begin that subculture. Each MOI was able to be expanded up to 44 population doublings, over 85 days, when cultures were stopped. Meanwhile, primary NHBEs achieved 18 population doublings over 44 days, however their proliferation capacity began to slow down around day 25 (Figure 3.1), after which the cells senesced, with no further observable cell division.

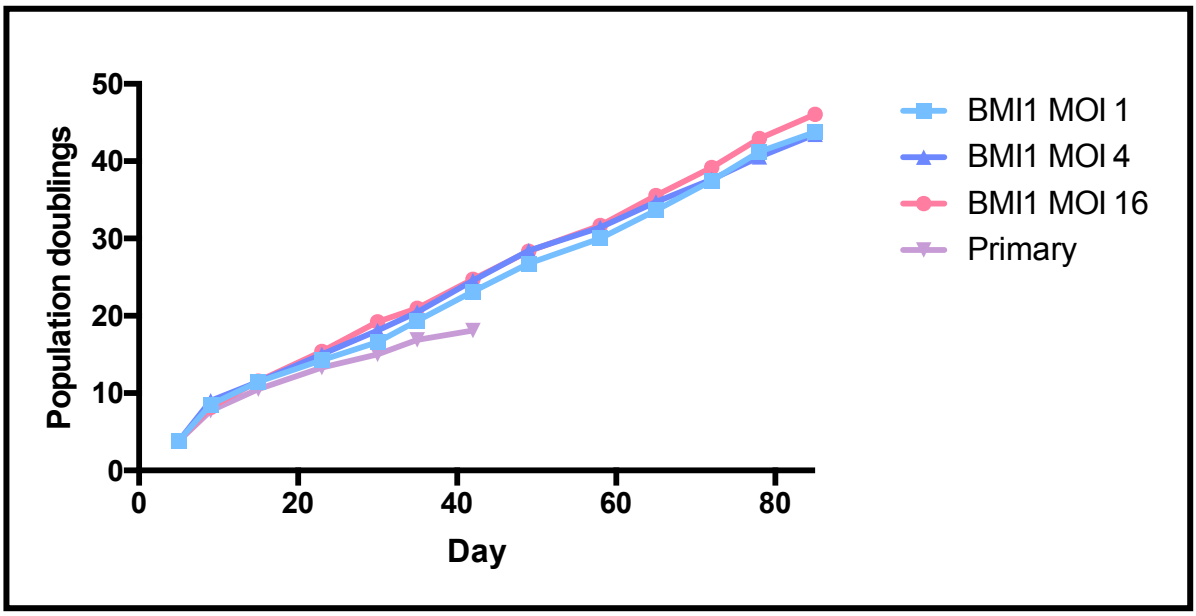


Figure 3.1. Replication kinetics to assess proliferative potential of NHBE BMI-1 cells. Population doubling capacity of primary NHBE cells was compared with NHBE BMI-1 cells at MOI 1, 4 and 16.

3.3.1.2 NHBE BMI-1 cell morphology

At P2,2, where P**2**,2 is the passage at which primary cells were transduced with *BMI-1* vector, and P2,**2** is the passage number of the transduced cells, each MOI maintained in submerged culture displays a characteristic cobblestone appearance (Figure 3.2). However, some cells have an elongated appearance (white arrows), whilst in others, squamous differentiation was evident (yellow arrows), suggesting these cells have not been successfully transduced by *BMI-1*. By P2,8, these differentiated, non-transduced cells have died off, whilst the culture maintains a normal cobblestone cell morphology, typical of basal cells.

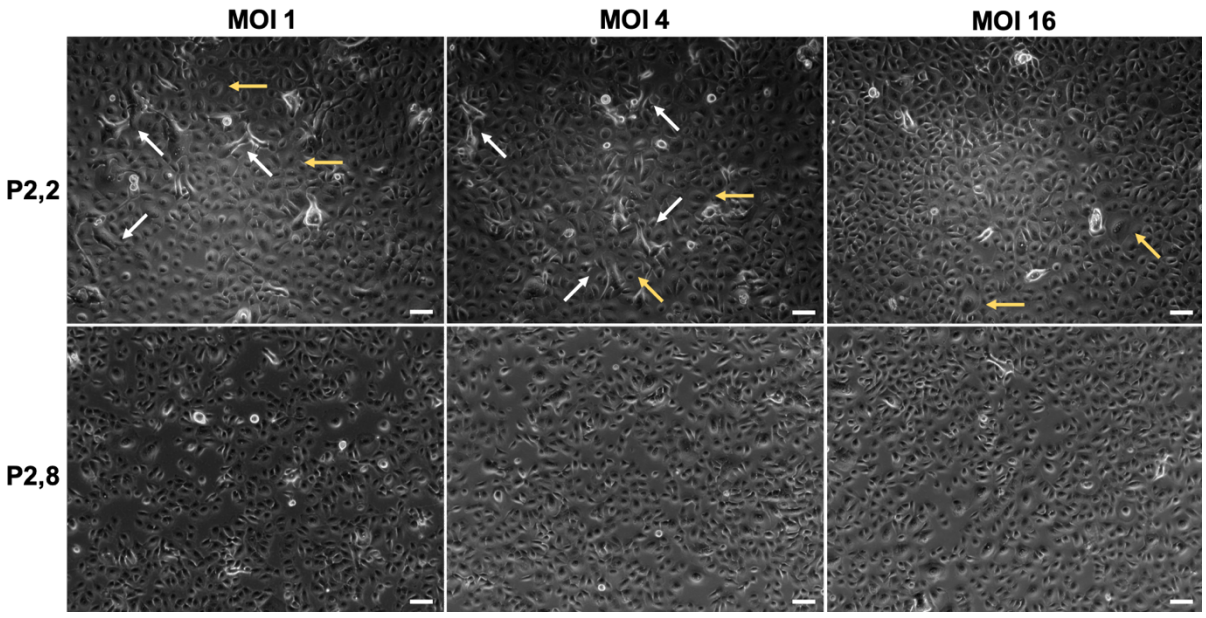


Figure 3.2. BMI-1 maintains healthy cell morphology in 2D culture. The morphology of NHBE-BMI-1 cells at MOI 1, 4 and 16 were observed under light microscopy at passage 2,2 and 2,8. White arrows highlight elongated cells and yellow arrows highlight squamous cells. Scale bars are 10 μm.

3.3.1.3 CFTR profiling in NHBE cells

CFTR expression was profiled initially in basal primary cells and at various time points during differentiation on ALI culture. As the gold standard for CF research, this provides an ideal reference point for any modified cellular model.

3.3.1.3.1 CFTR mRNA expression changes over time

We first profiled CFTR mRNA expression in primary NHBE cells. Cells were collected and total RNA extracted from submerged culture and ALI cultures at four time points (week 1, 2 and 3 and 4) for analysis by qRT-PCR. All cells in the starter cultures were passage 2. We found that CFTR mRNA expression increased over weeks 1 to 3, with a maximum x3,500 fold enrichment at week 3 in ALI culture as compared to submerged culture (*P≤0.0001).* There was then a significant decrease in CFTR expression between week 3 and week 4 in ALI culture (*P≤0.0001)* (Figure 3.3). These results have important implications for other studies using this model, such as determining the optimal time (when CFTR expression is highest) for functional analysis

of CFTR ion channel activity in Ussing chamber, or when to isolate protein for Western Blot analysis of CFTR.

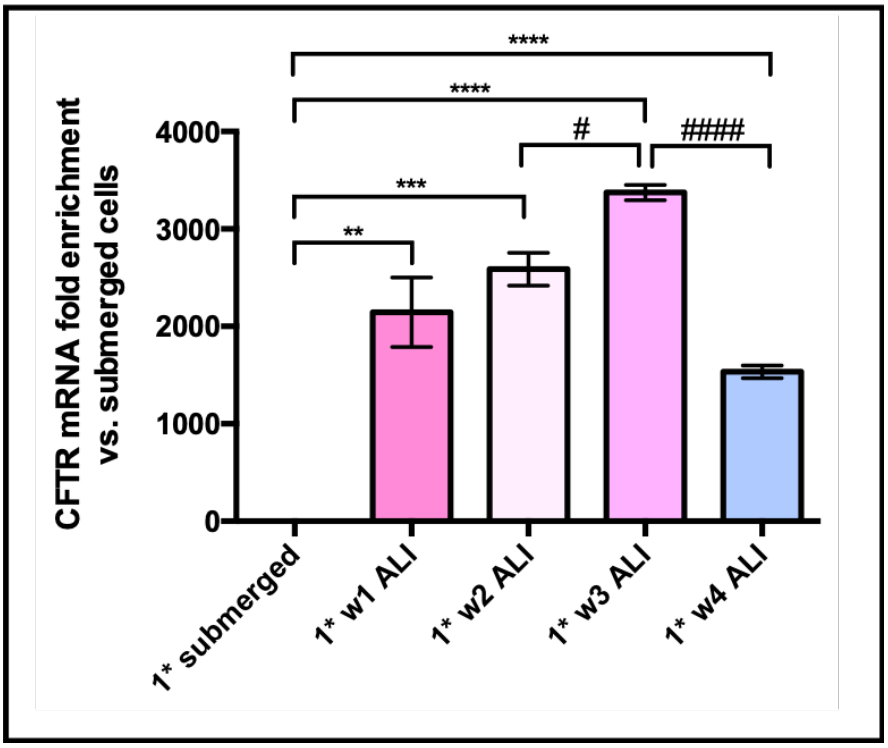


*Figure 3.3. qRT-PCR data of CFTR mRNA for primary NHBE cells grown on ALI. mRNA was isolated at weeks 1, 2, 3 and 4. All cells were passage 2. Fold difference in CFTR expression was normalised to primary NHBE cells in submerged culture (=1). All data was normalised on β actin transcripts. Results are presented as mean ± SD, n=3. Significance was calculated using One-way ANOVA with multiple comparisons- **, ## P≤0.01, ***P≤0.001 ****, #### P≤0.0001.*

3.3.1.3.2 CFTR protein expression in primary NHBEs

Having verified CFTR mRNA expression, we wanted to know whether the mRNA produced a functional protein that was successfully transported to the cell surface membrane primary NHBE cells. This was investigated by Western Blot, where a 'B-band' (140 kDa) is indicative of the presence of immature CFTR that has not been fully glycosylated, and a 'C-band' (160 kDa) indicates a mature, fully glycosylated form of the protein (Figure 3.4). Protein lysates were collected at different time points during the differentiation of primary NHBE cells (P3) on ALI culture (week 1, 2 and 3). CFTR was detected using the anti-CFTR UNC-596 Primary Antibody, while Na/K-ATPase acted as a loading control. There appeared to be expression of Band C in both week 2 and 3, with a small amount of B band expressed during week 2. There appeared

to be no expression of either band in week 1, although the loading control was also more faint (Figure.3.4). It was not possible to perform densitometry analysis on this Western Blot due to the poor resolution of the image. This data does not correlate exactly with qRT-PCR data, where CFTR mRNA expression peaks at week 3, however a similar trend is shown, with an increase in expression followed by a decrease (Figure 3.3).

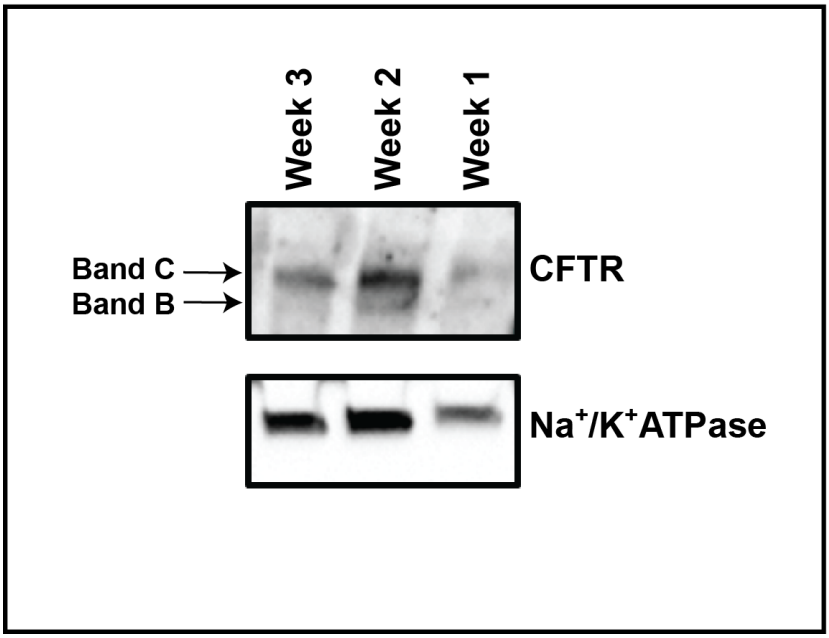


Figure 3.4 Western Blot analysis of CFTR in protein lysate taken at different time points during the differentiation of primary NHBE cells (P3) grown on ALI. *Cells were seeded on ALI, and lysates were collected at weeks 1, 2 and 3. Na+/K+-ATPase α-subunit was used as a housekeeping gene for protein normalisation. Band C = 160 kDa, Band B= 140 kDa, Na+/K+-ATPase = 100 kDa.*

3.3.1.4 CFTR profiling in NHBE BMI-1 cells

We next conducted the following experiments in *BMI-1* cell model. Cells were expanded to allow non-transduced cells to die off, before maintenance on ALI culture (P5) to profile CFTR.

3.3.1.4.1 CFTR mRNA expression in NHBE BMI-1 cells

After 4 weeks on ALI culture, P5 NHBE *BMI-1* cells (MOI 1, 4 and 16) were collected and RNA extracted for analysis by qRT-PCR. *BMI-1* MOI-1 cells expressed significantly higher levels of CFTR than MOI-4 or MOI-16 (*P≤0.05).* This work was performed before results from primary

NHBEs were obtained, which indicated that CFTR mRNA expression peaked at week 3; RNA from *BMI-1* cells was instead collected at week 4. However, we can conclude that there was an increase in CFTR mRNA expression of around x1000 fold in NHBE *BMI-1* MOI 1 cells at week 4 ALI compared to those in submerged culture (Figure 3.5). Importantly, this increase is significantly lower than that of primary NHBE cells (*P≤0.01),* highlighting a possible limitation of the *BMI-1* cell model.

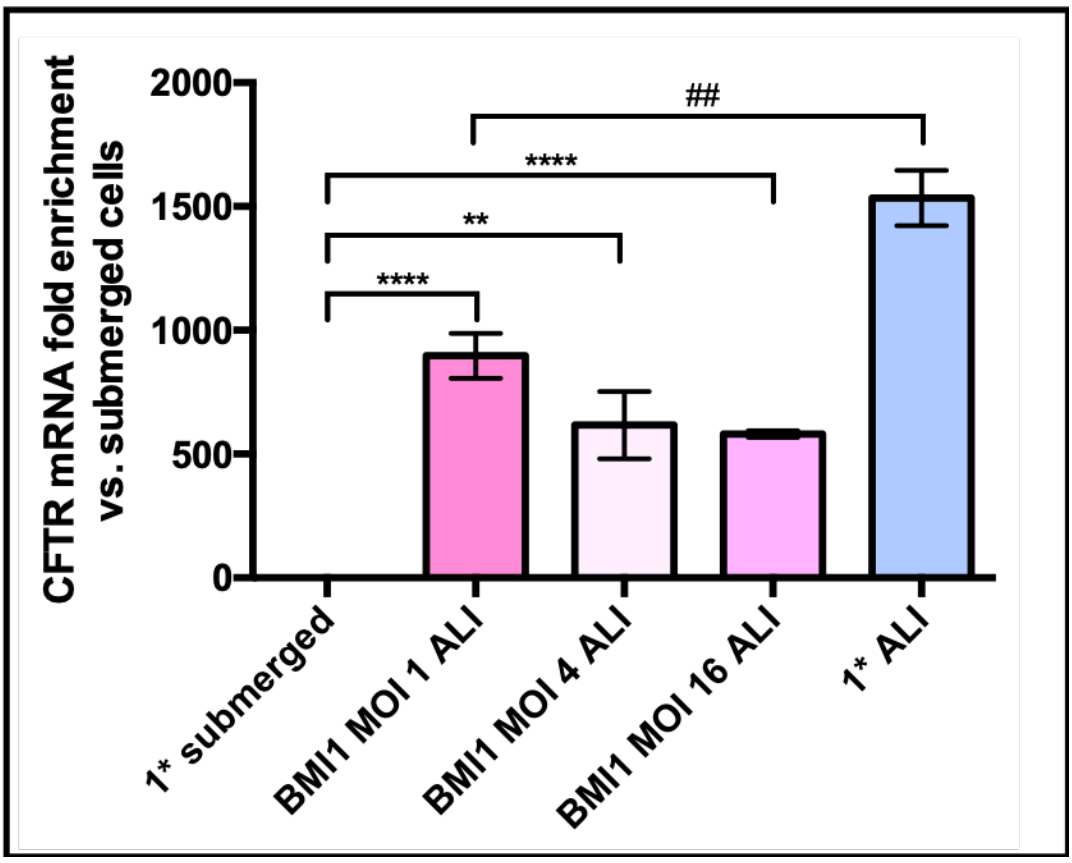


Figure 3.5. qRT-PCR data of CFTR mRNA for primary NHBE cells and NHBE BMI-1 cells grown on ALI (MOI 1, 4 and 16). *Primary cells were passage 2, BMI-1 cells were passage 5. mRNA was isolated at week 4. Fold difference in CFTR expression was normalised to primary NHBE cells in submerged culture (=1). All data was normalised on β actin transcripts. Results are presented as mean ± SD, n=3. Significance was calculated using One-way ANOVA with multiple comparisons **, ## P≤0.01,**** P≤0.0001.*

3.3.1.4.2 CFTR protein expression in NHBE BMI-1 cells

After 4 weeks on ALI culture, protein was extracted from NHBE *BMI-1* cells (MOI 1, 4 and 16). As before, CFTR was detected using the anti-CFTR UNC-596 Primary Antibody, while Na/K-

ATPase acted as a loading control. In contrast to Figure 3.4 , MOI-1 and MOI-16 have a more prominent B-band (Figure 3.6), suggesting the presence of more immature protein compared with the primary samples. MOI-4 doesn't appear to express CFTR, however we know protein is present from the loading control. From the qRT-PCR data shown above, we know that CFTR mRNA is present in MOI-4, so there was likely an issue with the culture.

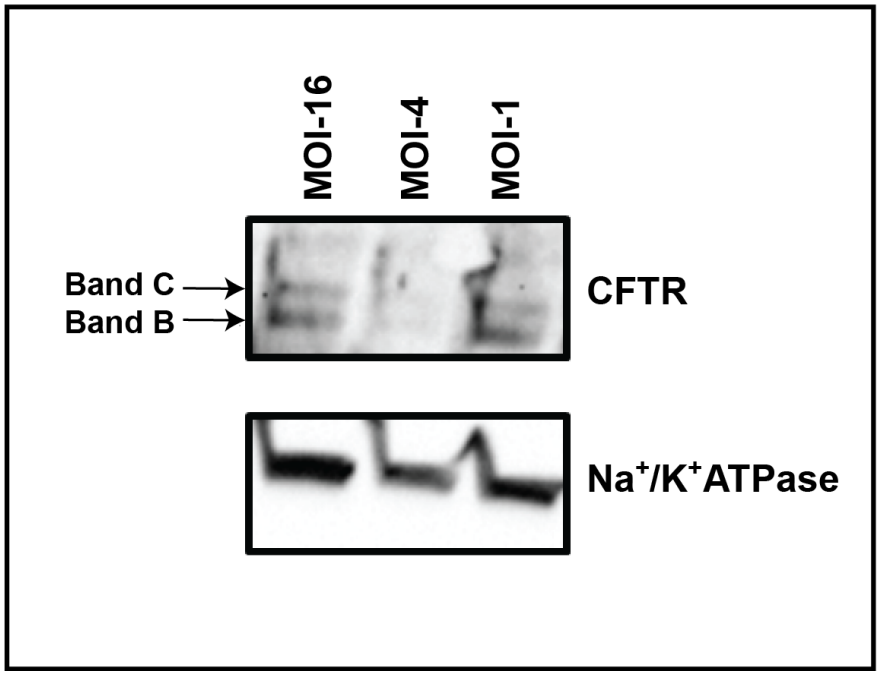


Figure 3.6. Western Blot analysis of CFTR in protein lysates of NHBE BMI-1 MOI 1, 4 and 16. *Cells were seeded on ALI, and lysates collected at week 4.Na+/K+-ATPase α-subunit was used as a loading control for protein normalisation. Band C = 160 kDa, Band B= 140 kDa, Na+/K+-ATPase 100 kDa.*

3.3.2 Receptor-targeted nanocomplexes for nucleic acid delivery in NHBE BMI-1 cells

The development and characterisation of the *BMI-1* cells provides a model in which to assess CRISPR/Cas9 delivery. Non-viral delivery systems are ideally suited for the delivery of CRISPR/Cas9 since transient expression minimises the risk of off-target effects. Our group has developed a novel receptor-targeted nanocomplex (RTN) comprised of cationic, receptor-targeting peptides and cationic lipids, which self-assemble upon mixing with nucleic acids at precise ratios. The RTNs are non-immunogenic, allowing for repeated delivery [379, 384], and

formulations have been successfully nebulised to the conducting airways of the murine and porcine lung, with efficient levels of gene expression [385, 386].

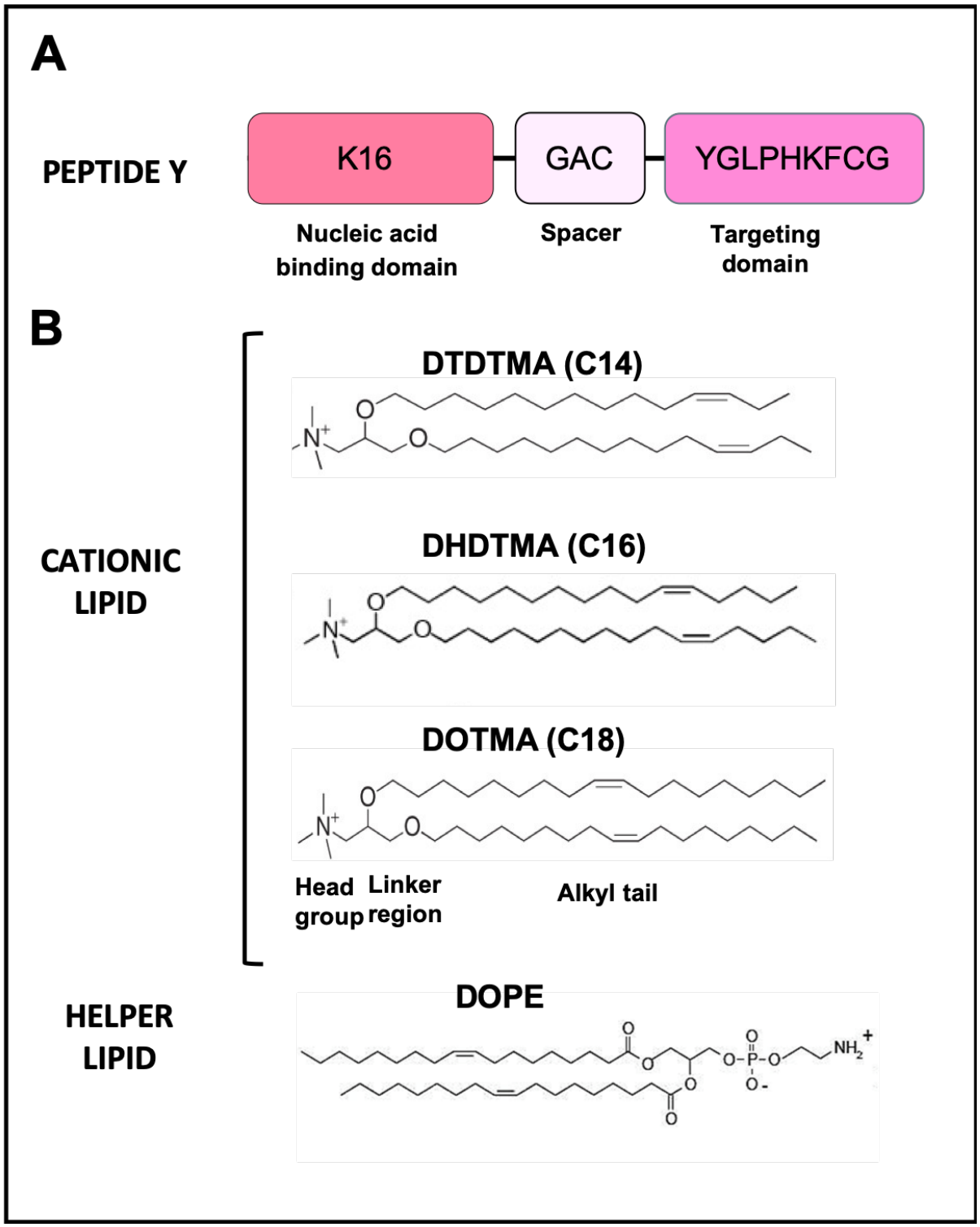


Figure 3.7 Components of receptor-targeted nanocomplexes (RTN). *The RTN which consists of nucleic acid cargo, for example DNA or mRNA, encapsulated with lipid and peptide. (A) The targeting peptide is comprised of three regions: nucleic acid binding domain, with a cationic*

K16 motif. At the opposite end, there is a receptor-targeting ligand for binding to the cell membrane. These regions are joined by a linker region, which provides flexibility necessary for binding. (B) The lipid moiety of the RTNs are generally comprised of a positively charged cationic lipid, such as DOTMA, and a neutral helper lipid such as DOPE.

3.3.2.1 DNA transfection

As mentioned, liposome based vectors have long been used for the delivery of therapeutic agents such as drugs, as well as nucleic acids. The phospholipids used in our RTNs comprise a positively-charged, hydrophilic head group and hydrophobic tail, where the positively-charged head groups can facilitate spontaneous electrostatic binding with negatively charged phosphate groups on DNA molecules, forming entropically favourable nanoparticles [181]. A neutral, 'helper' lipid such as DOPE (1,2-dioleoyl-sn-glycero-3-phosphoerhanolamine) is normally formulated with the cationic lipid because of its membrane-stabilising properties, which are believed to aid endosomal escape [184] (Figure 3.7). Previous studies have shown that cationic lipids with shorter alkyl tail lengths are better suited to fuse with the cell membrane and endosomal membranes than their longer chain counterparts, achieving higher transfection efficiency [387].

3.3.2.1.1 Determining optimal alkyl chain length for plasmid delivery

To investigate how varying alkyl chain length effects transfection efficiency, we chose 3 different cationic lipids, DTDTMA (1,2-Di-((Z)-tetradec-11-enyloxy)-N,N,N trimethylammonium propane chloride), DHDTMA (1,2-Di-((Z)-hexadec-11-enyloxy)-N,N,N trimethylammonium propane iodide), and DOTMA (1,2-Di-((Z)-octadec-9-enyloxy)-N,N,N trimethylammonium propane chloride), denoted C14, C16 and C18 respectively, formulated with DOPE at a 1:1 molar ratio. The lipids were mixed with Peptide Y and pEGFP-N1, a plasmid encoding the green fluorescent protein (GFP) reporter gene, at a weight ratio of 1:1:4 DNA: Lipid: Peptide. Nanocomplexes were used to transfect NHBE *BMI-1* cells, and compared to a commercially available reagent, Lipofectamine 3000™ (L3K), at a DNA: L3K weight ratio of 1:3. GFP expression analysed by Flow Cytometry 48 h after transfection. We found that all of our formulations performed significantly better than L3K, with a 4-fold increase in transfection efficiency compared to C16 ($P \leq 0.0001$). C16 DOPE performed better than either C14 DOPE

or C18 DOPE *(P≤ 0.05, P≤ 0.001 respectively),* with a transfection efficiency of 37% (Figure 3.8).

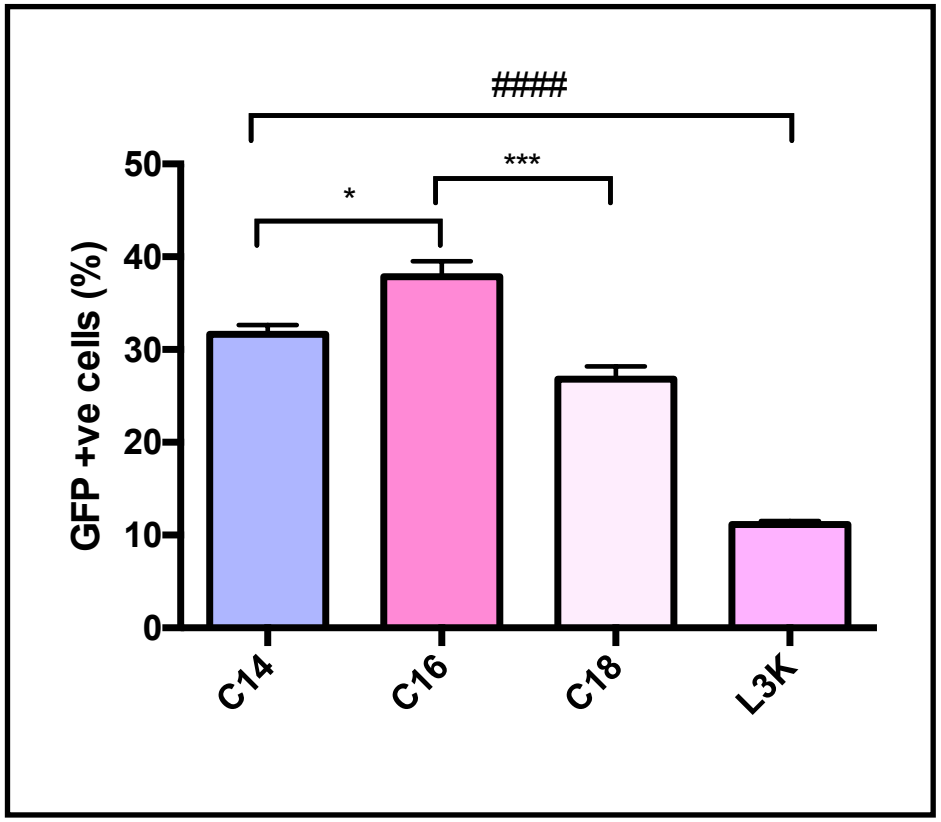


Figure 3.8: Delivery of GFP plasmid to NHBE BMI-1 cells *using C14, C14 or C18 DOPE liposomes formulated with Peptide Y at a weight ratio of 1:1:4 DNA: lipid: peptide, as compared to L3K. GFP expression was measured by Flow Cytometry 48 h post-transfection. All experiments are n=3; error bars are SEM. One-way ANOVA with multiple comparisons was used to assess the significance. *P≤ 0.05 *** P≤ 0.001 #### P≤ 0.0001.*

3.3.2.2 mRNA transfection

The transfection efficiency with pDNA in the previous experiment was almost 40% delivered as a RTN formulation. Here we want to investigate the possibility of enhancing the transfection efficiency using mRNA instead of pDNA. In general, mRNA delivery has advantages over pDNA delivery [388]. For example, mRNA does not require transfer into the

nucleus and there is no risk of integration into the genome, which is possible in pDNA transfection.

3.3.2.2.1 Determining optimal alkyl chain length for mRNA delivery

To determine optimal hydrocarbon chain length for mRNA delivery, C14, C16 and C18 DOPE were formulated with Peptide Y and GFP mRNA at a weight ratio of 1:3:4 RNA: L: P, with 250 ng of mRNA, the ratio having been determined previously in our group for mRNA delivery. As before, GFP expression analysed by Flow Cytometry 48 h after transfection. C18 was found to significantly outperform C14 and C16 (*P≤ 0.0001 P≤ 0.001* respectively, 1-way ANOVA with multiple comparisons) (Figure 3.9).

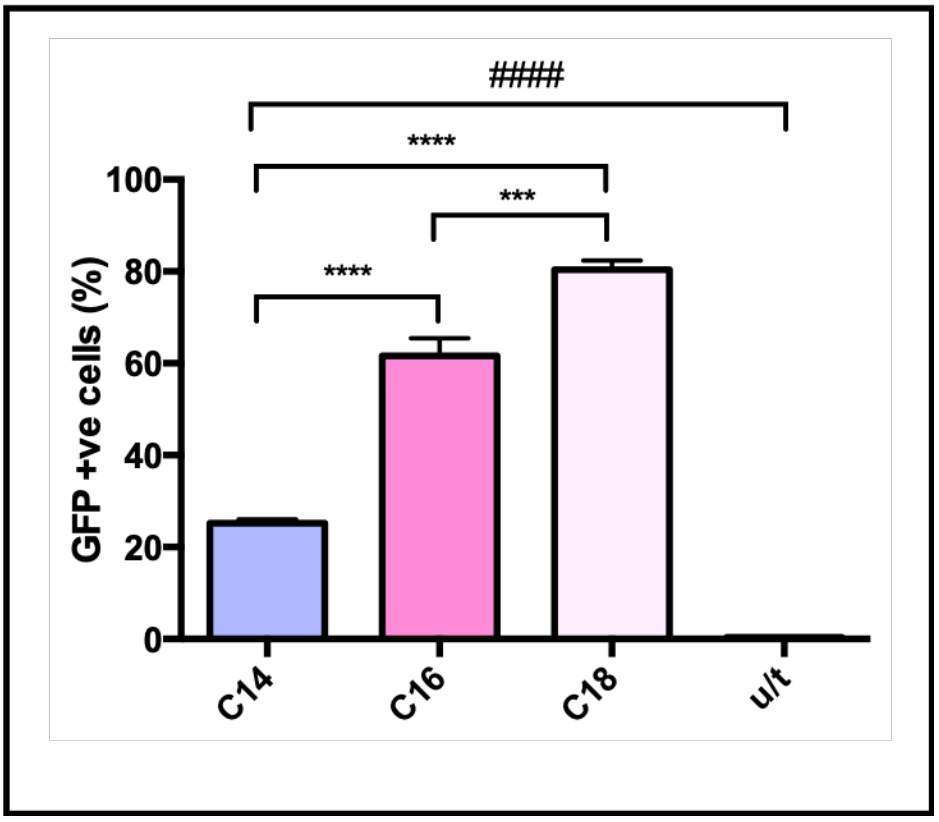


Figure 3.9: Delivery of GFP mRNA to NHBE BMI-1 cells *using C14, C14 or C18 DOPE liposomes formulated with Peptide Y at a weight ratio of 1:3:4 mRNA: lipid: peptide. GFP expression was measured by Flow Cytometry 48 h post-transfection. All experiments are n=3; error bars are*

*SEM. One-way ANOVA with multiple comparisons was used to assess the significance. *** P≤ 0.001 **** #### P≤ 0.0001.*

3.3.2.2.2 Comparison of different peptides for mRNA delivery

The targeting peptide used in our nanocomplexes consists of three regions: the cationic K16 motif, the linker region and the receptor-targeting ligand. The first part, necessary for nucleic acid binding, comprises a cationic K16 motif. At the opposite end, there is a receptor-targeting ligand for binding to the cell membrane. These regions are joined by a linker region, which provides flexibility necessary for binding. Previous work on lung epithelial cells has led to the development of two novel targeting peptides through a phage display technique. Peptide E, with the targeting ligand SERSMNF, shows the ability to bind to intercellular adhesion molecule-1 (ICAM-1) receptor on the airway epithelium, which interestingly is targeted by rhinovirus through a similar protein (SDRSMN) [389]. Peptide Y, with the targeting ligand YGLPHKF, shows sequence similarity with the intracellular pathogen *Legionella pneumophila* [390], however, the precise cellular receptor that peptide Y targets remains unknown. This peptide shows the ability to mediate targeting delivery of plasmid DNA and siRNA in nanocomplexes to airway cells as well as cells of neuronal origin [391, 392].

Initial experiments centred on the comparison of these peptides for mRNA delivery in NHBE *BMI-1* cells. GFP mRNA (R), liposome (L) and peptide (P) were formulated together at a weight ratio of 1:3:4, and used to transfect NHBE *BMI-1* cells. 48 h later, GFP expression was measured by Flow Cytometry. Peptide Y was found to give a three-fold increase in transfection efficiency (*P≤ 0.001*) when compared to Peptide E, as measured by Flow Cytometry (Figure 3.10A).

Different derivatives of Peptide Y (PY 31, PY 32 and PY 35), which contain the same targeting ligand, but different cleavable or hydrophobic linker regions, were then compared to see if small changes in peptide linker structure could increase transfection efficiency. There was found to be no significant difference between any of these peptide derivatives when GFP

expression was analysed by Flow Cytometry (Figure 3.10B). As a result of the above findings, Peptide Y was used for all subsequent experiments.

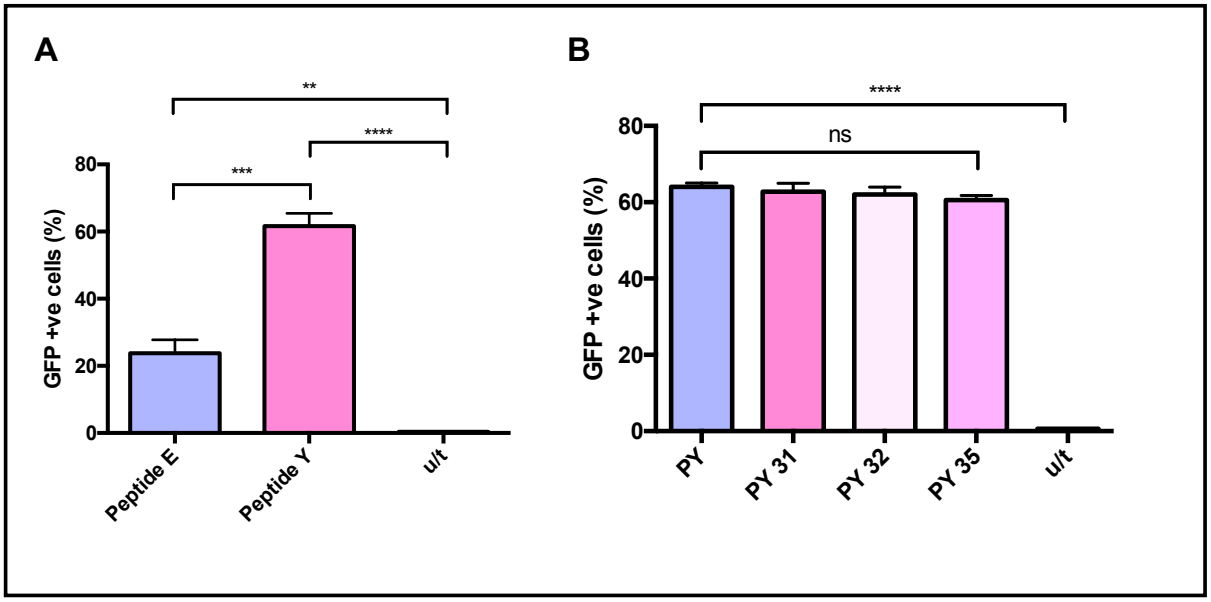


Figure 3.10: Delivery of GFP mRNA to NHBE BMI-1 *cells using C18 DOPE liposomes with either (A) Peptide Y or Peptide E, at a weight ratio of 1:3:4 mRNA: lipid: peptide. GFP expression was measured by Flow Cytometry 48 h post-transfection. Peptide Y performed significantly better than Peptide E. (B) Delivery using different derivatives of Peptide Y showed no significant difference in efficiency. All experiments are n=3; error bars are SEM. One-way ANOVA with multiple comparisons was used to assess the significance. ** P≤ 0.01 *** P≤ 0.001 ****P≤ 0.0001 ns- not significant.*

3.3.3 Development and validation of a cellular model for the optimisation of delivery of CRISPR/Cas9

Traditional methods used to determine CRISPR/Cas9 gene editing efficiency include the T7 endonuclease mismatch detection assay, a structure-selective enzyme that detects structural deformities in heteroduplexed DNA [393], and TIDE assay (Tracking of Indels by Decomposition), which requires parallel PCR reactions and capillary sequencing, followed by an R code software [394]. Both of these can be expensive and labour-intensive. We wished to develop an inexpensive, high-throughput bioassay that would allow us to optimise the delivery of the CRISPR/Cas9 technology. To achieve this, we integrated the GFP gene into the

genome of NHBE *BMI-1* cells, resulting in stable GFP expression. This was achieved by means of lentiviral transduction. Guides can then be designed to target GFP and, when transfected with Cas9, will result in a reduction in GFP fluorescence. Using the GFP model, cells can be conveniently analysed by Flow Cytometry, with GFP knockdown an indicator of editing efficiency. This editing efficiency is correlated to transfection efficiency, and so is an ideal system in which to optimise delivery.

3.3.3.1 Generation of NHBE *BMI-1* cells expressing GFP

NHBE-BMI1 cells were transduced with a LV-PGK-GFP at an MOI of 1. Seven days later, cells were analysed by Flow Cytometry and 21.4% were found to be expressing GFP (Figure 3.9). Cells were then expanded for a further week, before sorting for GFP expression on the MoFlo XDP Cell Sorter (Beckman Coulter, U.S). Sorted cells were found to be 92.5% GFP positive (Figure 3.11), and this was used as our polyclonal cell population in which to optimise the delivery of the CRISPR/Cas9 technology.

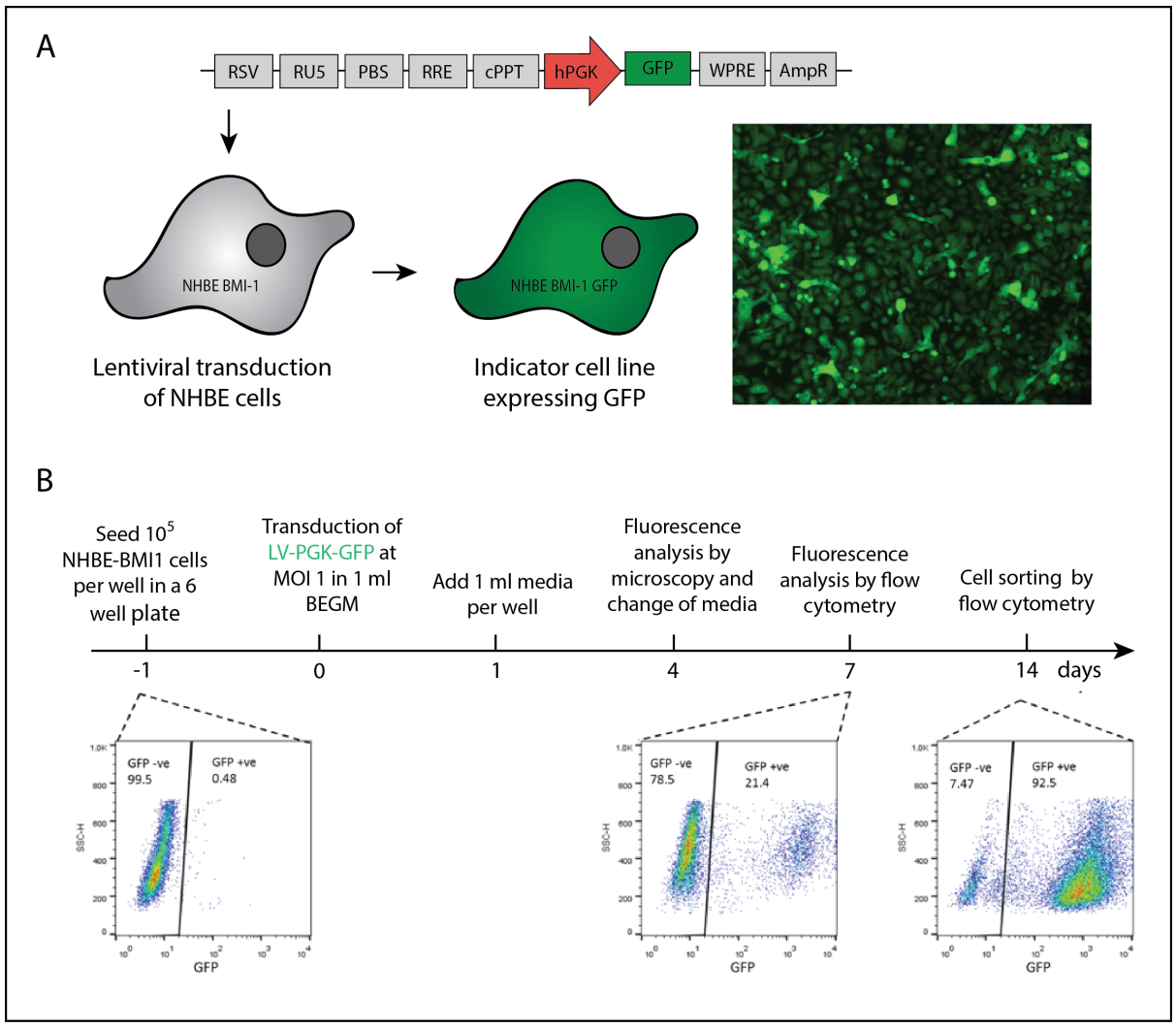


Figure 3.11. Establishment of NHBE-BMI-1 GFP reporter cell line to optimise delivery of CRISPR/Cas9. *(**A**) Schematic of LV-PGK-GFP vector and indicator line generation. (**B**) Timeline of the development of a GFP knockdown cellular model. Density plots of GFP expression as measured by Flow Cytometry are shown over various stages of development.*

3.3.3.2 Validation of NHBE-BMI1 GFP cellular model

The next step was to validate the NHBE *BMI1* GFP model by means of transfections with Cas9 mRNA and GFP targeting gRNAs. Three previously validated guide RNAs (gRNAs) targeting GFP were chosen to create DSBs in the GFP gene [395] (Table 3.1) (Figure 3.12). These DSBs are repaired by NHEJ, introducing indel mutations that disrupt the GFP open-reading frame and result in loss of fluorescence (Figure 3.13).

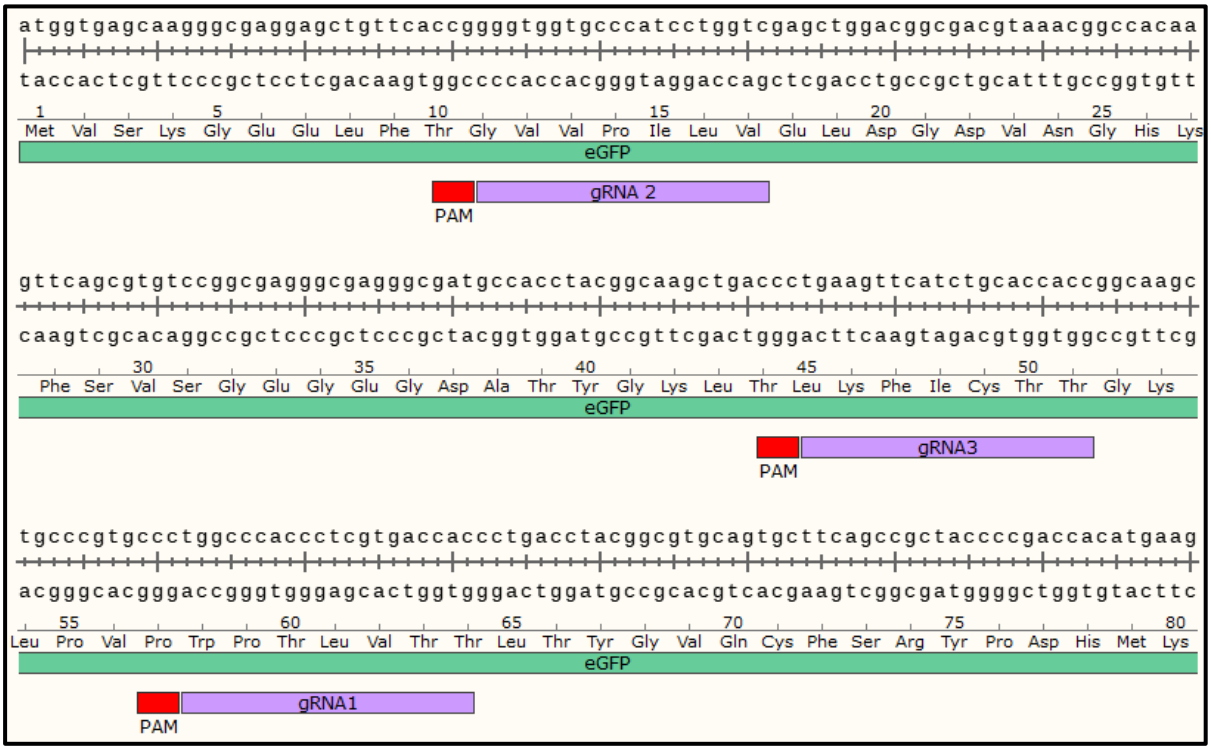


Figure 3.12. Section of GFP sequence in NHBE-BMI-1 GFP cells showing three gRNAs targeted to knock out the gene*. PAM sequences highlighted in red.*

Table 3.1. List of gRNAs targeting GFP gene. *The table shows gRNA sequence, strand targeted, the PAM motif sequence, the predicted efficiency score (on-target score) and specificity score (off-target score). The on-target and off-target scores were calculated by using the Benchling programme (https://benchling.com/crispr). Both scores are given /100, with higher scores suggestive of higher efficiency and specificity.*

gRNA	Sense	Sequence	PAM	Efficiency score	Specificity score
Guide 1	-	**GAATAGCTCAGAGGCCGAGG**	**GGG**	**64**	**54.1**
Guide 2	-	**GACCAGGATGGGCACCACCC**	**CGG**	**60.7**	**67.2**
Guide 3	-	**GGTGGTGCAGATGAACTTCA**	**GGG**	**57.8**	**66.8**

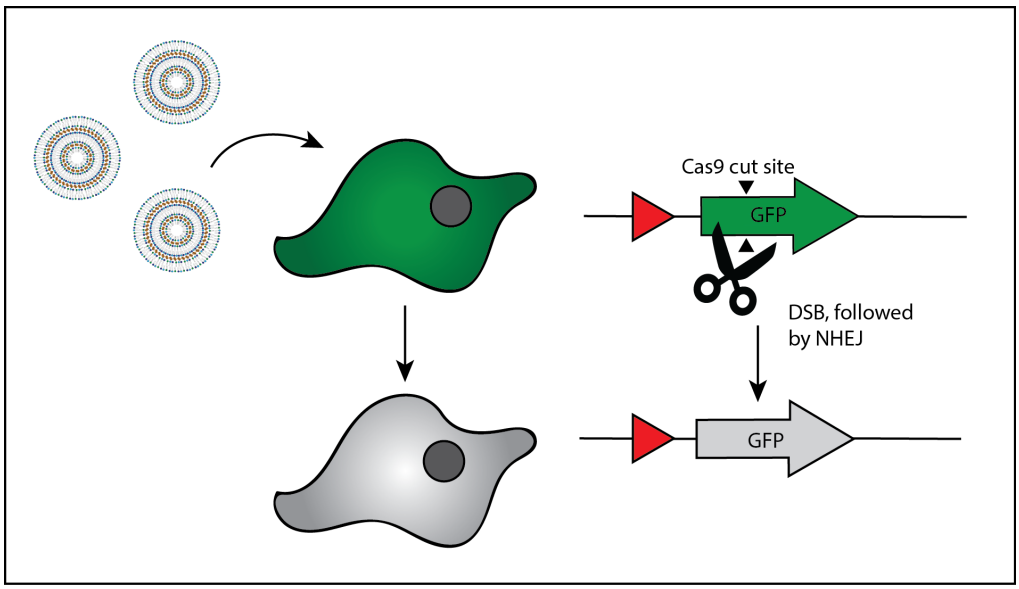


Figure 3.13. Principle of GFP knockout in NHBE BMI-1 reporter cell line. *Delivery of Cas9 and gRNA targeting GFP introduces DSBs into the GFP gene, which are repaired by NHEJ. This can result in indel mutations, disrupting the open-reading frame and causing loss of GFP protein expression in edited cells.*

Cas9 mRNA and GFP-targeting gRNAs (R) were formulated together with C18 DOPE and peptide Y, at a weight ratio of 1:3:4 R:L:P, and used to transfect NHBE *BMI-1* GFP cells. Eight days later, cells were assessed by Flow Cytometry for GFP knockout. All three gRNAs performed well, achieving levels between 50 and 60% knockout (Figure 3.14), as compared to Cas9 in the absence of gRNA but gRNA3 appeared to achieve greatest levels of knockout ($P \leq 0.05$) and so was used in all subsequent experiments.

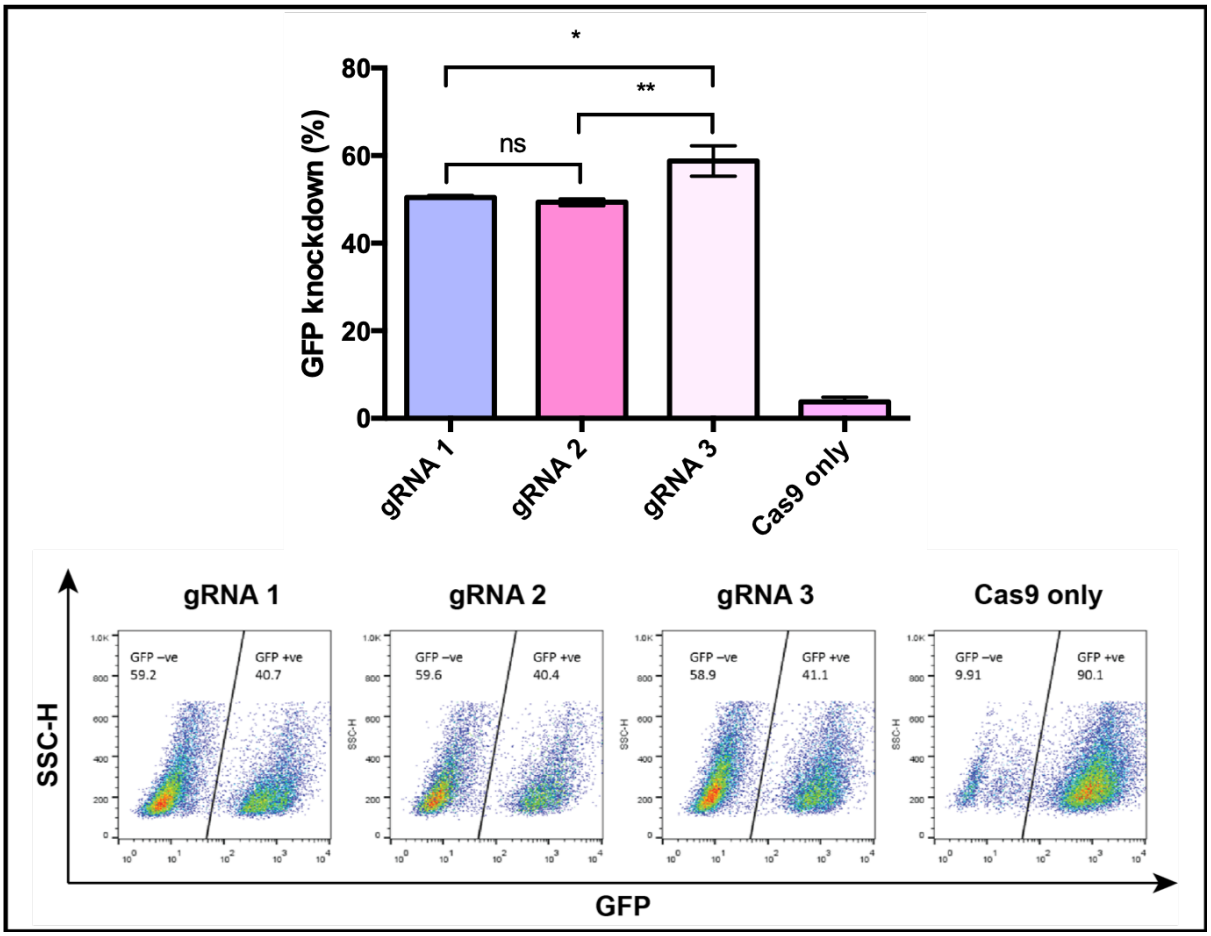


Figure 3.14. (A) Validation of GFP knockdown cell model for CRISPR/Cas9 delivery. *Three GFP-targeting guides were transfected with Cas9 mRNA into NHBE BMI-1 GFP cells using C18 DOPE & peptide Y, at a weight ratio of 1:3:4 (R:L:P). GFP fluorescence was analysed by Flow Cytometry 8 days later. Results are presented as mean ± SD, n=3. Significance was calculated using one-way ANOVA with multiple comparisons * P≤0.05, ** P≤0.01, ns- not significant. (**B**) Representative flow cytometry density plots of GFP knockdown with the 3 guides.*

3.3.3.3 Determining optimal hydrocarbon chain length for Cas9 mRNA delivery

To investigate the length of cationic lipid alkyl chain for delivery of Cas9 mRNA and gRNA, we utilised the GFP knockdown model described previously. Weight ratio between gRNA and Cas9 mRNA was 1:1.5. C14, C16 and C18 lipids were formulated with DOPE at a 1:1 molar ratio using the NanoAssemblr. These liposomes were then used to form nanocomplexes in combination with Peptide Y , and the Cas9 mRNA and gRNA targeting GFP, at a weight ratio of 1:3:4 R:L:P. Eight days later, GFP fluorescence was analysed by flow cytometry to assess knockdown of GFP expression. Similar to results achieved when delivering GFP mRNA, the

C18/DOPE cationic lipid performed significantly better than either C14/DOPE or C16/DOPE (*P≤0.001)* (Figure 3.15). In subsequent experiments, Cas9 mRNA and gRNA were delivered using C18 DOPE.

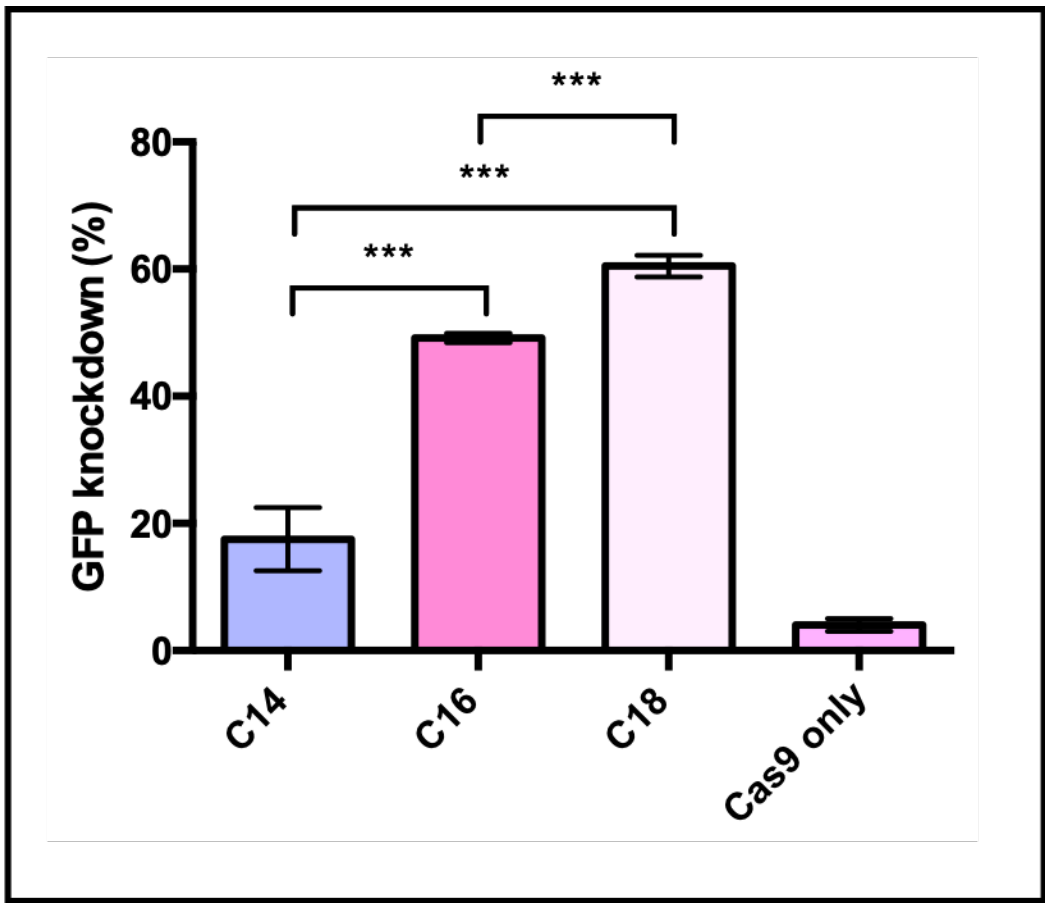


Figure 3.15. GFP knockdown (%) in NHBE BMI-1 GFP cells following transfection with Cas9 mRNA + gRNA using either C14, C16 or C18 liposomes with Peptide Y. *Ratio 1:3:4 R:L:P. GFP fluorescence was measured by Flow Cytometry 8 days after transfection. Results are presented as mean ± SD, n=3. Significance was calculated using one-way ANOVA with multiple comparisons *** P≤0.001.*

3.3.3.4 *Amount of Cas9 mRNA and gRNA affects editing efficiency*

Having optimised a nanoparticle for mRNA delivery (Peptide Y with C18 DOPE at a weight ratio of 1:3:4 R:L:P), the nanoparticle system could be further optimised by determining the quantities of Cas9 mRNA and gRNA that yielded the highest GFP knockdown. A series of transfections were therefore carried out, the first maintaining a constant concentration of gRNA (200 ng), with concentrations of Cas9 mRNA ranging from 50 ng to 300 ng while in the

second series, the Cas9 concentration remained constant (300 ng), with gRNA concentration ranging from 100 to 300 ng. We also wished to assess whether sequential delivery of Cas9 mRNA and gRNA 24 h apart would increase GFP knockdown, the rationale being that the Cas9 mRNA would have ample time to be translated into protein, allowing DSB cleavage of the genome as directed by the gRNA. NHBE *BMI-1* cells were therefore transfected with the above conditions and eight days later, GFP fluorescence was analysed by Flow Cytometry to assess knockdown.

In figure 3.16A, it was found that increasing the amount of Cas9 increased editing efficiency; 300 ng Cas9 + 200 ng gRNA gave the highest efficiency, with a 2.5-fold increase when compared to 50 ng Cas9 mRNA + 200 ng gRNA (p*≤0.0001*). (Figure 3.16A). All results were statistically different from Cas9-only controls (p*≤0.0001*). In Figure 3.16B, again the highest concentration- 300 ng Cas9 mRNA + 300 ng gRNA was found to give the greatest GFP knockdown compared to 200 ng and 100 ng of gRNA (p*≤0.01 and* p*≤0.0001 respectively).* There was a 5-fold decrease in GFP knockdown when using sequential delivery with Cas9 mRNA delivered first followed by gRNA 24 h later (p*≤0.0001*), validating our original protocol of Cas9 and gRNA co-transfection. All results were statistically different from Cas9-only controls (p*≤0.0001*). (Figure 3.16B).

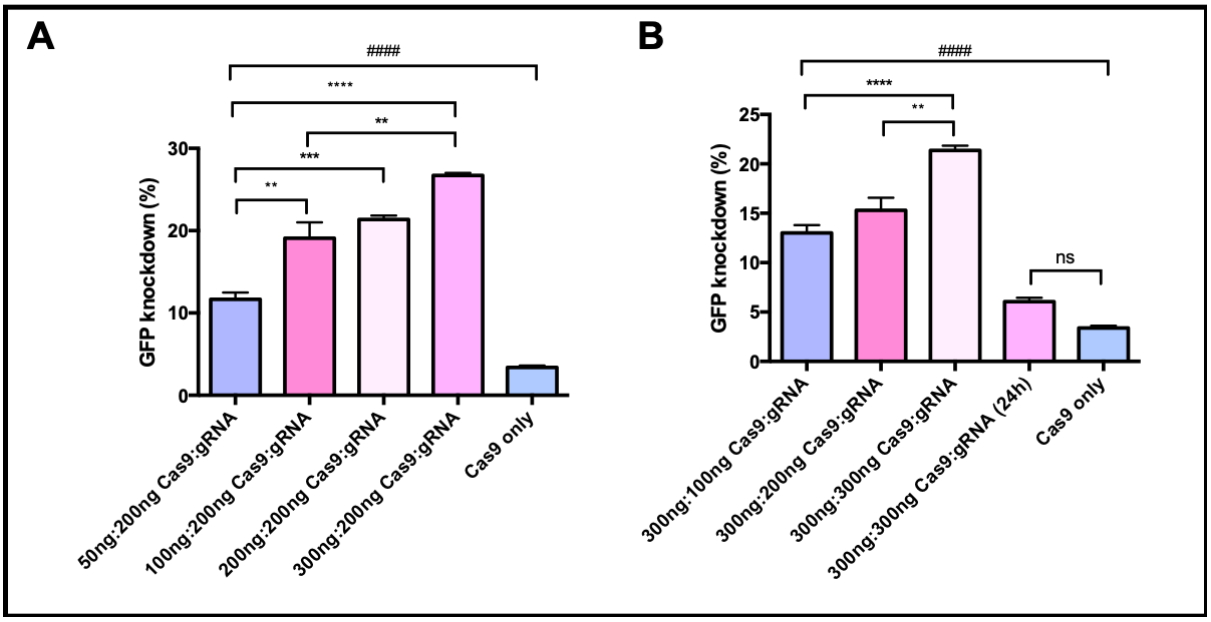


Figure 3.16. Assessment of Cas9 mRNA and gRNA concentration on GFP knockdown in NHBE BMI-1 GFP cells. *RTNs were C18 DOPE Peptide Y. (**A**) Amount of gRNA was 200 ng, whilst Cas9 increased from 50 to 300 ng. (**B**) Amount of Cas9 was 300 ng, whilst gRNA increased from 100*

*to 300 ng. (24h- 300 ng Cas9 was transfected followed by 300 ng gRNA 24 h later). GFP fluorescence was measured by Flow Cytometry 8 days after transfection. Results are presented as mean ± SD, n=3. One-way ANOVA with multiple comparisons was used to assess the significance** P≤0.01 *** P≤0.001 **** #### P≤0.0001*

3.3.4 Delivery of Cas9 protein and gRNA as a pre-assembled RNP complex

The delivery of Cas9 protein and guide RNA as a pre-assembled RNP complex offers several potential advantages over Cas9 mRNA or pDNA based systems, namely greater safety due to its transient nature, and therefore fewer off-target effects. Further, no transcription or translation is required, therefore increasing efficiency. However, delivery of the RNP remains a key challenge. We utilised our GFP knockdown model to assess the delivery of pre-assembled Cas9 + guide RNPs using our nanocomplexes, and compared this to a commercially available transfection reagent for CRISPR-Cas9 protein delivery, Lipofectamine™ CRISPRMAX™. Cas9 protein + gRNA were mixed at a 4:1 weight ratio (Molar ratio 1:8) to allow self-assembly, with a total concentration of either 500 ng or 1000 ng, before addition of lipid and peptide, at a weight ratio of 1:3:4 RNP: L: P. CRISPRMAX transfections were assembled as per manufacturer's instructions. Eight days after transfections, GFP fluorescence was measured by flow cytometry to assess GFP knockdown. It was observed that using the lower concentration of RNP resulted in a higher transfection efficiency (using C16 DOPE). C16 DOPE significantly out-performed both C14 and C18 DOPE ($p \leq 0.01$ and $p \leq 0.05$), with a 2-fold and 4-fold increase in GFP knockdown respectively. Notably, the knockdown efficiency of RNP prepared with 500 ng Cas9 protein and 125 ng gRNA was significantly better using C16 DOPE and peptide Y, than the same RNPs delivered by CRISPRMAX™, a reagent specifically optimised for RNP delivery, resulting in almost a 4-fold increase in GFP knockdown ($p \leq 0.01$) (Figure 3.17).

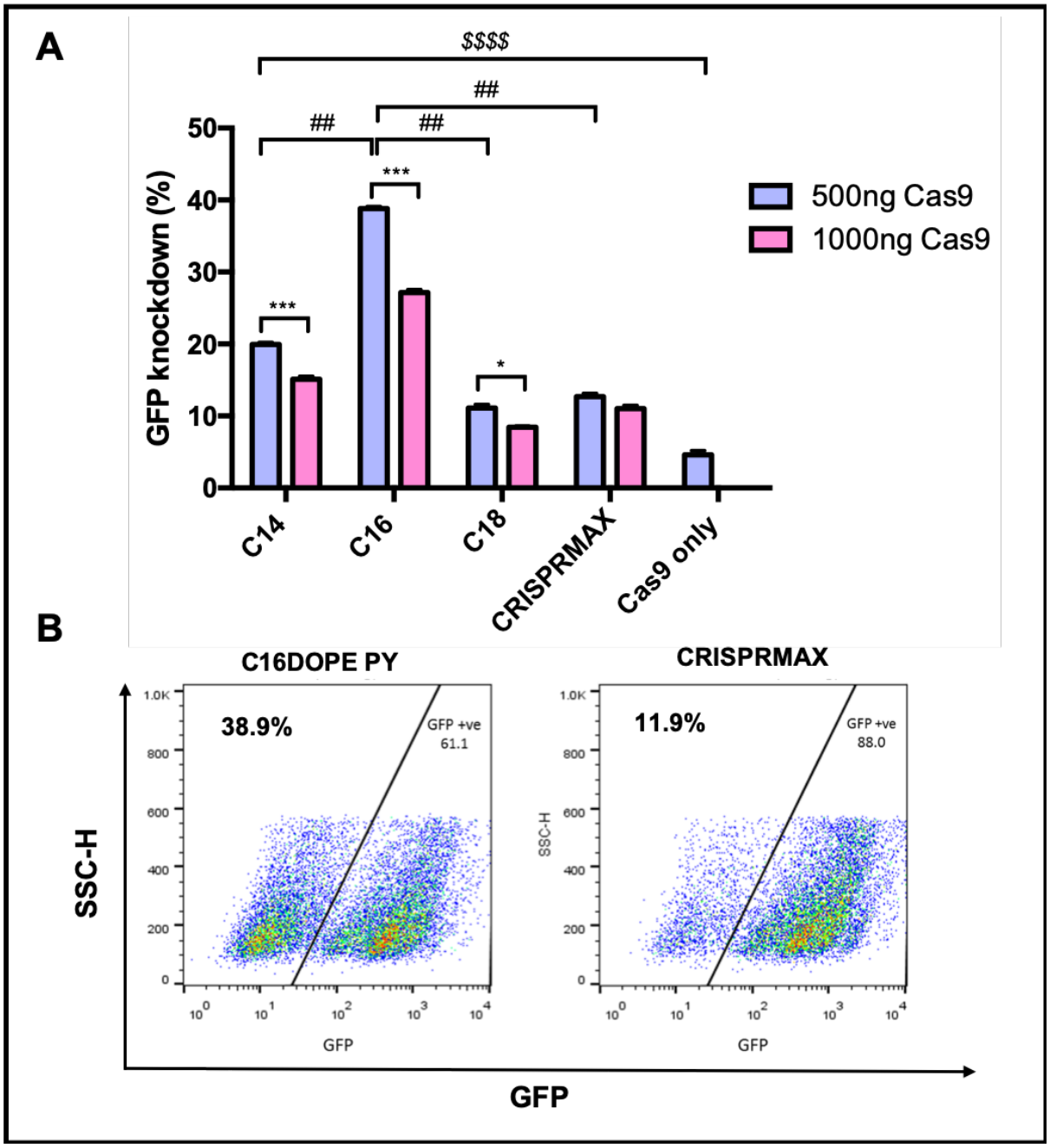


Figure 3.17 (A) GFP knockdown (%) in NHBE BMI-1 GFP cells following transfection with Cas9 RNP complex using either C14, C16 or C18 DOPE liposomes and peptide Y, at a weight ratio of 1:3:4 RNP:L:P. *Two concentrations of RNP were used: 500ng Cas9 protein with 125 ng gRNA (purple) and 1000 ng Cas9 protein with 250 ng gRNA (pink). GFP fluorescence was measured by Flow Cytometry 8 days after transfection. (**B**) Representative density plots of GFP knockdown with C16 DOPE PY and CRISPRMAX (500 ng Cas9 protein). Results are presented as mean ± SD, n=3. Student's T-test was used to measure significance between amounts of Cas9 protein * P≤0.05 *** P≤0.001. One-way ANOVA with multiple comparisons was used to assess the significance between transfection conditions ## P≤0.01 $$$$ P≤0.0001.*

3.3.4.1 Comparison of Cas9 mRNA vs. Cas9 RNP on editing efficiency

Having determined the most efficient nanoformulation for RNP delivery (C16DOPE), we wanted to compare this with the optimal delivery formulation for Cas9 mRNA (C18DOPE). We also chose to compare two different gRNAs, gRNA 1 and gRNA 3 (Section 3.3.3.2) to confirm that gRNA 3 works best in the RNP format as well as with Cas9 mRNA. Both sets of nanocomplexes were formulated at a weight ratio of 1:3:4 mRNA/RNP: L: P. Eight days post-transfection, GFP knockdown was measured by flow cytometry. Cas9 RNP significantly out-performed Cas9 mRNA, with an almost 2-fold increase in editing efficiency ($P \leq 0.001$) when formulated with gRNA 3. gRNA 3 also significantly out-performed gRNA 1 in RNP format ($P \leq 0.0001$). All results were statistically significant when compared to Cas9 only ($P \leq 0.0001$). Cas9 protein with gRNA 3 in RNP format was used in all subsequent experiments (Figure 3.18).

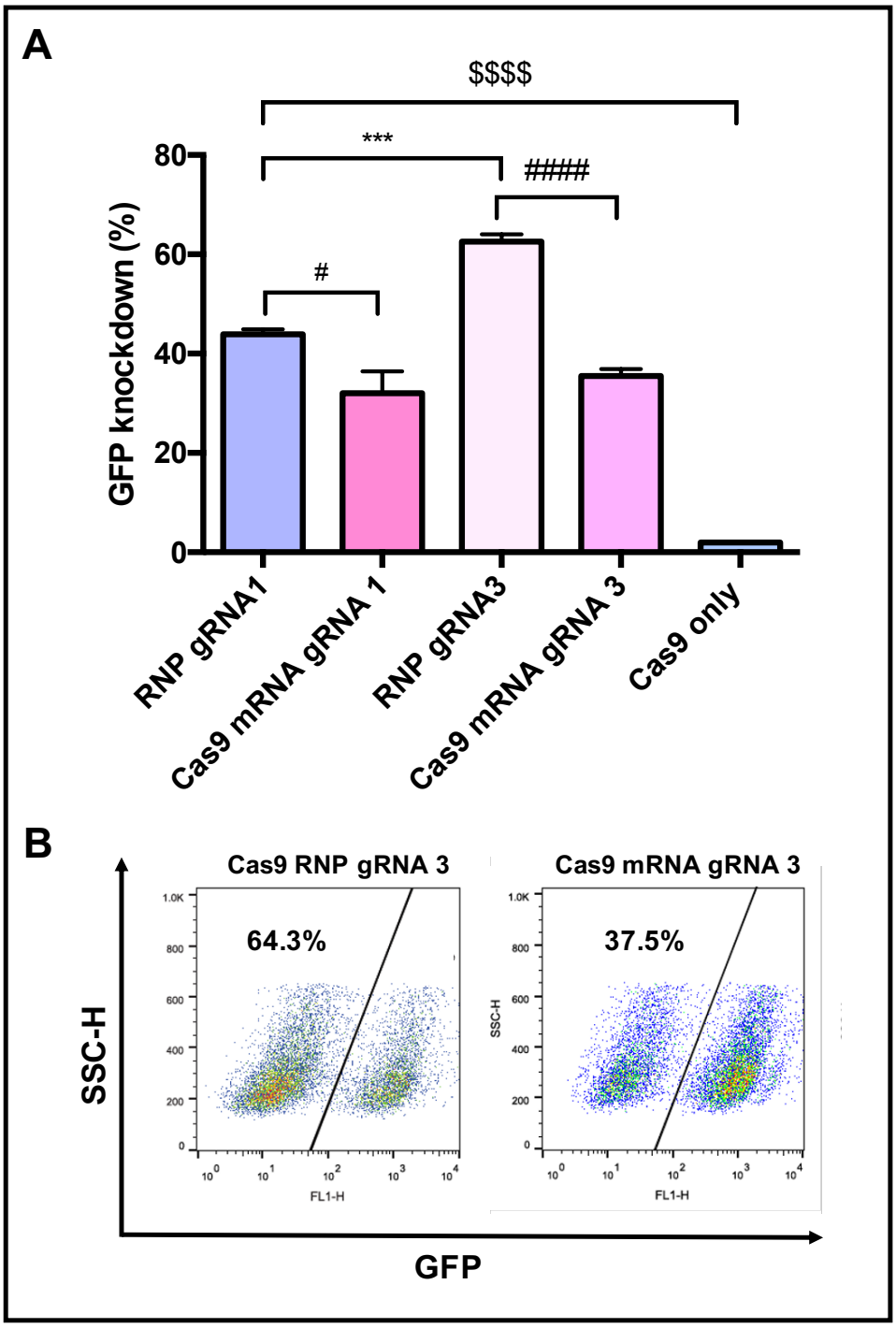


Figure 3.18. Assessment of Cas9 mRNA vs. Cas9 and gRNA on GFP knockdown in NHBE BMI-1 GFP cells. *Two different gRNAs were used, gRNA 1 and gRNA 3. RTNs were C18 DOPE PY for mRNA and C16 DOPE PY for RNP. (**A**) GFP fluorescence was measured by Flow Cytometry 8 days after transfection. (**B**) Representative density plots of GFP knockdown with Cas9 RNP and Cas9 mRNA RTNs. Results are presented as mean ± SD, n=3. One-way ANOVA with multiple comparisons was used to assess the significance between Cas9 format and gRNAs ** P≤0.01 *** P≤0.001 $$$$ #### P≤0.0001.*

3.3.4.2 Comparison of lipid ratios for RNP delivery

Having found that Cas9 and gRNA in an RNP format can elicit a greater GFP knockout in NHBE *BMI-1* cells than Cas9 mRNA and gRNA, we wanted to determine the optimal ratio of each of the lipid and peptide components in our RTNs. Initially, we used the same ratios found to be optimal for mRNA delivery, however, using lower amounts of the components in our nanocomplexes confers several advantages, namely a lower amount of material being delivered to the cell, therefore limiting any toxicity. Previous papers have shown that lower amounts of lipid in our RTNs can increase transfection efficiency of luciferase DNA [396]. We therefore investigated much lower ratios of C18 DOPE, ranging from 1:0.1:4 to 1:1:4 RNP: L: P. GFP knockdown efficiency increased as the amount of lipid increased (Figure 3.19) with the 1:1:4 ratio achieving a 72% GFP knockdown efficiency *(P≤0.0001).* All results were statistically different from each other, except for the difference between 1:05:4 and 1:0.75:4, as measured by one-way ANOVA with multiple comparisons.

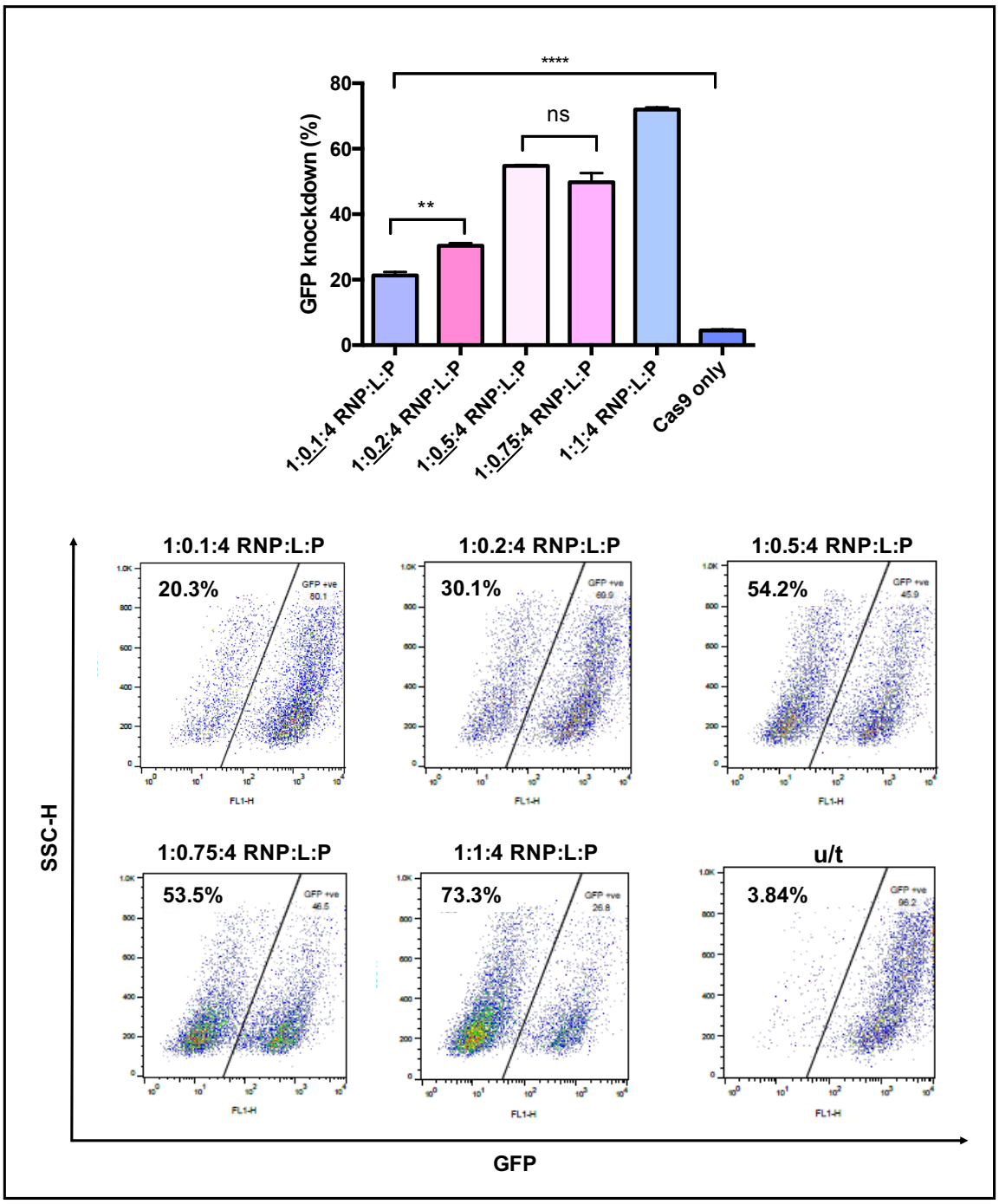


Figure 3.19. GFP knockdown (%) in NHBE BMI-1 GFP cells following transfection with Cas9 + gRNA RNP using either C18DOPE liposomes with Peptide Y, with lipid ratios ranging from 01. to 1. *Fluorescence was measured by Flow Cytometry 8 days after transfection. Results are presented as mean ± SD, n=3. One-way ANOVA with multiple comparisons was used to assess the significance. **P≤0.01 **** P≤0.0001.*

3.3.4.3 Optimising peptide ratios for RNP delivery

We also investigated how lower ratios of peptide Y would effect transfection efficiency of RNP complexes by comparing peptide ratios ranging from 1:3:1 to 1:3:4 RNP: L: P. It was observed that lower ratios of peptide achieved a higher transfection efficiency, with the ratio 1:3:1 RNP: L: P achieving over a 2-fold as compared to 1:3:4 *(P≤0.0001).* All results were statistically significant from each other and the as measured by one-way ANOVA with multiple comparisons (Figure 3.20). This efficiency was lower than that achieved by the 1:1:4 RNP: L: P formulation achieved in Figure 3.19 (72 % vs. 40 % GFP knockdown respectively), and so the 1:1:4 formulation was used for future experiments.

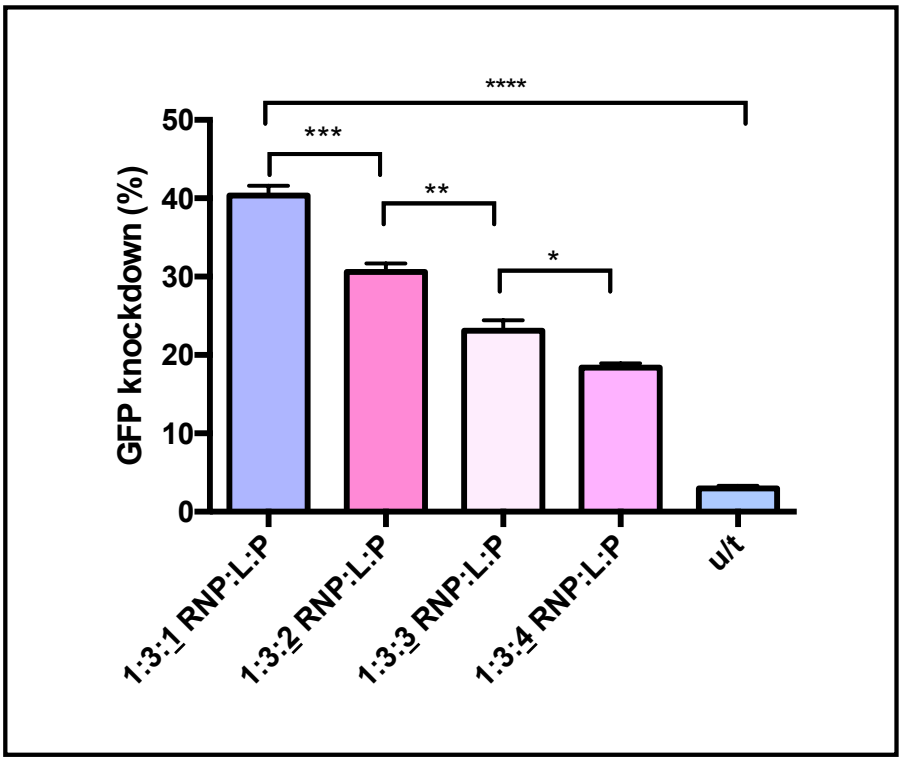


Figure 3.20. GFP knockdown (%) in NHBE BMI-1 GFP cells following transfection with Cas9 + gRNA RNP using either C18DOPE liposomes with Peptide Y, with peptide ratios ranging from 1 to 4. *Fluorescence was measured by Flow Cytometry 8 days after transfection. Results are presented as mean ± SD, n=3. One-way ANOVA with multiple comparisons was used to assess the significance. *P≤0.05** P≤0.01 *** P≤0.001 **** P≤0.0001.*

3.3.4.4 Cell viability in RNP transfected cells

In order to access cell viability after transfection of RNP nanocomplexes, the cell viability dye ToPro3 was used to stain dead cells. Live cells are impermeable to ToPro3 to but it penetrates the compromised membranes characteristic of dead cells, and undergoes a spectral shift upon association with DNA. Cell suspensions were collected each day for up to one week post-transfection, and incubated with the ToPro3 counterstain, before analysis by flow cytometry. As expected, our RTNs caused very little cell death across the week long period, with approximately 97% cell viability (Figure 3.21). In parallel, we also assessed decrease in GFP expression, and found a steady decrease over the 7 days, from 3% on day 1 to 57.6% by day 7.

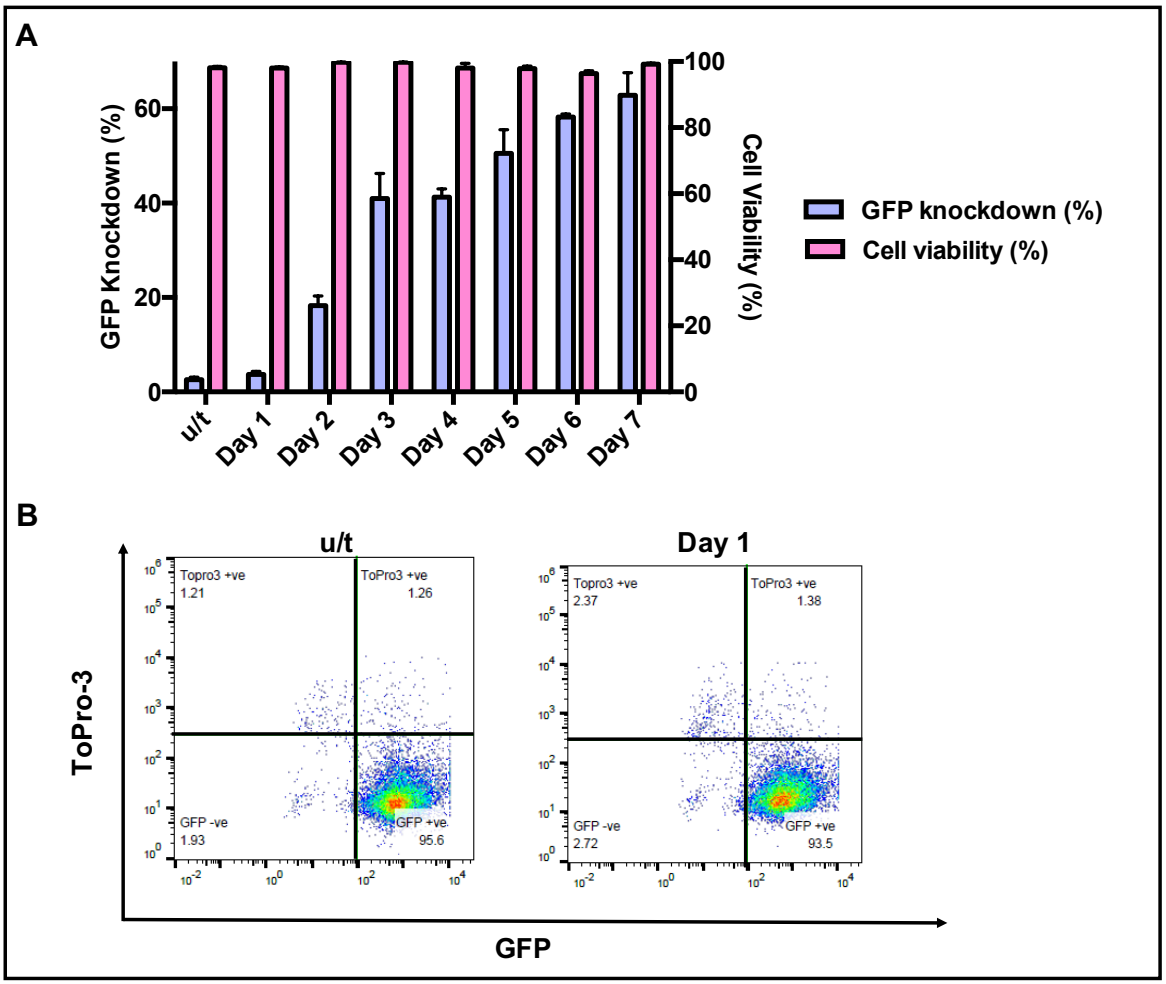


Figure 3.21. ToPro3 cytotoxicity assay for RNP nanocomplexes. *(**A**) Cell viability was measured each day for 1 week post-transfection, and GFP knockdown was measured over the same week. GFP knockdown (%) presented on left y-axis (purple), and cell viability (%)*

*presented on right y-axis. (**B**) Representative flow cytometry density plots of un-transfected cells vs. RNP treated cells on Day 1. Results are presented as mean ± SD, n=3.*

3.3.5 Biophysical characterisation of Receptor Targeted Nanocomplexes

Differences in the morphology, size and surface charge of RTNs can influence the cellular uptake of the particles. An understanding of these characteristics is crucial in order to develop more effective nanoparticle systems, and is essential should they be applied in an *in vivo* setting.

3.3.5.1 Dynamic-Light Scattering

Dynamic Light Scattering (DLS) measurement is a non-invasive technique for measuring size and distribution of submicronic particles. Complexes were prepared as previously described, however, formulated in nuclease-free water rather than OptiMEM, and analysed to determine their hydrodynamic size (nm), polydispersity index, and zeta potential (mV) using a Malvern Nano ZS Zetasizer. All particles were formed with Peptide Y, at a weight ratio of 1:3:4 RNA: L: P and 1:1:4 RNP: L: P. For the Cas9 mRNA RTNs, C14 DOPE and C16 DOPE had similar hydrodynamic sizes (134.2± 3.1 nm and 128.1± 4.5 nm respectively) and PDIs (0.33 and 0.31), however, C14 DOPE nanoparticles were less positively charged, at +20.8 ±2.8 mV, compared to 33.8± 1 mV for C16 DOPE-containing nanoparticles. C18 DOPE had a larger Z-Ave value of 177.3±1.3 nm, however a lower PDI of 0.28, suggesting the particles were more uniform, and a surface charge of +36.8± 3.4 mV.

In comparison, Cas9 RNP particles comprised of C16 and C18 DOPE were significantly smaller than their Cas9 mRNA counterparts, with Z-Ave values of 95.7± 0.7 nm and 90.0±0.5 nm respectively. The particles also had a more uniform size distribution, with PDIs of 0.21 and 0.18. C14 DOPE RNP particles were closer to the size of the Cas9 mRNA particles, with a hydrodynamic diameter of 142.6± 3.0 nm, and the PDI value of 0.43 indicated a heterogenous population. C14 and C18 DOPE RNP particles had a highly positive surface charge of 61.0± 5.9 mV and 60.4±4.7 mV, while C16 DOPE surface charge was 35.5 ±2.3 mV (Table 3.2).

Table 3.2. Biophysical characterisation of Cas9 mRNA and RNP nanoparticles. *Table shows hydrodynamic size (nm), polydispersity index (PDI) and charge (mV), as measured by dynamic*

light scattering (DLS) using a Malvern Nano Zetasizer (Malvern Instruments, England) Peptide Y was used in all measurements. Three repeat measurements were made and standard deviation calculated.

Nanoformulation		Z-Ave (d.nm)	PDI	Charge (mV)
Cas9 mRNA	C14 DOPE	134.2± 3.1	0.33	20.8 ±2.8
	C16 DOPE	128.1± 4.5	0.31	33.8 ±1.0
	C18 DOPE	177.3± 1.3	0.28	36.8 ±3.4
Cas9 RNP	C14 DOPE	*142.6*± 3.0	*0.43*	*61.0* ±5.9
	C16 DOPE	*95.7*± 0.7	*0.21*	*35.5* ±2.3
	C18 DOPE	*90.0*± 0.5	*0.18*	*60.4* ±4.7

3.3.5.2 Transmission electron cryomicroscopy (CryoTEM)

Cryo-TEM imaging was used to determine the morphology of the RTNs. The order of mixing of RTN components is as follows: lipid, mRNA/RNP then peptide. We therefore used cryo-TEM to determine the morphology of the RTNs at each stage of component addition. Weight ratios used were 1:3:4 mRNA: L:P and 1:1:4 RNP: L: P. C18 DOPE and C16 DOPE alone, for mRNA and RNP delivery respectively, formed small, spherical, unilamellar vesicles. The population was relatively heterogenous in size, with particle diameters ranging from ~10 - 100 nm (Figure 3.22**A** and **D**). Upon addition of Cas9 mRNA, several of the smaller empty liposomes fuse together, with the nucleic acid cargo residing between lipid leaflets, in very compact, multi-lamellar structures. Most particles were ~100 nm, with some larger aggregates (Figure 3.22**B**). Finally, with addition of peptide, the morphology of the particles completely changed, with smaller (<40 nm), less uniform particles observed, with higher polydispersity (Figure 3.22 **C**). Upon addition of Cas9 RNP to the liposomes, some multi-lamellar structures were observed, with the RNP seemingly residing between lipid bilayers (yellow arrows). Particles were relatively uniform, with the majority ~100 nm. Some empty liposomes were apparent, suggesting the amount of lipid was in excess relative to the RNP needed to form complexes (Figure 3.22**E**). The addition of peptide, again, completely distorted the complexes and, in the case of the RNP complexes, caused large aggregates over

200 nm in diameter, and a high polydispersity (Figure 3.22**F**). We hypothesise that the darker regions within the larger structures are where the peptides have inserted themselves into the lipid bilayers (white arrows). Smaller particles of ~100 nm were observed, and are believed to be more representative of the entire population, due to the particle sizes observed in section 3.3.5.1.

Cryo-TEM was performed under the guidance of Jayesh Kulkarni, UBC, Vancouver.

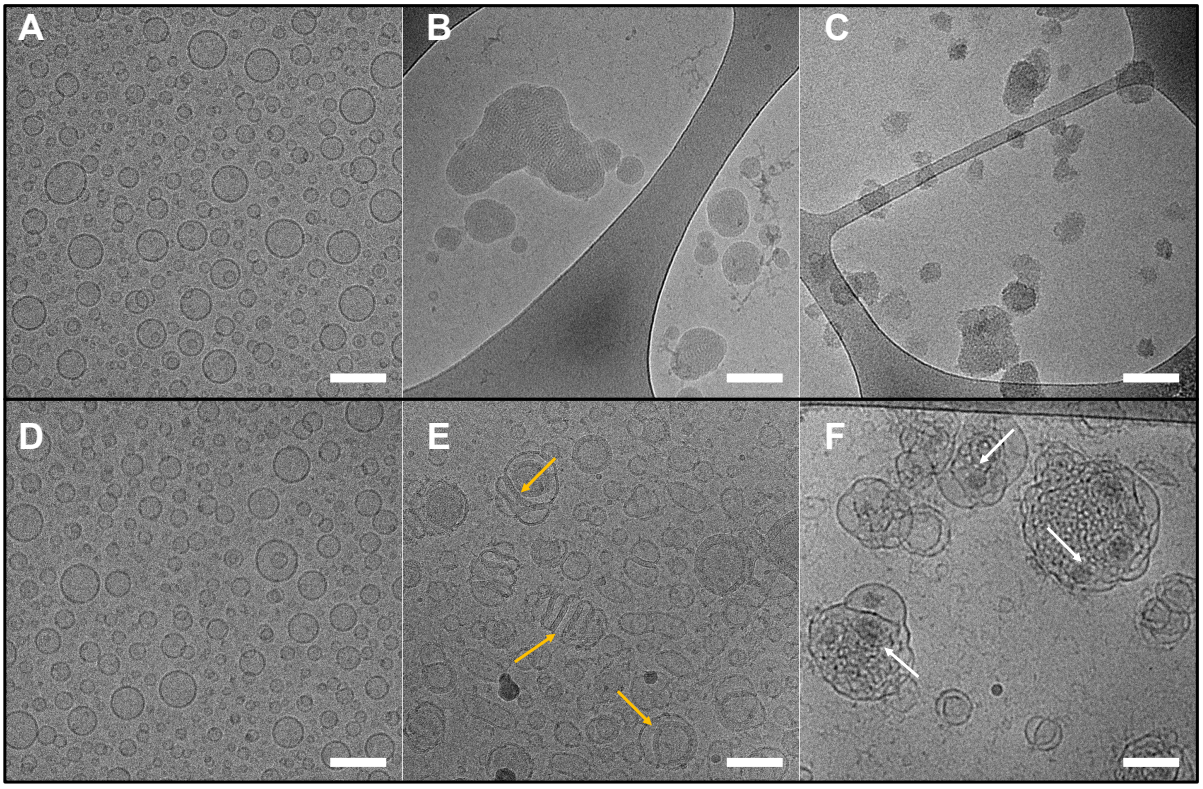


Figure 3.22. Cryo-TEM images of Cas9 mRNA and RNP RTNs. *(**A**) empty liposomes before addition of Cas9 mRNA. (**B**) Lipid + Cas9 mRNA components. (**C**) Fully formed RTNs after addition of Peptide Y. (**D**) empty liposomes before addition of Cas9RNP. (**E**) Lipid + Cas9 RNP components. (**F**) Fully formed RTNs after addition of Peptide Y. All samples were imaged at a 55,000× magnification with a nominal under-focus of 1–2 µm to enhance contrast. Yellow arrows indicate RNPs residing between bilayer. White arrows indicate where peptide has inserted into bilayer structure. Scale bar is 100 nm.*

3.3.5.3 Nucleic acid condensation and packaging properties of RTNs

Distinctions in morphology, size and surface charge indicates differences in nucleic acid condensation and packaging. These properties were assessed using the gel retardation assay, which utilizes SYBR Safe to detect mRNA presence. Nanocomplexes are too large to travel through the pores of a 1% agarose gel and so mobility of RNA, when condensed and packaged into nanocomplexes, is retarded, and this can be observed by electrophoresing complexed RNA alongside naked RNA on an agarose gel. In addition to assessing packaging capabilities, RNA release can be assessed by using a highly anionic species, such as heparan sulphate, that will compete with RNA to interact with cationic species within nanocomplexes thereby releasing the RNA [397, 398].

C14 DOPE, C16 DOPE and C18 DOPE nanocomplexes with Cas9 mRNA were prepared as previously described and incubated for 30 min. Complexes were then incubated with heparin sulphate for 30 min. As expected, there was no difference in RNA mobility between untreated and heparin treated naked mRNA. C16 and C18 DOPE nanocomplexes completely retarded the mobility of RNA, as shown by retention in gel wells, with strong SYBR Safe staining. C14 DOPE appeared to be less efficient in packaging than C16 and C18, with a smear of nucleic acid present in the lane, however there was still strong retention in the well (Fig. 3.23). In the presence of heparin, dissociation of RTNs was observed, with RNA bands shown in the gel suggesting liberation of Cas9 mRNA from the RTNs. However, some staining was still evident in the wells of the gel, suggesting that the complexes had not completely dissociated.

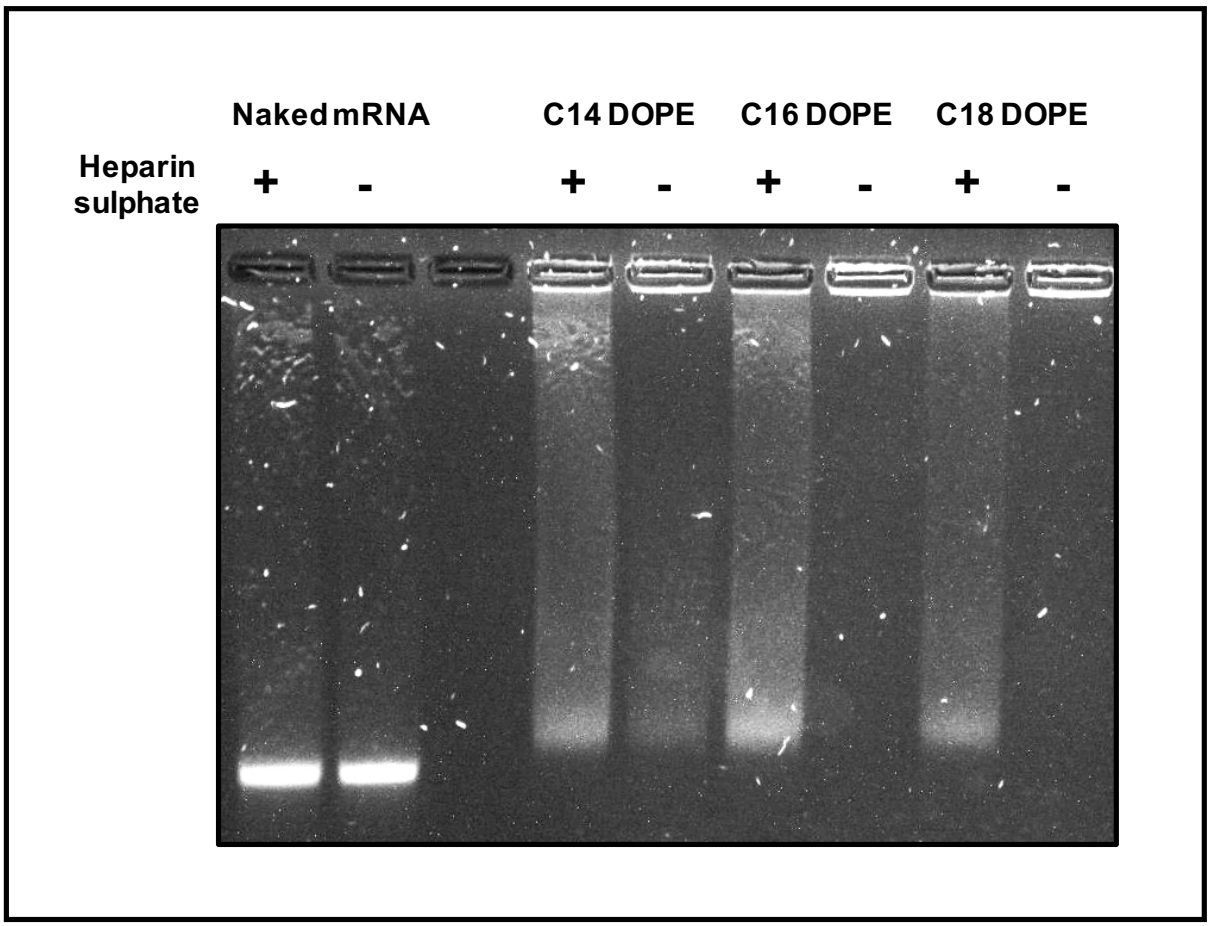
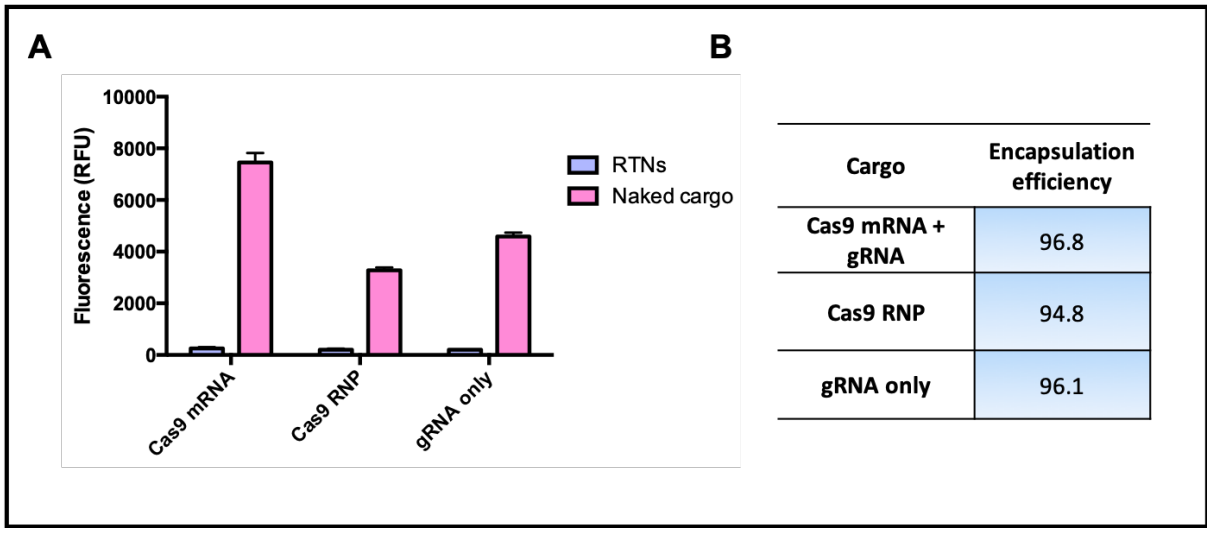


Figure 3.23 nucleic acid packaging and release by RTNs, as measured by a gel retardation assay. *400ng Cas9 mRNA + gRNA were complexed with C14, C16 and C18 DOPE and Peptide Y, followed by 30 min incubation in the presence or absence of heparin sulphate. RTNs were compared to naked Cas9 mRNA.*

Encapsulation efficiency was further investigated in a Ribogreen quenching assay, a fluorescent dye used for detection and quantification of nucleic acids. We chose to compare optimised Cas9 mRNA and RNP nanoparticles- C18 DOPE Peptide Y and C16 DOPE Peptide Y respectively. RTNs containing Cas9 (1) mRNA + gRNA, (2) Cas9 RNP complexes and (3) gRNA alone were formulated as previously described. RTNs were then incubated with Ribogreen, in parallel to naked Cas9 mRNA, RNP or gRNA, before measurement of Ribogreen fluorescence. Encapsulation efficiency was measured as fluorescence in RTN incubation as a fraction of fluorescence in the presence of free nucleic acid, and expressed as a percentage. RTNs were able to encapsulate Cas9 mRNA + gRNA with 96% efficiency, with 94% efficiency for Cas9 RNPs (Figure 3.24A). The lower fluorescence in the presence of naked Cas9 RNPs as compared to

gRNA only suggests that Cas9 protein alone in part protects the gRNA when forming an RNP complex, and so the gRNA is partly shielded from the Ribogreen dye (Figure 3.24B).



Cargo	Encapsulation efficiency
Cas9 mRNA + gRNA	96.8
Cas9 RNP	94.8
gRNA only	96.1

Figure 3.24 Ribogreen quenching assay to determine nucleic acid encapsulation efficiency. *(**A**) Ribogreen fluorescence in the presence of naked nucleic acid cargo vs. encapsulated RTNs. RFU = relative fluorescence units. (**B**) Encapsulation efficiency was calculated as follows: fluorescence in RTN incubation as a fraction of fluorescence in the presence of naked nucleic acid, and expressed as a percentage.*

3.4 Discussion

Gene therapy targeted to the airway epithelium has been an objective for CF researchers for over two decades, and several clinical trials have taken place, most recently a Phase IIb in 2012, which showed a modest improvement in FEV1 of 3.7% [205]. The development of the CRISPR/Cas9 genome editing technology has generated huge excitement, affording researchers the ability of precisely altering the genome with much greater ease and lower cost than previous gene editing techniques [399]. Indeed, CRISPR has been used to correct several myriad of CF-causing mutations *in vitro* [101, 368-370]. However, delivery of the CRISPR components into the lung will remain a challenge, which must be overcome for translational application. A delivery method must be both safe and efficient, in order to realise the potential of this robust and precise technology.

Suitable *in vitro* models must be developed in order to test the efficiency and delivery of these technologies into the appropriate cells. Of course, primary epithelial cells isolated from the lungs of CF patients are a gold standard, and these cells, grown on ALI culture replicate many of the features of a CF airway including production of thick, sticky mucus, and depleted PCL [87, 88]. However, the problem with these basal cells is that they lose their differentiation capacity after two passages, necessitating frequent sampling from CF donors. Moreover, gene editing strategies are likely to take longer than the few passages primary cells can undergo before losing their differentiation capacity, rendering these cells unworkable for longer experiments.

Recent advancements have provided new ways of expanding the proliferative potential of primary cells, including culture on a feeder layer [91, 92], or culture with dual inhibition of SMAD signalling pathways [94]. A former member of the Hart laboratory expanded proliferative capacity by transduction of the anti-senescent, proto-oncogene B-cell-specific Moloney murine leukaemia virus integration site 1 (BMI-1) [96], allowing cells to be expanded for at least 20 population doublings, whilst maintaining genetic stability.

It is essential that these modified cells maintain key properties of primary epithelial cells; for example, if we are to execute a CFTR correction strategy by means of CRISPR/Cas9, it is crucial that CFTR expression is maintained despite modification. In the first part of this chapter, we characterised *BMI-1* transduced normal human bronchial epithelial cells, firstly by assessing their proliferative potential. We found that cells transduced with *BMI-1* at MOI

1, 4 and 16 could all be expanded for at least 44 population doublings, when the cultures were stopped. Interestingly, when comparing to un-modified primary cells, we found that the primary cells were able to be expanded for approximately 18 population doublings over 44 days. Despite growth slowing at around 15 population doublings, this is still a higher proliferative capacity than the literature would suggest [400]. When comparing cell morphology between early and later passage, we found that both passages maintained the cobblestone appearance, typical of primary epithelial cells.

We next investigated CFTR expression in primary bronchial epithelial cells by means of q-RTPCR and Western Blot, at various stages of ALI culture, in order to determine when CFTR is most highly expressed. CFTR mRNA expression peaked at week 3, whilst it appeared that protein expression was highest at week 2. When these experiments were repeated on *BMI-1* modified cells, we found that MOI-1 had the highest CFTR mRNA expression, while protein expression appeared similar between MOI-1 and 16. Interestingly, there appeared to be a more prominent B-band in the *BMI-1* modified cells compared to primary cells. The CFTR protein can exist in three different molecular weights of approximately 127 (A-band), 131 (B-band) and 160 kDa (C-band), representative of non-glycosylated, core-glycosylated or fully mature CFTR respectively [401]. The presence of B-band indicates endoplasmic reticulum (ER) processing, while the mature C-band with complex glycosylation indicated Golgi processing [402], and therefore a less mature form of the protein. In any case, the experiments would need to be repeated to confirm there is a significant difference between the glycosylation of CFTR in primary vs. *BMI-1* modified cells, however this is something to consider when employing a CFTR editing strategy within this model.

Our group has previously used receptor-targeted nanocomplexes (RTNs) for delivery of plasmid and siRNA, to a variety of cell types and tissues as well as *in vivo* in the mouse lung [375-380]. The next part of this chapter focused on the optimisation of these particles to deliver CRISPR/Cas9 components. In order to assess RTN delivery efficiency, we developed a GFP reporter model to assess DSB repair by NHEJ; the CRISPR components mediate a DSB within the GFP sequence, which is repaired by the NHEJ pathway within the cell, causing small indels at the DSB site and corrupting the GFP reading frame, hence impairing GFP protein expression. Interestingly, the model could additionally be used to monitor HDR events. Changing one amino acid sequence (TAC to CAT at position 67) within the coding sequence of GFP alters the protein's fluorescent colour from green to blue (BFP) [403]. By designing a

gRNA to target this sequence, and providing a donor template which encodes histidine instead of tyrosine, it would be possible to measure HDR efficiency by means of flow cytometry [404]. The reporter system could therefore be applied in assessing RTNs for donor template delivery in repair strategies that so require it.

Using this model, we first optimised RTNs for Cas9 mRNA delivery. Our group had previously optimised the particles for siRNA delivery [391, 405-408], showing that C18 DOPE delivered most efficient silencing. However, siRNA and mRNA vary dramatically in their size and molecular weights, which may effect formulation. We therefore tested a range of lipids (C14, C16 and C18) and several different peptides (Peptide E, Peptide Y, and 4 derivatives of Peptide Y) in order to establish an optimal formulation for mRNA delivery, concluding that C18 DOPE peptide Y formulations were most efficient. Moreover, we tested different ratios between Cas9 mRNA and gRNA, finding that 300 ng Cas9 : 200 ng gRNA achieved highest GFP knockout and therefore cutting efficiency.

The RTNs had never before been used to deliver any sort of protein cargo, however initial experiments we performed showed that RNP delivery by RTNs achieved a significantly higher editing efficiency than the leading commercial product for that purpose, CRISPRMAX. When comparing Cas9 mRNA delivery and Cas9 RNP delivery, we found that there was a significant increase in editing efficiency by the RNP format, and could achieve upwards of 65% GFP knockout. After optimising the ratios of lipid and peptide components within the RTNs, a knockout rate of 73% was achieved. Interestingly, we found that using lower ratios of both lipid and peptide component resulted in higher transfection efficiencies. This confers several advantages, chiefly a lower amount of material entering the cell, limiting any toxicity.

Differences in the biophysical properties of the RTNs can influence the cellular uptake and therefore transfection efficiency. Characterisation of these properties affords a great insight as to how the particles will interact with and within the target cells. With respect to transfection, several key steps must take place to allow for efficient expression of the CRISPR/Cas9 machinery, including i) formation of stable particles that can protect the cargo from degradation, ii) attachment of the nanoparticle vector to cells, iii) cellular uptake of the nanoparticle and iv) endosomal escape. In the case of Cas9 mRNA, it must be translated into a protein by translational machinery, where with pre-formed RNP complexes, this step is omitted (Figure 3.25). The components of our RTNs impart specific properties that help overcome these barriers. As shown in the gel retardation assay and Ribogreen encapsulation

assays, our lipopolyplex formulations, containing both lipid and peptide moieties, can successfully form stable particles, effectively shielding the nucleic acid cargo and protecting it from degradation. Meanwhile, the receptor-targeting properties of the complex are afforded by the peptide component, comprised of a receptor-targeting domain to target airway epithelial cells [379]. The positive surface charges imparted on our RTNs through inclusion of cationic lipids and targeting peptides allow the RTNs to be internalised by means of endocytosis, allowing successful cellular uptake of the particles. Endosomal escape following cellular uptake is essential in order to evade lysosomal degradation of the RTNs. This is mostly achieved through destabilisation of the endosome bilayer through the use of fusogenic liposomes or peptides [409]. The presence of DOPE within our RTNs imparts a mechanism for endosomal escape within our complexes.

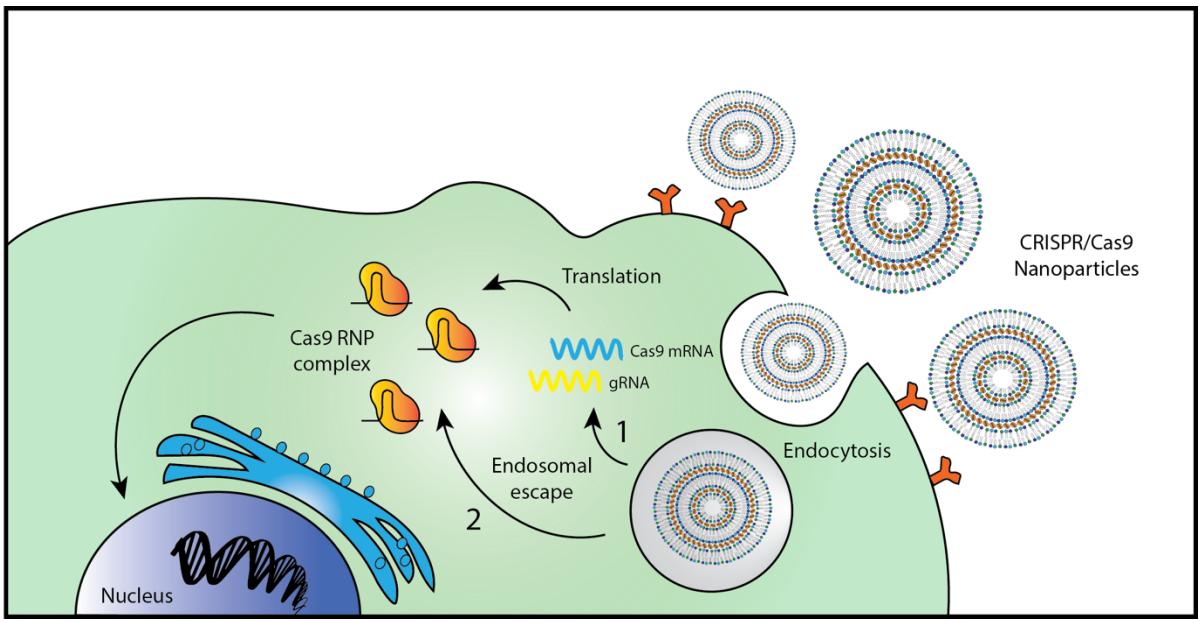


Figure 3.25. Simplified overview of cellular entry and intracellular trafficking pathways and mechanisms of RTNs. *Attachment of the nanoparticle vectors to the cell is followed by cellular entry though the endocytic pathway. After internalisation, nanoparticles are typically trafficked to an endosome, from which they must escape. The nucleic acid cargo is released from the RTNs; In the case of Cas9 mRNA (1), it must be translated into a protein by translational machinery, while Cas9 RNPs can directly enter the nucleus via the nuclear pores (2).*

Nanoparticle size and surface properties are key parameters influencing a delivery system in terms of safety, release kinetics and cellular internalization [410] [411]. It is widely reported that efficient cellular internalisation by endocytosis requires particles of less than 500 nm [412], while larger particles are reportedly phagocytosed by macrophages [413] that reside in the CF lung where they contribute to the pathophysiology of the disease [414]. Moreover, in the context of CF, previous studies suggest that nanocomplexes must have diameters smaller than the mucus mesh pores, which are ~140 nm in CF mucus [415]. A look at the DLS data shows that RNP RTN formulations has a significantly smaller hydrodynamic size and PDI than their Cas9 mRNA counterparts (size 95.7± 0.7 nm vs. 177± 1.3 nm, and PDI 0.21 vs 0.28 for RNP and Cas9 mRNA respectively). It is suggested that PDI values below 0.3 are desired in order to ensure monodisperse nanoparticle formulations [410]; both of our optimised RTN complexes meet this requirement.

With respect to the cryo-TEM data, when Cas9 mRNA and lipid are complexed together, they appear as stacked lamellar aggregates. However, through the addition of peptide Y, small discrete particles formed, suggesting this is primarily a consequence of peptide/RNA interactions, suggesting they more efficiently condense the nucleic acid than cationic liposomes alone, in agreement with other studies [416]. A similar reorganisation of particles can be observed upon addition of Peptide Y in the RNP RTNs, with a complete distortion of particle morphology. However, peptide Y addition also appeared to produce large aggregates of RTNs, which isn't seen with Cas9 mRNA RTNs. With reference to the DLS data, we can see that the size and PDI of RNP RTNs suggest small, monodisperse particles. The process of cryoTEM requires particles to be extremely concentrated (20mg/ml total liposome). We therefore hypothesise that this concentration has had an effect on the particles, contributing to this aggregation. Nevertheless, there are several molecules which could be added to the RTNs to impart different properties within the nanocomplex. PEGylation of the particles (through addition of poly(ethylene glycol)) can form protective hydrophilic layer on the cationic RTN surface, changing the surface properties and preventing aggregation of particles [187, 417, 418]. Meanwhile, cholesterol offers an increase the stability of liposomes, and can also prevent aggregation of liposomes [419]. The addition of either of these molecules could be considered for Cas9 RNP delivery, however we have chosen the optimal formulation of 1:1:4 RNP: L: P with C16 DOPE and Peptide Y to take forward into future *in vivo* work.

Chapter 4
Results

Developing NHEJ-mediated CRISPR/Cas9 correction strategies for treatment of CF

4.1 Introduction

A significant fraction of CF mutations (13%) affect the pre-mRNA splicing of the gene, by disrupting or generating intronic splicing motifs, critical for exon recognition (http://www.genet.sickkids.on.ca/cftr). Additionally, mutations can disrupt exonic splice sites, which can also affect splicing pattern. Splice mutations can be divided into two groups based on their effect on splicing. In group one, mutations completely abolish exon recognition resulting in complete absence of correctly spliced transcripts, such as 621 +1 G>T and 711 +1 G>T [420]. The second group includes mutations that generate both aberrantly and correctly spliced transcripts. One such mutation, 3849 +10kb C>T, was initially discovered in patients that have chronic lung disease similar to that of a CF patient, but have normal sweat chloride levels [46], making it difficult to diagnose. The CFTR gene was sequenced in these patients, and a C > T point mutation was detected deep within intron 19. This mutation generates a cryptic GT splice donor site, leading to the insertion of an 84 bp pseudoexon between exons 19 and 20. This pseudoexon contains an in-frame TAA stop codon, which generates a truncated, non-functional protein [46]. Studies using RT-PCR showed considerable variation in the level of aberrantly spliced RNA transcribed from the mutation, with low levels of correct transcripts (0-28%) identified in the nasal epithelium from 10 patients [47]. Increased levels of aberrant mRNA correlated significantly with disease severity, as measured by FEV1. Further studies have shown that levels of aberrant splicing can vary among organs of the same patient, with these individuals having extremely variable disease expression ranging from minimal lung disease, male fertility and pancreatic sufficiency, to severe disease in all affected organs [421].

It was demonstrated that an increase in correctly spliced mRNA could be achieved by means of overexpressing the splice factors SC35 and Htra2-β1 which promote skipping of the 84 bp pseudoexon, restoring the levels of full-length CFTR transcripts [422]. A targeted therapy reducing aberrant splicing may then reduce disease phenotype, with valuable implications for this sub-population of CF patients. Based on the observation that NHEJ repair of two DSBs would result in the targeted deletion of the genomic sequence between the cut sites, it was shown that CRISPR/Cas9 was able to delete the 3849 +10kb C>T mutation as well as two other CFTR deep-intronic splice mutations, as evaluated in a mini-gene assay in

HEK293T cells[423]. In the study, gRNA pairs were designed to make two targeted DSBs, upstream and downstream of the mutation. Following NHEJ-repair, editing efficiencies of up to 45% were demonstrated, restoring normal splicing of CFTR mRNA.

Cell lines such as HEK293T, whilst simple to culture and easy to transfect, often differ phenotypically and genetically to their tissue of origin, with an altered morphology [424]. Further, HEK293 cells, do not endogenously express CFTR [425]. In order to use a more clinically relevant cell model, we wished to conduct similar experiments with *BMI-1* transduced primary bronchial epithelial cells homozygous for the 3849+10kbC>T (referred to hereafter as CFBE3849 cells), and functionally validate correction and restoration of CFTR expression.

While this NHEJ-mediated strategy could be applied to a specific cohort of CF-causing splice mutations, therapies with wider potential for CF patients will require the delivery of a donor template. Several HDR-mediated repair strategies have been used in cellular models for CFTR genetic repair, [101, 363-365, 368], however efficiency tends to be low, limiting clinical application. Moreover, the HDR pathway is not highly active in human cells, particularly post-mitotic, terminally differentiated cells, such as the ciliated airway epithelial cells that would be targeted *in vivo*. Therefore, an NHEJ-mediated strategy is likely to have much higher efficiency. Belmonte and colleagues described in 2017 a homology-independent targeted integration method (HITI), allowing for efficient and precise knock-ins in both dividing and, critically, non-dividing cells both *in vitro* and *in vivo* [373]. Instead of utilising HDR, HITI exploits the NHEJ pathway to repair DSBs created in the genome. The HITI donor template lacks homology arms, but rather includes a Cas9 cleavage site flanking the donor sequence. Upon addition of Cas9 nuclease, DSBs are mediated simultaneously in both the target genomic sequence and the exogenously provided donor template, and the template can be used for repair *via* the NHEJ pathway, allowing for genomic integration. After incorporation in the desired orientation, the Cas9 target site is disrupted, preventing further Cas9 cutting. However, should the template be inserted in the reverse orientation, the Cas9 target site remains intact, and so a second round of Cas9 cutting can occur, mediating the removal of the integrated donor, and allowing a second opportunity for desired integration direction, and allowing for insertion in a predetermined direction [373]. In the final part of the chapter we investigate this NHEJ-mediated strategy for insertion of the GFP transgene

into the AAVS1 locus of HEK293T cells, as proof-of-concept for replacing a larger part of the CFTR gene as a method of correcting a larger number of CF-causing mutations.

4.2 Aims

In this chapter, we aim to:

1. Validate gRNAs and the targeted excision of region of interest in HEK293T cell model.

2. Excise the 3949 + 10 kb C>T mutated region in *BMI-1* modified primary lung epithelial cells homozygous for this mutation using an NHEJ-mediated strategy delivered by receptor-targeted nanocomplexes.

3. Molecularly and functionally validate CFTR correction in CFBE3849 cells.

4. Develop an NHEJ-mediated HITI editing strategy for GFP knock-in in a HEK293T cell line.

4.3 Results

4.3.1 Targeted excision of a deep intronic splice mutation using an NHEJ-mediated editing strategy

Here, we focus on using a pair of gRNAs for the targeted excision of the 3849 + 10 kb C>T mutation using RTN delivery of CRISPR/Cas9.

4.3.1.1 Validation of gRNAs in HEK293T cells

Using an online tool that provides specificity and efficiency scores for gRNAs through a built-in algorithm (https://benchling.com/crispr), guides were chosen that target upstream and downstream of our mutated region of interest. gRNA locations and PAM sites are shown in Figure 4.1 and summarised in Table 4.1.

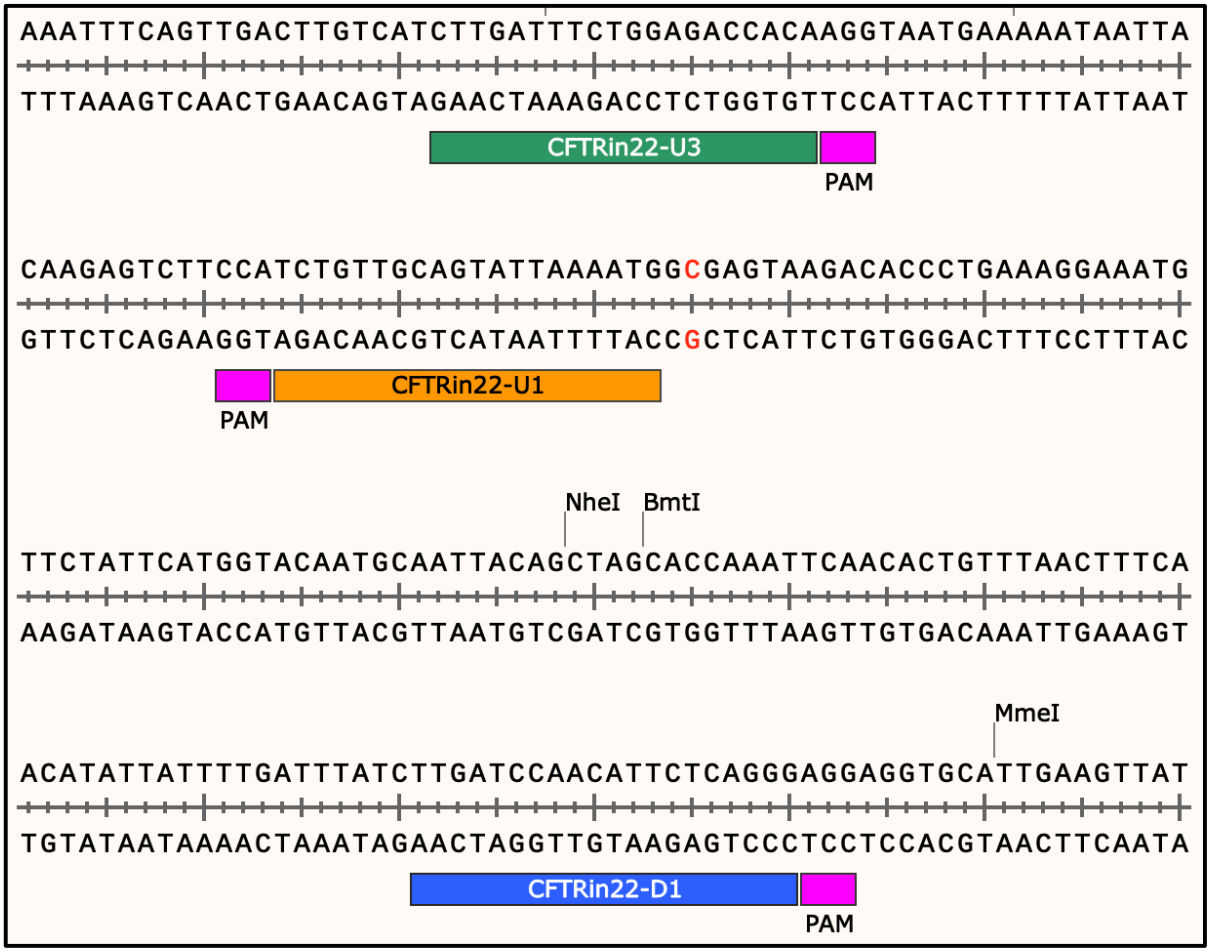
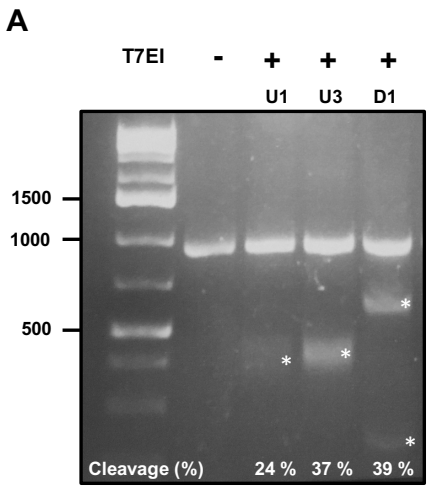


Figure 4.1. Section of CFTR sequence surround CFBE3849 + 10 kb C>T mutation *in CFBE3849 BMI-1 cells showing three gRNAs targeted to excise the mutated region. PAM sequences highlighted in pink. Mutation highlighted in red.*

Table 4.1. List of gRNAs targeting the CFB3849 C>T mutation. *The table shows gRNA sequence, strand targeted, the PAM motif sequence, the predicted efficiency score (on-target score) and specificity score (off-target score). The on-target and off-target scores were calculated by using the Benchling programme (https://benchling.com/crispr).*

gRNA	Sense	Sequence	PAM	Efficiency score	Specificity score
upstream U1	-	CATTTTAATACTGCAACAGA	TGG	74	61.1
upstream U3	+	CTTGATTTCTGGAGACCACA	AGG	63.7	79.2
downstream D1	+	TTGATCCAACATTCTCAGGG	AGG	59.8	72.8

To confirm the functioning of these guides *in vitro*, HEK293T cells were chosen as an easy to transfect model to assess gRNA cutting efficiency. Cells were seeded in 24-well plates 24 hours prior to transfection. gRNAs were complexed with Cas9 protein before addition of Lipofectamine 2000® (L2K), a commercially available transfection reagent, at a weight ratio of 1:3 RNP:L2K. 48 hours later, genomic DNA from the cells was harvested, followed by PCR amplification of ~1kb region of interest surrounding target area. The T7EI endonuclease assay was used to evaluate indel formation identified by the presence of an extra band on the gel (Figure 4.2). Densitometry was performed on the agarose gel, and the efficiency of each gRNA was calculated using the following formula: % gene modification = 100 x (1 – (1- fraction cleaved)$^{1/2}$). All three gRNAs successfully created indels within the region of interest, the most efficient being gRNA D1, with 39% gene modification.



B

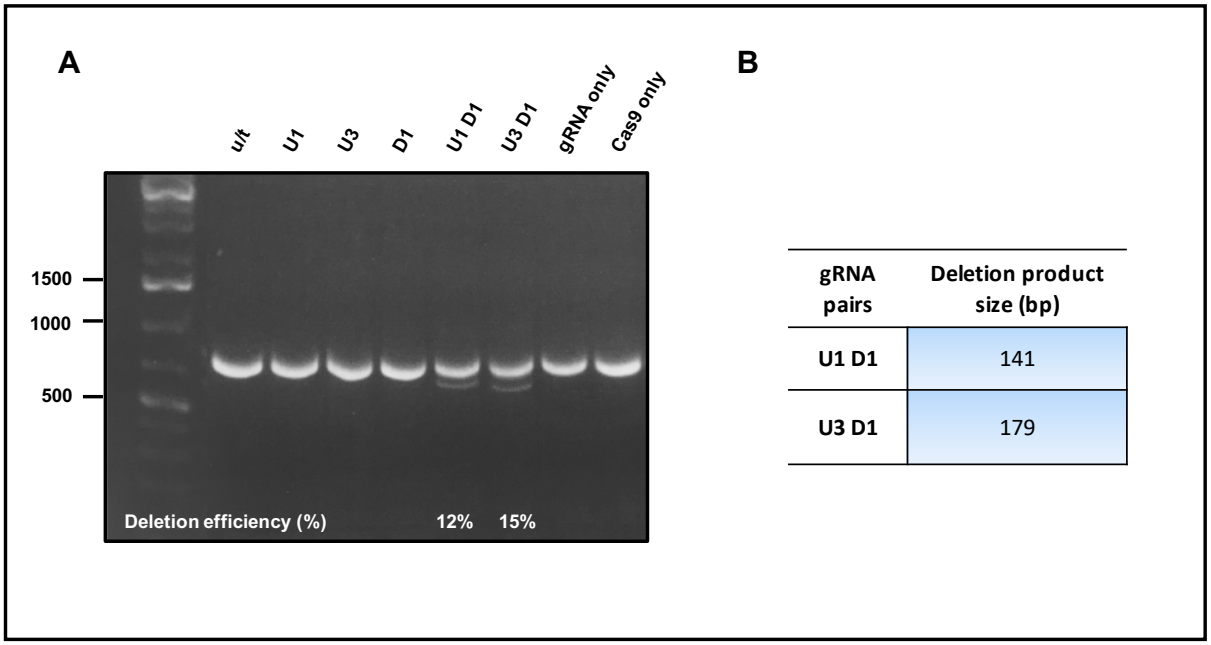
gRNA	Cleavage site	Fragment size 5'	3'
U1	522	552	471
U3	484	484	509
D1	663	663	330

Figure 4.2. gRNA validation in HEK293T cells*. (**A**) T7 Endonuclease I mismatch assay products from CFTR intron 22 in HEK293 cells treated with gRNA Cas9 RNP using a Lipofectamine® 2000. Products were resolved on a 1% agarose gel and visualised under UV light. Molecular weight referenced by 1 Kb Plus DNA Ladder. Cleaved product bands are indicated with an asterisk. (**B**) Expected gRNA-guided Cas9 cleavage site and resulting fragment sizes.*

4.3.1.2 Targeted excision using gRNA pairs in HEK293T cells

Having validated that each gRNA upstream and downstream of the mutation can successfully create a DSB, the next step was to co-transfect pairs of gRNA in order to create an NHEJ-mediated targeted deletion within the genome. As before, cells were seeded 24 hours prior to transfection. Pairs of gRNAs were complexed with Cas9 protein before addition of L2K, at a weight ratio of 1:3 RNP:L2K. 48 hours later, genomic DNA from the cells was harvested, followed by PCR amplification of ~1kb region of interest surrounding target area. The PCR product was then resolved by agarose gel electrophoresis. Any lower molecular weight bands are indicative of sequences that contain deletions. As expected, single bands were apparent in un-transfected cells and in cells transfected with only one gRNA, as well as negative controls, transfected with Cas9 only, or gRNA only (Figure 4.3a). However, when pairs of gRNAs were co-transfected (U1 D1 and U3 D1), a smaller product was visible below the parental band, suggesting a deletion event. Using densitometry, deletion efficiency was calculated to be 12% and 15% for U1 D1 and U3 D1 pairs respectively. Expected deletion

product size is given in Figure 4.3b, however precise product sizes within the gel were difficult to determine. PCR products containing a targeted excision were sent for Sanger sequencing, along with an un-transfected control. Sequence disruption is present downstream of gRNA cut-sites, indicative of a gRNA specific-editing event (Figure 4.4).



gRNA pairs	Deletion product size (bp)
U1 D1	141
U3 D1	179

Figure 4.3. Targeted excision in CFTR intron 22 induced by pairs of gRNA-Cas9*. (**A**) HEK293T cells were treated with Cas9 and individual, or paired upstream and downstream gRNAs in RNP format. All transfections were performed using Lipofectamine® 2000. Genomic DNA was extracted for amplification of region of interest before resolution on a 1% agarose gel and visualisation under UV light. Lower molecular weight species represent products of targeted excision. DNA band quantification was performed via ImageJ and deletion efficiency given as cleaved product as a % of total product. Molecular weight referenced by 1 Kb Plus DNA ladder. (**B**) Expected deletion size induced by paired gRNAs.*

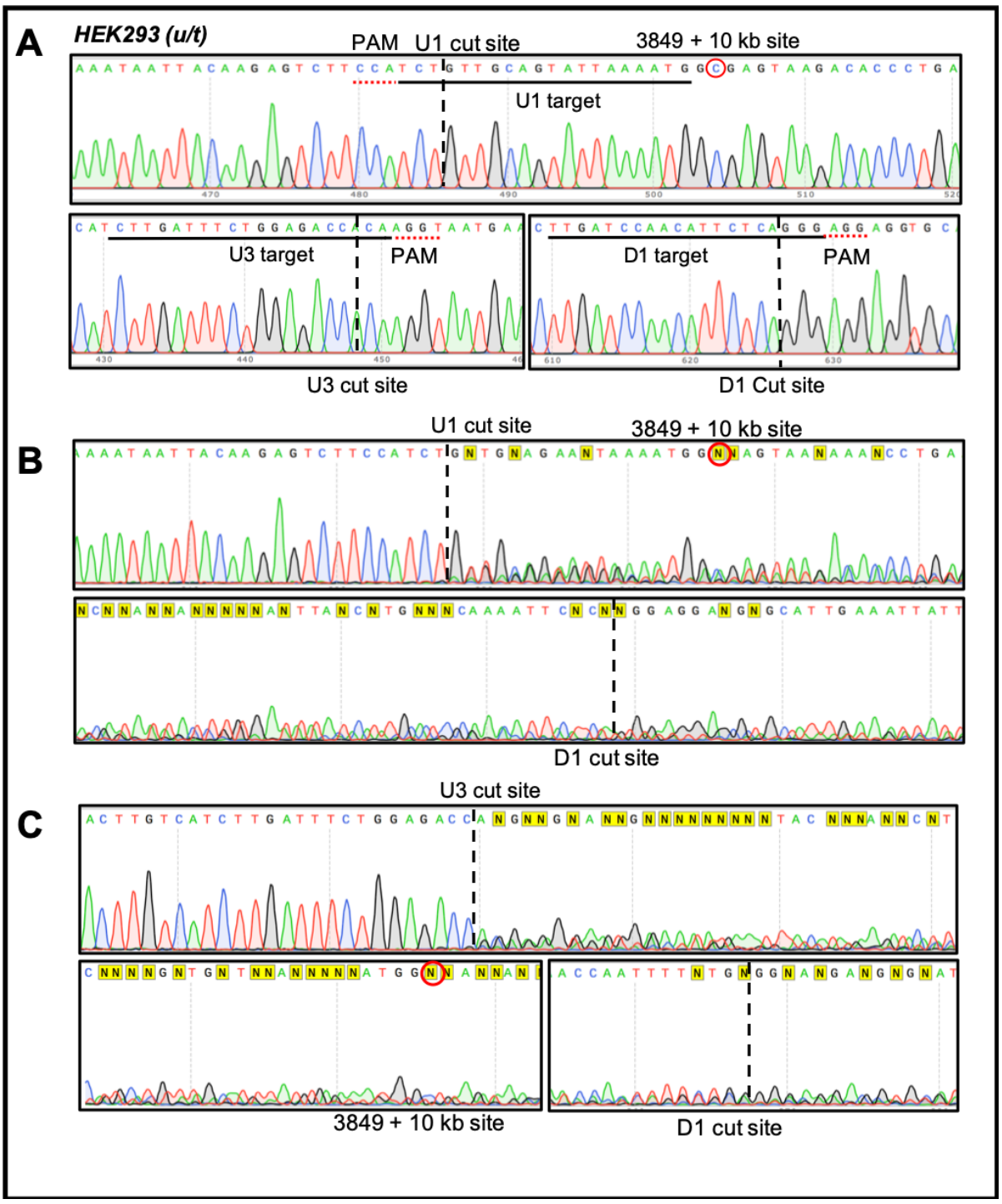


Figure 4.4. NHEJ-mediated editing of the CFTR 3840+10kb site in HEK293T cells using pairs of gRNAs. *HEK293T cells were **(A)** untreated, **(B)** treated with U1 D1 Cas9 RNP, or **(C)** U3 D1 Cas9 RNP using L2K. A ~1kb region of genomic DNA was PCR amplified around the region of interest sent for Sanger sequencing. gRNA-specific sequence disruption occurs in both treated populations, exhibiting a heterogeneous base population and disrupted 3849+10kb site, as circled in red and denoted by 'N', indicating the software is unable to identify the base present at this location.*

4.3.1.3 Validation of gRNAs in CFBE3849 cells

After successful validation of gRNA pairs in HEK293T cells, we moved the strategy to a more clinically relevant cell type. Primary bronchial epithelial cells homozygous for the 3849+10kbC>T mutation were *BMI-1* transduced as previously described in section 3.3.1, at an MOI of 1, providing a model with sufficient proliferative capacity to carry out an editing strategy and functional analysis of corrected cells.

Initially, we tested the cutting efficiency of individual gRNAs in CFBE3849 cells. Rather than using L2K as a transfection reagent, our RTNs were instead used, after showing in chapter 3 a significant increase in delivery efficiency for RNP complexes (Section 3.3.4). Single gRNAs and Cas9 protein were formulated together as an RNP, along with DOTMA (C18)/DOPE lipids and Peptide Y at a weight ratio of 1:3:4 RNP: L: P. 48 h later, genomic DNA from 50% of the cells was isolated, while the remainder of the cells were re-plated for a subsequent round of transfection. Following genomic DNA extraction after round two, PCR amplification was performed across the region of interest. The T7EI endonuclease assay was performed to identify indels caused by the gRNAs, identified by the presence of an extra band on the gel (Figure 4.5A). After the first transfection, extra bands on the gel appeared to be very faint, and we were unable to perform densitometry. However, following the second transfection, cleavage products were clearly visible on the gel. All three gRNA successfully created indels within the region of interest, the most efficient being gRNA U3 and D1, with 54% and 45% gene modification respectively (Figure 4.5A). Sanger sequencing was also performed to assess the editing efficiency with greater precision. Consistent with the T7 endonuclease results, using TIDE software, U1 had the lowest cutting efficiency (8%), whilst U3 and D1 efficiencies were 55.9% and 37.5% respectively (Figure 4.5B). The TIDE software was also used to analyse the indel spectrum of each of the gRNAs. 43% of sequences from U3 RNP treated cells had a 1 bp deletion (Figure 4.6B) while the most common occurrence in D1 treated cells was a 1 bp insertion (Figure 4.6C). No significant indels were found in U1 treated cells (Figure 4.6 A) and so U3 and D1 guide pairs were used in all subsequent transfections (Figure 4.6 B and C).

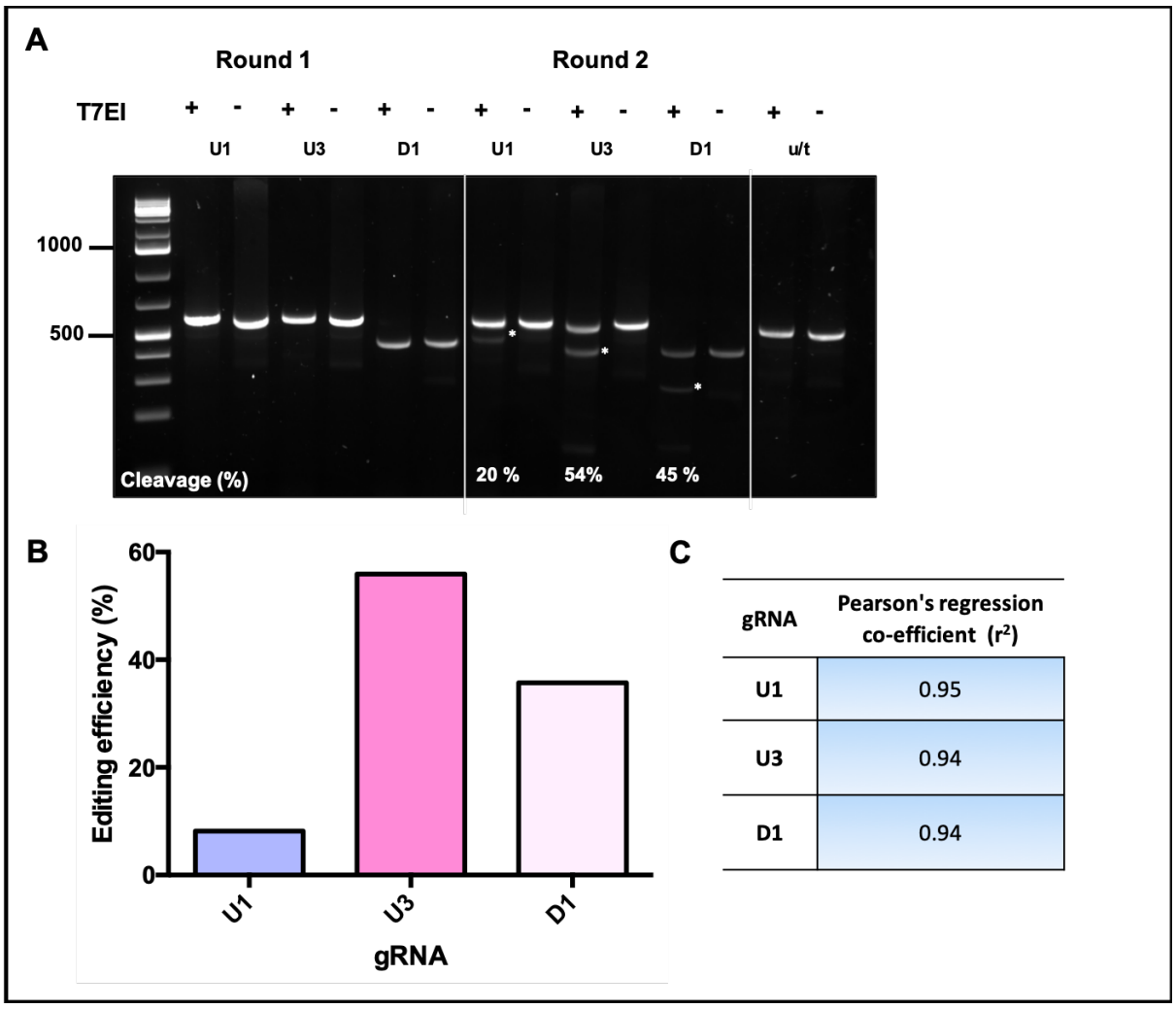


Figure 4.5. gRNA validation in CFBE3849 cells. *(**A**) T7 Endonuclease I mismatch assay products from CFTR intron 22 in CFBE3849 cells treated with individual gRNA Cas9 RNP using a C18DOPE peptide Y nanocomplex. Products were resolved on a 1% agarose gel and visualised under UV light. Molecular weight referenced by 100 bp DNA Ladder. Cleaved product bands are indicated with an asterisk. + - indicates presence or absence of T7 endonuclease. (**B**) Editing efficiency of individual gRNAs as assessed by TIDE software. (**C**) R^2 values indicate fraction of variation in the edited and unedited sequences accounted for by the models.*

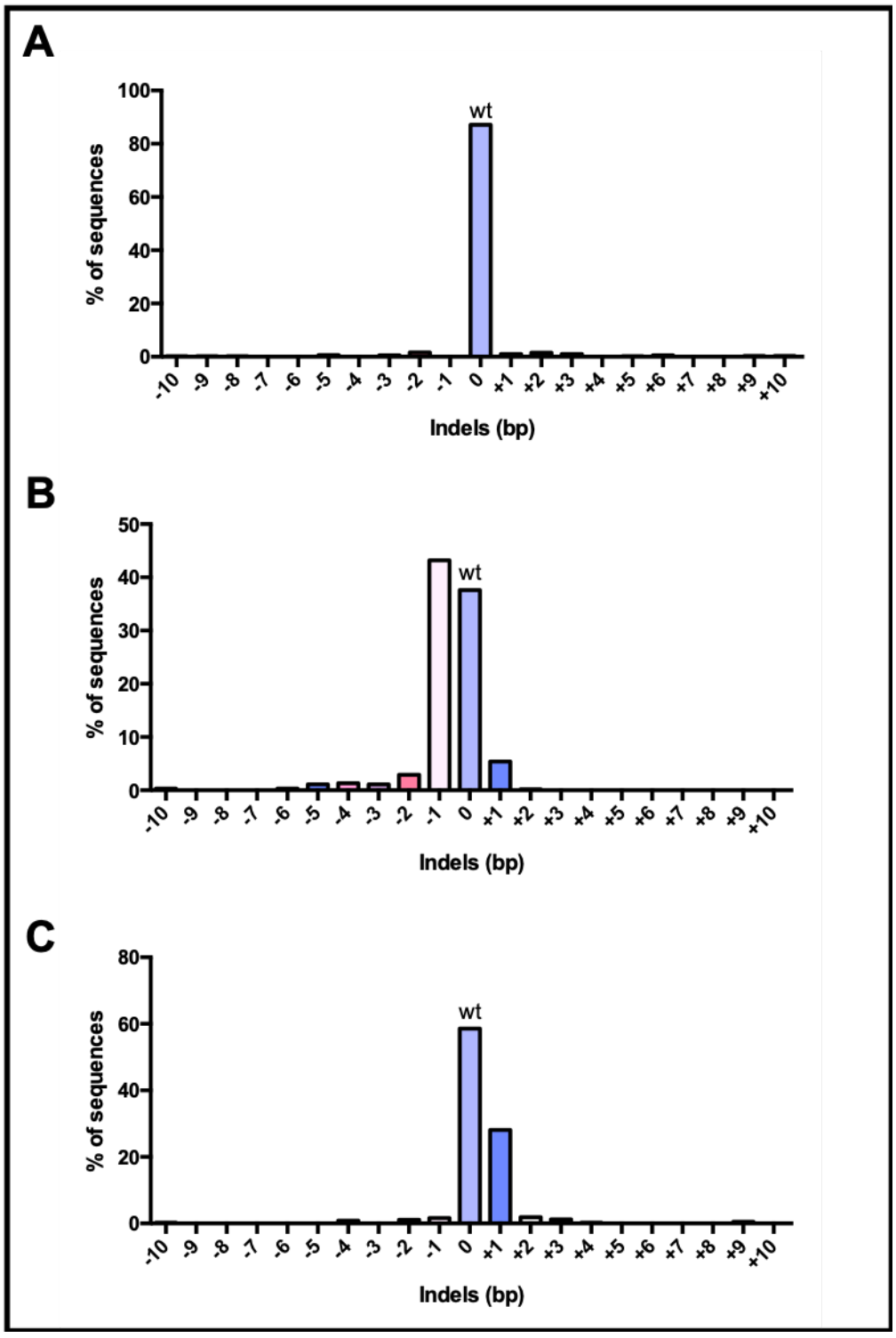
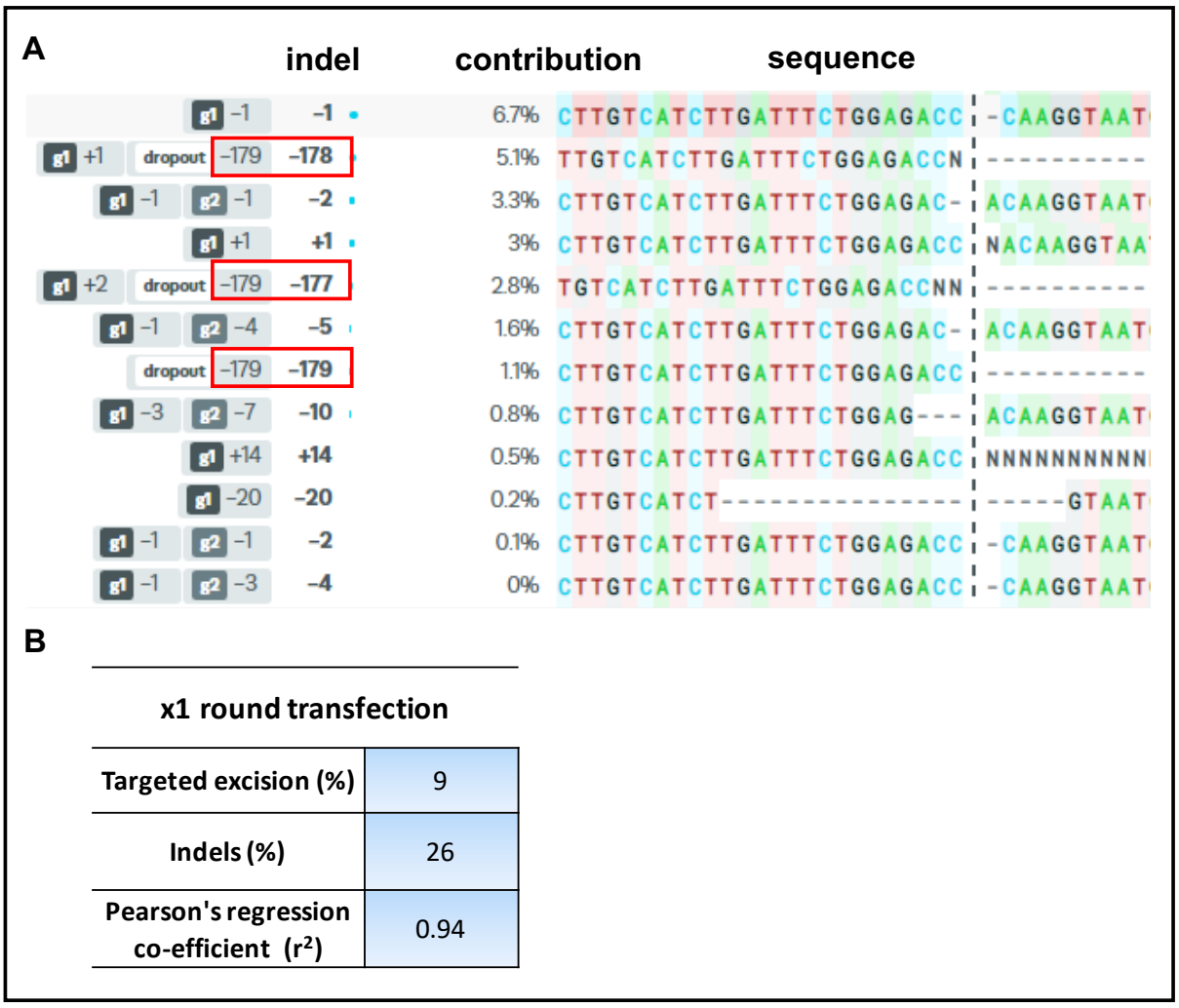


Figure 4.6. The indel spectrum of CFBE3849 cells treated with individual gRNA Cas9 RNP using C18DOPE peptide Y*. The % of sequences in the population with indels (±10bp) was assessed by TIDE from the ~1kb Sanger sequenced DNA region containing the 3849+10kb C>T site in cells treated with gRNA (**A**) U1, (**B**) U3 or (**C**) D1. R^2 values are described in Figure 4.6.*

4.3.1.4 Targeted excision using gRNA pairs in CFBE3949 cells

Having validated that each gRNA upstream and downstream of the mutation can successfully create a DSB in CFBE3849 cells, and determining that guides U3 and D1 had the highest cutting efficiencies, we next co-transfected this pair of gRNAs in order to create an NHEJ-mediated targeted deletion within the genome. As before, cells were seeded 24 hours prior to transfection. U3 and D1 gRNAs were formulated together with Cas9 protein as an RNP, along with C18/DOPE and Peptide Y at a weight ratio of 1:3:4 RNP: L: P. 48 post transfection, genomic DNA from the cells was harvested, followed by PCR amplification of ~1kb region of interest surrounding target area. PCR products were sent for Sanger sequencing, along with an un-transfected control. Sequencing results were uploaded to the Synthego software tool ICE (Inference of CRISPR Edits), which, compared to TIDE, has the added benefit of being able to detect larger deletions within genomic sequences, as well as indels. After one round of transfection, 9% of sequences analysed had a deletion of ~178 bp, the expected excision size if both gRNAs have successfully cut the genomic DNA, as highlighted in red boxes (Figure 5.7A). Small additional indels at individual gRNA cut sites can account for the varying deletions from 177-179 bp . The total indel percentage was 26 %, while the R^2 value, indicating how well the proposed indel distribution fits the Sanger sequence data of the edited population, was 94% (Figure 5.7B). Sanger sequence traces from edited and control samples are shown in Figure 5.8, where a heterogeneous base population is apparent in edited samples close to the gRNA cut sites (shown with a dotted line). This is indicative of indel formation due to NHEJ repair.



x1 round transfection	
Targeted excision (%)	9
Indels (%)	26
Pearson's regression co-efficient (r^2)	0.94

Figure 4.7 Inference of CRISPR Edits (ICE) analysis of CFBE3849. *(**A**) Cells were transfected with U3 D1 gRNAs Cas9 RNP complex using a C18DOPE peptide Y nanocomplex.* ***(A)*** *Sequences highlighted in red boxes show a deletion of ~178 bp, indicative of successful cutting by both gRNAs upstream and downstream of mutation.* ***(B)*** *Targeted excision rate was 9%, with 26% total indels, and an r^2 value of 0.94.*

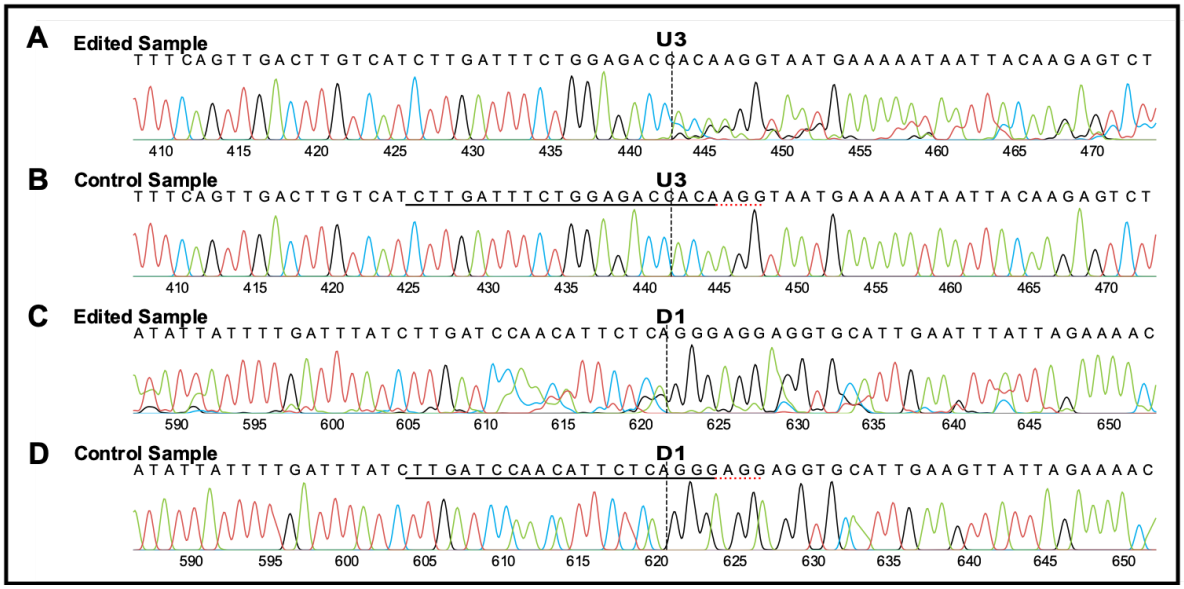


Figure 4.8. Sanger sequence analysis of edited CFBE3849 cells. *Cells were transfected with U3 D1 gRNAs Cas9 RNP complex using a C18DOPE peptide Y nanocomplex, before genomic DNA extraction and PCR amplification of region of interest. Amplicons were Sanger sequenced along with an un-transfected control and analysed using ICE software. Dotted lines indicated gRNA cutting sites, where A and B show gRNA U3 cut site, and C and D show gRNA D1 cut site. A heterogeneous base population in edited samples indicates indel formation due to NHEJ repair.*

4.3.1.5 Sequential transfections to increase targeted excision efficiency in CFBE3849 cells

In order to increase targeted excision efficiency, repeat transfections were performed on the same pool of cells. This was repeated four times using the same transfection reagents as before (4.3.4), with genomic DNA isolated for analysis after every round of transfection. DNA from the cells was harvested, followed by PCR amplification of the region of interest. PCR products were sent for Sanger sequencing, along with an un-transfected control, and again analysed by ICE. After four rounds of transfection, 60.9% of sequences analysed had the expected ~178 bp deletion as highlighted in red boxes (Figure 4.9A). The total indel percentage was 83 %, while the R^2 value was 0.88, (Figure 4.9B). Sanger sequence traces from edited and control samples are shown in Figure 4.10, where the base population is extremely heterogenous close to the gRNA cut sites in edited samples, more so than in Figure 4.8, where cells had only undergone one round of transfection. Figure 4.11 shows the targeted excision

(%) and indels (%) in each transfection, with efficiency of both increasing across the 4 rounds of transfection.

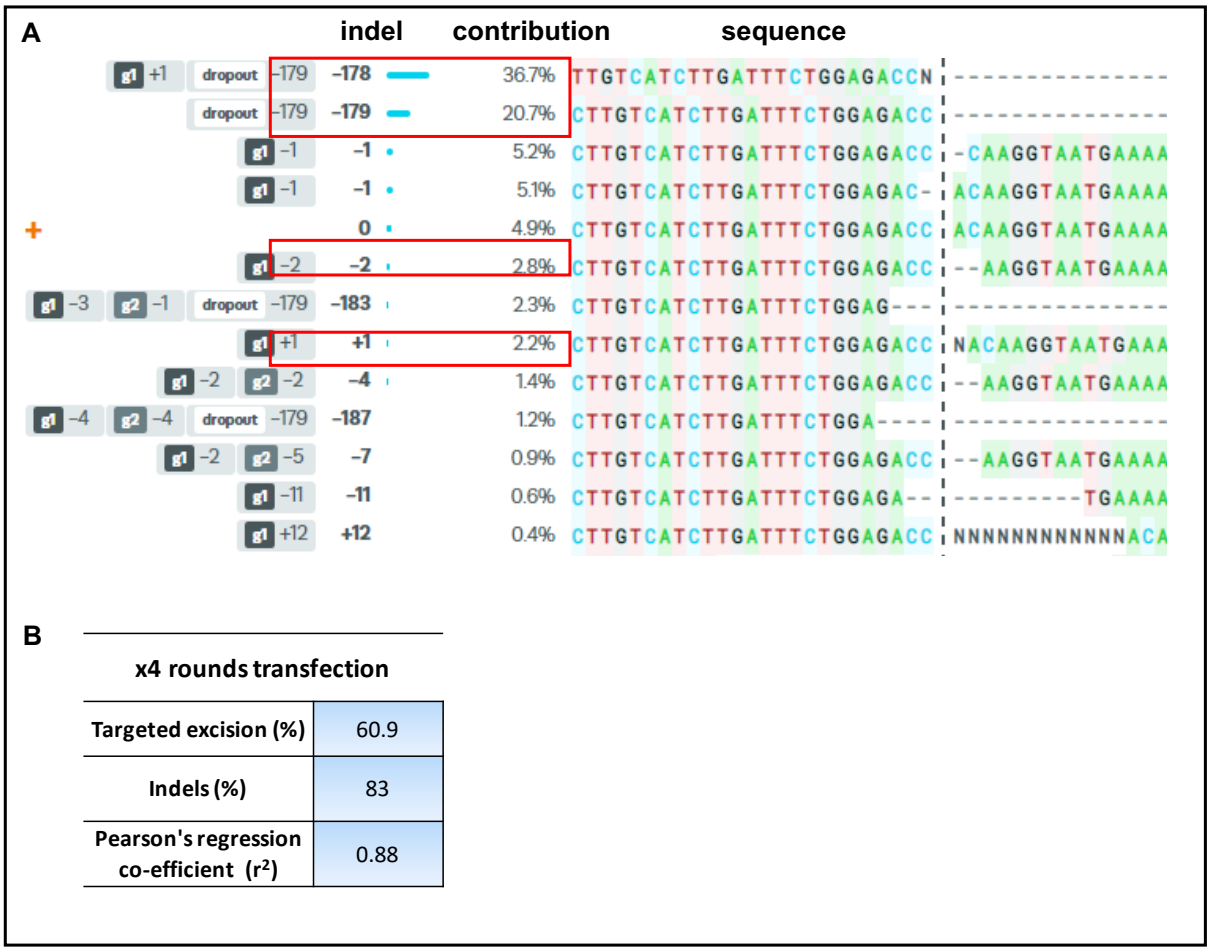


Figure 4.9. Inference of CRISPR Edits (ICE) analysis of CFBE3849. *(**A**) Cells were sequentially transfected four times with U3 D1 gRNAs Cas9 RNP complex using a C18DOPE peptide Y nanocomplex.* ***(A)*** *Sequences highlighted in red boxes show a deletion of ~178 bp, indicative of successful cutting by both gRNAs upstream and downstream of mutation.* ***(B)*** *Targeted excision rate after four transfections was 60.9%, with 83% total indels, and an r^2 value of 0.88.*

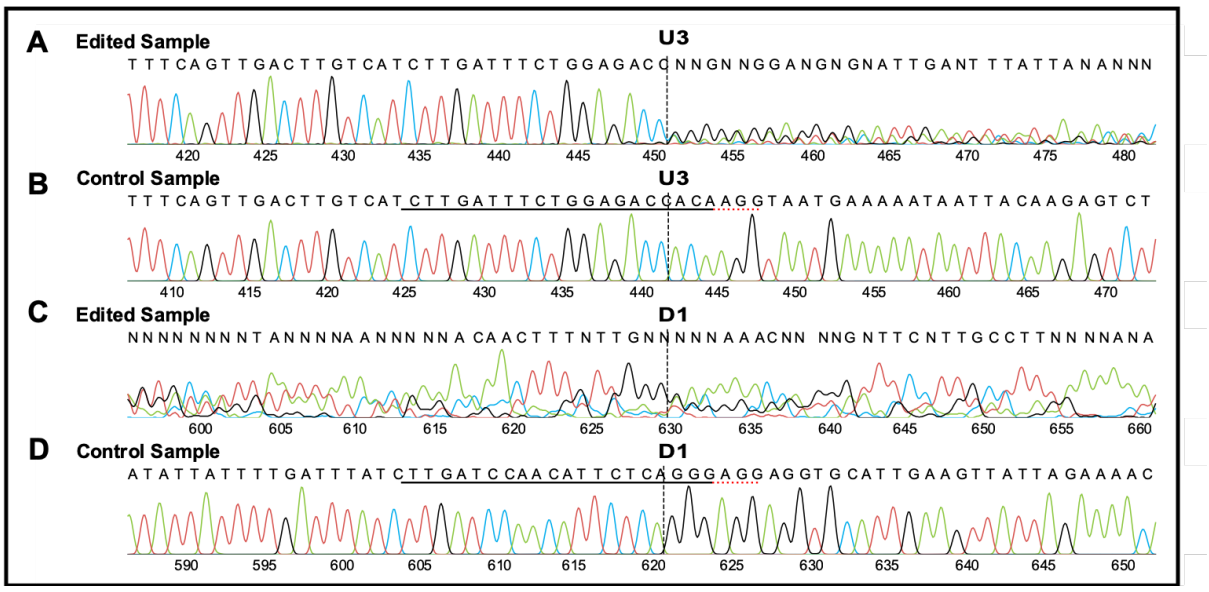


Figure 4.10. Sanger sequence analysis of sequentially edited CFBE3849 cells. *Cells were transfected 4 times with U3 D1 gRNAs Cas9 RNP complex using a C18DOPE peptide Y nanocomplex, before genomic DNA extraction and PCR amplification of region of interest. Amplicons were Sanger sequenced along with an un-transfected control and analysed using ICE software. Dotted lines indicated gRNA cutting sites, where A and B show gRNA U3 cut site, and C and D show gRNA D1 cut site. A heterogeneous base population in edited samples indicates indel formation due to NHEJ repair.*

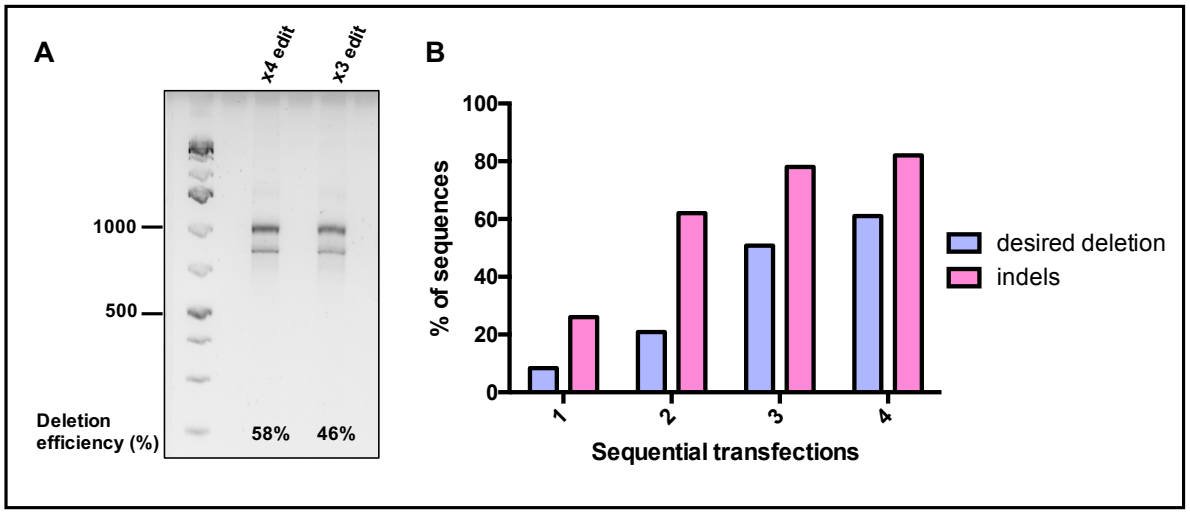


Figure 4.11.Targeted excision and indel efficiency increases over sequential transfections. *Cells were transfected up to 4 times with U3 D1 gRNAs Cas9 RNP complex using a C18DOPE peptide Y nanocomplex. (**A**) Genomic DNA was extracted for amplification of region of interest before resolution by standard gel electrophoresis on a 1% agarose gel and visualisation under*

*UV light. Lower molecular weight species represent products of targeted excision. DNA band quantification was performed via ImageJ and deletion efficiency given as cleaved product as a % of total product. Molecular weight referenced by 100 bp DNA ladder. (**B**)* Desired deletion efficiency of ~178 bp (%), and indel efficiency (%), across 4 transfections.

4.3.1.6 Single cell cloning of corrected CFBE3849 cells

To precisely assess editing event , single cells were seeded into 2x 96 well plates using serial dilution. 15 days later, plates were checked for successful growth of clones. From the two plates, 45 of the clones had proliferated, 20 of which were chosen for expansion with 6 clones successfully expanded to a T75 cm^2 flask. The remaining clones ceased proliferation at various stages during expansion. Genomic DNA was isolated from the remaining clones, followed by PCR amplification of the region of interest. PCR product was visualised on a 1% agarose gel. Of the 6 clones, clone 5 showed two bands on the gel, suggestive of a 178 bp deletion in one allele of the cells, while clone 6 showed only the lower band product, suggestion a homozygous deletion in both alleles of the cells (Figure 4.12 A). This was confirmed when the product was sent for Sanger sequencing, showing a 179 bp deletion in the edited clone as compared to a non-transfected control (Figure 4.12 C).

We next attempted to expand these corrected cells for growth on ALI culture to enable functional analysis, however they were unable to be sufficiently expanded. A similar problem was encountered when trying to generate a monoclonal GFP population in chapter 3. It appeared that the clones growing most quickly stopped replicating at a certain time point, while the slower clones were able to continue to replicate successfully, suggesting that different cells have different clonogenicity potential. In section 3.3.1.1, we investigated the proliferative potential of *BMI-1* modified cells, showing the cells could be expanded for at least 44 population doublings. It was estimated that these cells have undergone upwards of 70 population doublings, and that this is limiting the differentiation capabilities of the corrected clones.

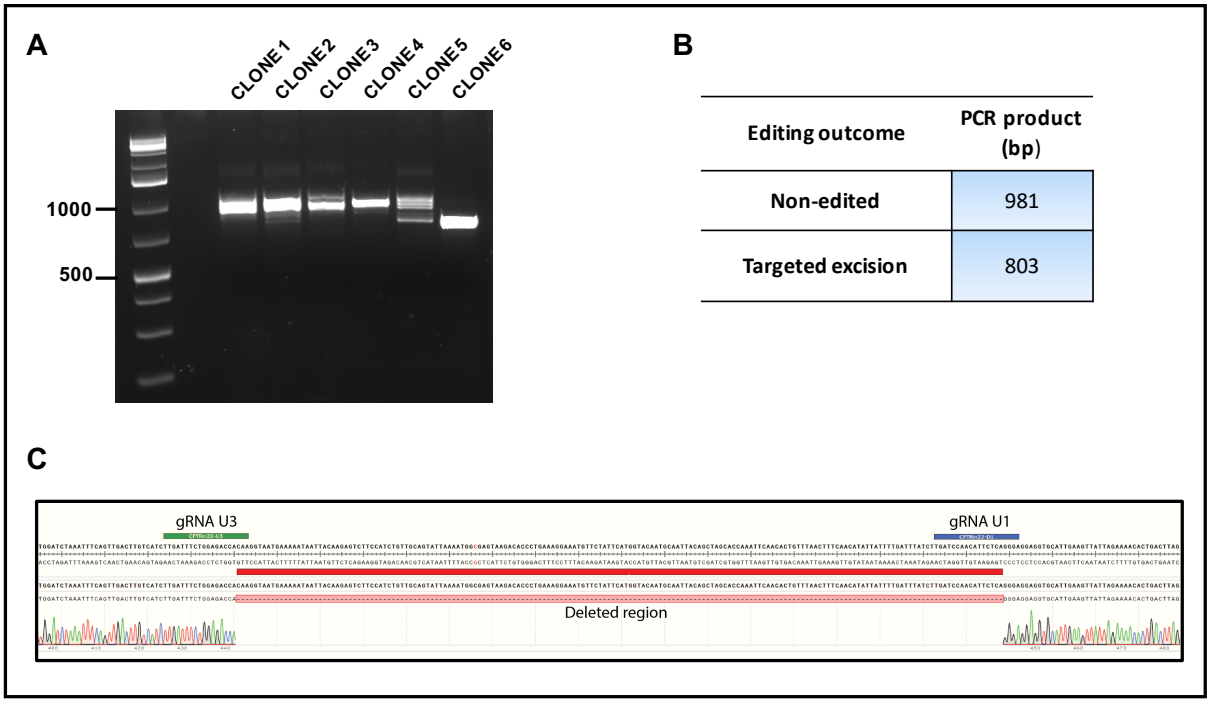


Figure 4.12. Single cell clone analysis of CFBE3849 cells *(**A**) Genomic DNA from single cell clones was extracted for amplification of region of interest before resolution by standard gel electrophoresis on a 1% agarose gel and visualisation under UV light. Lower molecular weight species represent products of targeted excision. Clone 6 shows a single band shift of predicted size, indicative of a homozygous edited cell population with a 178 bp deletion in both alleles. (**B**) Expected product sizes of non-edited vs. edited genomic DNA. (**C**) Sanger sequencing trace of CFBE single cell clone 6, with expected 178 bp deletion of genomic DNA between U3 and D1 cut sites, as compared to reference sequence. Deleted region highlighted in red.*

4.3.1.7 CFTR mRNA splicing analysis in CFBE3849 cells

In order to test if targeted excision of the 3849 +10kb C>T mutation could rescue splicing of CFTR, we first had to assess aberrant splicing in non-corrected cells. Studies in the literature show considerable variation in levels of aberrantly spliced mRNA, with levels varying between patients, and even in organs of the same patients [421]. Generally, there are low levels of a correct CFTR transcript [47].

CFBE3849 cells were expanded and seeded onto ALI filters, where they were differentiated into a pseudostratified monolayer. At week 3, RNA was extracted from the cells and reverse-transcribed using SuperScript™ Double-Stranded cDNA Synthesis Kit. CFTR mRNA was amplified by PCR using a FAM-labelled forward primer and unlabelled reverse primer.

PCR product was analysed by agarose gel electrophoresis (Figure 4.13A) before sending to Eurofins for fragment length analysis (Figure 4.13B). CFBE3849 cells were found to contain 25.2% wt CFTR transcript, with 74.9% being mutated. As expected, NHBE cells had 100% wt CFTR transcripts (Figure 4.13C).

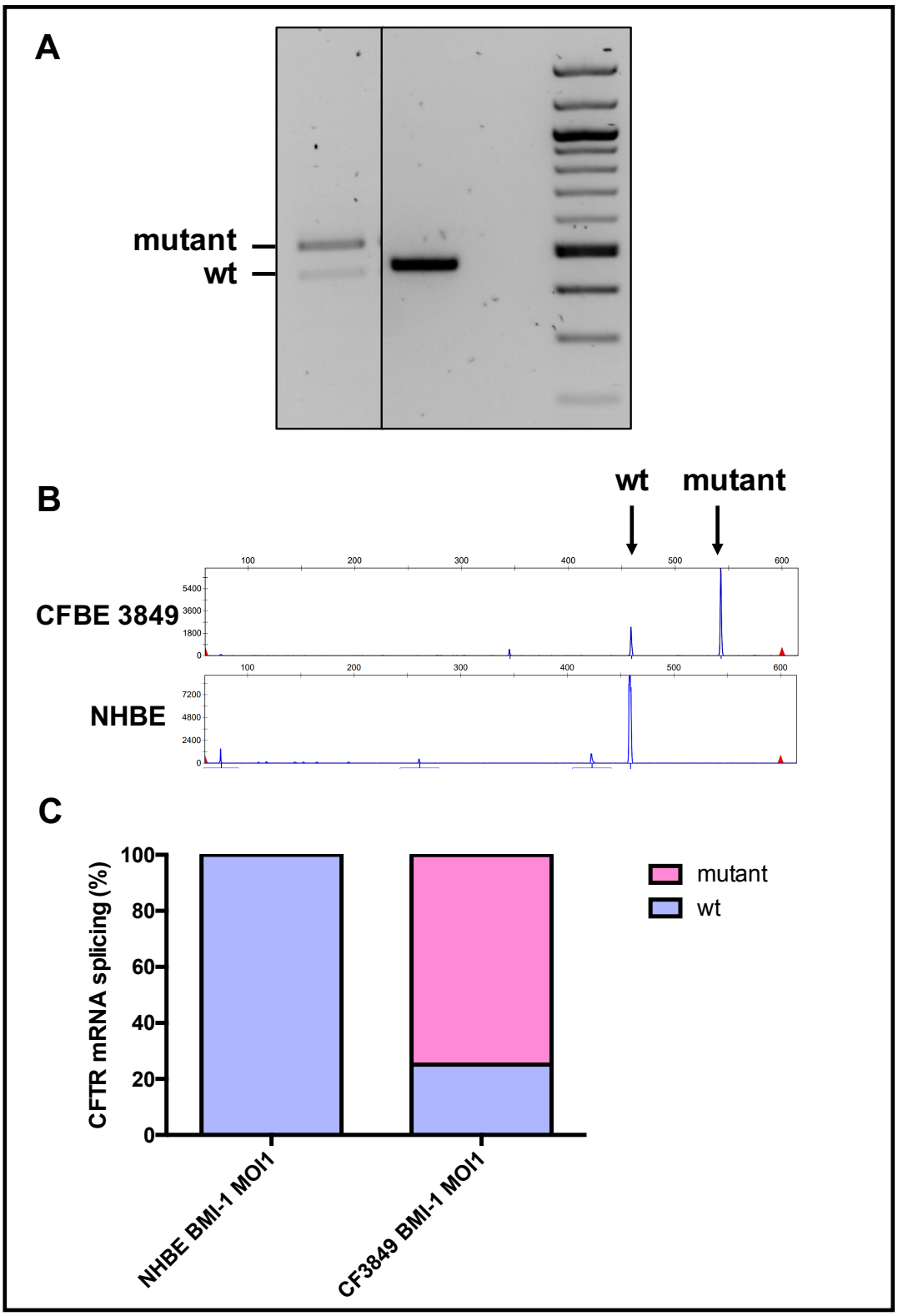


Figure 4.13. CFTR mRNA splicing analysis in CFBE3849 cells. *(**A**) Agarose gel electrophoresis analysis of CFBE 3849 vs wt NHBE cells. The top band in CFBE3849 cells indicates the mutant CFTR transcript, while the lower band is wt CFTR. (**B**) Electropherogram analysis of RT-PCR products of wt NHBE cells and CFBE3849 cells. Wildtype and mutant CFTR transcripts indicated with arrows. (**C**) Quantification of CFTR mRNA splicing in NHBE and CFBE3849 cells.*

4.3.1.8 CFTR mRNA splicing analysis in corrected CFBE3849 cells

The corrected CFBE3849 cells, found to have a cumulative targeted excision efficiency of 60.9%, were expanded and seeded onto ALI filters. At week 3, RNA was extracted from the cells and reverse-transcribed, before amplification by PCR using a FAM-labelled forward primer and unlabelled reverse primer. PCR product was analysed by agarose gel electrophoresis (Figure 4.14A) before fragment length analysis (Figure 4.14 B and C). Un-corrected CFBE3849 cells were found to contain 36.6% wt CFTR transcript, with 63.4% being mutated. Corrected cells had 93.5 % wt CFTR transcripts and only 6.5 % mutated transcript, suggesting that CRISPR/Cas9 mediated deletion of the deep-intronic splice mutation can restore CFTR splicing close to wt levels (Figure 4.14 C).

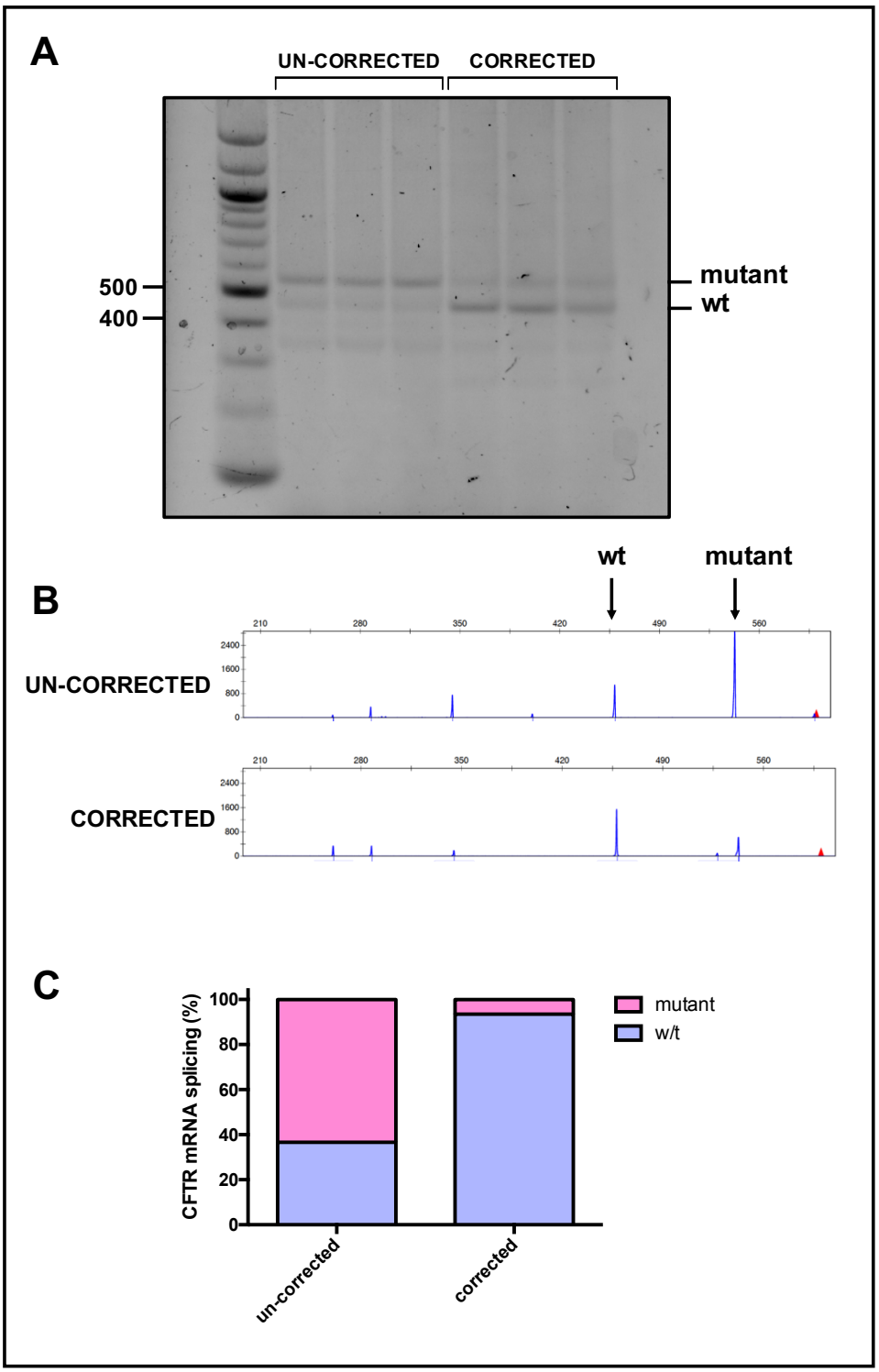


Figure 4.14. CFTR mRNA splicing analysis in corrected CFBE3849 cells. *(**A**) Agarose gel electrophoresis analysis of corrected and un-corrected CFBE3849 cells. The top band in CFBE3849 cells indicates the mutant CFTR transcript, while the lower band is wt CFTR. (**B**) Electropherogram analysis of RT-PCR products of corrected and un-corrected CFBE3849 cells.. Wildtype and mutant CFTR transcripts indicated with arrows. (**C**) Quantification of CFTR mRNA splicing.*

4.3.1.9 Electrophysiology studies in an Ussing chamber

We next examined whether targeted excision of the 3849 +10 kb C>T mutation would restore the activity of the CFTR chloride channel (Figure 4.15). Non-corrected and 4x corrected CFBE3849 cells were expanded and seeded onto ALI filters, where they were maintained for three weeks to allow monolayer formation and polarisation. Transepithelial electrical resistance (TEER) was measured in the cultures. TEER reflects the ionic conductance of the paracellular pathway in the epithelial monolayer, and is a non-invasive indicator of the integrity of the tight-junctions between cells forming the monolayer [426]. Filters with TEER values above 350 $\Omega.cm^2$ were taken forward for ion transport studies using an Ussing chamber. Cl^- concentration in the Ringer's solution buffer at both apical and basolateral sides are symmetric. 10 µM amiloride was added to the apical side of the filter to block the ENaC channel. Both non-corrected and corrected cells showed a decrease in current, confirming the expression of ENaC channels on the apical surface of the cell membrane. Cells were then treated with 25 µM Forskolin and 100 µM 3-isobutyl-1-methylxanthine (IBMX) at both apical and basolateral sides. As cAMP agonists, both agents activate CFTR and therefore stimulate transepithelial Cl^- and bicarbonate transport. Both non-corrected and corrected cells showed a small increase in current, indicative of the presence of CFTR (1.4 and 2.9 $\mu A/cm^2$ respectively), however, there was no statistically significant difference between the two (Figure 4.15 B). Once the CFTR-dependent current was stabilized, 10 µM CFTRinh-172 was added. The amplitude of the current decrease elicited by the inhibitor reflects the extent of CFTR function within the monolayer. There was a statistically significant difference in current decrease between non-corrected and corrected cells, with a decrease of -2.5 and -5.8 µA/cm2 respectively (Student's unpaired T-test, $p \leq 0.05$). This suggests functional restoration of the CFTR channel in 4x corrected CFBE3849 cells (Figure 4.15C). Finally, ATP was added to elicit anion secretion through Ca^{2+} dependent chloride channel, TMEM16A, causing a large spike in current, at 21.5 and 38.2 $\mu A/cm^2$ respectively.

Ussing chamber analysis was performed with the help of Max Woodall, St George's University, London.

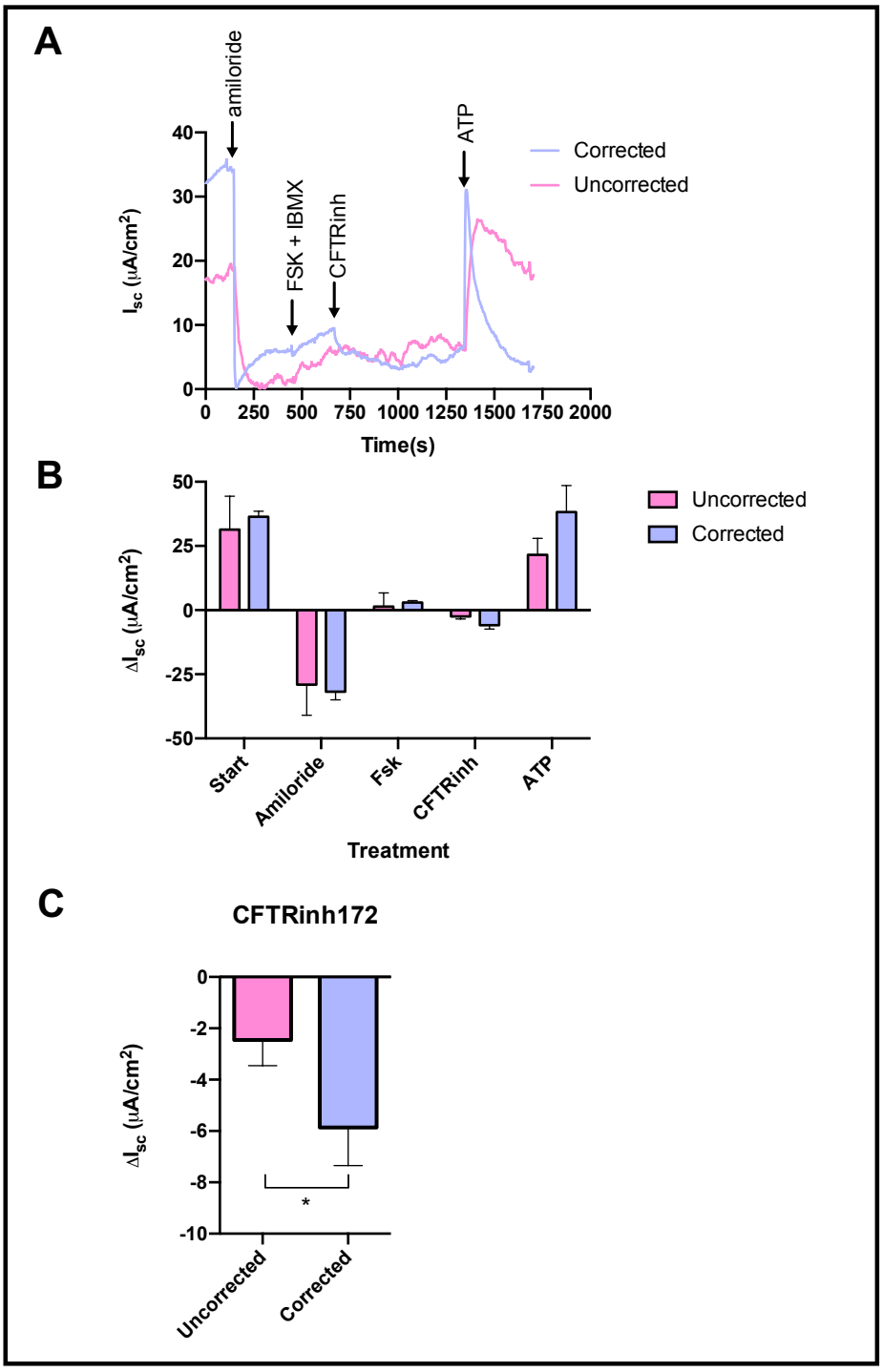


Figure 4.15. Electrophysiology studies in 4x correct CFBE3849 cells. *(**A**)The short-circuit current (Isc) traces of non-corrected and corrected CFBE3849 cells and their response to apically and basolaterally applied forskolin and IBMX, followed by CFTR inhibitor-172, recorded using the Ussing chamber. (**B**) Summary of short-circuit data from non-corrected and corrected CFBE3849 cells. (**C**) short-circuit current (Isc) data of cells after the addition of CFTRinh172. Results are presented as mean ± SD, n=3. Significance was calculated using the Student's T.Test- * P≤0.05.*

4.3.2 Investigating a NHEJ-mediated HITI editing strategy for targeted knock-in of GFP

The aim of the final part of this chapter was to develop an NHEJ-mediated HITI editing strategy for GFP knock-in in a HEK293T cell line. For a therapeutic strategy, we could insert a donor template into the endogenous CFTR locus, which could enable restoration of CFTR function close to physiological levels. In this way, researchers have previously inserted a partial cDNA encompassing exons 11-27 of wt CFTR after creating a DSB at exon 11 of the CFBE41o- cell line, restoring CFTR mRNA expressed from the endogenous CFTR promoter, ion transport as measured by Ussing chamber. However, the regulation of CFTR expression is not fully understood and this approach could affect positioning of certain enhancer elements in the CFTR gene [427]. Conversely, a safe harbour site can allow for strong transgene expression without disrupting nearby endogenous gene expression [428]. A well characterized safe harbour site called AAVS1, located on chromosome 19, is known as a site at which wild-type AAV has tendency to insert at and thus, due to the low pathology associated with AAV infection, it is thought to be a safe site [429]. The site lies between exon 1 and intron 1 of the PPP1R12C gene; the function of which is not fully known, however no adverse effects have been reported from its disruption as of yet, although this requires further investigation [430]. The aim of the following experiments is to integrate the GFP gene into the AAVS1 locus of HEK293T by designing a HITI construct that uses NHEJ-mediated repair pathway for targeted integration of a donor template. The cells will be co-transfected with an editing plasmid, containing Cas9 and a gRNA targeting the AAVS1 locus, as well as a plasmid donor containing the GFP gene without a promoter, and the gene will be flanked with the same gRNA sequences and PAMs, but in inverted direction. The template can be integrated in either a forward or reverse orientation, however reverse integration recreates an intact gRNA sequence which can then undergo additional cutting by Cas9 and so enrich for the forward integration [373] (Figure 4.16).

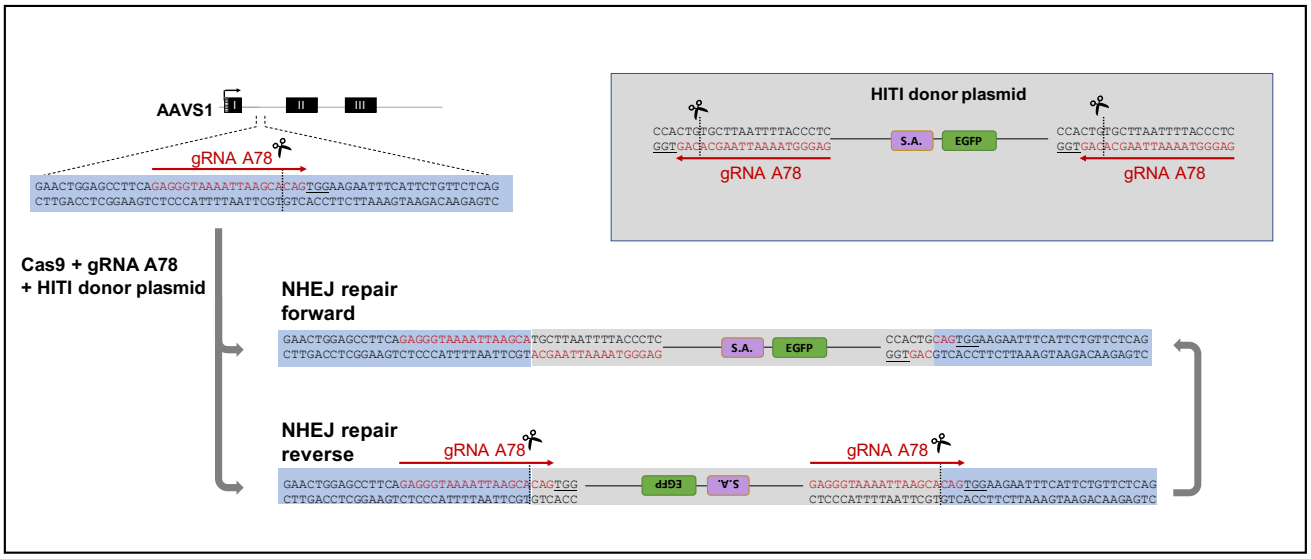


Figure 4.16. Homology-independent targeted integration (HITI) strategy. *(**A**) The donor template is designed to include a gRNA target and PAM site in the inverted orientation. Upon addition of Cas9 + gRNA, genomic DNA and donor templates are both cleaved. Integration of donor template in the forward direction results in loss of intact gRNA target site, preventing Cas9 from re-cutting. Should integration of template be in the reverse direction, target sites are reformed and can undergo additional cutting by Cas9, preferentially directing unidirectional knock-in. (**B**) HITI donor plasmid with inverted gRNA target and PAM sites.*

4.3.2.1 Construction of HITI donor plasmid and edting plasmid

To construct the donor plasmid, we designed an oligonucleotide containing the gRNA target sequence and PAM recognition site flanked by extra bases in order to proceed with cloning. The pMC.T7_SA-EGFP_T7 plasmid was then linearised using *Aat*II and *Pvu*I restriction enzymes, followed by a ligation, for insertion of inverted A78 target and PAM site at the 5' end, as described in Material and Methods. For insertion at 3' end, creating a donor template with two cut sites, the plasmid was the linearised with *Sac*I and *Psp*OMI restriction enzymes for insertion of inverted A78 and PAM (Figure 4.17).

Next, gRNA A78 was cloned into 65323_AAVS1 plasmid using *Avr*II and *Xba*I restriction enzymes. The plasmid contains a BFP fluorophore to assess transfection efficiency, as well as a chicken β hybrid (Cbh promoter) and Cas9. Original plasmids were constructed by Ileana Guerrini, a former lab member.

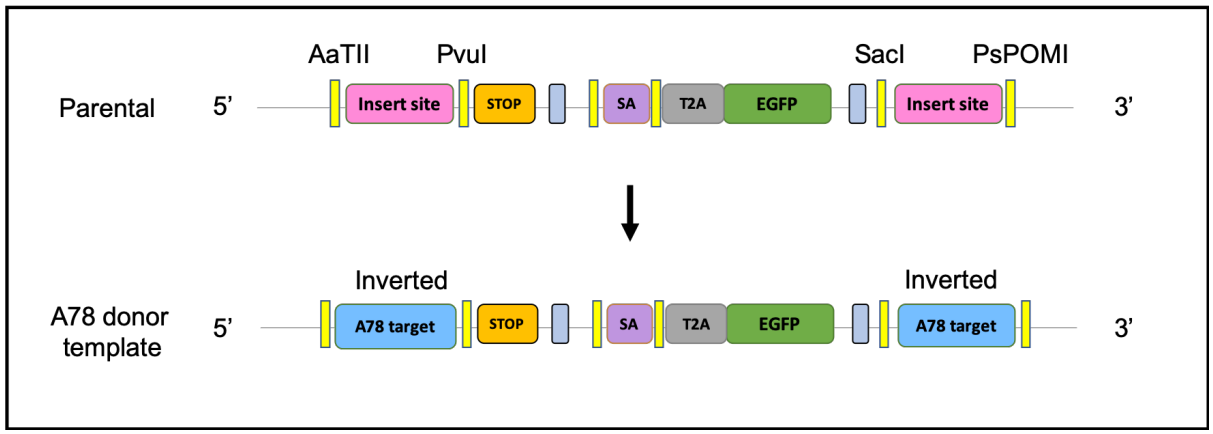


Figure 4.17. Construction GFP HITI donor template with A78 gRNA. *The original parental pMC.T7_SA-EGFP_T7 contained the GFP transgene flanked with two sites for cloning. Using AatII and PVuI restriction enzymes, the inverted gRNA A78 target and PAM sequence was cloned into the 5' end of the plasmid. The same sequences were then cloned into the 3' end, using SacI and PspOMI, creating a donor template with two cut sites.*

4.3.2.2 Transient transfection efficiency of RTNs containing donor and editing plasmids in HEK293T cells

HEK293T cells were transfected with RTNs comprised of C14, C16, C18 DOPE, along with Peptide Y. A 1:1 weight ratio of editing plasmid (~9.2 kb) and donor template (~5.3 kb) was used with 250 ng of each plasmid at a final weight ratio of 1:1:4 DNA: lipid: peptide. The primary outcome at 48 hours was percentage of cells simultaneously expressing BFP and GFP, indicative of successful delivery of editing plasmid and donor template respectively. C16 DOPE PY RTNs achieved the highest delivery efficiency of the two plasmids, with 41.7% of cells expressing both BFP and GFP, as measured by flow cytometry. All efficiencies were statistically significant from each other as well an un-transfected control, and donor only control ($P \leq 0.0001$, $P \leq 0.01$ between C14 and C16 DOPE, one-way ANOVA). C18 DOPE RTNs were least efficient, with 21.7% cells expressing both GFP and BFP (Figure 4.18).

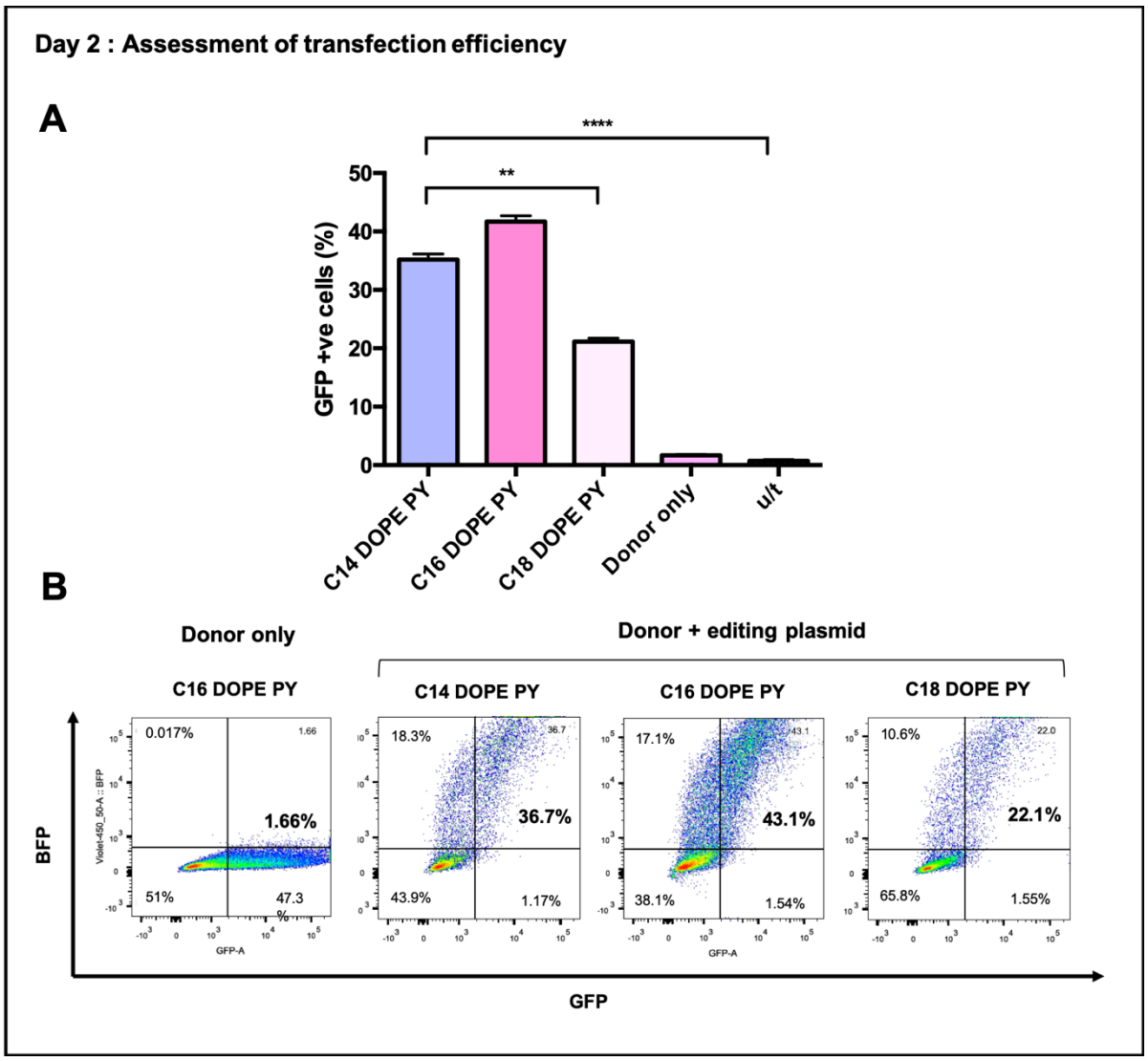


Figure 4.18. Transfection efficiency of editing plasmid + HITI donor template delivered by RTNs. *HEK293T cells were co-transfected with an editing plasmid, containing A78 gRNA and Cas9, as well as a BFP fluorophore, and the donor plasmid, containing the A78 target site, PAM and GFP transgene, using C14, C16 and C18 DOPE PY RTNs. (**A**) Cells were analysed by flow cytometry after 48 h to assess transfection efficiency of the RTNs. BFP + GFP +ve cells indicated successful integration of both plasmids to the cells. (**B**) Representative density plots of cells transfected with donor plasmid only, or donor + editing plasmid using different RTNs Results are presented as mean ± SD, n=3. Significance was calculated using One-way ANOVA with multiple comparisons ** P≤0.01 **** P≤0.0001.*

4.3.2.3 *Efficiency of HITI-mediated GFP integration in HEK293T cells 12 days post-transfection*

Simultaneous expression of BFP and GFP indicates the cells that have potential for HITI to have occurred. As C16 DOPE PY RTNs gave the highest transfection efficiencies, cells transfected with these particles were expanded over a 12 day period before flow cytometry analysis of GFP expression, indicative of permanent integration of GFP plasmid (Figure 4.19). We found that 14.3% of cells showed GFP expression, as compared to 2.14% of cells transfected only with the donor template (*P≤0.0001).*

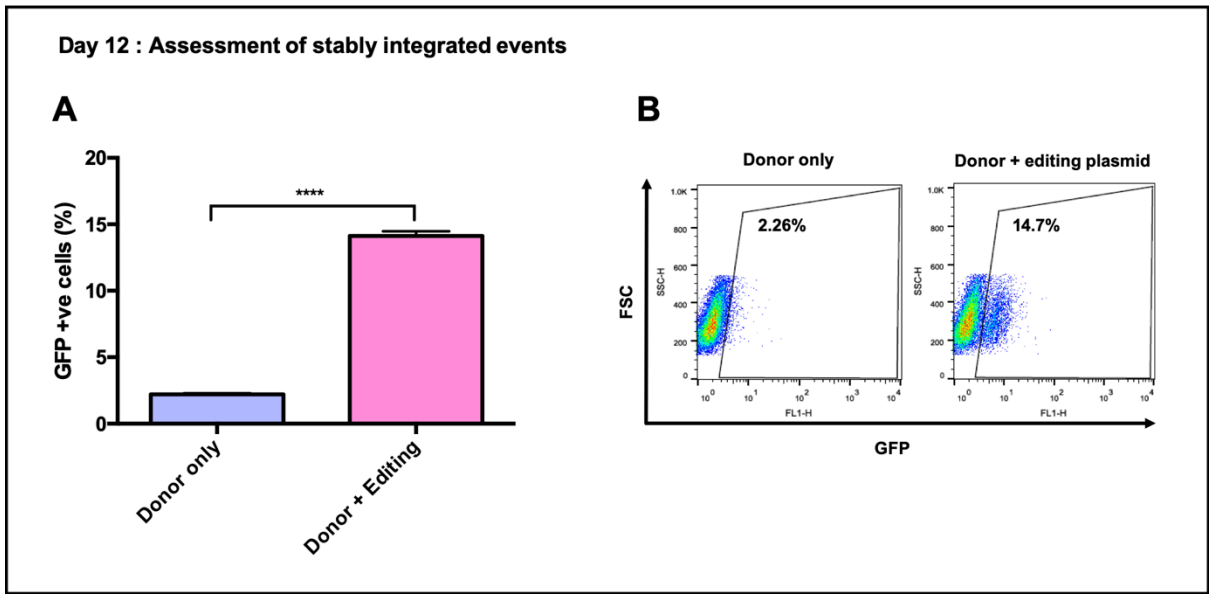


Figure 4.19. Stable integration of GFP*. HEK293T cells were co-transfected with a HITI editing plasmid, containing A78 gRNA and Cas9, as well as a BFP fluorophore, and the donor plasmid, containing the A78 target site, PAM and GFP transgene, using C16DOPE PY RTNs. (**A**) Cells were analysed by flow cytometry 12 days post transfection to assess stable integration of GFP. (**B**) Representative density plots of cells transfected with donor plasmid only, or donor + editing plasmid. Results are presented as mean ± SD, n=3. Significance was calculated using the Student's T.Test- **** P≤0.0001.*

4.3.2.4 *Molecular analysis of HITI-mediated integration events via junction PCR*

In order to molecularly assess HITI integration in HEK293T cells, two PCRs were carried out. First, a genomic PCR was performed to amplify a region between introns I and II at the AAVS1 site (Figure 4.20 A). This was predicted to give an amplicon of 630 bp without the GFP insert.

Secondly, a 5' junction PCR was carried out which used the same forward primer targeted to a sequence after intron I, and a reverse primer targeting the beginning of the transgene GFP sequence within the donor template, giving a predicted product of 250 bp (Figure 4.20 C). Therefore, this latter PCR product would only appear if the GFP sequence had integrated correctly at the AAVS1 site.

As expected, all conditions produced a 630bp,indicative of the unedited alleles in the whole cell population (Figure 4.20 B). When the Cas9 editing plasmid and donor plasmid were co-transfected, a 250 bp band was visible on the gel, indicative of successful GFP construct integration within the genome (Figure 4.20 D). This band was not present in un-transfected cells, or cells that received only the donor template.

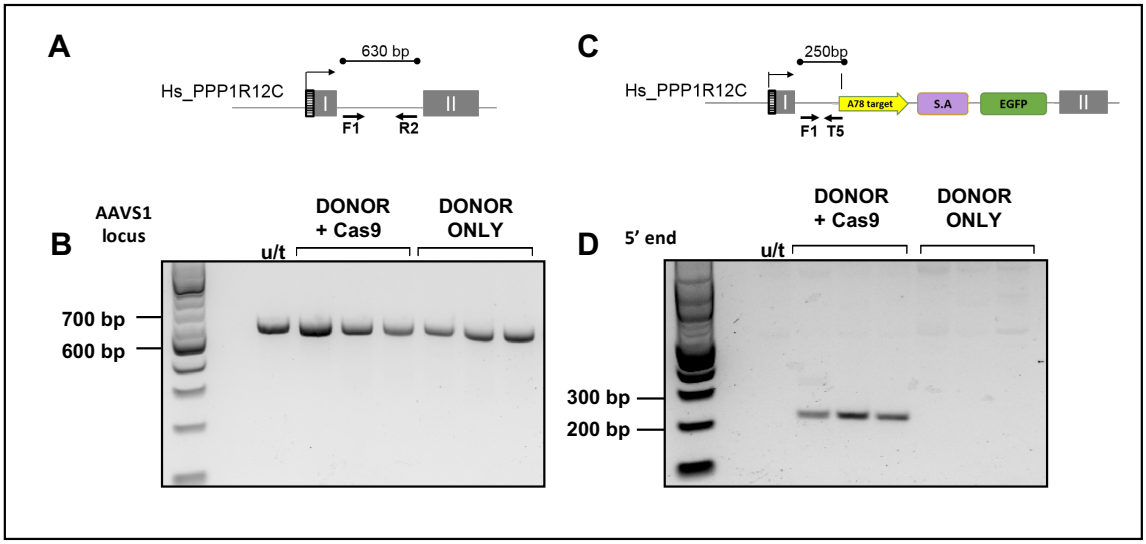


Figure 4.20. *(**A**) Schematic of PCR using genome-specific forward and reverse primers, where forward primer is located and the end of exon I and reverse primer is located at the beginning of exon II. Expected PCR product size is 630 bp. (**B**) Genomic DNA was extracted for amplification of region of interest before resolution by standard gel electrophoresis on a 1% agarose gel and visualisation under UV light. 630 bp product is seen for all conditions, indicating no transgene integration. (**C**) Schematic of 5' junction PCR to confirm plasmid insertion into genome. Forward primer is genome-specific, located at the end of exon I, while reverse primer is insert-specific, located at the beginning of the donor template. Expected PCR product size is 250bp. (**D**) As expected, 250 bp product only visible in samples that received donor template, indicating successful transgene integration. Molecular weight referenced by 100 bp DNA ladder.*

4.3.2.5 *Single cell clone molecular analysis of HITI-mediated integration events via junction PCR*

To further assess molecular integration events, transfected cells were FACs sorted for GFP+ve cells, and single cells were seeded into 192 wells in two 96-well plates by serial dilution. 15 days later, plates were checked for successful growth of clones, and 22 clones that showed the most successful proliferation were chosen for expansion, along with two un-transfected controls. Genomic DNA was isolated from the 22 clones, followed by PCR amplification of the AAVS1 locus and the 5'-junction between genome and template, as described in 4.3.10.4, to assess the integration events of the HITI template. PCR products were then visualised on a 1% agarose gel. Of the 22 clones, 15 produced a 250 bp band when the junction PCR was performed, suggesting successful integration of the donor template in the correct orientation in 68 % of cells. Of these, 9 clones also showed a 630 bp product in the genomic PCR, indicating that 60% of successfully edited clones had monoallelic correction, and 40% had biallelic insertion of template (Figure 4.21).

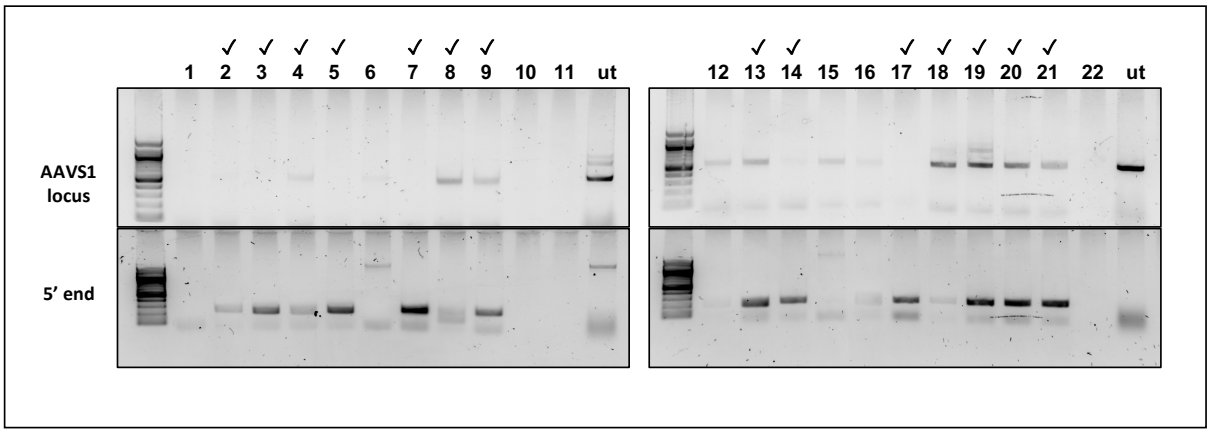


Figure 4.21. Single cell clone analysis of GFP+ve cells after HITI integration*. GFP +ve cells were cell sorted by FACs and seeded for single cell cloning by serial dilution. Genomic DNA was isolated from successful clones for genome-specific and 5' junction PCRs to assess integration events. Of 22 clones, 14 had template integration in the correct orientation, with the expected 250 bp product present in the 5' junction PCR. Molecular weight referenced by 100 bp DNA ladder.*

4.3 Discussion

The aim of this chapter was to use a pair of gRNAs to make two DSBs upstream and downstream of the 3849 + 10kb C>T mutation in cells isolated from CF patients homozygous for this mutation. In this way, we could excise the deep-intronic splice mutation and restore CFTR expression. Molecular correction of the cryptic splice site was confirmed by CFTR mRNA fragment length analysis and functional validation of CFTR chloride channel activity measured by Ussing chamber.

Prediction of the cutting efficiency of gRNAs before transfection can be difficult, even with the use of specially designed software such as Benchling or Synthego's algorithms. We therefore wanted to validate cutting efficiency of our chosen gRNAs in HEK293T cells, an easy to culture and transfect cell line, before moving into a more clinically relevant CF cell model. The deep-intronic location of the 3849 + 10 kb C>T mutation, which lies outside of classical canonical splicing signals, essentially meant there were no restrictions when selecting gRNAs upstream and downstream of the mutation. We therefore chose the same 3 gRNAs reported previously [423] and confirmed that all 3 gRNAs successfully induced a DSB within the DNA, with an average efficiency of 33%, as assessed by a T7EI endonuclease assay. Co-transfection of gRNA pairs to create a targeted deletion within the genome was also performed in HEK293T cells, achieving deletion efficiencies of ~14%. Sanger sequencing confirmed a heterogenous base population around gRNA cut sites indicative of DSB and NHEJ-mediate repair.

The editing strategy was then carried forward into a more clinically relevant cell type-*BMI-1* modified lung epithelial cells isolated from a CF patient homozygous for the 3849 + 10kb C>T mutation. RTNs were used for delivery of the RNP complexes, shown in chapter 3 to achieve a greater delivery efficiency over commercially available reagents. After two repeat transfections, a maximum cutting efficiency of 54% was achieved by gRNA U3, as measured by the T7EI assay. Inherently, densitometry analysis is prone to subjective bias due to manual quantification. However, a recent study compared the T7EI endonuclease and TIDE assays, as well as next-generation sequencing (NGS) as the gold standard for detecting the indels events. They found that the T7EI assay generally under-reported the number of indels because it depended on the diversity of the heteroduplex formation. For example, a specific gRNA showed an indel rate of 96% as measured by NGS, however a T7EI assay gave only a 10% indication of indels; this is because the majority of indels were a 1 bp deletion [431]. The

paper suggested that the TIDE assay was the most cost-effective technique, giving similar results to NGS data, and so in subsequent experiments we mainly used the TIDE assay, or the similar ICE software, which confers the benefit of detecting the larger deletions expected in this particular strategy. Here, the TIDE assay indicated a cutting efficiency of 8% vs. 20% with the T7EI assay for gRNA U1, while results for U3 and U1 were largely similar (60% vs. 54%, and 37.7% vs. 45% respectively). However, analysis by TIDE assay allows us to analyse the indel spectrum of each of the gRNAs. As no significant indels were found in gRNA U1 treated cells, U3 and D1 guide pairs were used in all subsequent transfections.

Co-transfection of U3 and U1 achieved a targeted excision rate of ~9% in CFBE3849 cells after one round of transfection. However, through the use of RTNs as a delivery vector, with less associated toxicity, we were able implement a repeat dosing strategy, where cells received multiple doses of the RNP complexes, resulting in a targeted excision rate of 60.9%, and an indel efficiency rate of 83%. For comparison, CFTR correction of as little as 5-10% of cells in the airway epithelium *in vitro* was sufficient in restoring the electrical properties of the airway epithelium [149], although restoration of mucus transport and airway surface liquid height required as much as 25% correction [150]. Studies in ALI culture have shown that epithelia generated with 20% wt cells and 80% CF cells can generate upwards of 70% of the transepithelial Cl^- current of a fully wt epithelium [149].

In order to confirm that successful excision of the 3849 +10kb C>T mutation rescued splicing of CFTR, fragment length analysis was performed on non-corrected and corrected cells, and we found that 96% of wt CFTR transcripts were observed in corrected cells, as compared to 35% in non-corrected cells.

Finally, to define functional relevance of the successful deleted mutation, electrophysiology experiments were carried out using an Ussing chamber. Encouragingly, we found a significant difference in current decrease upon addition of CFTRinh-172 between non-corrected and corrected cells, confirming functional restoration of the CFTR channel in 4x corrected CFBE3849 cells.

In contrast to classical gene editing strategies, which generally involve the addition of template DNA by means of HDR, our editing strategy exploits the NHEJ-mediated repair pathway. The HDR pathway is not highly active in post-mitotic cells such as the ciliated airway epithelium, and so an NHEJ-bases strategy is likely to have much higher efficiency. Additionally, delivery of the CRISPR/Cas9 platform is much simpler without the requirement

of a DNA donor template. Recently, another group developed an alternative editing strategy to correct the 3849 + 10 kb C>T mutation, and an additional splice mutation, 3272–26 A>G [432]. The researchers utilised Cas12a and a single gRNA and were able to discriminate between wt and mutant sequences. Their approach is based on small deletions within the intronic sequence that remove essential splicing regulatory elements. They verified their strategy in intestinal organoids, showing functional recovery of the CFTR channel. This research demonstrates the increasing application of NHEJ-mediated repair for editing strategies in addition to creating knockout models.

It is of course necessary to test this approach in an *in vivo* CF model, however this is not currently possible due to CFTR sequence differences between human and mouse. One option would be to redesign gRNAs to target CFTR in mice, although it may be more informative to study humanised animal models of CF, several of which are in development [433]. Currently, however, the issue remains that despite high similarity between human and murine CFTR sequence, disease pathophysiology differs between the two species, and lung disease is not evident in mice [112]. Therefore, whilst it would be possible to show proof-of-concept of targeted excision within a mouse lung, it is unlikely to have an effect on CFTR function. Indeed, a proof-of-principle study using gRNA pairs to excise a region of DNA *in vivo* is performed in chapter 5 of this thesis.

Whilst CFTR modulators such as the potentiator Ivacaftor are now routinely used in clinic for patients harbouring splice mutations, the drugs are associated with side effects, and their efficacy strongly depends on the residual correct splicing of CFTR, which varies significantly between patients [47, 421, 434]. There is therefore an urgent, unmet need for permanent correction of CFTR, the basis of which is examined in this chapter.

In addition to the 3849 + 10 kb C>T mutation, there are several other CF-causing mutations to which this editing strategy could be applied. C.1811 + 1.6kb A>G creates a very strong splice donor site in intron 12 of CFTR, with 99% of transcripts containing a 49 bp pseudoexon, containing an in-frame TAA stop codon causing premature termination of the CFTR protein [435]. mRNA analysis showed that 1811 + 1.6kb A>G mRNA was 5-10-fold less abundant than F508del mRNA, and so is classified as a severe CF-causing mutation [435]. Another mutation, 3272-26 A>G creates a splice acceptor site 26 bp before exon 20, extending the exon by 25 bp. The resultant frameshift leads to premature termination of CFTR at a TGA stop codon [436]. Low levels of normal CFTR transcript explain the milder CF phenotype when

compared to 1811 + 1.6kb A>G. Additionally c.1584+18672A>G [437], c.1680-877G>T [438] and c3717+40A>G, which extends existing exons [438], could also be amenable to a targeted excision strategy. Together, these six mutations represent ~1.6% of individuals with CF. The high efficiency and simplicity of the approach, namely that a donor template is not required, could form the basis of a potential therapeutic intervention. Likewise, this approach may be applicable in more than 75 other genetic disorders where deep-intronic mutations have been identified as disease causing [439], including monogenic diseases such as β-Thalassemia [440], Leber congenital amaurosis 10 [441] and Amyotrophic lateral sclerosis 1 [442]. Moreover, diseases that affect vision, such as gyrate atrophy of choroid and retina (GA) retinitis pigmentosa 11[442], could be particularly responsive to a targeted excision strategy delivered by RTNs due to the relative accessibility of the eye and non-necessity of systemic delivery.

While the correction strategy we have described is applicable to a small cohort of CF patients and several other disorders, therapies for a larger number of CF patients will require the delivery of a donor template. As discussed, NHEJ is the predominant DSB repair pathway in terminally differentiated cells, such as those residing in the lung. Until recently, insertion of a donor DNA template was thought to require the HDR repair pathway, however, in 2017 the homology-independent targeting integration (HITI) method was reported, allowing for targeted integration using the NHEJ-mediated repair pathway, without the use of homology arms [373]. HITI enabled ~60% correction efficiency in transfected neurons *in vitro* and an efficiency of 3.5% in adult mice brain cells *in vivo* and up to 10% in younger mice. Here, we investigated the use of HITI for insertion on the GFP transgene into the AAVS1 locus of HEK293T cells, and were able to achieve 14.3% targeted transgene integration. In future work, it will be necessary to sequence the single cells clones to molecularly assess the integration site for indels within the genomic DNA and construct, as well as investigating the frequency of reverse integration events.

We have shown previously in the chapter that multiple transfection dosages is an efficient technique to maximise gene editing; this will be useful *in vivo* to increase gene editing efficiency. We have not investigated multiple doses of the HITI constructs described here, however it is likely that integration efficiency would have increased further upon repeated transfections.

The donor plasmid for the HITI technique could be optimised by using a minicircle plasmid, a supercoiled DNA molecule that lacks all the prokaryotic parts required for replication, making it both smaller in size and less inflammatory, as the bacterial elements, rich in unmethylated CpG-repeats, are removed [443]. In the HITI paper previously described, the researchers use a minicircle plasmid for targeted insertion, and show superiority over a conventional plasmid (58% vs. 43.5% respectively). Interestingly however, they found the GFP transgene expression delivered by a minicircle as template dropped by 25 % 85 days post-transfection, showing that despite increased efficiency, transcriptional silencing still occurs *[444]*. This result shows that, despite the efficiency of minicircle for expression with devoid CpG-motifs, they do not always prevent transcriptional silencing, in agreement with the literature [445, 446]. Our lab has investigated the use of minicircles as compared to plasmid for transfections in vitro and in vivo, in terms of improving the gene expression as well as reducing the inflammatory response. Transfection efficiency the luciferase gene in an epithelial cell line (16HBE14o-) using RTNs was 10-fold greater using minicircle as compared to plasmid when equivalent dose was used, while in the mouse lung, expression was 4-fold greater *[447]*.

Mini-intronic Plasmid (MIP) is another plasmid propagation process in which the antibiotic resistance gene is removed from the plasmid along with un-essential prokaryotic parts, however essential bacterial elements for plasmid replication and selection are placed within an engineered intron contained within the eukaryotic expression cassette. The MIP is able to overcome transgene silencing and shows improved transgene expression, between 2 to 10-fold in comparison with the minicircle in both in vitro and in vivo transfection *[448, 449]*. In future work, involving the replacement of a larger part of the CF-gene, or indeed a 'superexon' for delivery of CFTR cDNA into the AAVS1 locus, minicircle or MIPs will be employed. We will also move this strategy into a more clinically relevant cell model, CFBE BMI-1 cells, and assess the HITI efficiency using Cas9 gRNA in an RNP format.

In our experiments, we have chosen to use a 'gene trap' system, inserting a vector containing a transgene that is only expressed upon integration in a functional gene. This is achieved by means of a promoter-less construct that includes a splice acceptor site upstream of transgene, meaning that expression of the GFP transgene is under the endogenous AAVS1 promoter. Our lab has also investigated the use of HITI constructs that include a promoter for insertion of a GFP construct, however lower GFP expression was achieves (work not

published). A side by side comparison is necessary to determine which strategy is most appropriate for template insertion, particularly for a larger construct including the CFTR cDNA.

In summary, we have investigated two separate NHEJ-mediated gene editing strategies. We have shown that CRISPR Cas9/gRNA NHEJ-mediated strategy using pairs of gRNAs can successfully and precisely excise a small region of the CFTR intron 22 containing a disease causing mutation. Targeted excision occurred in ~60% of *BMI-1* modified primary cells, a level of editing which is over 20-fold higher than previous studies of HDR-based editing with Cas9/gRNA in the same locus in a HEK293T minigene assay [450]. Moreover, this rate of correction restored close to normal CFTR splicing patterns, assessed by fragment length analysis of the CFTR transcript, and CFTR ion channel function as measured by Ussing chamber. We have also demonstrated the potential of a HITI-based knock-in strategy as an alternative to the cell-cycle dependent HDR pathway. This provides proof-of-concept for replacing a larger part of the CFTR gene or insertion of CFTR cDNA as a method of correcting an unlimited number of CF-causing mutations.

Chapter 5
Results

In vivo delivery of CRISPR/Cas9 using receptor-targeted nanocomplexes

5.1 Introduction

The work carried out in chapter 4 of this thesis showed functional correction of the 3849+ 10kb C>T mutation and described 5 other CF causing mutations to which the editing strategy could be applied, representing ~1.6% of individuals with CF. Owing to its deep intronic location, pairs of gRNAs could be designed to create two DSBs upstream and downstream of the 3849 + 10kb C>T mutation in cells isolated from CF patients homozygous for this mutation, excising the mutation and restoring normal splicing of CFTR. To confirm functional correction of CFTR, chloride channel activity was assessed by short circuit current measurements in an Ussing chamber.

In order to advance this editing strategy towards the clinic, it is necessary to test this approach in an in vivo CF model. Since the discovery of the CF gene in 1989, several different murine models have been developed, including knockout mice, and mice with specific mutations such as F508del and G551D [105-111]. However, despite the high similarity of CFTR sequences between humans and mice (78% identical sequence), the pathophysiology of CF in mice differs to that in humans, with no evidence of lung disease [112]. Mice over-expressing the β unit of ENaC present with thickened mucus, poor airway clearance and bronchiectasis and so are often used as a surrogate model for CF [114, 115]. However, this model would be unsuitable for our correction strategy as their murine cftr is normal. Several humanised mouse models of CF , expressing human CFTR with known CF-causing mutations, are currently in development but are not yet available for use, and again, are unlikely to develop lung disease [433].

Transgenic reporter mice are useful for assessing delivery efficiency of CRISPR in different formats and with different kinds of nanoparticle formulations. The Ai9 mouse reporter model contains the tdTomato gene (red fluorescent protein variant) in an inactive state, where transcription is prevented by a STOP cassette located upstream of the gene, located in the Gt(ROSA)26Sor locus [451]. The STOP cassette is flanked by two loxP sequences and so, upon addition of Cre recombinase, the cassette can be removed, and tdTomato fluorescence can be driven by a CAG promoter. We hypothesise that co-delivery of two gRNAs, targeting upstream and downstream of the STOP cassette along with Cas9 protein would also result in release of the cassette, allowing for CAG promoter-driven tdTomato fluorescence (Figure 5.1). In this way, we can provide proof-of concept for our editing

strategy and investigate the delivery efficiency of our RNP RTNs for in vivo lung delivery for the first time.

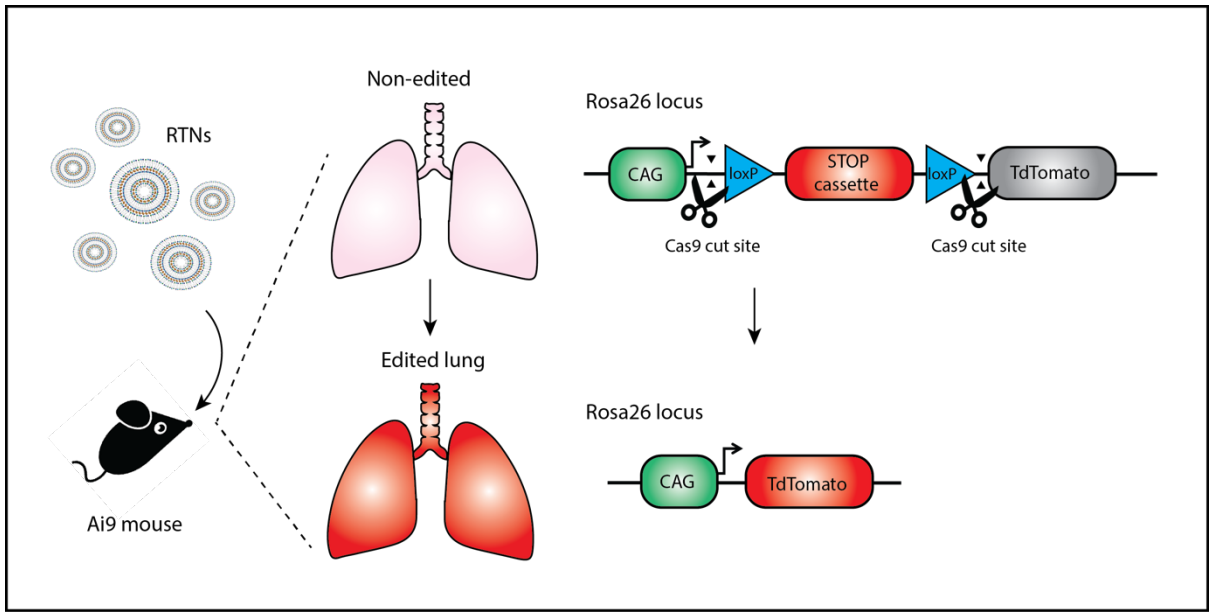


Figure 5.1 Ai9 Cre reporter system. *Ai9 mice have a loxP-flanked STOP cassette preventing transcription of a CAG promoter-driven tdTomato, all inserted into the Gt(ROSA)26Sor locus. RTN delivery of 2 gRNAs targeting upstream and downstream of the STOP cassette as a pre-formed RNP complex will result in release of the cassette, allowing for tdTomato fluorescence.*

5.2 Aims

In this chapter, we aim to:

1. Assess RTN delivery capabilities *in vivo* using Luciferase mRNA.

2. Deliver Cre mRNA or Cas9 RNP using RTNs in Ai9 mice and assess successful Cas9-mediated DSBs and therefore tdTomato+ve cells via flow cytometry and immunohistochemistry.

3. Histologically analyse Cre mRNA or Cas9 RNP treated Ai9 mice for signs of toxicity associated with RTN delivery.

5.3 Results

5.3.1 RTNs for delivery of Luciferase mRNA to the murine lung

To confirm oropharyngeal delivery efficiency of our RTNs, we first assessed their ability to delivery Luciferase mRNA (*Luc* mRNA) to the lungs of 12 week old C56B/6J mice. 15 ug *Luc* mRNA was delivered using the RTN formulation C18 DOPE, with a weight ratio of 1:3:4 mRNA: L: P. A luciferase assay on the lung extracts 24 h post instillation showed a significantly higher relative fluorescence unit (RFU) as compared to an untreated control ($p \leq 0.05$), with an average value of 5000 RLU/mg protein. Two of the mice had considerably higher expression that the third mouse (5937 and 5232 vs 3838 RLU/mg protein), suggestion the instillation was less successful in one mouse (Figure 5.2). DLS data for the *Luc* mRNA RTNs is included to assess the biophysical characteristics of the articles (Table 5.1).

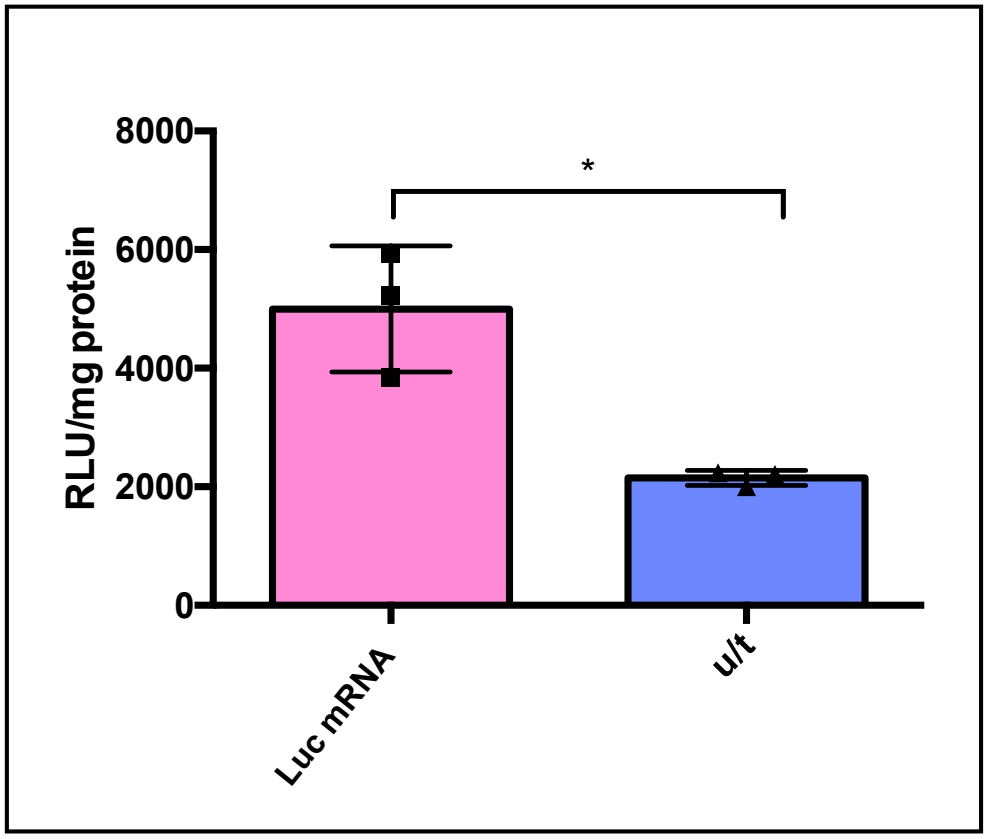


Figure 5.2. Luciferase transgene expression following in vivo mRNA transfections. *Luciferase activity following oropharyngeal instillation of C56BL/6N mice with RTN complexes containing Luc mRNA. Lung tissue was harvested 24 h post instillation. N=3 in each group. Significance assessed using One-way ANOVA. u/t= un-transfected.*

Table 5.1. Biophysical characterisation of nanoparticles used for in vivo delivery of Luciferase mRNA . *Table shows hydrodynamic size (nm), polydispersity index (PDI) and charge (mV), as measured by dynamic light scattering (DLS) using a Malvern Nano Zetasizer (Malvern Instruments, England). Three repeat measurements were made and standard deviation calculated.*

Nanocomplex	Z-Ave (d.nm)	PDI	ZP (mv)
Luc mRNA	**166±0.6**	**0.17**	**38.8±0.8**

5.3.2 RTNs delivery for Cre mRNA or Cas9 RNP in the Ai9 mouse

5.3.2.1 gRNAs targeting the STOP cassette in Ai9 mice

Having validated RTNs for *in vivo* oropharyngeal delivery of *Luc mRNA*, we moved to the Ai9 mouse model for Cas9 RNP delivery. Two previously validated gRNAs were chosen to create DSB within the Gt(ROSA)26Sor locus, targeting upstream and downstream of the STOP cassette which prevents transcription of the tdTomato transgene (Figure 5.3) (Table 3.2).

Figure 5.3. *Lox-P flanked STOP cassette preventing transcription of a CAG promoter-driven red fluorescent protein variant (tdTomato) in Ai9 mice. gRNA sequences target each end of the cassette in order to remove it, allowing tdTomato expression. PAM sequences highlighted in red.*

Table 5.2. List of gRNAs targeting Ai9 stop cassette*. The table shows gRNA sequence, strand targeted, the PAM motif sequence, the predicted efficiency score (on-target score) and specificity score (off-target score). The on-target and off-target scores were calculated by using the Benchling programme (https://benchling.com/crispr)*

gRNA	Sense	Sequence	PAM	Efficiency score	Specificity score
T1	+	AAAGAATTGATTTGATACCG	CGG	78.84	18.37
T4	-	CTAGCTTGGGCTGCAGGTCG	AGG	75.26	1.43

5.3.2.2 Delivery of Cas9 RNPs and Cre mRNA via RTNs

In Chapter 4, we achieved indel effieciency rates upwards of 80% following 4 sequential transfections. We therefore chose to use the same RTN formulation for *in vivo* delivery, namely C16DOPE Peptide Y, at a weight ratio of 1:1:4 RNP: L: P. As a positive control, we used Cre recombinase mRNA, delivered by RTN formulations fould optimal for mRNA deliery, C18DOPE Peptide Y, at a weight ratio of 1:3:4 RNA: L: P.

The *in vivo* experiment included 15 mice (6 weeks of age) divided into 3 groups: 1) control, where no RTNs were administered (*n=5),* 2) Cre recombinase group, where 15 μg Cre mRNA was formulated as described above *(n=5)* and, 3) RNP group, where 15 μg of RNP (12 μg Cas9 protein, and 1.5 μg gRNA T1 and T4) was administered *(n=5).* Each mouse was given 3 sequential treatments, with a dosing regimen of every 48 hours. On day 7, 24 h post final dosing, lung tissue was harvested.

5.3.2.3 Weight of treated mice over time-course of experiment

It is important to moniter the weight of mice, as any loss of weight can indicate adverse effects of the treatment. As per the Animals Scientific Procedures Act 1986, mice that are approaching 20 % weight loss should be humanely killed. All mice maintained their weight over the 4 measurements. One female mouse treated with Cas9 RNP had decreased in weight on day 7, however it was still within 20% of the starting weight.

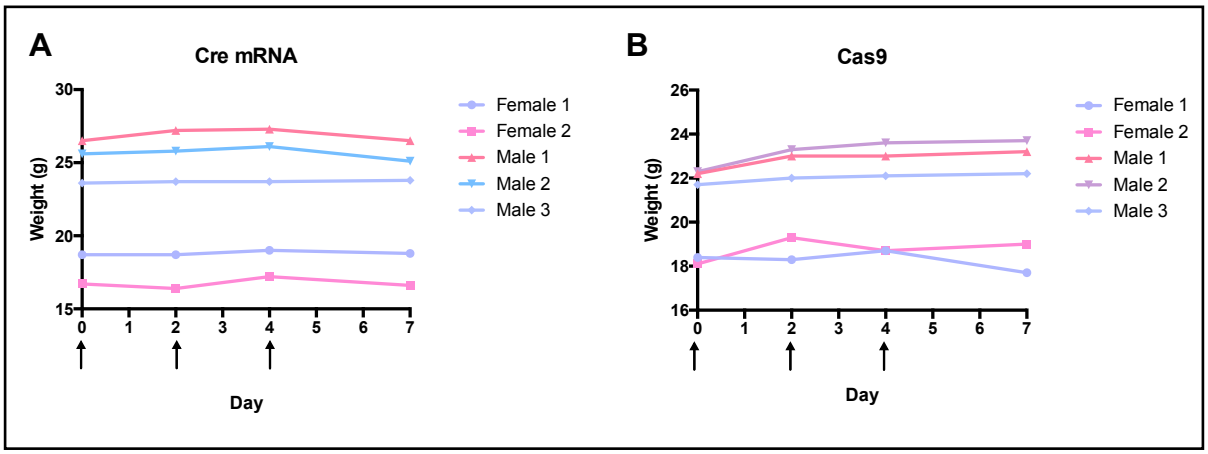


Figure 5.4. *Change in body weight of mice after 3 oropharyngeal instillations of (**A**) Cre mRNA and (**B**) Cas9 RNPs via RTNS, at a weight ratio of 1:3:4 mRNA: L: P and 1:1:4 RNP: L: P. Arrows indicate dosing days. Lung tissue was isolated on day 7 for flow cytometry and IHC analysis.*

5.3.2.4 Flow cytometry analysis to assess tdTomato+ve cells

Lung tissue was harvested 3 days after final administration, and processed for Flow cytometry analysis. Cells were treated with DNAse/collagenase D in order to obtain a single cell suspension, before analysis by flow cytometry. Gates were created to isolate healthy, single cells (Figure 5.5 A and B) before gating tdTomato +ve cells in the whole lung cell population (Figure 5.5 C). We next gated for epithelial cells by isolating cells that were negative for CD45 (a leukocyte marker) and positive for CD326 (Ep-CAM, Epithelial cell adhesion molecule) (Figure 5.5 D) and finally tdTomato +ve epithelial cells (Figure 5.5 E).

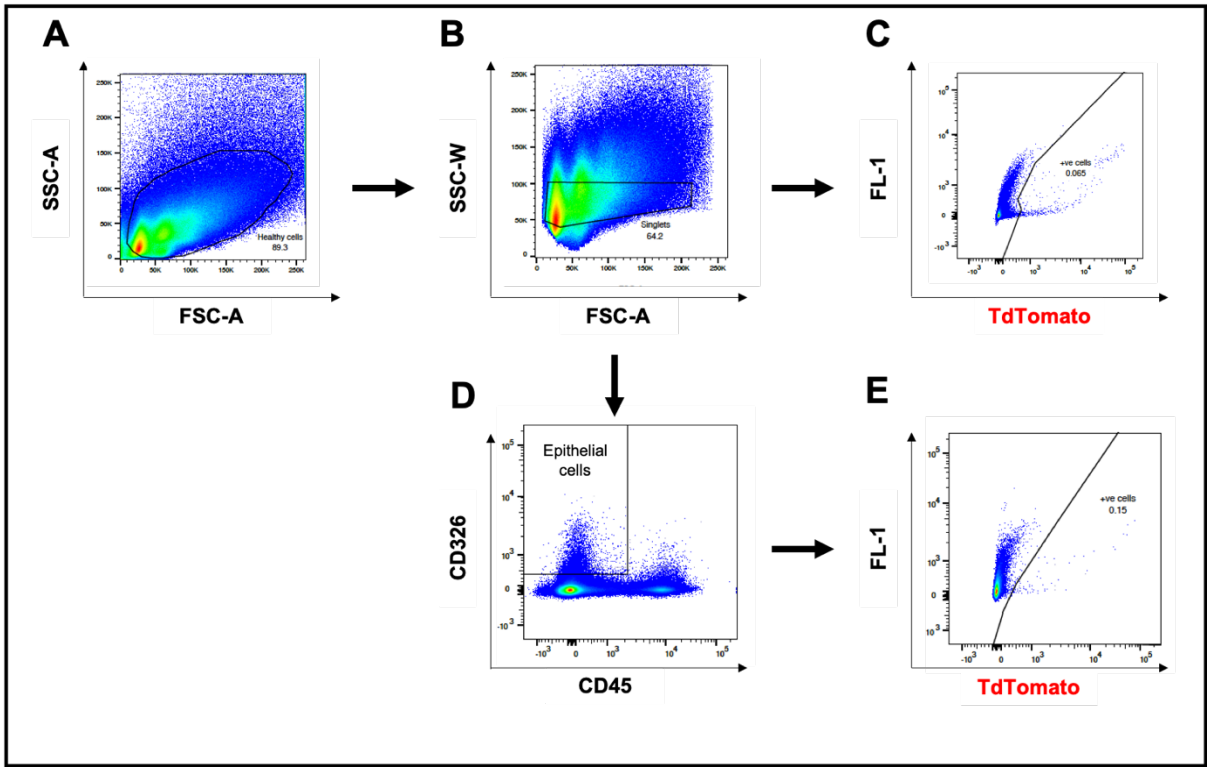


Figure 5.5. Flow cytometry density plots of Ai9 mice treated with RNP or Cre recombinase mRNA nanocomplexes. *Mice received 3 doses of RTNs containing either Cas9 RNP or Cre mRNA, over the course of 1 week.* ***(A)*** *Cell debris was excluded by using forward scatter area (x-axis) and side scatter area (y-axis).* ***(B)*** *Single cells were gated with forward scatter height (x-axis) and side scatter width (y-axis).* ***(C)*** *Gating on cells that express tdTomato.* ***(D)*** *Lung epithelial cells (negative for CD45 and positive for CD326 markers) were isolated.* ***(E)*** *Gating on cells that express tdTomato (Cre mRNA mouse 2 is depicted here).*

The analyses of the whole lung cell population showed on average 0.05 ±0.02 % tdTomato+ve cells after cre mRNA treatment, while RNP treated miced showed 0.007 ±0.001% tdTomato +ve cells, with no statistically significance between the two (Figure 5.6A). We found that epithelial cells (CD45-,CD326+) constituted around 1% of the whole lung cell population, consistent with results in the literature [452]. Of these, 0.09% ±0.03 of the Cre mRNA- treated group were tdTomato+ve, while 0.032% ±0.002 of lung epithelial cells were postive after RNP treatment, with no statistically significant difference between the two groups (Figure 5.6B). These results are 1.8-fold and 3.4-fold higher as compared to editing in the whole lung cell population for Cre mRNA and RNP respectively. However, tdTomato expression was considerably lower than results achieved in previous experiments conducted with a former lab member [453] (Figure 5.7).

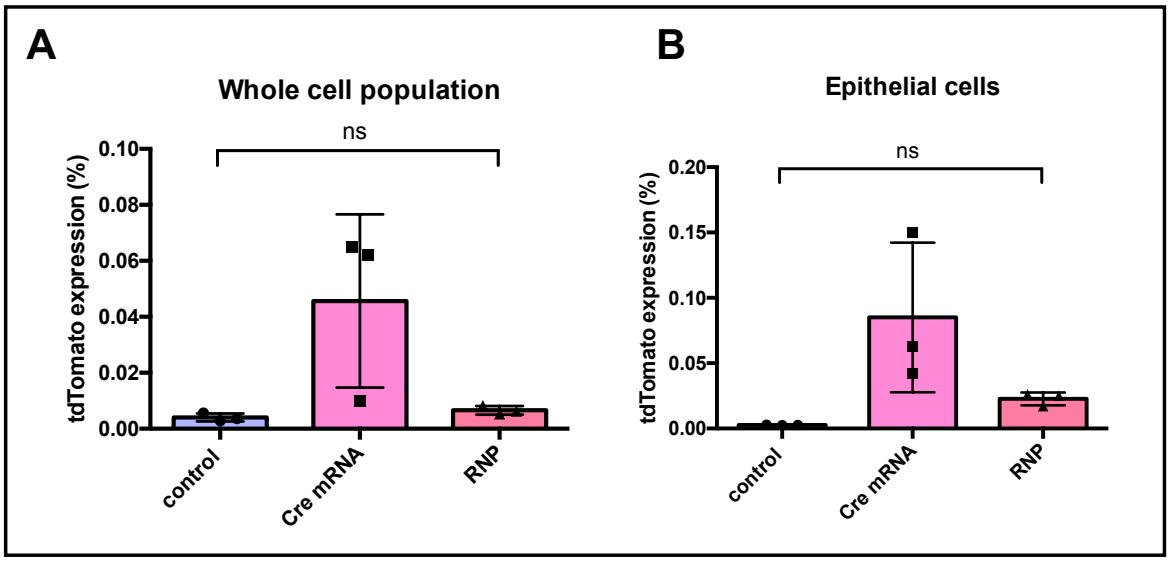


Figure 5.6. tdTomato expression in Ai9 mice with RNP or Cre recombinase mRNA nanocomplexes. *Mice received 3 doses of RTNs containing either Cas9 RNP (15 μg) or Cre mRNA (15 μg) , over the course of 1 week, before analysis by flow cytometry. (**A**) tdTomato expression was in the whole lung cell population and (**B**) in isolated epithelial cells, as explained in Figure 5.5.*

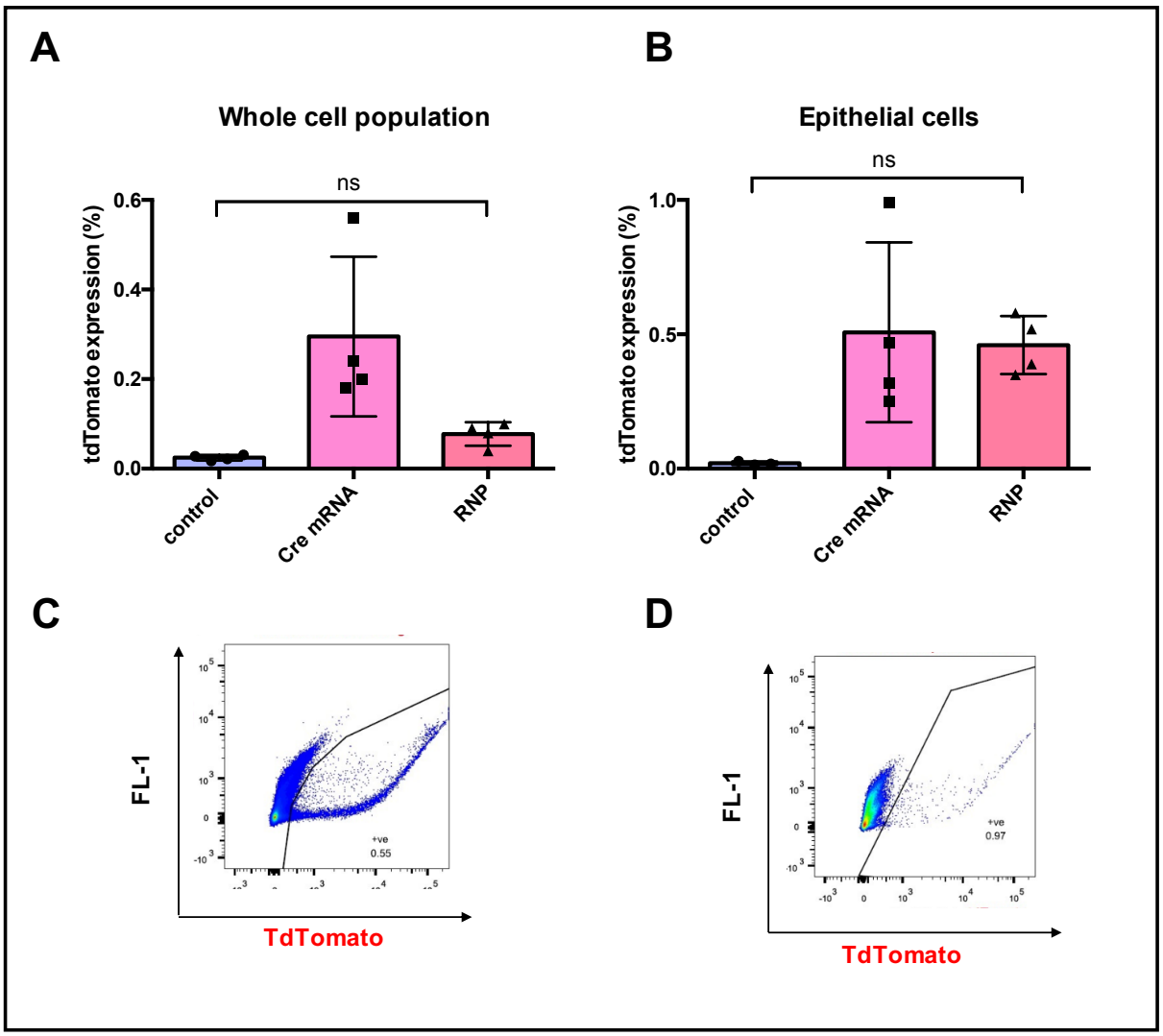


Figure 5.7. tdTomato expression in Ai9 mice transfected with RNP or Cre recombinase mRNA nanocomplexes. *Mice received 3 doses of RTNs containing either Cas9 RNP (15 μg) or Cre mRNA (15 μg) , over the course of 1 week, before analysis by flow cytometry. (A) tdTomato expression was in the whole lung cell population and (B) in isolated epithelial cells, as explained in Figure 5.5. Representative density plots of tdTomato +ve cells in (C) whole cell population and (D) lung epithelial cells after Cre mRNA treatment.*

5.3.2.5 Immunohistochemistry to assess tdTomato +ve cells

5.3.2.5.1 TdTomato embryo as a +ve control for IHC

Due to analysis by flow cytometry showing considerably lower tdTomato +ve cells than anticipated, immunohistochemistry was performed on lung sections to confirm the expression of tdTomato and identify transfected cell types and their location in the lung. Due to low levels of tdTomato expression in RTN treated mice, a good positive control is essential

for any IHC experiments. However we were unable to obtain any mouse lung tissue positive for tdTomato expression. Instead, a +9.5 day mouse embryo with cytoplasmic tdTomato expression, a kind gift from Dr Gabriel Galea, was used as a positive control. Fluorescent proteins, including tdTomato, are easily quenched through processing [454]. Therefore, an anti-RFP antibody was employed to locate tdTomato cells within the mouse airways. Following embedding in paraffin wax, 5 µm sections of tissue were cut and used to assess various methods of antigen retrieval, as well as concentration of antibody required. Both 1-in-50 and 1-in-100 dilutions of anti-RFP antibody in sodium citrate and Triton-X 100 buffers showed successful straining of tdTomato (Figure 5.8 A-B, D-E). Isotype control at dilution of 1 in 100 showed no staining of tdTomato (Figure 5.8 F) while a 1-in-300 dilution in sodium citrate buffer (p.H 6.0) failed to detect tdTomato (Figure 5.8 C). We therefore chose to use a 1-in-100 dilution in sodium citrate buffer (p.H 6.0) to detect tdTomato +ve cells in Ai9 mice, to minimise background signal.

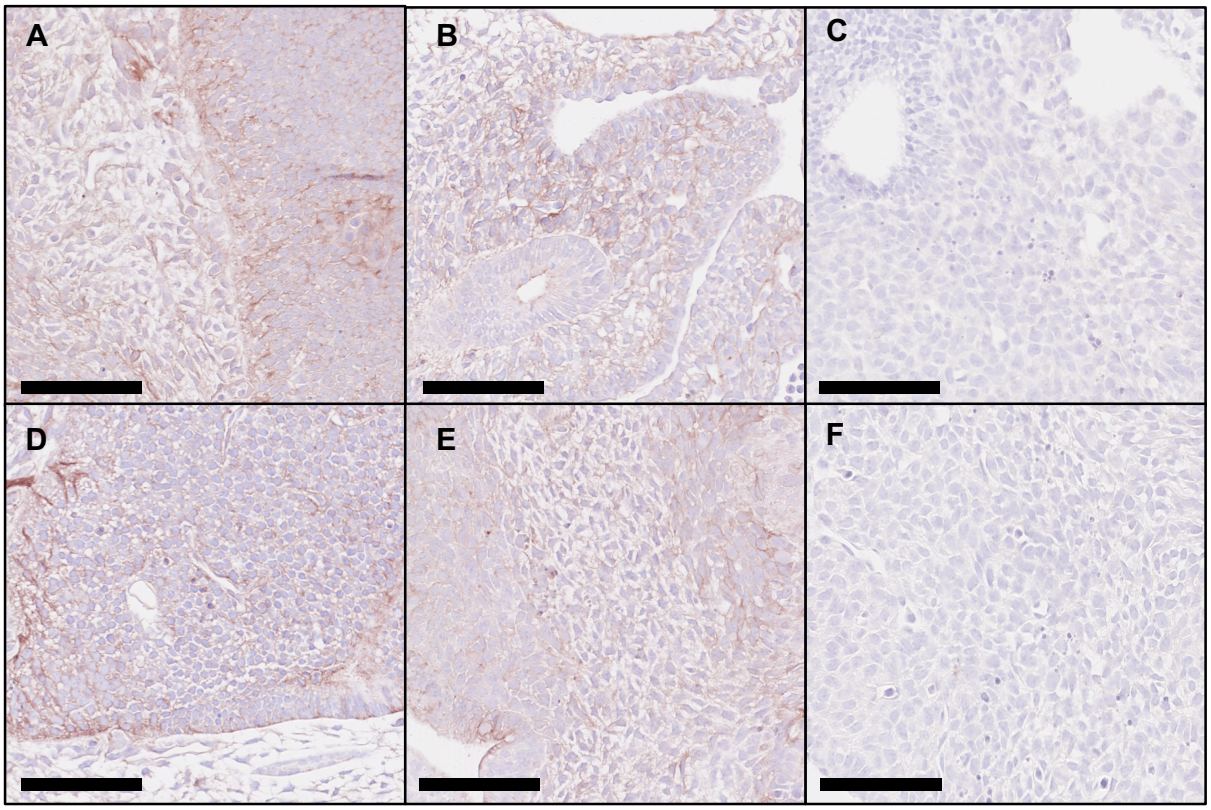


Figure 5.8. Antigen retrieval buffers for heat-induced epitope retrieval of tdTomato. ICH images show the detection of tdTomato in paraffin-embedded tdTomato +ve day 9.5 mouse embryos. All conditions were (A-C) Antigen retrieval solution was sodium citrate Buffer (pH, 6.0) with anti-RFP antibody dilutions of 1 in 50 (A), 1 in 100 (B), and 1 in 300 (C). (D-E) Antigen

retrieval solution was Triton-X 100 at antibody dilutions of 1 in 50 (D) and 1 in 100 (E). (F) shows isotype control antibody with a dilution of 1 in 100 in sodium citrate buffer (p.H 6.0). Scale bar = 100 uM.

5.3.1.1.1 IHC for detection of tdTomato +ve cells in Ai9 mice

Lungs were fixed by perfusion with 4 % PFA, before embedding in paraffin wax. Lung sections from each region within the lung were then cut and dried onto microscope slides. Following processing, anti-RFP antibody, at a dilution of 1-in-100, was then used to identify tdTomato +ve lung cells. Encouragingly, tdTomato +ve cells were found throughout the airways of both groups of mice, with tracheal (blue arrows), bronchial (black arrows) and even alveolar expression (yellow arrows) (Figure 5.9). Panel A-C shows RNP treated mice, while D-F shows Cre recombinance mRNA treated mice. G & H are from the control group, with no tdTomato expression shown, while I is the +9.5 day mouse embryo as a positive control. Positive staining was observed across all animals, with all sections and lobes of the lung staining positive for tdTomato expression. The Cre mRNA-treated mice exhibited small clusters of positive cells while, except for in the trachea, positive cells in RNP treated mice seemed to appear as single cells, and were less frequent that the Cre mRNA counterparts.

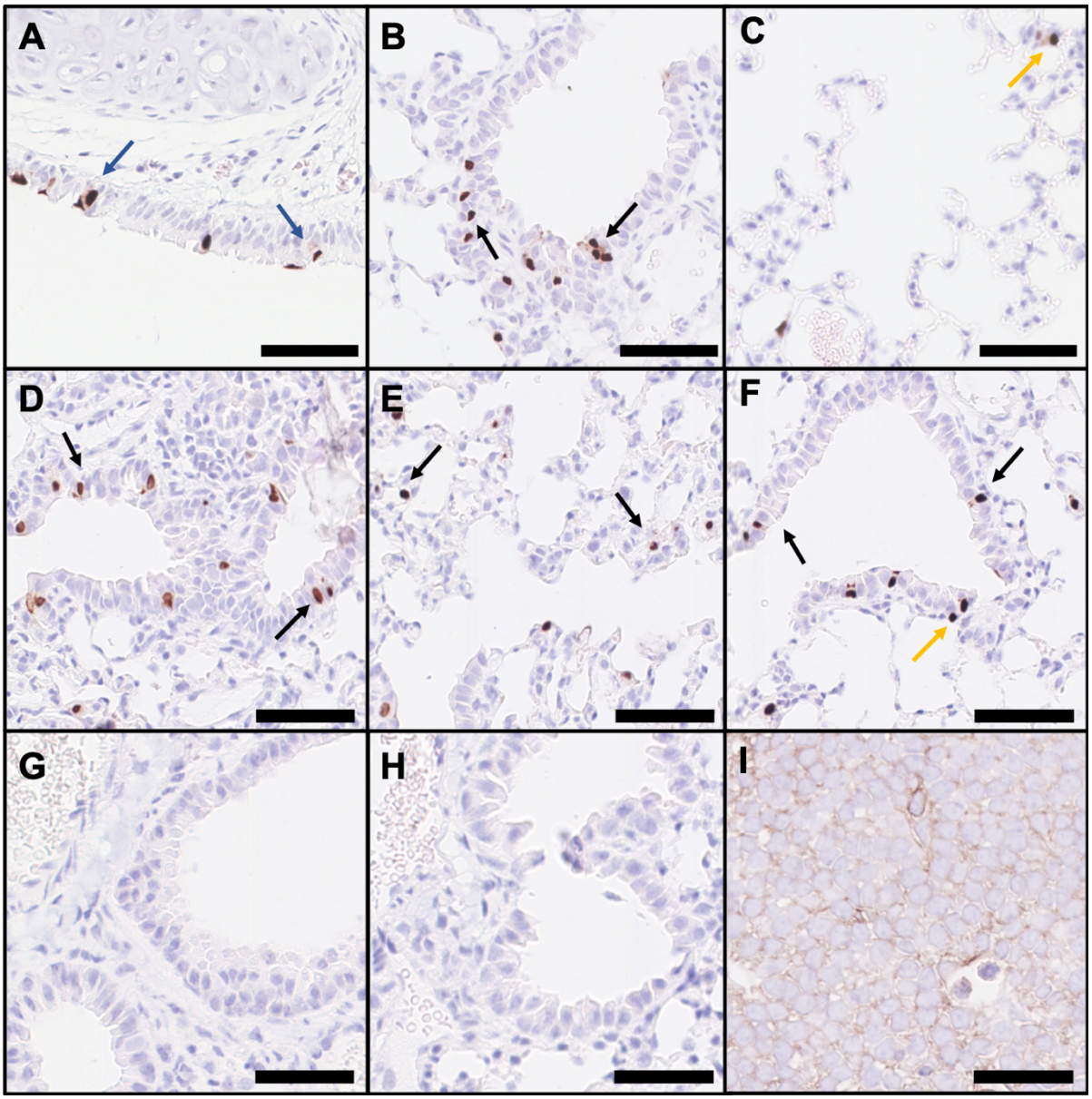


Figure 5.9. Immunohistochemistry of Ai9 mice treated with RNP (A-C) or Cre recombinase mRNA (D-F) nanocomplexes. *TdTomato +ve cells were identified using an anti-RFP antibody, and are observable in the trachea (blue arrows), bronchioles (black arrows) and alveoli (yellow arrows).* ***G + H*** *show control-treated mice.* ***I*** *shows IHC staining of a TdTomato +ve mouse embryo (day 9.5) as a positive control, with cytoplasmic TdTomato expression. Images are representative of the overall staining of the sections. Scale bar = 50 uM.*

5.3.2.6 Histology to assess in vivo toxicity of RTNs

Twenty four hours after the final administration, lungs were fixed by perfusion with 4 % PFA and processed for H&E staining, as described in Materials and Methods. Lungs of mice receiving the Cas9 RNP RTNs showed no sign of inflammation or immune cell infiltration

(Figure 5.10A), with histology of the lung similar to that of an un-treated control (Figure 5.10 C). In contrast, lungs of mice transfected with Cre recombinase mRNA showed some localised consolidation of the peribronchial parenchyma with mild inflammatory infiltrate, residing between the two bronchioles and in the solid mass of cells at the bottom right corner of Figure 5.10 B, as indicated with arrows. However, this inflammation was only found in one of the 5 Cre mRNA- dosed mice, while all other lungs looked normal.

IHC and histology was performed under the guidance of David Pearce, UCL, London.

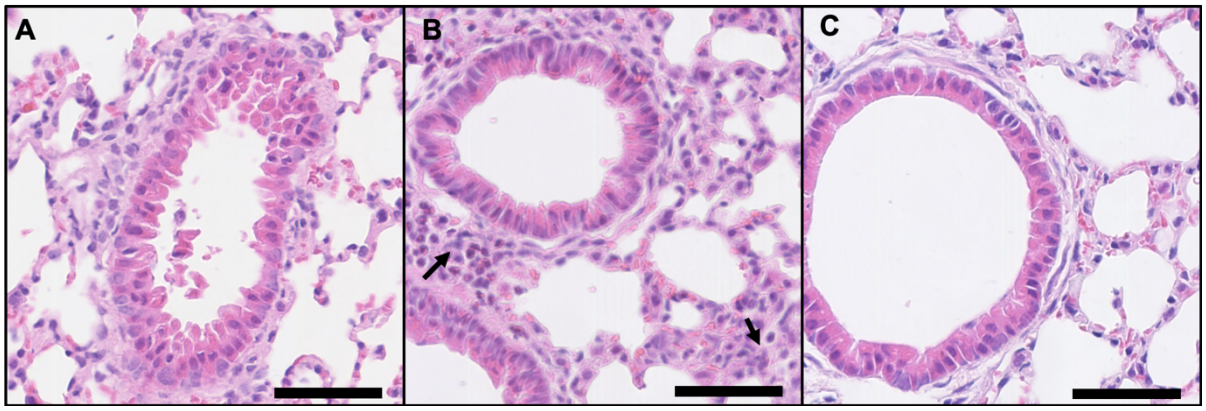


Figure 5.10. Representative H&E stained Ai9 mice lung sections following 3 consecutive oropharyngeal instillations *of (**A**) Cas9 RNP RTNs (**B**) Cre recombinase mRNA and (**C**) untreated control. RNP RTNs were formulated at a weight ratio of 1:1:4 RNP: L: P, while mRNA ratios were 1:3:4 mRNA:L:P. Mice were dosed every 48 h and lung tissue isolated on Day 7. Scale bar = 50 uM. Arrows point at inflammatory cell infiltration.*

5.3.2.7 Dynamic Light Scattering for biophysical characterisation of RTNs

Across the 3 treatments, the average size for the RNP and Cre mRNA RTNs was 255.2 ± 21.9 nm and 167.9 ± 42.1 nm, while both had PDIs of 0.35. Surface charge was +23.2± 4.1 mV and 29.1± 3.4 mV for RNP and Cre mRNA respectively. Size and PDI of the RTNs were significantly higher than those described in section 3.3.5.1, where we were able to formulate particles with a hydrodynamic diameter of ~90 nm. This is likely due to the higher concentration of particles required for *in vivo* delivery, causing the RTNs to aggregate.

Table 5.2. Biophysical characterisation of nanoparticles used for in vivo delivery of Cas9 RNP or Cre mRNA. *Table shows hydrodynamic size (nm), polydispersity index (PDI) and charge (mV), as measured by dynamic light scattering (DLS) using a Malvern Nano Zetasizer (Malvern Instruments, England). Three repeat measurements were made and standard deviation calculated.*

Treatment	Dose	Z-Ave (d.nm)	PDI	Charge (mV)
Cas9 RNP	1	271.43 ± 34.41	0.44	26.67 ± 4.05
	2	242.43 ± 7.37	0.30	21.27 ± 2.97
	3	251.83 ± 6.25	0.33	24.19 ± 0.15
Cre mRNA	1	216.00 ± 22.61	0.37	31.00 ± 3.74
	2	149.30 ± 5.98	0.33	31.67 ± 2.43
	3	132.97 ± 2.76	0.35	23.14 ± 0.15

5.4 Discussion

In chapter 4, we successfully corrected a deep-intronic, splice mutation using CRISPR/Cas9 and functionally rescued CFTR, as measured by Ussing chamber in *BMI-1* modified primary epithelial cells homozygous for the mutation. This was achieved by using a pair of gRNAs with each guide targeting upstream and downstream of the mutation site. While this work provides a preliminary proof-of-concept which could be broadened not only to include 5 other CF-causing mutations [435-438], but to treat up to 75 other genetic diseases caused by splicing alterations [439], it is essential to test the strategy in an *in vivo* setting to assess its translational potential.

Currently available mouse models of CF *[105-111]* are inappropriate for testing this strategy, and while humanised CF mouse models are currently under development, there are none currently available containing splice mutations. We therefore employed the Ai9 mouse model in order to assess our editing strategy of deleting a region in between two guides flanking a target site in an in vivo setting. By creating targeted DSBs using two gRNAs upstream and downstream of the STOP cassette, we aimed to release the cassette, allowing CAG promoter-driven tdTomato transcription and therefore fluorescence within the lung.

We first assessed our RTNs for Luc mRNA delivery in the lungs of C56B/6J mice, and found a significant increase in RFU / mg protein as compared to an untreated control (Figure 5.2). However, this level of luminescence was considerably lower than levels achieved previously within our group with Luc minicircle DNA [455] and even plasmid DNA [456]. There could be several factors contributing to this, including the age of the mice used in our experiments (12) weeks, twice the age of mice used in previous publications. Plasmid DNA delivery requires nuclear localisation and transcription, and so could be assumed to be to be less efficient than mRNA delivery. However, the inclusion of a strong promoter can increase transgene expression, which contributes to the high levels on expression previously found [456]. Regardless, mRNA delivery affords several advantages over its DNA counterpart, namely exclusion of unexpected insertional mutagenesis, a greater ability of gene transfer in non-dividing cells, and less inflammation by means of chemical modification, making mRNA-mediated gene delivery a more likely candidate in a clinical trial setting.

The Ai9 model is a useful tool to investigate the efficiency of gene editing both *in vitro* and *in vivo*. Frame-disrupting mutations in the DMD gene, encoding dystrophin, compromise

myofiber integrity and drive muscle deterioration in Duchenne muscular dystrophy (DMD) [457]. By removing one or more exons from the mutated transcript, it is possible to produce an in-frame mRNA and a truncated but still functional protein. The translational potential of this strategy is supported by the relatively milder phenotype of Becker Muscular Dystrophy (BMD) seen in patients with in-frame deletions within the *DMD* gene [458, 459]. The potential efficacy of this strategy was first tested in primary satellite cells isolated from the Ai9 mouse, by delivering a plasmid encoding spCas9 and 2 gRNAs, targeting upstream and downstream of the STOP cassette, and found 10% tdTomato +ve cells as measured by flow cytometry [460]. For *in vivo* work, they chose to use SaCas9, whos smaller size permits packaging in an AAV vector, along with two gRNAs, and were able to achieve 4% and 3.8% of tdTomato +ve gene editing satellite cells after intraperitoneal or intramuscular injection respectively.

This exon skipping strategy was also investigated by Lee and colleagues, who first used the Ai9 mouse model to assess their CRISPR-GOLD nanoparticles, comprised of gold nanoparticles (GNPs) complexed with donor DNA, Cas9 RNP and the endosomal disruptive polymer poly(*N*-(*N*-(2-aminoethyl)-2-aminoethyl) aspartamide) (PAsp(DET)) [461]. In their experiments, they delivered Cas9 RNP and gRNAs using CRISPR-GOLD nanoparticles via intramuscular injection. After two weeks, tdTomato expression was assessed in the tibialis anterior muscle section, millimetres from the injection site, where tdTomato expression was observed across a broad area of muscle, however quantification of editing was not performed. Encouraged by these results, they followed up with experiments in *Mdx* mice, a model for DMD and showed that delivery of CRISPR-GOLD particles, containing a donor DNA sequence designed to repair a nonsense mutation in the dystrophin gene, was able to enhance animal strength and agility.

Staahl and colleagues employed the Ai9 model to assess Cas9 RNP-mediated gene editing within the mouse brain, choosing various stereotaxic injection sites, including the hippocampus and striatum, in male mice age 14 weeks [462]. Rather than using a pair of gRNAs to release the STOP cassette, the group used a single gRNA to create multiple cuts within the cassette, therefore sufficiently disrupting the reading frame and allowing tdTomato transcription. 14 days post-injection, brain sections were analysed by confocal microscopy, showing that tdTomato fluorescence co-localised with neuron marker proteins. Increasing doses of Cas9 RNPs significantly increased the number of tdTomato +ve gene-edited cells within the striatum. This work supports the use of the Ai9 mouse as a relevant

model in which to determine the translational potential of gene editing strategies and their delivery.

Our *in vivo* work involved 3 repeat doses of RTNs containing Cas9 RNP or Cre mRNA via oropharyngeal instillation. 3 days after the last dose, lung tissue was harvested for analysis. Disappointingly, flow cytometry analysis showed levels of ~0.09 % tdTomato +ve cells, with slightly higher levels of positive cells when isolating for epithelial cells, which we hypothesise indicates the effect of the targeting peptide in our RTNs that has affinity for airway epithelial cells [391, 392]. These results were around 5-fold lower than previously achieved in experiments performed with a former lab member [453]. There are several reasons why this could be; in prior experiments, only female mice were used in the cohort, with lower average body weights than their male counterparts. It is also likely that the oropharyngeal instillation procedure was more efficient in the first experiments. In particular, Cre mRNA RTNs delivered to one particular mouse achieved levels of almost 1% tdTomato+ve lung epithelial cells, indicating that the procedure is more successful in some mice than others. Whether this is an effect of a more successful instillation technique itself remains to be investigated. Finally, lung tissue in the first experiments was isolated 24 h after final administration, while in the experiments described here we waited 3 day. It could be the case that edited cells cannot tolertate modifications and undergoing apoptosis and being cleared from the lung before analysis.

Nevertheless, in immunohistochemistry performed on lung sections from both RNP and Cre mRNA treated mice in the second set of experiments, tdTomato +ve cells observed throughout the lung in all sections and lobes, from the upper airways down to the alveoli. As mentioned, expression appeared higher in Cre mRNA treated mice as compared to RNP treated mice, suggesting that Cre recombination of the *Lox P* sequences is more efficient than Cas9-mediated DSBs within the genome.

Despite low levels of tdTomato +ve cells after RTN instillation, it is possible that NHEJ efficiency may be higher than indicated. This model depends on DSBs occurring both upstream and downstream of the STOP cassette for successful excision and therefore transcription of tdTomato. However, it is likely that DSBs are occurring at only one of these locations, with cells therefore remaining negative for tdTomato expression as the cassette has not been removed. Further, it will be important to assess fluorescence levels after only one dose of RTNs, and isolate the lung tissue 24 h post instillation, in comparison to 3 doses

of RTNs with tissue isolation 3 days after the final dose, to assess whether edited cells are in fact being cleared from the lung.

Importantly, histological analysis showed mice receiving the RNP RTNs showed no signs of inflammation or cellular infiltration, with lungs looking similar to un-treated mice. One of the Cre mRNA-treated mice showed some signs of pathology, with some mild inflammatory infiltrate in peribronchiolar region, however all other mice in this cohort appeared largely normal. It is not possible to determine from this data whether the mild inflammation observed was acute or represented chronic changes.

In regard to the larger hydrodynamic size and PDI values of the RTNs used for *in vivo* delivery as compared to those achieved in section 3.3.5.1, the volume in which the complexes must be formed for oropharyngeal instillation is 50 μl, in comparison to the 200 μl we formulated for *in vitro* transfections. Further, the concentration of the particles is around 30-fold higher (15μg Cas9 RNP *in vivo* compared with ~0.5 μg *in vitro*). This, along with the 4-fold reduction in volume, means the RTNs are over 120-fold more concentrated. We hypothesise that this may lead to particle aggregation and account for the larger sizes we see in the DLS data, rather than individual complexes increasing in size. In the context of the CF lung, particles of this size are likely to be too large to penetrate through the mucus mesh pores, which are ~140 nm in CF mucus [415]. Previous work using our RTNs containing Cy3-labelled siRNA has shown that they can successfully penetrate pig gastric mucus, as well as mucus from non-CF and CF airways, at a rate significantly faster than commercially available, fluorescent cationic, polystyrene (PS) nanoparticles [377]. Several modifications could be made to the nanocomplexes in order to reduce aggregation and increase transfection efficiency, such as PEGylation and the addition of cholesterol [187, 417-419].

Taken together, this preliminary data suggests our RTNs are a safe and feasible option for repeat delivery of the CRISPR/Cas9 system within the lung, however further optimisation of RTNs and CRISPR/Cas9 system is needed to enhance *in vivo* gene editing efficiencies.

Chapter 6
General Discussion and Conclusion

6.1 Discussion

Cystic fibrosis (CF) is an autosomal recessive disease, which affects around 1 in 2,500 babies born in the UK [463], with an estimated 70,000 sufferers worldwide [464]. The clinical manifestations of CF are caused by mutations in the gene encoding the cystic fibrosis transmembrane regulator (CFTR), of which over 2,000 mutant variants have been identified, 281 of which are suggested to be disease causing [34]. Treatment of the disease requires a multidisciplinary approach provided in specialist CF care centres to tackle the multi-faceted disease, with primary objectives maintaining lung function and preventing and controlling infection, as well as nutritional therapy. Recently, there has been rapid progress in the treatment of CF with the emergence of CFTR modulator therapies, including 'potentiators' such as Ivacaftor, and 'corrector' drugs such as Lumacaftor [126]. While the different modulator drugs are useful for treating CFTR mutations that produce a defective protein, there are several classes of CF mutations for which these drugs are not suitable. Moreover, the high cost of modulator therapies, coupled with their 'moderate' benefit in F508del patients, means that the drugs are not currently funded in several countries, however as of 24th October 2019, they have been recently approved in the UK [138]. So, in spite of the recent advancements in drug therapies, there is still an urgent need for new treatments that are amenable to a larger cohort of CF patients.

Since the discovery of the CFTR gene in 1989 [12], several gene therapy approaches have been evaluated by delivery of CFTR cDNA by viral or non-viral vectors, however no clinically effective therapy has emerged to date [152]. Major challenges include limited levels of gene transfer, and persistence of transgene expression. In recent years, there has been rapid advancement in the development of genome editing techniques including ZFNs and TALENs but in particular CRISPR/Cas9. As a monogenic disease, CF is an ideal candidate for CRISPR/Cas9 gene correction, and the system is being widely adopted in the CF field [101, 368-370, 450], with hopes that precise gene editing may be more effective than cDNA addition. Further, the technology has the potential to treat all CF mutations, rather than certain cohorts of patients.

A suitable model is required to test the delivery and efficiency of any gene editing strategies, and to assess the functional consequences of editing. To this end, chapter 3 of this

thesis describes the transduction of primary bronchial epithelial cells with the proto-oncogene *BMI-1* to produce a semi-immortalised cell line capable of extended passaging whilst maintaining its mucociliary differentiation capacity. We next integrated a *GFP* gene into these *BMI-1* modified cells, resulting in a stably-expressing GFP cell model in which to access the delivery efficiency of our RTNs for delivery of the CRISPR/Cas9 components. There are several different formats in which the CRISPR components can be delivered. Originally, for laboratory use, the Cas9 nuclease was encoded in a plasmid along with a guide RNA. However, the necessity of nuclear translocation of plasmid DNA limits efficiency *in vivo*, particularly in the non-dividing, terminally differentiated lung epithelium. Instead, in this chapter we successfully modified our RTNs for Cas9 mRNA and Cas9 RNP delivery, which confers several advantages; no need for transcription of mRNA, or indeed translation of protein in the case of RNP delivery. Further, protein-based Cas9 delivery offers transient exposure of the host genome to the nuclease, which may result in reduced off-target effects and minimal potential immunogenicity. The benefits conferred by RNP delivery, as well as the higher editing efficiencies achieved by RNP as compared to Cas9 mRNA, led use to choose this CRISPR format for all subsequent work.

The correction of 3849+ 10 kb C>T mutations is investigated in the chapter 4 of this thesis, and the reasons for choosing this mutation are severalfold. Critically, the correction strategy for this mutation does not require the presence of a donor template, and, therefore, exploits the NHEJ-mediated repair pathway. Not only is this pathway more efficient than the HDR repair pathway, it is also the more active pathway in non-dividing cells. Moreover, delivery of the CRISPR/Cas9 platform is much simpler without the requirement of a DNA donor template. This strategy therefore offered a 'low hanging fruit', allowing us to provide a proof of concept of CRISPR-mediated gene editing in *BMI-1* modified CF lung epithelial cells. Through delivery of Cas9 RNP using our RTNs, we were able to precisely excise a small region of the CFTR intron 22 containing the disease causing mutation. Excision occurred in ~60% of *BMI-1* modified primary cells, restoring close to normal CFTR splicing patterns and CFTR ion channel function as measured by Ussing chamber. Also exploiting the NHEJ-mediated repair pathway, we showed a 14% GFP transgene knock-in efficiency in a HEK293T cell model using a HITI-based method, providing proof-of-concept for a gene editing strategy applicable to a wider cohort of CF-causing mutations.

In chapter 5, we assessed the capabilities of this editing strategy *in vivo,* employing the Ai9 mouse model. Despite achieving lower than anticipated editing efficiency, tdTomato +ve cells were observed in all mice, in all sections and lobes, from the upper airways down to the alveoli, validating the editing strategy and the use of RTNs for RNP delivery *in vivo*. Although we histologically assessed the lungs of mice dosed with our RTNs, we could further test the immunogenicity of Cas9 protein delivery. As mentioned, the natural turnover of lung epithelial cell populations necessitates repeated administration of the CRISPR/Cas9 components, and so it is essential that the therapy will not elicit an immune response. We could assess this by measuring the systemic cytokine profile in the time period after transfection of the nanoparticles. This has been described previously to assess the immune response of Cas9 RNPs delivered intramuscularly using gold nanoparticles [461]. It was shown that there was no up-regulation of inflammatory cytokines in the blood plasma, nor any weight loss in the mice, after multiple injections, suggesting multiple doses of Cas9 RNP complexes can be delivered safely [461].

While we have shown an editing strategy that can restore CFTR function with the potential to treat ~1.6% of individuals with CF by simply removing a section of deep-intronic DNA, therapies for a larger cohorts of CF patients will require the delivery of a donor template. Transfecting terminally-differentiated cells such as the airway epithelium, presents two problems, the first of which is that plasmid-based, CFTR gene transfer is inefficient due to the highly selective nature of uptake through pores in the nuclear envelope limiting nuclear uptake of plasmid in post-mitotic cells [233, 234]. This barrier is likely to be one of the main reasons why clinical gene therapy studies with liposomal nanoparticles have demonstrated such low levels of transgene expression. Even by employment of a HITI strategy that triggers the NHEJ repair pathway, active in all stages of the cell cycle [247], there is still inefficiency associated with the delivery of the donor template to the nucleus. When considering this strategy for the knock-in of full length CFTR cDNA, using a smaller size plasmid such as the minicircle plasmid [447] or a mini-intronic plasmid (MIP) [465] where the bacteria backbone has been removed, is a potential approach to improving delivery of such a large donor. It has been widely reported that the use of single-stranded oligonucleotides as a DNA donor template for HDR repair in gene editing increases editing efficiency [466, 467]. By rationally designing single-stranded DNA (ssDNA) donors, HDR rates in HEK293 cells were increased to up to 60% [468]. The use of single-stranded DNA templates circumvents the need of a large

DNA plasmid or even minicircle or MIPs, with the belief that these smaller templates can be delivered to the nucleus of cells more efficiently than their larger counterparts, leading to an increase in editing efficiency. However, the drawback to using ssODNs is that they are restricted in length to a few hundred bases, so insertion size is limited, making them more amenable to correcting single mutations over creating large insertions.

Whilst we used a splice acceptor in our construct for expression under the endogenous AAVS1 promoter, to enhance transgene expression it may be necessary to use stronger promoter and enhancer elements, such as the use of CMV or SV40 viral promoters. Use of the promoters from the human polyubiquitin C (UbC) and the elongation factor 1α (EF1α) genes resulted in persistent, high-level reporter activity detectable 6 months after a single administration of naked plasmid via intra-nasal dosing in mice, compared to the more transient expression obtained with viral promoters CMV, RSV and SV40 [203].

It is likely that future editing strategies requiring a donor template will employ both viral and non-viral delivery methods, with non-viral delivery of Cas9 nuclease and gRNA, by means of electroporation or nanoparticles, conferring several advantages including an unlimited packaging capacity and transient nature potentially reducing off-target effects. Meanwhile, the use of AAV vectors for template delivery significantly increases delivery efficiency and are being increasingly used in CRISPR/Cas9-based editing strategies. A recent approach has been described by the He lab termed CRISPR-READI (CRISPR RNP electroporation and AAV donor infection), combining AAV-mediated HDR donor delivery with Cas9/sgRNA RNP electroporation [469]. In this way, the group were able to engineer large site-specific modifications in the mouse genome, including the addition of a fluorescent reporter, a 2.1 kb CreERT2 driver, and a 3.3 kb expression cassette into endogenous loci in mouse zygotes, with HDR efficiencies of up to 69% for insertion of the CreERT2 driver. When these edited embryos were transferred to pseudo pregnant mice, 2 of the 5 resulting pups were also edited [469]. In the CF field, the latest publication from the Matthew Porteus lab showed precise editing in up to 43% of alleles in primary airway basal stem cells obtained from F508del homozygous patients, after electroporation of the cells for RNP delivery, followed by immediate transduction of donor template by AAV6 vector [369]. While these editing strategies using a combination of AAV and non-viral delivery methods are becoming more frequently employed, the use of a viral vector relinquishes the possibility of repeat

dosing afforded by a non-viral delivery method such as our RTNs. In addition, the scalability of viral production remains an issue.

Despite these meaningful increases in HDR efficiency levels, practical limitations for conventional HDR-based knock-in strategies still exist because HDR is mainly active in dividing cells. Belmonte and colleagues, who first published the HITI editing method, recently described a novel gene editing method of targeted gene knock-in mediated by both the NHEJ and HDR pathways [470]. The strategy requires a DSB induction site within a single stretch of homologous sequence within the donor template, and is termed 'intercellular linearized Single homology Arm donor mediated intron-Targeting Integration (SATI)'. SATI, therefore, allows for a single homology arm mediated knock-in *via* HDR, or via the NHEJ-based HITI method, permitting targeting of a broad range of mutations or cell types [470]. The paper details the insertion of a minigene in the intron of a target gene locus using an intracellularly linearized single homology arm donor and demonstrate its applicability through *in vivo* correction of a point mutation in a progeria mouse model. This new editing system overcomes some of the issues associated with HITI; while HITI can insert DNA at a precise location within the genome, it cannot repair point or frame shift mutations, and is therefore unsuitable for targeting gain-of-function, dominant mutations. However, limitations associated with the NHEJ-mediated template insertion still exist with SATI, namely the creation of small indels at the junction between the genomic locus and inserted DNA, which may lead to an out-of-frame shift causing cessation of gene function. To circumvent this, the group targeted intronic regions upstream of the appropriate exon. By including a splice acceptor, relevant downstream exons, and a 3'UTR, this would, in theory, result in transcription of the donor exons in place of endogenous exons. Importantly, small indels introduced into the intron have reduced probability of affecting target gene function, something we also took advantage of in the 3849 editing strategy described in Chapter 4. Before moving to clinically relevant *in vivo* models, the group first evaluated their approach *in vitro* in post mitotic neurons, through insertion of GFP. When either only the genomic sequence or the DSB induction site within the single homology arm were cleaved, low levels of integration were observed, however, the junction site where the donor template integrates remained intact, showing no features of HITI. The authors postulated, therefore, that an unknown, non-canonical HDR pathway was taking place in post-mitotic neurons, allowing for successful, albeit low-levels of integration. They term this one-armed HDR (oaHDR), distinguishing it from conventional HDR which

utilises two homology arms. When both the genomic sequence and homology template are cleaved simultaneously, much higher editing efficiencies are achieved, comparable to a classical HITI donor template with no homology (~36%). Interestingly, they found that with the SATI construct only ~5% of integration was mediated by HITI, attributing the other 95% to their newly described oaHDR, which they claim requires the participation of elements of both NHEJ- and HDR-related pathways. This was not the case in dividing cells, where NHEJ-mediated HITI was the predominant insertion pathway [470].

In the progeria mouse model, the group used AAV-mediated delivery containing the SATI-mediated gene-correction donor with one 1.9 kb homology arm for delivery to adult mouse brains. They observed that oa-HDR mediated integration was the predominant repair pathway in neuronal cells, with some evidence of HITI correction. The constructs were then delivered systemically at postnatal day 1, and mouse tissue isolated at day 100 for assessment of gene correction. Levels of 2.07% in the liver and 0.14% in the heart were observed; while these efficiencies seem low, they were associated with some levels of phenotypic rescue, with slowdown of progressive weight loss, increased epidermal thickness, and an extension of median survival time by 1.45-fold [470]. The authors concluded that the SATI method could potentially be used to correct many dominant mutations *in vivo* through the targeting of multiple tissues via systemic delivery, with over 90% of human genes having open reading frames <4 kb and therefore within the capacity of AAV-based packaging [471]. Crucially however, this strategy in its current form would be unsuitable for CF, with a cDNA of 4.4 kb [15], as well as a strong promoter and poly-A signal required.

The recently described technique base editing (BEs) represents a new dimension of CRISPR/Cas-mediated precise editing [472, 473]. Base editing has the capacity to generate single-nucleotide changes in DNA or RNA independently of double-strand breaks and homology-directed repair. The first type of base editor created by the lab of David Liu was a cytosine base editor (CBE), capable of targeted C >T conversions, and therefore G > A in the opposite strand. The CBE is comprised of a cytosine deaminase that catalyses the conversion of C to U, a modified Cas9 (nCas9/dCas9) that binds target DNA without creating a DSB, a sgRNA which direct Cas9-cytosine deaminase to bind the target locus, and a uracil DNA glycosylase inhibitor (UGI) that prevents the subsequent U:G mismatch from being repaired back to a C:G base pair [472, 473]. Adenine base editors (ABE) were created soon after, able to convert a targeted A > G (T > C in the opposite strand) [472]. Theoretically, fusing an

adenosine deaminase with dCas9/nCas9 would give rise to ABE, however, no enzyme was known to deaminate adenine in DNA. The Liu group overcame this by, evolving a transfer RNA adenosine deaminase to operate on DNA through extensive directed evolution, fused to a catalytically impaired CRISPR–Cas9 mutant by protein engineering, greatly expanding the scope of the base editing technology. Base editing yields a high efficiency of editing with very low rates of indel formation, reducing off-target effects . Critically, BEs can work in both dividing and non-dividing cells [472, 473].

The latest gene editing method from the Liu lab has been shown to create more than 175 different edits in various human cell lines, including targeted insertions up to 44 bp, deletions up to 80 bp, and 119 point mutations, without requiring DSBs or donor DNA templates in dividing and post-mitotic cells [349]. The technique is termed prime editing, and uses a catalytically impaired Cas9 fused to an engineered reverse transcriptase, programmed with a prime editing guide RNA (pegRNA) that both specifies the target site and encodes the desired edit. Significantly, distance from PAM site does not impede insertion efficiency, unlike the case with HDR techniques. Like BEs, prime editing offers much lower off-targeting than HDR due to lack of any DSBs within the genome, and substantially expands the scope and capabilities of genome editing. Because 85-99% of insertions, deletions, indels, and duplications in ClinVar are ≤30 bp, in principle, prime editing can correct up to ~89% of known pathogenic human genetic variants (https://www.ncbi.nlm.nih.gov/clinvar/).

To achieve maximal CFTR restoration by gene editing *in vivo*, it is important to target the cells that express CFTR. Expression appears most evident in submucosal glands and in the ciliated airway surface epithelium [228]. Recently, two studies identified the presence of novel, rare cell type celled a ‘pulmonary ionocyte’, which are a major source of transcript of CFTR in both human and mice [474, 475]. Of all these cell types, the ciliated epithelium are likely to be the most easily accessible in terms of an RTN-mediated delivery strategy within the lung. However, airway epithelial cells are terminally differentiated with a half-life of approximately 6 months in mice [232], although epithelial turnover may be shorter in the human CF lung, as a result of chronic inflammation [476]. Periodically, therefore, any therapy will have to be re-administered in order to have a sustained effect, which is afforded by our RTNs [218-220, 379].

Nebulised delivery has been used extensively for the administration of medication by inhalation, typically generating aerosol particles less than 5 µm in diameter that can reach the

lower respiratory tract [477]. Nanoparticles must be able to be concentrated to high doses and also withstand shearing forces generated during aerosolisation whilst maintaining their stability and biophysical characterisation. Nebulized delivery has been adopted in human clinical trials for inhaled delivery of CFTR-DNA to cystic fibrosis patients [478]. Recently, it was shown that hyperbranched poly(beta amino esters) nanoparticles were able to successfully deliver mRNA to the lungs of mice [479]. This approach has broad therapeutic applicability as it permits temporal and dose-dependent control of encoded protein expression. Impressively, in an Ai14 reporter mouse model, similar to the Ai9 mouse used in our experiments, (with loxP-flanked STOP cassette preventing expression of tdTomato), they were able to show tdTomato expression in 24.6% of lung epithelial cells after a single dose of Cre mRNA [479]. It was shown previously that RTNs, similar to those used in this study, are able to be successfully nebulised and used for delivery of a *lacZ* plasmid to the lungs of mice [385]. This is something we aim to assess in the Ai9 mouse model for Cas9 RNP delivery, however, it has been estimated that only ~0.5% of the total nebulized dose is deposited in the lung using whole body chambers [480, 481], with most of the aerosolised product depositing in the fur of the mice, and was therefore unfeasible in this project.

This thesis has optimised RTNs for the delivery of the CRISPR/Cas9 platform, with the intention that these particles could be delivered *in vivo* for the correction of cells in the lung *in situ*. An alternative therapeutic approach for CF could be the correction of a patients' own stem cells *ex vivo* and re-administering them, offering the prospect for autologous cell therapy. This approach could have several advantages over an *in vivo* gene editing approach, including the ability of sorting cells before engraftment, followed by expansion of corrected cells, therefore improving the safety and efficacy of the therapy. For proof of concept, basal lung epithelial cells from CF patients could be edited and sorted, before engraftment in immunodeficient NOD-SCID-gamma (NSG) mice, widely used for human stem cells xenotransplantation [482]. Cells could be administered via intravenous (I.V) injection or directly via oropharyngeal instillation. A phase I safety study (NCT02866721) is currently underway to assess I.V delivery of allogenic MSCs and their effect on CF disease state. If successful, this would support and *ex vivo* manipulation strategy of a patient's stem cells followed by I.V re-administration. Oropharyngeal instillation of corrected stem cells would require preconditioning to create a niche for engraftment of engineered basal cells with agents such as polidocanol [235], by irradiation [236] or sulphur dioxide treatment [237],

which will ablate a proportion of the airway epithelial cells. Whilst this sort of conditioning is not appropriate in humans, the influenza virus naturally strips the airway epithelium [483], which could be used as a strategy for ablation of cells before a cellular therapy. It would be important to assess the differentiation capacity of the basal cells into ciliated and secretory cell types, their CFTR activity, and their capacity to proliferate and function as multipotent airway stem cells. Human basal stem cells have been previously shown to restore a well-differentiated and functional airway epithelium in cells [238]. Similarly, a mixture of human tracheobronchial epithelial tissue stem cells and progenitor cells were able to repopulate an injured airway epithelium in mice [239]. Two weeks post-transplant, the cells had differentiated into the three major epithelial cell types; basal, secretory and ciliated.

In this thesis, we have not considered the off-target effects of our gene editing. Several methods of assessing off-target cutting were discussed in section 1.7.3.4 of the introduction. Identification of putative off-target sites followed by PCR amplification of these sites is a method that could be easily performed on our corrected cells. Methods involving whole genome sequencing, including Digenome-Seq and GUIDE-Seq and DISCOVER-Seq [358, 360, 361] could also be employed. This type of analysis is essential before moving any gene editing techniques toward the clinic.

Should a gene editing therapy make it to clinical trials for CF, it is likely that many of the patients recruited to the trial will already be on CFTR modulator therapies, with a significant benefit to their health. A washout period of the drug could greatly decrease lung function within a patient and would be considered unethical, therefore, it is likely that patients would remain on modulator drugs throughout the duration of the trial. Ivacaftor has been shown to work on wild-type CFTR as well as mutated transcripts [484] and could lead to a synergistic effect with the modulator drugs, boosting the effect of increased levels of correctly processed CFTR through gene editing. It would therefore be useful to test the effect of the modulators in vitro on our corrected cells, and see if they further increase CFTR expression and function.

Cost is certainly another issue that needs to be addressed should gene editing therapies make it to the clinic for CF. In 2017, the FDA-approved gene therapy product, Luxturna by Spark Therapeutics, was made available for treatment of inherited retina dystrophy caused by biallelic mutations in RPE65, with a cost of $425,000 per eye [155]. The newly approved gene therapy drug Zolgensma for treatment of spinal muscular atrophy has

a price tag of $2.125 million, which may provide an indication of cost of gene editing techniques. The number of people currently living with CF in the UK is over 10,500 [1]. Considering the four year battle between Vertex and NHS England over the price of modulator drugs, the cost of a future CF gene therapy it is likely to be a significant issue for the NHS. However, gene editing can potentially provide a curative therapy and therefore, considering the cost of treating a patient throughout their lifetime, may actually be the more economical treatment. Moreover, costs will decrease dramatically as gene editing becomes more widely applied; a review by MIT recently examined projections from the existing pipeline of cell and gene therapies, and estimate 40-60 product launches by 2030 [485].

While we are still far from the clinic, a number of advances towards a gene editing therapy for CF has been described in this thesis. This includes the development of a cellular model for optimisation of RTNs containing the CRISPR/Cas9 editing platform, the modification and characterisation RTNs for RNP delivery *in vitro* and *in vivo*, the functional correction of the 10th most common CF-causing mutation, with the potential application to a number of other splice mutations, and provided proof-of concept for a HITI-based editing strategy that could be applied to correct all CF-causing mutations. The technologies developed in this thesis project also have broader applications beyond CF gene therapy.

1. *UK Cystic Fibrosis Registry. Annual data report 2018*. 2018.
2. Ioannou, L., et al., *Population-based carrier screening for cystic fibrosis: a systematic review of 23 years of research.* Genetics In Medicine, 2013. **16**: p. 207.
3. Busch, R., *On the history of cystic fibrosis.* Acta Univ Carol Med (Praha), 1990. **36**(1-4): p. 13-5.
4. Quinton, P.M., *Physiological Basis of Cystic Fibrosis: A Historical Perspective.* Physiological Reviews, 1999. **79**(1): p. S3-S22.
5. ANDERSEN, D.H., *CYSTIC FIBROSIS OF THE PANCREAS AND ITS RELATION TO CELIAC DISEASE: A CLINICAL AND PATHOLOGIC STUDY.* JAMA Pediatrics, 1938. **56**(2): p. 344-399.
6. Farber, S., H. Shwachman, and C.L. Maddock, *PANCREATIC FUNCTION AND DISEASE IN EARLY LIFE. I. PANCREATIC ENZYME ACTIVITY AND THE CELIAC SYNDROME.* J Clin Invest, 1943. **22**(6): p. 827-38.
7. ANDERSEN, D.H. and R.G. HODGES, *CELIAC SYNDROME: V. Genetics of Cystic Fibrosis of the Pancreas With a Consideration of Etiology.* JAMA Pediatrics, 1946. **72**(1): p. 62-80.
8. Kessler, W.R. and D.H. Andersen, *Heat prostration in fibrocystic disease of the pancreas and other conditions.* Pediatrics, 1951. **8**(5): p. 648-56.
9. Di Sant'Agnese, P.A., et al., *Abnormal electrolyte composition of sweat in cystic fibrosis of the pancreas; clinical significance and relationship to the disease.* Pediatrics, 1953. **12**(5): p. 549-63.
10. Quinton, P.M., *Chloride impermeability in cystic fibrosis.* Nature, 1983. **301**(5899): p. 421-2.
11. Boucher, R.C., et al., *Epithelial dysfunction in cystic fibrosis lung disease.* Lung, 1983. **161**(1): p. 1-17.
12. Kerem, B., et al., *Identification of the cystic fibrosis gene: genetic analysis.* Science, 1989. **245**(4922): p. 1073-80.
13. Riordan, J.R., et al., *Identification of the cystic fibrosis gene: cloning and characterization of complementary DNA.* Science, 1989. **245**(4922): p. 1066-73.
14. Rommens, J.M., et al., *Identification of the cystic fibrosis gene: chromosome walking and jumping.* Science, 1989. **245**(4922): p. 1059-65.
15. Ellsworth, R.E., et al., *Comparative genomic sequence analysis of the human and mouse cystic fibrosis transmembrane conductance regulator genes.* Proceedings of the National Academy of Sciences, 2000. **97**(3): p. 1172-1177.
16. Hyde, S.C., et al., *Structural model of ATP-binding proteing associated with cystic fibrosis, multidrug resistance and bacterial transport.* Nature, 1990. **346**(6282): p. 362-365.
17. Cheng, S.H., et al., *Phosphorylation of the R domain by cAMP-dependent protein kinase regulates the CFTR chloride channel.* Cell, 1991. **66**(5): p. 1027-36.
18. Anderson, M.P., et al., *Nucleoside triphosphates are required to open the CFTR chloride channel.* Cell, 1991. **67**(4): p. 775-84.
19. Smith, P.C., et al., *ATP Binding to the Motor Domain from an ABC Transporter Drives Formation of a Nucleotide Sandwich Dimer.* Molecular Cell, 2002. **10**(1): p. 139-149.
20. Procko, E., et al., *The mechanism of ABC transporters: general lessons from structural and functional studies of an antigenic peptide transporter.* Faseb j, 2009. **23**(5): p. 1287-302.

21. Rowe, S.M., S. Miller, and E.J. Sorscher, *Cystic Fibrosis.* New England Journal of Medicine, 2005. **352**(19): p. 1992-2001.
22. Ratjen, F. and G. Doring, *Cystic fibrosis.* Lancet, 2003. **361**(9358): p. 681-9.
23. Hoiby, N., *Hemophilus influenzae, Staphylococcus aureus, Pseudomonas cepacia, and Pseudomonas aeruginosa in patients with cystic fibrosis.* Chest, 1988. **94**(2 Suppl): p. 97s-103s.
24. Widdicombe, J.H., *Regulation of the depth and composition of airway surface liquid.* J Anat, 2002. **201**(4): p. 313-8.
25. Matsui, H., et al., *Evidence for Periciliary Liquid Layer Depletion, Not Abnormal Ion Composition, in the Pathogenesis of Cystic Fibrosis Airways Disease.* Cell, 1998. **95**(7): p. 1005-1015.
26. Widdicombe, J.G., *Airway Surface Liquid: Concepts and Measurements*, in *Airway Mucus: Basic Mechanisms and Clinical Perspectives*, D.F. Rogers and M.I. Lethem, Editors. 1997, Birkhäuser Basel: Basel. p. 1-17.
27. Smith, J.J., et al., *Cystic fibrosis airway epithelia fail to kill bacteria because of abnormal airway surface fluid.* Cell, 1996. **85**(2): p. 229-36.
28. Goldman, M.J., et al., *Human beta-defensin-1 is a salt-sensitive antibiotic in lung that is inactivated in cystic fibrosis.* Cell, 1997. **88**(4): p. 553-60.
29. Knowles, M.R., et al., *Ion composition of airway surface liquid of patients with cystic fibrosis as compared with normal and disease-control subjects.* J Clin Invest, 1997. **100**(10): p. 2588-95.
30. Matsui, H., et al., *Reduced three-dimensional motility in dehydrated airway mucus prevents neutrophil capture and killing bacteria on airway epithelial surfaces.* J Immunol, 2005. **175**(2): p. 1090-9.
31. Matsui, H., et al., *Evidence for periciliary liquid layer depletion, not abnormal ion composition, in the pathogenesis of cystic fibrosis airways disease.* Cell, 1998. **95**(7): p. 1005-15.
32. Verdugo, P., et al., *Molecular mechanism of mucin secretion: I. The role of intragranular charge shielding.* J Dent Res, 1987. **66**(2): p. 506-8.
33. Quinton, P.M., *Cystic fibrosis: impaired bicarbonate secretion and mucoviscidosis.* The Lancet, 2008. **372**(9636): p. 415-417.
34. Foundation, C.F. *The Clinical and Functional TRanslation of CFTR (CFTR2)*. 2017 [cited 2017 August]; Available from: https://www.cftr2.org.
35. De Boeck, K. and M.D. Amaral, *Progress in therapies for cystic fibrosis.* The Lancet Respiratory Medicine, 2016. **4**(8): p. 662-674.
36. Pettit, R.S. and C. Fellner, *CFTR modulators for the treatment of cystic fibrosis.* Pharmacy and Therapeutics, 2014. **39**(7): p. 500.
37. Dork, T., et al., *Characterization of a novel 21-kb deletion, CFTRdele2,3(21 kb), in the CFTR gene: a cystic fibrosis mutation of Slavic origin common in Central and East Europe.* Hum Genet, 2000. **106**(3): p. 259-68.
38. Farinha, C.M. and M.D. Amaral, *Most F508del-CFTR is targeted to degradation at an early folding checkpoint and independently of calnexin.* Molecular and cellular biology, 2005. **25**(12): p. 5242-5252.
39. Osborne, L., et al., *Incidence and expression of the N1303K mutation of the cystic fibrosis (CFTR) gene.* Hum Genet, 1992. **89**(6): p. 653-8.

40. Federici, S., et al., *[CFTR gene analyis in 207 patients with cystic fibrosis in southwest France: high frequency of N1303K and 1811+1.6bA>G mutations].* Arch Pediatr, 2001. **8**(2): p. 150-7.
41. Liu, F., et al., *Molecular Structure of the Human CFTR Ion Channel.* Cell, 2017. **169**(1): p. 85-95.e8.
42. Berger, A.L., et al., *Mutations That Change the Position of the Putative γ-Phosphate Linker in the Nucleotide Binding Domains of CFTR Alter Channel Gating.* Journal of Biological Chemistry, 2002. **277**(3): p. 2125-2131.
43. Bompadre, S.G., M. Li, and T.-C. Hwang, *Mechanism of G551D-CFTR (Cystic Fibrosis Transmembrane Conductance Regulator) Potentiation by a High Affinity ATP Analog.* Journal of Biological Chemistry, 2008. **283**(9): p. 5364-5369.
44. Bobadilla, J.L., et al., *Cystic fibrosis: a worldwide analysis of CFTR mutations--correlation with incidence data and application to screening.* Hum Mutat, 2002. **19**(6): p. 575-606.
45. Sheppard, D.N., et al., *Mutations in CFTR associated with mild-disease-form Cl- channels with altered pore properties.* Nature, 1993. **362**(6416): p. 160-164.
46. Highsmith, W.E., et al., *A Novel Mutation in the Cystic Fibrosis Gene in Patients with Pulmonary Disease but Normal Sweat Chloride Concentrations.* New England Journal of Medicine, 1994. **331**(15): p. 974-980.
47. Chiba-Falek, O., et al., *The Molecular Basis of Disease Variability among Cystic Fibrosis Patients Carrying the 3849+10 kb C→T Mutation*. Genomics, 1998. Vol. 53, 276-83.
48. Kerem, E., et al., *A cystic fibrosis transmembrane conductance regulator splice variant with partial penetrance associated with variable cystic fibrosis presentations.* Am J Respir Crit Care Med, 1997. **155**(6): p. 1914-20.
49. Haardt, M., et al., *C-terminal truncations destabilize the cystic fibrosis transmembrane conductance regulator without impairing its biogenesis. A novel class of mutation.* J Biol Chem, 1999. **274**(31): p. 21873-7.
50. Silvis, M.R., et al., *A Mutation in the Cystic Fibrosis Transmembrane Conductance Regulator Generates a Novel Internalization Sequence and Enhances Endocytic Rates.* Journal of Biological Chemistry, 2003. **278**(13): p. 11554-11560.
51. Ahmed, N., et al., *Molecular consequences of cystic fibrosis transmembrane regulator (CFTR) gene mutations in the exocrine pancreas.* Gut, 2003. **52**(8): p. 1159-1164.
52. Rogan, M.P., D.A. Stoltz, and D.B. Hornick, *Cystic fibrosis transmembrane conductance regulator intracellular processing, trafficking, and opportunities for mutation-specific treatment.* Chest Journal, 2011. **139**(6): p. 1480-1490.
53. Lopes-Pacheco, M., *CFTR Modulators: Shedding Light on Precision Medicine for Cystic Fibrosis.* Frontiers in Pharmacology, 2016. **7**: p. 275.
54. Hanukoglu, I. and A. Hanukoglu, *Epithelial sodium channel (ENaC) family: Phylogeny, structure–function, tissue distribution, and associated inherited diseases.* Gene, 2016. **579**(2): p. 95-132.
55. Waldmann, R., et al., *Molecular cloning and functional expression of a novel amiloride-sensitive Na+ channel.* J Biol Chem, 1995. **270**(46): p. 27411-4.
56. Edelheit, O., et al., *Novel mutations in epithelial sodium channel (ENaC) subunit genes and phenotypic expression of multisystem pseudohypoaldosteronism.* Clin Endocrinol (Oxf), 2005. **62**(5): p. 547-53.
57. Hughey, R.P., et al., *Epithelial Sodium Channels Are Activated by Furin-dependent Proteolysis.* Journal of Biological Chemistry, 2004. **279**(18): p. 18111-18114.

58. Vuagniaux, G., et al., *Activation of the amiloride-sensitive epithelial sodium channel by the serine protease mCAP1 expressed in a mouse cortical collecting duct cell line.* J Am Soc Nephrol, 2000. **11**(5): p. 828-34.
59. Bhalla, V. and K.R. Hallows, *Mechanisms of ENaC Regulation and Clinical Implications.* Journal of the American Society of Nephrology, 2008. **19**(10): p. 1845.
60. Tarran, R., et al., *The Relative Roles of Passive Surface Forces and Active Ion Transport in the Modulation of Airway Surface Liquid Volume and Composition.* The Journal of General Physiology, 2001. **118**(2): p. 223.
61. Tarran, R., et al., *Normal and Cystic Fibrosis Airway Surface Liquid Homeostasis: The Effects of Phasic Shear Stress and Viral Infections.* Journal of Biological Chemistry, 2005. **280**(42): p. 35751-35759.
62. Mall, M., et al., *Increased airway epithelial Na+ absorption produces cystic fibrosis-like lung disease in mice.* Nature Medicine, 2004. **10**(5): p. 487-493.
63. Yan, W., et al., *Cystic fibrosis transmembrane conductance regulator differentially regulates human and mouse epithelial sodium channels in Xenopus oocytes.* J Biol Chem, 2004. **279**(22): p. 23183-92.
64. Suaud, L., et al., *Genistein restores functional interactions between Delta F508-CFTR and ENaC in Xenopus oocytes.* J Biol Chem, 2002. **277**(11): p. 8928-33.
65. Berdiev, B.K., et al., *Molecular proximity of cystic fibrosis transmembrane conductance regulator and epithelial sodium channel assessed by fluorescence resonance energy transfer.* J Biol Chem, 2007. **282**(50): p. 36481-8.
66. Ismailov, II, et al., *Role of actin in regulation of epithelial sodium channels by CFTR.* Am J Physiol, 1997. **272**(4 Pt 1): p. C1077-86.
67. Snyder, P.M., et al., *Multiple WW domains, but not the C2 domain, are required for inhibition of the epithelial Na+ channel by human Nedd4.* J Biol Chem, 2001. **276**(30): p. 28321-6.
68. Reddy, M.M. and P.M. Quinton, *Functional interaction of CFTR and ENaC in sweat glands.* Pflugers Arch, 2003. **445**(4): p. 499-503.
69. Sanders, K.M., S.D. Koh, and S.M. Ward, *Interstitial cells of cajal as pacemakers in the gastrointestinal tract.* Annu Rev Physiol, 2006. **68**: p. 307-43.
70. Large, W.A. and Q. Wang, *Characteristics and physiological role of the Ca(2+)-activated Cl- conductance in smooth muscle.* Am J Physiol, 1996. **271**(2 Pt 1): p. C435-54.
71. Melvin, J.E., et al., *Regulation of fluid and electrolyte secretion in salivary gland acinar cells.* Annu Rev Physiol, 2005. **67**: p. 445-69.
72. Caputo, A., et al., *TMEM16A, A Membrane Protein Associated with Calcium-Dependent Chloride Channel Activity.* Science, 2008. **322**(5901): p. 590.
73. Yang, Y.D., et al., *TMEM16A confers receptor-activated calcium-dependent chloride conductance.* Nature, 2008. **455**: p. 1210.
74. Yu, K., et al., *Explaining Calcium-Dependent Gating of Anoctamin-1 Chloride Channels Requires a Revised Topology.* Circulation Research, 2012. **110**(7): p. 990-999.
75. Terashima, H., A. Picollo, and A. Accardi, *Purified TMEM16A is sufficient to form Ca2+-activated Cl− channels.* Proceedings of the National Academy of Sciences, 2013. **110**(48): p. 19354.
76. Yang, T. and H.M. Colecraft, *Calmodulin regulation of TMEM16A and 16B Ca(2+)-activated chloride channels.* Channels (Austin, Tex.), 2016. **10**(1): p. 38-44.

77. Rock, J.R., et al., *Transmembrane protein 16A (TMEM16A) is a Ca2+-regulated Cl- secretory channel in mouse airways.* The Journal of Biological Chemistry, 2009. **284**(22): p. 14875-14880.
78. Seidler, U., et al., *The role of the NHERF family of PDZ scaffolding proteins in the regulation of salt and water transport.* Ann N Y Acad Sci, 2009. **1165**: p. 249-60.
79. Benedetto, R., et al., *Epithelial Chloride Transport by CFTR Requires TMEM16A.* Scientific Reports, 2017. **7**(1): p. 12397-12397.
80. Faria, D., R. Schreiber, and K. Kunzelmann, *CFTR is activated through stimulation of purinergic P2Y2 receptors.* Pflügers Archiv - European Journal of Physiology, 2008. **457**(6): p. 1373.
81. Grubb, B.R., R.N. Vick, and R.C. Boucher, *Hyperabsorption of Na+ and raised Ca(2+)-mediated Cl- secretion in nasal epithelia of CF mice.* Am J Physiol, 1994. **266**(5 Pt 1): p. C1478-83.
82. Ruffin, M., et al., *Anoctamin 1 dysregulation alters bronchial epithelial repair in cystic fibrosis.* Biochimica et Biophysica Acta (BBA) - Molecular Basis of Disease, 2013. **1832**(12): p. 2340-2351.
83. Knowles, M.R., L.L. Clarke, and R.C. Boucher, *Activation by Extracellular Nucleotides of Chloride Secretion in the Airway Epithelia of Patients with Cystic Fibrosis.* New England Journal of Medicine, 1991. **325**(8): p. 533-538.
84. Tarran, R., et al., *Regulation of Murine Airway Surface Liquid Volume by CFTR and Ca²⁺-activated Cl[−] Conductances.* The Journal of General Physiology, 2002. **120**(3): p. 407.
85. Gruenert, D.C., et al., *Established cell lines used in cystic fibrosis research.* Journal of Cystic Fibrosis, 2004. **3**: p. 191-196.
86. Lechner, J.F., et al., *Clonal growth of normal adult human bronchial epithelial cells in a serum-free medium.* In Vitro, 1982. **18**(7): p. 633-42.
87. Astrand, A.B., et al., *Linking increased airway hydration, ciliary beating, and mucociliary clearance through ENaC inhibition.* Am J Physiol Lung Cell Mol Physiol, 2015. **308**(1): p. L22-32.
88. Choi, H.C., C.S. Kim, and R. Tarran, *Automated acquisition and analysis of airway surface liquid height by confocal microscopy.* Am J Physiol Lung Cell Mol Physiol, 2015. **309**(2): p. L109-18.
89. Corvol, H., et al., *Genome-wide association meta-analysis identifies five modifier loci of lung disease severity in cystic fibrosis.* Nat Commun, 2015. **6**: p. 8382.
90. Strug, L.J., et al., *Cystic fibrosis gene modifier SLC26A9 modulates airway response to CFTR-directed therapeutics.* Hum Mol Genet, 2016.
91. Miyashita, H., et al., *Long-term maintenance of limbal epithelial progenitor cells using rho kinase inhibitor and keratinocyte growth factor.* Stem Cells Transl Med, 2013. **2**(10): p. 758-65.
92. Liu, X., et al., *ROCK inhibitor and feeder cells induce the conditional reprogramming of epithelial cells.* The American journal of pathology, 2012. **180**(2): p. 599-607.
93. Butler, C.R., et al., *Rapid Expansion of Human Epithelial Stem Cells Suitable for Airway Tissue Engineering.* Am J Respir Crit Care Med, 2016. **194**(2): p. 156-68.
94. Mou, H., et al., *Dual SMAD Signaling Inhibition Enables Long-Term Expansion of Diverse Epithelial Basal Cells.* Cell stem cell, 2016. **19**(2): p. 217-231.
95. Fulcher, M.L., et al., *Novel human bronchial epithelial cell lines for cystic fibrosis research.* Am J Physiol Lung Cell Mol Physiol, 2009. **296**(1): p. L82-91.

96. Munye, M.M., et al., *BMI-1 extends proliferative potential of human bronchial epithelial cells whilst retaining their mucociliary differentiation capacity.* Am J Physiol Lung Cell Mol Physiol, 2016: p. ajplung 00471 2016.
97. Fatehullah, A., S.H. Tan, and N. Barker, *Organoids as an in vitro model of human development and disease.* Nat Cell Biol, 2016. **18**(3): p. 246-54.
98. Dye, B.R., et al., *In vitro generation of human pluripotent stem cell derived lung organoids.* Elife, 2015. **4**.
99. Takasato, M., et al., *Kidney organoids from human iPS cells contain multiple lineages and model human nephrogenesis.* Nature, 2015. **526**(7574): p. 564-8.
100. Paşca, S.P., *Assembling human brain organoids.* Science, 2019. **363**(6423): p. 126.
101. Schwank, G., et al., *Functional repair of CFTR by CRISPR/Cas9 in intestinal stem cell organoids of cystic fibrosis patients.* Cell Stem Cell, 2013. **13**(6): p. 653-8.
102. Yui, S., et al., *Functional engraftment of colon epithelium expanded in vitro from a single adult Lgr5(+) stem cell.* Nat Med, 2012. **18**(4): p. 618-23.
103. Hynds, R.E., et al., *Expansion of Human Airway Basal Stem Cells and Their Differentiation as 3D Tracheospheres.* Methods Mol Biol, 2016.
104. Guimbellot, J.S., et al., *Nasospheroids permit measurements of CFTR-dependent fluid transport.* JCI insight, 2017. **2**(22): p. e95734.
105. Snouwaert, J.N., et al., *An Animal Model for Cystic Fibrosis Made by Gene Targeting.* Science, 1992. **257**(5073): p. 1083.
106. Ratcliff, R., et al., *Production of a severe cystic fibrosis mutation in mice by gene targeting.* Nature Genetics, 1993. **4**(1): p. 35-41.
107. Rozmahe, R., et al., *Modulation of disease severity in cystic fibrosis transmembrane conductance regulator deficient mice by a secondary genetic factor.* Nature Genetics, 1996. **12**(3): p. 280-287.
108. Hasty, P., et al., *Severe phenotype in mice with termination mutation in exon 2 of cystic fibrosis gene.* Somatic Cell and Molecular Genetics, 1995. **21**(3): p. 177-187.
109. Zeiher, B.G., et al., *A mouse model for the delta F508 allele of cystic fibrosis.* The Journal of Clinical Investigation, 1995. **96**(4): p. 2051-2064.
110. van Doorninck, J.H., et al., *A mouse model for the cystic fibrosis delta F508 mutation.* The EMBO Journal, 1995. **14**(18): p. 4403-4411.
111. Delaney, S.J., et al., *Cystic fibrosis mice carrying the missense mutation G551D replicate human genotype-phenotype correlations.* The EMBO journal, 1996. **15**(5): p. 955-963.
112. Ellsworth, R.E., et al., *Comparative genomic sequence analysis of the human and mouse cystic fibrosis transmembrane conductance regulator genes.* Proceedings of the National Academy of Sciences of the United States of America, 2000. **97**(3): p. 1172-1177.
113. Davidson, D.J. and M. Rolfe, *Mouse models of cystic fibrosis.* Trends Genet, 2001. **17**(10): p. S29-37.
114. Mall, M., et al., *Increased airway epithelial Na+ absorption produces cystic fibrosis-like lung disease in mice.* Nat Med, 2004. **10**(5): p. 487-93.
115. Mall, M.A., et al., *Development of chronic bronchitis and emphysema in beta-epithelial Na+ channel-overexpressing mice.* Am J Respir Crit Care Med, 2008. **177**(7): p. 730-42.
116. Ostedgaard, L.S., et al., *The DeltaF508 mutation causes CFTR misprocessing and cystic fibrosis-like disease in pigs.* Sci Transl Med, 2011. **3**(74): p. 74ra24.

117. Rogers, C.S., et al., *Production of CFTR-null and CFTR-DeltaF508 heterozygous pigs by adeno-associated virus-mediated gene targeting and somatic cell nuclear transfer.* J Clin Invest, 2008. **118**(4): p. 1571-7.
118. Stoltz, D.A., et al., *Intestinal CFTR expression alleviates meconium ileus in cystic fibrosis pigs.* J Clin Invest, 2013. **123**(6): p. 2685-93.
119. Sun, X., et al., *Disease phenotype of a ferret CFTR-knockout model of cystic fibrosis.* J Clin Invest, 2010. **120**(9): p. 3149-60.
120. Lavelle, G.M., et al., *Animal Models of Cystic Fibrosis Pathology: Phenotypic Parallels and Divergences.* BioMed research international, 2016. **2016**: p. 5258727-5258727.
121. Mogayzel, P.J., et al., *Improving chronic care delivery and outcomes: the impact of the cystic fibrosis Care Center Network.* BMJ Quality & Safety, 2014. **23**(Suppl 1): p. i3.
122. Donaldson, S.H., et al., *Mucus clearance and lung function in cystic fibrosis with hypertonic saline.* N Engl J Med, 2006. **354**(3): p. 241-50.
123. Konstan, M.W. and F. Ratjen, *Effect of dornase alfa on inflammation and lung function: potential role in the early treatment of cystic fibrosis.* J Cyst Fibros, 2012. **11**(2): p. 78-83.
124. Ermund, A., et al., *OligoG CF-5/20 normalizes cystic fibrosis mucus by chelating calcium.* Clin Exp Pharmacol Physiol, 2017. **44**(6): p. 639-647.
125. Lyczak, J.B., C.L. Cannon, and G.B. Pier, *Lung Infections Associated with Cystic Fibrosis.* Clinical Microbiology Reviews, 2002. **15**(2): p. 194.
126. Fajac, I. and K. De Boeck, *New horizons for cystic fibrosis treatment.* Pharmacol Ther, 2017. **170**: p. 205-211.
127. Whiting, P., et al., *Ivacaftor for the treatment of patients with cystic fibrosis and the G551D mutation: a systematic review and cost-effectiveness analysis.* Health Technol Assess, 2014. **18**(18): p. 1-106.
128. Partridge, M., *FDA Approves KALYDECO® (ivacaftor) for More Than 600 People Ages 2 and Older With Cystic Fibrosis Who Have Certain Residual Function Mutations*. 2017, Vertex Pharmaceuticals Vertex Pharmaceuticals Incorporated.
129. Partridge, M., *FDA Approves KALYDECO® (ivacaftor) as First and Only CFTR Modulator to Treat Eligible Infants with CF as Early as Six Months of Age*. 2019, Vertex Pharmaceuticals Incorporated: Vertex.
130. Van Goor, F., et al., *Rescue of DeltaF508-CFTR trafficking and gating in human cystic fibrosis airway primary cultures by small molecules.* Am J Physiol Lung Cell Mol Physiol, 2006. **290**(6): p. L1117-30.
131. Farinha, C.M. and M.D. Amaral, *Most F508del-CFTR is targeted to degradation at an early folding checkpoint and independently of calnexin.* Mol Cell Biol, 2005. **25**(12): p. 5242-52.
132. Jensen, T.J., et al., *Multiple proteolytic systems, including the proteasome, contribute to CFTR processing.* Cell, 1995. **83**(1): p. 129-35.
133. Lukacs, G.L., et al., *The delta F508 mutation decreases the stability of cystic fibrosis transmembrane conductance regulator in the plasma membrane. Determination of functional half-lives on transfected cells.* J Biol Chem, 1993. **268**(29): p. 21592-8.
134. Sharma, M., et al., *Conformational and temperature-sensitive stability defects of the delta F508 cystic fibrosis transmembrane conductance regulator in post-endoplasmic reticulum compartments.* J Biol Chem, 2001. **276**(12): p. 8942-50.

135. Clancy, J.P., et al., *Results of a phase IIa study of VX-809, an investigational CFTR corrector compound, in subjects with cystic fibrosis homozygous for the F508del-CFTR mutation.* Thorax, 2012. **67**(1): p. 12-8.
136. Wainwright, C.E., et al., *Lumacaftor-Ivacaftor in Patients with Cystic Fibrosis Homozygous for Phe508del CFTR.* N Engl J Med, 2015.
137. Wainwright, C.E., et al., *Lumacaftor-Ivacaftor in Patients with Cystic Fibrosis Homozygous for Phe508del CFTR.* N Engl J Med, 2015. **373**(3): p. 220-31.
138. England, N., *NHS England concludes wide-ranging deal for cystic fibrosis drugs*. 2019.
139. Taylor-Cousar, J.L., et al., *Tezacaftor-Ivacaftor in Patients with Cystic Fibrosis Homozygous for Phe508del.* N Engl J Med, 2017. **377**(21): p. 2013-2023.
140. Talamo Guevara, M. and S.A. McColley, *The safety of lumacaftor and ivacaftor for the treatment of cystic fibrosis.* Expert Opin Drug Saf, 2017. **16**(11): p. 1305-1311.
141. Martiniano, S.L., et al., *Highlights from the 2017 North American Cystic Fibrosis Conference.* Pediatric Pulmonology, 2018. **53**(7): p. 979-986.
142. *FDA approves new breakthrough therapy for cystic fibrosis*, in *Treatment approved for approximately 90% of patients with cystic fibrosis, many of whom had no approved therapeutic options*, F.n. release, Editor. 2019, FDA: U.S Food and Drug Administration
143. McDonald, C.M., et al., *Ataluren in patients with nonsense mutation Duchenne muscular dystrophy (ACT DMD): a multicentre, randomised, double-blind, placebo-controlled, phase 3 trial.* Lancet, 2017. **390**(10101): p. 1489-1498.
144. Zhang, S., et al., *Sinupret activates CFTR and TMEM16A-dependent transepithelial chloride transport and improves indicators of mucociliary clearance.* PLoS One, 2014. **9**(8): p. e104090.
145. Anagnostopoulou, P., et al., *SLC26A9-mediated chloride secretion prevents mucus obstruction in airway inflammation.* J Clin Invest, 2012. **122**(10): p. 3629-34.
146. Mall, M.A. and L.J. Galietta, *Targeting ion channels in cystic fibrosis.* J Cyst Fibros, 2015. **14**(5): p. 561-70.
147. Haq, I.J., et al., *Airway surface liquid homeostasis in cystic fibrosis: pathophysiology and therapeutic targets.* Thorax, 2016. **71**(3): p. 284-7.
148. Koshland, D.E., *The Cystic Fibrosis Gene Story.* Science, 1989. **245**(4922): p. 1029.
149. Johnson, L.G., et al., *Efficiency of gene transfer for restoration of normal airway epithelial function in cystic fibrosis.* Nat Genet, 1992. **2**(1): p. 21-5.
150. Zhang, L., et al., *CFTR delivery to 25% of surface epithelial cells restores normal rates of mucus transport to human cystic fibrosis airway epithelium.* PLoS Biol, 2009. **7**(7): p. e1000155.
151. Farmen, S.L., et al., *Gene transfer of CFTR to airway epithelia: low levels of expression are sufficient to correct Cl- transport and overexpression can generate basolateral CFTR.* Am J Physiol Lung Cell Mol Physiol, 2005. **289**(6): p. L1123-30.
152. Griesenbach, U., K.M. Pytel, and E.W. Alton, *Cystic Fibrosis Gene Therapy in the UK and Elsewhere.* Hum Gene Ther, 2015. **26**(5): p. 266-75.
153. Ku, C.A. and M.E. Pennesi, *Retinal Gene Therapy: Current Progress and Future Prospects.* Expert Rev Ophthalmol, 2015. **10**(3): p. 281-299.
154. Perrin, G.Q., R.W. Herzog, and D.M. Markusic, *Update on clinical gene therapy for hemophilia.* Blood, 2018: p. blood-2018-07-820720.
155. RELEASE, F.N., *FDA approves novel gene therapy to treat patients with a rare form of inherited vision loss*. 2017.

156. Button, B., et al., *A periciliary brush promotes the lung health by separating the mucus layer from airway epithelia.* Science, 2012. **337**(6097): p. 937-41.
157. Duncan, G.A., et al., *The Mucus Barrier to Inhaled Gene Therapy.* Mol Ther, 2016. **24**(12): p. 2043-2053.
158. Boucher, R.C., et al., *Gene therapy for cystic fibrosis using E1-deleted adenovirus: a phase I trial in the nasal cavity. The University of North Carolina at Chapel Hill.* Hum Gene Ther, 1994. **5**(5): p. 615-39.
159. Welsh, M.J., et al., *Cystic fibrosis gene therapy using an adenovirus vector: in vivo safety and efficacy in nasal epithelium.* Hum Gene Ther, 1994. **5**(2): p. 209-19.
160. Cavazzana-Calvo, M., et al., *Gene therapy of human severe combined immunodeficiency (SCID)-X1 disease.* Science, 2000. **288**(5466): p. 669-72.
161. Hacein-Bey-Abina, S., et al., *A serious adverse event after successful gene therapy for X-linked severe combined immunodeficiency.* N Engl J Med, 2003. **348**(3): p. 255-6.
162. McCarthy, M., *US FDA rejects defence of gene-therapy trial.* The Lancet, 2000. **355**(9208): p. 997.
163. Miyoshi, H., et al., *Development of a self-inactivating lentivirus vector.* J Virol, 1998. **72**(10): p. 8150-7.
164. Montini, E., et al., *The genotoxic potential of retroviral vectors is strongly modulated by vector design and integration site selection in a mouse model of HSC gene therapy.* J Clin Invest, 2009. **119**(4): p. 964-75.
165. Farrow, N., et al., *Airway gene transfer in a non-human primate: lentiviral gene expression in marmoset lungs.* Sci Rep, 2013. **3**: p. 1287.
166. Miah, K.M., S.C. Hyde, and D.R. Gill, *Emerging gene therapies for cystic fibrosis.* Expert Review of Respiratory Medicine, 2019. **13**(8): p. 709-725.
167. Cooney, A.L., et al., *Lentiviral-mediated phenotypic correction of cystic fibrosis pigs.* JCI Insight, 2016. **1**(14): p. e88730.
168. Griesenbach, U., J.C. Davies, and E. Alton, *Cystic fibrosis gene therapy: a mutation-independent treatment.* Curr Opin Pulm Med, 2016. **22**(6): p. 602-9.
169. Alton, E.W., et al., *Preparation for a first-in-man lentivirus trial in patients with cystic fibrosis.* Thorax, 2017. **72**(2): p. 137-147.
170. *Oxford BioMedica, UK Cystic Fibrosis Gene Therapy Consortium, Boehringer Ingelheim and Imperial Innovations Form Partnership to Develop a Novel Gene Therapy Treatment for Cystic Fibrosis*. 2018, Oxford Biomedica Oxford Biomedica.
171. Yang, J., et al., *Concatamerization of adeno-associated virus circular genomes occurs through intermolecular recombination.* J Virol, 1999. **73**(11): p. 9468-77.
172. Moss, R.B., et al., *Repeated adeno-associated virus serotype 2 aerosol-mediated cystic fibrosis transmembrane regulator gene transfer to the lungs of patients with cystic fibrosis: a multicenter, double-blind, placebo-controlled trial.* Chest, 2004. **125**(2): p. 509-21.
173. Moss, R.B., et al., *Repeated aerosolized AAV-CFTR for treatment of cystic fibrosis: a randomized placebo-controlled phase 2B trial.* Hum Gene Ther, 2007. **18**(8): p. 726-32.
174. Kurosaki, F., et al., *Optimization of adeno-associated virus vector-mediated gene transfer to the respiratory tract.* Gene Ther, 2017. **24**: p. 290-297.
175. McClain, L.E., et al., *Vector serotype screening for use in ovine perinatal lung gene therapy.* J Pediatr Surg, 2016. **51**(6): p. 879-84.
176. Masat, E., G. Pavani, and F. Mingozzi, *Humoral immunity to AAV vectors in gene therapy: challenges and potential solutions.* Discov Med, 2013. **15**(85): p. 379-89.

177. Halbert, C.L., et al., *Prevalence of neutralizing antibodies against adeno-associated virus (AAV) types 2, 5, and 6 in cystic fibrosis and normal populations: Implications for gene therapy using AAV vectors.* Hum Gene Ther, 2006. **17**(4): p. 440-7.
178. Yan, Z., et al., *A novel chimeric adenoassociated virus 2/human bocavirus 1 parvovirus vector efficiently transduces human airway epithelia.* Molecular Therapy : the journal of the American Society of Gene Therapy, 2013. **21**(12): p. 2181-2194.
179. Yan, Z., et al., *Human Bocavirus Type-1 Capsid Facilitates the Transduction of Ferret Airways by Adeno-Associated Virus Genomes.* Hum Gene Ther, 2017. **28**(8): p. 612-625.
180. Yan, Z., et al., *Establishment of a High-Yield Recombinant Adeno-Associated Virus/Human Bocavirus Vector Production System Independent of Bocavirus Nonstructural Proteins.* Hum Gene Ther, 2019. **30**(5): p. 556-570.
181. Margineanu, D.G., *Equilibrium and non-equilibrium approaches in biomembrane thermodynamics.* Arch Int Physiol Biochim, 1987. **95**(5): p. 381-422.
182. Khalil, I.A., et al., *Uptake pathways and subsequent intracellular trafficking in nonviral gene delivery.* Pharmacol Rev, 2006. **58**(1): p. 32-45.
183. Felgner, P.L., et al., *Lipofection: a highly efficient, lipid-mediated DNA-transfection procedure.* Proceedings of the National Academy of Sciences of the United States of America, 1987. **84**(21): p. 7413-7417.
184. Farhood, H., N. Serbina, and L. Huang, *The role of dioleoyl phosphatidylethanolamine in cationic liposome mediated gene transfer.* Biochimica et Biophysica Acta (BBA) - Biomembranes, 1995. **1235**(2): p. 289-295.
185. Coderch, L., et al., *Influence of cholesterol on liposome fluidity by EPR. Relationship with percutaneous absorption.* J Control Release, 2000. **68**(1): p. 85-95.
186. Semple, S.C., A. Chonn, and P.R. Cullis, *Influence of Cholesterol on the Association of Plasma Proteins with Liposomes.* Biochemistry, 1996. **35**(8): p. 2521-2525.
187. Mishra, S., P. Webster, and M.E. Davis, *PEGylation significantly affects cellular uptake and intracellular trafficking of non-viral gene delivery particles.* European Journal of Cell Biology, 2004. **83**(3): p. 97-111.
188. Huang, F.-W., et al., *PEGylated PEI-based biodegradable polymers as non-viral gene vectors.* Acta Biomaterialia, 2010. **6**(11): p. 4285-4295.
189. Troiber, C. and E. Wagner, *Nucleic acid carriers based on precise polymer conjugates.* Bioconjugate chemistry, 2011. **22**(9): p. 1737-1752.
190. Caplen, N.J., et al., *Liposome-mediated CFTR gene transfer to the nasal epithelium of patients with cystic fibrosis.* Nature Medicine, 1995. **1**(1): p. 39-46.
191. McLachlan, G., et al., *Laboratory and clinical studies in support of cystic fibrosis gene therapy using pCMV-CFTR-DOTAP.* Gene Ther, 1996. **3**(12): p. 1113-23.
192. Gill, D.R., et al., *A placebo-controlled study of liposome-mediated gene transfer to the nasal epithelium of patients with cystic fibrosis.* Gene Ther, 1997. **4**(3): p. 199-209.
193. Porteous, D.J., et al., *Evidence for safety and efficacy of DOTAP cationic liposome mediated CFTR gene transfer to the nasal epithelium of patients with cystic fibrosis.* Gene Ther, 1997. **4**(3): p. 210-8.
194. Knowles, M.R., et al., *A double-blind, placebo controlled, dose ranging study to evaluate the safety and biological efficacy of the lipid-DNA complex GR213487B in the nasal epithelium of adult patients with cystic fibrosis.* Hum Gene Ther, 1998. **9**(2): p. 249-69.

195. Hyde, S.C., et al., *Repeat administration of DNA/liposomes to the nasal epithelium of patients with cystic fibrosis.* Gene Ther, 2000. **7**(13): p. 1156-65.
196. Noone, P.G., et al., *Safety and biological efficacy of a lipid-CFTR complex for gene transfer in the nasal epithelium of adult patients with cystic fibrosis.* Mol Ther, 2000. **1**(1): p. 105-14.
197. Ziady, A.G., et al., *Transfection of airway epithelium by stable PEGylated poly-L-lysine DNA nanoparticles in vivo.* Mol Ther, 2003. **8**(6): p. 936-47.
198. Middleton, P.G., et al., *Nasal application of the cationic liposome DC-Chol:DOPE does not alter ion transport, lung function or bacterial growth.* Eur Respir J, 1994. **7**(3): p. 442-5.
199. Hyde, S.C., et al., *Repeat administration of DNA/liposomes to the nasal epithelium of patients with cystic fibrosis.* Gene Therapy, 2000. **7**(13): p. 1156-1165.
200. Hyde, S.C., et al., *CpG-free plasmids confer reduced inflammation and sustained pulmonary gene expression.* Nat Biotechnol, 2008. **26**(5): p. 549-51.
201. Lee, E.R., et al., *Detailed analysis of structures and formulations of cationic lipids for efficient gene transfer to the lung.* Human gene therapy, 1996. **7**(14): p. 1701-1717.
202. Chadwick, S., et al., *Safety of a single aerosol administration of escalating doses of the cationic lipid GL-67/DOPE/DMPE-PEG 5000 formulation to the lungs of normal volunteers.* Gene therapy, 1997. **4**(9): p. 937.
203. Gill, D.R., et al., *Increased persistence of lung gene expression using plasmids containing the ubiquitin C or elongation factor 1alpha promoter.* Gene Ther, 2001. **8**(20): p. 1539-46.
204. Alton, E.W., et al., *A phase I/IIa safety and efficacy study of nebulized liposome-mediated gene therapy for cystic fibrosis supports a multidose trial.* American journal of respiratory and critical care medicine, 2015. **192**(11): p. 1389-1392.
205. Alton, E.W., et al., *Repeated nebulisation of non-viral CFTR gene therapy in patients with cystic fibrosis: a randomised, double-blind, placebo-controlled, phase 2b trial.* The Lancet Respiratory Medicine, 2015. **3**(9): p. 684-691.
206. Johler, S.M., et al., *Nebulisation of IVT mRNA Complexes for Intrapulmonary Administration.* PLoS One, 2015. **10**(9): p. e0137504.
207. Bangel-Ruland, N., et al., *Cystic fibrosis transmembrane conductance regulator-mRNA delivery: a novel alternative for cystic fibrosis gene therapy.* J Gene Med, 2013. **15**(11-12): p. 414-26.
208. Anderson, B.R., et al., *Nucleoside modifications in RNA limit activation of 2'-5'-oligoadenylate synthetase and increase resistance to cleavage by RNase L.* Nucleic Acids Res, 2011. **39**(21): p. 9329-38.
209. Strenkowska, M., et al., *Cap analogs modified with 1,2-dithiodiphosphate moiety protect mRNA from decapping and enhance its translational potential.* Nucleic Acids Res, 2016. **44**(20): p. 9578-9590.
210. Sahin, U., K. Kariko, and O. Tureci, *mRNA-based therapeutics--developing a new class of drugs.* Nat Rev Drug Discov, 2014. **13**(10): p. 759-80.
211. Mauro, V.P. and S.A. Chappell, *A critical analysis of codon optimization in human therapeutics.* Trends Mol Med, 2014. **20**(11): p. 604-13.
212. Kormann, M.S., et al., *Expression of therapeutic proteins after delivery of chemically modified mRNA in mice.* Nat Biotechnol, 2011. **29**(2): p. 154-7.

213. Heda, G.D., M. Tanwani, and C.R. Marino, *The Delta F508 mutation shortens the biochemical half-life of plasma membrane CFTR in polarized epithelial cells.* Am J Physiol Cell Physiol, 2001. **280**(1): p. C166-74.
214. Probst, J., et al., *Spontaneous cellular uptake of exogenous messenger RNA in vivo is nucleic acid-specific, saturable and ion dependent.* Gene Ther, 2007. **14**(15): p. 1175-80.
215. Zeyer, F., et al., *mRNA-Mediated Gene Supplementation of Toll-Like Receptors as Treatment Strategy for Asthma In Vivo.* PLoS One, 2016. **11**(4): p. e0154001.
216. Mahiny, A.J., et al., *In vivo genome editing using nuclease-encoding mRNA corrects SP-B deficiency.* Nat Biotechnol, 2015. **33**(6): p. 584-6.
217. Antony, J.S., et al., *Modified mRNA as a new therapeutic option for pediatric respiratory diseases and hemoglobinopathies.* Mol Cell Pediatr, 2015. **2**(1): p. 11.
218. Manunta, M.D.I., et al., *Nebulisation of Receptor-Targeted Nanocomplexes for Gene Delivery to the Airway Epithelium.* PLoS ONE, 2011. **6**(10).
219. Manunta, M.D., et al., *Airway deposition of nebulized gene delivery nanocomplexes monitored by radioimaging agents.* Am J Respir Cell Mol Biol, 2013. **49**(3): p. 471-80.
220. Davies, L.A., et al., *Aerosol delivery of DNA/liposomes to the lung for cystic fibrosis gene therapy.* Hum Gene Ther Clin Dev, 2014. **25**(2): p. 97-107.
221. Boucher, R.C., *New concepts of the pathogenesis of cystic fibrosis lung disease.* Eur Respir J, 2004. **23**(1): p. 146-58.
222. Birket, S.E., et al., *A functional anatomic defect of the cystic fibrosis airway.* Am J Respir Crit Care Med, 2014. **190**(4): p. 421-32.
223. Tang, X.X., et al., *Acidic pH increases airway surface liquid viscosity in cystic fibrosis.* J Clin Invest, 2016. **126**(3): p. 879-91.
224. Shah, P.L., et al., *In vivo effects of recombinant human DNase I on sputum in patients with cystic fibrosis.* Thorax, 1996. **51**(2): p. 119-25.
225. Suk, J.S., et al., *N-acetylcysteine enhances cystic fibrosis sputum penetration and airway gene transfer by highly compacted DNA nanoparticles.* Mol Ther, 2011. **19**(11): p. 1981-9.
226. Castellani, S., et al., *Magnetofection Enhances Lentiviral-Mediated Transduction of Airway Epithelial Cells through Extracellular and Cellular Barriers.* Genes (Basel), 2016. **7**(11).
227. Suk, J.S., et al., *Lung gene therapy with highly compacted DNA nanoparticles that overcome the mucus barrier.* J Control Release, 2014. **178**: p. 8-17.
228. Engelhardt, J.F., et al., *Submucosal glands are the predominant site of CFTR expression in the human bronchus.* Nature Genetics, 1992. **2**: p. 240-248.
229. Pilewski, J.M., et al., *Adenovirus-mediated gene transfer to human bronchial submucosal glands using xenografts.* Am J Physiol, 1995. **268**(4 Pt 1): p. L657-65.
230. MacVinish, L.J., et al., *Normalization of ion transport in murine cystic fibrosis nasal epithelium using gene transfer.* Am J Physiol, 1997. **273**(2 Pt 1): p. C734-40.
231. Ziady, A.-G., et al., *Functional evidence of CFTR gene transfer in nasal epithelium of cystic fibrosis mice in vivo following lumenal application of DNA complexes targeted to the serpin-enzyme complex receptor.* Mol Ther, 2002. **5**(4): p. 413-419.
232. Rawlins, E.L. and B.L. Hogan, *Ciliated epithelial cell lifespan in the mouse trachea and lung.* Am J Physiol Lung Cell Mol Physiol, 2008. **295**(1): p. L231-4.
233. Liu, G., et al., *Nanoparticles of compacted DNA transfect postmitotic cells.* J Biol Chem, 2003. **278**(35): p. 32578-86.

234. Meng, Q., et al., *Efficient transfection of non-proliferating human airway epithelial cells with a synthetic vector formulated with EGTA.* The Journal of Gene Medicine, 2004. **6**: p. 210-221.
235. Gui, L., et al., *Efficient intratracheal delivery of airway epithelial cells in mice and pigs.* Am J Physiol Lung Cell Mol Physiol, 2015. **308**(2): p. L221-8.
236. Rosen, C., et al., *Preconditioning allows engraftment of mouse and human embryonic lung cells, enabling lung repair in mice.* Nat Med, 2015. **21**(8): p. 869-79.
237. Zuo, W., et al., *p63(+)Krt5(+) distal airway stem cells are essential for lung regeneration.* Nature, 2015. **517**(7536): p. 616-20.
238. Hajj, R., et al., *Basal cells of the human adult airway surface epithelium retain transit-amplifying cell properties.* Stem Cells, 2007. **25**(1): p. 139-48.
239. Ghosh, M., et al., *Transplantation of Airway Epithelial Stem/Progenitor Cells: A Future for Cell-Based Therapy.* Am J Respir Cell Mol Biol, 2017. **56**(1): p. 1-10.
240. Slavin, S., et al., *Nonmyeloablative Stem Cell Transplantation and Cell Therapy as an Alternative to Conventional Bone Marrow Transplantation With Lethal Cytoreduction for the Treatment of Malignant and Nonmalignant Hematologic Diseases.* Blood, 1998. **91**(3): p. 756.
241. Horwitz, E.M., et al., *Isolated allogeneic bone marrow-derived mesenchymal cells engraft and stimulate growth in children with osteogenesis imperfecta: Implications for cell therapy of bone.* Proceedings of the National Academy of Sciences, 2002. **99**(13): p. 8932.
242. Mehta, A. and J.E. Haber, *Sources of DNA double-strand breaks and models of recombinational DNA repair.* Cold Spring Harbor perspectives in biology, 2014. **6**(9): p. a016428-a016428.
243. Lindahl, T., *Instability and decay of the primary structure of DNA.* Nature, 1993. **362**(6422): p. 709-15.
244. Vilenchik, M.M. and A.G. Knudson, *Endogenous DNA double-strand breaks: production, fidelity of repair, and induction of cancer.* Proc Natl Acad Sci U S A, 2003. **100**(22): p. 12871-6.
245. Sonoda, E., et al., *Reverse genetic studies of homologous DNA recombination using the chicken B-lymphocyte line, DT40.* Philos Trans R Soc Lond B Biol Sci, 2001. **356**(1405): p. 111-7.
246. Bhattacharjee, S. and S. Nandi, *Choices have consequences: the nexus between DNA repair pathways and genomic instability in cancer.* Clinical and Translational Medicine, 2016. **5**(1): p. 45.
247. Aylon, Y., B. Liefshitz, and M. Kupiec, *The CDK regulates repair of double-strand breaks by homologous recombination during the cell cycle.* The EMBO journal, 2004. **23**(24): p. 4868-4875.
248. Walker, J.R., R.A. Corpina, and J. Goldberg, *Structure of the Ku heterodimer bound to DNA and its implications for double-strand break repair.* Nature, 2001. **412**(6847): p. 607-14.
249. Mari, P.O., et al., *Dynamic assembly of end-joining complexes requires interaction between Ku70/80 and XRCC4.* Proc Natl Acad Sci U S A, 2006. **103**(49): p. 18597-602.
250. Moshous, D., et al., *Artemis, a Novel DNA Double-Strand Break Repair/V(D)J Recombination Protein, Is Mutated in Human Severe Combined Immune Deficiency.* Cell, 2001. **105**(2): p. 177-186.

251. Ramsden, D.A., *Polymerases in nonhomologous end joining: building a bridge over broken chromosomes.* Antioxidants & redox signaling, 2011. **14**(12): p. 2509-2519.
252. Lieber, M.R., *The Mechanism of Human Nonhomologous DNA End Joining.* Journal of Biological Chemistry, 2008. **283**(1): p. 1-5.
253. Gu, J., et al., *Single-stranded DNA ligation and XLF-stimulated incompatible DNA end ligation by the XRCC4-DNA ligase IV complex: influence of terminal DNA sequence.* Nucleic Acids Res, 2007. **35**(17): p. 5755-62.
254. Wang, H., et al., *Biochemical evidence for Ku-independent backup pathways of NHEJ.* Nucleic acids research, 2003. **31**(18): p. 5377-5388.
255. Wang, M., et al., *PARP-1 and Ku compete for repair of DNA double strand breaks by distinct NHEJ pathways.* Nucleic acids research, 2006. **34**(21): p. 6170-6182.
256. Hopkins, B.B. and T.T. Paull, *The P. furiosus Mre11/Rad50 Complex Promotes 5′ Strand Resection at a DNA Double-Strand Break.* Cell, 2008. **135**(2): p. 250-260.
257. Limbo, O., et al., *Ctp1 Is a Cell-Cycle-Regulated Protein that Functions with Mre11 Complex to Control Double-Strand Break Repair by Homologous Recombination.* Molecular Cell, 2007. **28**(1): p. 134-146.
258. McVey, M. and S.E. Lee, *MMEJ repair of double-strand breaks (director's cut): deleted sequences and alternative endings.* Trends in genetics : TIG, 2008. **24**(11): p. 529-538.
259. Simsek, D., et al., *DNA Ligase III Promotes Alternative Nonhomologous End-Joining during Chromosomal Translocation Formation.* PLOS Genetics, 2011. **7**(6): p. e1002080.
260. Dueva, R. and G. Iliakis, *Alternative pathways of non-homologous end joining (NHEJ) in genomic instability and cancer.* Translational Cancer Research, 2013. **2**(3): p. 163-177.
261. Thompson, L.H. and D. Schild, *Homologous recombinational repair of DNA ensures mammalian chromosome stability.* Mutation Research/Fundamental and Molecular Mechanisms of Mutagenesis, 2001. **477**(1): p. 131-153.
262. Takata, M., et al., *Homologous recombination and non-homologous end-joining pathways of DNA double-strand break repair have overlapping roles in the maintenance of chromosomal integrity in vertebrate cells.* The EMBO Journal, 1998. **17**(18): p. 5497-5508.
263. Li, X. and W.-D. Heyer, *Homologous recombination in DNA repair and DNA damage tolerance.* Cell research, 2008. **18**(1): p. 99-113.
264. Heyer, W.-D., K.T. Ehmsen, and J. Liu, *Regulation of homologous recombination in eukaryotes.* Annual review of genetics, 2010. **44**: p. 113-139.
265. Stracker, T.H. and J.H.J. Petrini, *The MRE11 complex: starting from the ends.* Nature Reviews Molecular Cell Biology, 2011. **12**: p. 90.
266. Sung, P. and H. Klein, *Mechanism of homologous recombination: mediators and helicases take on regulatory functions.* Nature Reviews Molecular Cell Biology, 2006. **7**(10): p. 739-750.
267. Pâques, F. and J.E. Haber, *Multiple pathways of recombination induced by double-strand breaks in Saccharomyces cerevisiae.* Microbiology and molecular biology reviews : MMBR, 1999. **63**(2): p. 349-404.
268. Heller, R.C. and K.J. Marians, *Replisome assembly and the direct restart of stalled replication forks.* Nature Reviews Molecular Cell Biology, 2006. **7**(12): p. 932-943.
269. Capecchi, M., *Gene targeting. How efficient can you get?* Nature, 1990. **348**(6297): p. 109.

270. Jasin, M., *Genetic manipulation of genomes with rare-cutting endonucleases.* Trends Genet, 1996. **12**(6): p. 224-8.
271. Rouet, P., F. Smih, and M. Jasin, *Expression of a site-specific endonuclease stimulates homologous recombination in mammalian cells.* Proc Natl Acad Sci U S A, 1994. **91**(13): p. 6064-8.
272. Smih, F., et al., *Double-strand breaks at the target locus stimulate gene targeting in embryonic stem cells.* Nucleic Acids Res, 1995. **23**(24): p. 5012-9.
273. Miller, J., A.D. McLachlan, and A. Klug, *Repetitive zinc-binding domains in the protein transcription factor IIIA from Xenopus oocytes.* Embo j, 1985. **4**(6): p. 1609-14.
274. Pavletich, N.P. and C.O. Pabo, *Zinc finger-DNA recognition: crystal structure of a Zif268-DNA complex at 2.1 A.* Science, 1991. **252**(5007): p. 809-17.
275. Li, L., L.P. Wu, and S. Chandrasegaran, *Functional domains in Fok I restriction endonuclease.* Proc Natl Acad Sci U S A, 1992. **89**(10): p. 4275-9.
276. Waugh, D.S. and R.T. Sauer, *Single amino acid substitutions uncouple the DNA binding and strand scission activities of Fok I endonuclease.* Proc Natl Acad Sci U S A, 1993. **90**(20): p. 9596-600.
277. Kim, Y.G., J. Cha, and S. Chandrasegaran, *Hybrid restriction enzymes: zinc finger fusions to Fok I cleavage domain.* Proc Natl Acad Sci U S A, 1996. **93**(3): p. 1156-60.
278. Bitinaite, J., et al., *FokI dimerisation is required for DNA cleavage.* Proceedings of the National Academy of Sciences, 1998. **95**(18): p. 10570.
279. Bibikova, M., et al., *Stimulation of Homologous Recombination through Targeted Cleavage by Chimeric Nucleases.* Molecular and Cellular Biology, 2001. **21**(1): p. 289.
280. Smith, J., et al., *Requirements for double-strand cleavage by chimeric restriction enzymes with zinc finger DNA-recognition domains.* Nucleic Acids Res, 2000. **28**(17): p. 3361-9.
281. Rebar, E.J. and C.O. Pabo, *Zinc finger phage: affinity selection of fingers with new DNA-binding specificities.* Science, 1994. **263**(5147): p. 671-3.
282. Jamieson, A.C., S.H. Kim, and J.A. Wells, *In vitro selection of zinc fingers with altered DNA-binding specificity.* Biochemistry, 1994. **33**(19): p. 5689-95.
283. Choo, Y. and A. Klug, *Toward a code for the interactions of zinc fingers with DNA: selection of randomized fingers displayed on phage.* Proc Natl Acad Sci U S A, 1994. **91**(23): p. 11163-7.
284. Wolfe, S.A., et al., *Beyond the "recognition code": structures of two Cys2His2 zinc finger/TATA box complexes.* Structure, 2001. **9**(8): p. 717-23.
285. Maeder, M.L., et al., *Oligomerized pool engineering (OPEN): an open-source protocol for making customized zinc-finger arrays.* Nature Protocols, 2009. **4**: p. 1471.
286. Sander, J.D., et al., *Selection-free zinc-finger-nuclease engineering by context-dependent assembly (CoDA).* Nat Methods, 2011. **8**(1): p. 67-9.
287. Miller, J.C., et al., *A TALE nuclease architecture for efficient genome editing.* Nature Biotechnology, 2010. **29**: p. 143.
288. Kay, S. and U. Bonas, *How Xanthomonas type III effectors manipulate the host plant.* Curr Opin Microbiol, 2009. **12**(1): p. 37-43.
289. Kay, S., et al., *A bacterial effector acts as a plant transcription factor and induces a cell size regulator.* Science, 2007. **318**(5850): p. 648-51.
290. Moscou, M.J. and A.J. Bogdanove, *A simple cipher governs DNA recognition by TAL effectors.* Science, 2009. **326**(5959): p. 1501.

291. Boch, J., et al., *Breaking the code of DNA binding specificity of TAL-type III effectors.* Science, 2009. **326**(5959): p. 1509-12.
292. Bedell, V.M., et al., *In vivo genome editing using a high-efficiency TALEN system.* Nature, 2012. **491**(7422): p. 114-8.
293. Carlson, D.F., et al., *Efficient TALEN-mediated gene knockout in livestock.* Proc Natl Acad Sci U S A, 2012. **109**(43): p. 17382-7.
294. Hockemeyer, D., et al., *Genetic engineering of human pluripotent cells using TALE nucleases.* Nature biotechnology, 2011. **29**(8): p. 731-734.
295. Mak, A.N., et al., *The crystal structure of TAL effector PthXo1 bound to its DNA target.* Science, 2012. **335**(6069): p. 716-9.
296. Deng, D., et al., *Structural basis for sequence-specific recognition of DNA by TAL effectors.* Science, 2012. **335**(6069): p. 720-3.
297. Engler, C., et al., *Golden Gate Shuffling: A One-Pot DNA Shuffling Method Based on Type IIs Restriction Enzymes.* PLOS ONE, 2009. **4**(5): p. e5553.
298. Reyon, D., et al., *FLASH assembly of TALENs for high-throughput genome editing.* Nature biotechnology, 2012. **30**(5): p. 460-465.
299. Naitou, A., et al., *Heterodimeric TALENs induce targeted heritable mutations in the crustacean <em>Daphnia magna</em&gt.* Biology Open, 2015. **4**(3): p. 364.
300. Jansen, R., et al., *Identification of genes that are associated with DNA repeats in prokaryotes.* Mol Microbiol, 2002. **43**(6): p. 1565-75.
301. Ishino, Y., et al., *Nucleotide sequence of the iap gene, responsible for alkaline phosphatase isozyme conversion in Escherichia coli, and identification of the gene product.* Journal of Bacteriology, 1987. **169**(12): p. 5429.
302. Mojica, F.J.M., G. Juez, and F. Rodriguez-Valera, *Transcription at different salinities of Haloferax mediterranei sequences adjacent to partially modified PstI sites.* Molecular Microbiology, 1993. **9**(3): p. 613-621.
303. Mojica, F.J.M., et al., *Intervening Sequences of Regularly Spaced Prokaryotic Repeats Derive from Foreign Genetic Elements.* Journal of Molecular Evolution, 2005. **60**(2): p. 174-182.
304. Makarova, K.S., et al., *A DNA repair system specific for thermophilic Archaea and bacteria predicted by genomic context analysis.* Nucleic Acids Research, 2002. **30**(2): p. 482-496.
305. Makarova, K.S., et al., *A putative RNA-interference-based immune system in prokaryotes: computational analysis of the predicted enzymatic machinery, functional analogies with eukaryotic RNAi, and hypothetical mechanisms of action.* Biology Direct, 2006. **1**(1): p. 7.
306. Bolotin, A., et al., *Clustered regularly interspaced short palindrome repeats (CRISPRs) have spacers of extrachromosomal origin.* Microbiology, 2005. **151**(8): p. 2551-2561.
307. Barrangou, R., et al., *CRISPR Provides Acquired Resistance Against Viruses in Prokaryotes.* Science, 2007. **315**(5819): p. 1709-1712.
308. Brouns, S.J.J., et al., *Small CRISPR RNAs Guide Antiviral Defense in Prokaryotes.* Science, 2008. **321**(5891): p. 960-964.
309. Marraffini, L.A. and E.J. Sontheimer, *CRISPR Interference Limits Horizontal Gene Transfer in Staphylococci by Targeting DNA.* Science, 2008. **322**(5909): p. 1843-1845.
310. Garneau, J.E., et al., *The CRISPR/Cas bacterial immune system cleaves bacteriophage and plasmid DNA.* Nature, 2010. **468**: p. 67.

311. Wright, Addison V., James K. Nuñez, and Jennifer A. Doudna, *Biology and Applications of CRISPR Systems: Harnessing Nature's Toolbox for Genome Engineering.* Cell, 2016. **164**(1): p. 29-44.
312. Deltcheva, E., et al., *CRISPR RNA maturation by trans-encoded small RNA and host factor RNase III.* Nature, 2011. **471**(7340): p. 602-7.
313. Jinek, M., et al., *A programmable dual-RNA-guided DNA endonuclease in adaptive bacterial immunity.* Science, 2012. **337**(6096): p. 816-21.
314. Gasiunas, G., et al., *Cas9–crRNA ribonucleoprotein complex mediates specific DNA cleavage for adaptive immunity in bacteria.* Proceedings of the National Academy of Sciences, 2012. **109**(39): p. 15539.
315. Mali, P., et al., *RNA-guided human genome engineering via Cas9.* Science, 2013. **339**(6121): p. 823-6.
316. Cong, L., et al., *Multiplex genome engineering using CRISPR/Cas systems.* Science, 2013. **339**(6121): p. 819-23.
317. Nishimasu, H., et al., *Crystal Structure of Cas9 in Complex with Guide RNA and Target DNA.* Cell, 2014. **156**(5): p. 935-949.
318. Zhang, H.-X., Y. Zhang, and H. Yin, *Genome Editing with mRNA Encoding ZFN, TALEN, and Cas9.* Molecular Therapy, 2019. **27**(4): p. 735-746.
319. Koonin, E.V., K.S. Makarova, and F. Zhang, *Diversity, classification and evolution of CRISPR-Cas systems.* Current opinion in microbiology, 2017. **37**: p. 67-78.
320. S, S., *A short guide to the human genome*. 2008, Cold Spring Harbor, NY: Cold Spring Harbor Laboratory Press.
321. Kleinstiver, B.P., et al., *Engineered CRISPR-Cas9 nucleases with altered PAM specificities.* Nature, 2015. **523**: p. 481.
322. Hu, J.H., et al., *Evolved Cas9 variants with broad PAM compatibility and high DNA specificity.* Nature, 2018. **556**: p. 57.
323. Friedland, A.E., et al., *Characterization of Staphylococcus aureus Cas9: a smaller Cas9 for all-in-one adeno-associated virus delivery and paired nickase applications.* Genome Biol, 2015. **16**: p. 257.
324. Kim, E., et al., *In vivo genome editing with a small Cas9 orthologue derived from Campylobacter jejuni.* Nature communications, 2017. **8**: p. 14500-14500.
325. Zetsche, B., et al., *Cpf1 Is a Single RNA-Guided Endonuclease of a Class 2 CRISPR-Cas System.* Cell, 2015. **163**(3): p. 759-771.
326. Shmakov, S., et al., *Diversity and evolution of class 2 CRISPR-Cas systems.* Nat Rev Microbiol, 2017. **15**(3): p. 169-182.
327. Kim, H.K., et al., *In vivo high-throughput profiling of CRISPR-Cpf1 activity.* Nat Methods, 2017. **14**(2): p. 153-159.
328. Kleinstiver, B.P., et al., *High-fidelity CRISPR-Cas9 nucleases with no detectable genome-wide off-target effects.* Nature, 2016. **529**(7587): p. 490-495.
329. Qi, L.S., et al., *Repurposing CRISPR as an RNA-guided platform for sequence-specific control of gene expression.* Cell, 2013. **152**(5): p. 1173-1183.
330. Perez-Pinera, P., et al., *RNA-guided gene activation by CRISPR-Cas9-based transcription factors.* Nature methods, 2013. **10**(10): p. 973-976.
331. Balboa, D., et al., *Conditionally Stabilized dCas9 Activator for Controlling Gene Expression in Human Cell Reprogramming and Differentiation.* Stem Cell Reports, 2015. **5**(3): p. 448-59.

332. Egorova, T.V., et al., *CRISPR/Cas9-generated mouse model of Duchenne muscular dystrophy recapitulating a newly identified large 430 kb deletion in the human DMD gene.* Disease Models & Mechanisms, 2019. **12**(4): p. dmm037655.
333. Gao, Y., et al., *Single Cas9 nickase induced generation of NRAMP1 knockin cattle with reduced off-target effects.* Genome Biology, 2017. **18**(1): p. 13.
334. Hsu, Patrick D., Eric S. Lander, and F. Zhang, *Development and Applications of CRISPR-Cas9 for Genome Engineering.* Cell, 2014. **157**(6): p. 1262-1278.
335. Shalem, O., et al., *Genome-scale CRISPR-Cas9 knockout screening in human cells.* Science, 2014. **343**(6166): p. 84-87.
336. Ousterout, D.G., et al., *Multiplex CRISPR/Cas9-based genome editing for correction of dystrophin mutations that cause Duchenne muscular dystrophy.* Nature communications, 2015. **6**: p. 6244-6244.
337. Richardson, C.D., et al., *Enhancing homology-directed genome editing by catalytically active and inactive CRISPR-Cas9 using asymmetric donor DNA.* Nat Biotechnol, 2016. **34**(3): p. 339-44.
338. Lin, S., et al., *Enhanced homology-directed human genome engineering by controlled timing of CRISPR/Cas9 delivery.* eLife, 2014. **3**: p. e04766.
339. Maruyama, T., et al., *Increasing the efficiency of precise genome editing with CRISPR-Cas9 by inhibition of nonhomologous end joining.* Nat Biotechnol, 2015. **33**(5): p. 538-42.
340. Aird, E.J., et al., *Increasing Cas9-mediated homology-directed repair efficiency through covalent tethering of DNA repair template.* Communications Biology, 2018. **1**(1): p. 54.
341. Pickar-Oliver, A. and C.A. Gersbach, *The next generation of CRISPR–Cas technologies and applications.* Nature Reviews Molecular Cell Biology, 2019. **20**(8): p. 490-507.
342. Paquet, D., et al., *Efficient introduction of specific homozygous and heterozygous mutations using CRISPR/Cas9.* Nature, 2016. **533**: p. 125.
343. Komor, A.C., et al., *Programmable editing of a target base in genomic DNA without double-stranded DNA cleavage.* Nature, 2016. **533**(7603): p. 420-4.
344. Nishida, K., et al., *Targeted nucleotide editing using hybrid prokaryotic and vertebrate adaptive immune systems.* Science, 2016. **353**(6305): p. aaf8729.
345. Gaudelli, N.M., et al., *Programmable base editing of A*T to G*C in genomic DNA without DNA cleavage.* Nature, 2017. **551**(7681): p. 464-471.
346. Rees, H.A., et al., *Improving the DNA specificity and applicability of base editing through protein engineering and protein delivery.* Nat Commun, 2017. **8**: p. 15790.
347. Song, C.-Q., et al., *Adenine base editing in an adult mouse model of tyrosinaemia.* Nature Biomedical Engineering, 2019.
348. Li, G., et al., *Highly efficient and precise base editing in discarded human tripronuclear embryos.* Protein Cell, 2017. **8**(10): p. 776-779.
349. Anzalone, A.V., et al., *Search-and-replace genome editing without double-strand breaks or donor DNA.* Nature, 2019.
350. Maeder, M.L., et al., *Development of a gene-editing approach to restore vision loss in Leber congenital amaurosis type 10.* Nature Medicine, 2019. **25**(2): p. 229-233.
351. *ClinicalTrials.gov [Internet].* 2000 Feb 29, National Library of Medicine (US): Bethesda (MD).
352. Fu, Y., et al., *High-frequency off-target mutagenesis induced by CRISPR-Cas nucleases in human cells.* Nature biotechnology, 2013. **31**(9): p. 822-826.

353. Iyer, V., et al., *No unexpected CRISPR-Cas9 off-target activity revealed by trio sequencing of gene-edited mice.* PLOS Genetics, 2018. **14**(7): p. e1007503.
354. Akcakaya, P., et al., *In vivo CRISPR editing with no detectable genome-wide off-target mutations.* Nature, 2018. **561**(7723): p. 416-419.
355. Tang, L., et al., *CRISPR/Cas9-mediated gene editing in human zygotes using Cas9 protein.* Molecular Genetics and Genomics, 2017. **292**(3): p. 525-533.
356. Fu, Y., et al., *Improving CRISPR-Cas nuclease specificity using truncated guide RNAs.* Nat Biotechnol, 2014. **32**(3): p. 279-284.
357. Hsu, P.D., et al., *DNA targeting specificity of RNA-guided Cas9 nucleases.* Nat Biotechnol, 2013. **31**(9): p. 827-32.
358. Wienert, B., et al., *Unbiased detection of CRISPR off-targets in vivo using DISCOVER-Seq.* Science, 2019. **364**(6437): p. 286.
359. Cho, S.W., et al., *Heritable gene knockout in Caenorhabditis elegans by direct injection of Cas9-sgRNA ribonucleoproteins.* Genetics, 2013. **195**(3): p. 1177-80.
360. Kim, D., et al., *Digenome-seq: genome-wide profiling of CRISPR-Cas9 off-target effects in human cells.* Nature Methods, 2015. **12**: p. 237.
361. Tsai, S.Q., et al., *GUIDE-seq enables genome-wide profiling of off-target cleavage by CRISPR-Cas nucleases.* Nat Biotechnol, 2015. **33**(2): p. 187-197.
362. Li, H., et al., *In vivo genome editing restores haemostasis in a mouse model of haemophilia.* Nature, 2011. **475**(7355): p. 217-21.
363. Bednarski, C., et al., *Targeted integration of a super-exon into the CFTR locus leads to functional correction of a cystic fibrosis cell line model.* PloS one, 2016. **11**(8): p. e0161072.
364. Lee, C.M., et al., *Correction of the ΔF508 Mutation in the Cystic Fibrosis Transmembrane Conductance Regulator Gene by Zinc-Finger Nuclease Homology-Directed Repair.* BioResearch open access, 2012. **1**(3): p. 99-108.
365. Crane, A.M., et al., *Targeted correction and restored function of the CFTR gene in cystic fibrosis induced pluripotent stem cells.* Stem cell reports, 2015. **4**(4): p. 569-577.
366. Merkert, S., et al., *Generation of a gene-corrected isogenic control iPSC line from cystic fibrosis patient-specific iPSCs homozygous for p.Phe508del mutation mediated by TALENs and ssODN.* Stem Cell Research, 2017. **23**: p. 95-97.
367. Xia, E., et al., *TALEN-Mediated Gene Targeting for Cystic Fibrosis-Gene Therapy.* Genes, 2019. **10**(1): p. 39.
368. Ruan, J., et al., *Efficient Gene Editing at Major CFTR Mutation Loci.* Molecular Therapy. Nucleic acids, 2019. **16**: p. 73-81.
369. Vaidyanathan, S., et al., *Highly Efficient Repair of the ΔF508 Mutation in Airway Stem Cells of Cystic Fibrosis Patients with Functional Rescue of the Differentiated Epithelia.* bioRxiv, 2019: p. 561183.
370. Maule, G., et al., *Allele specific repair of splicing mutations in cystic fibrosis through AsCas12a genome editing.* Nature communications, 2019. **10**(1): p. 3556-3556.
371. Fan, Z., et al., *A sheep model of cystic fibrosis generated by CRISPR/Cas9 disruption of the CFTR gene.* JCI insight, 2018. **3**(19): p. e123529.
372. Sakuma, T., et al., *MMEJ-assisted gene knock-in using TALENs and CRISPR-Cas9 with the PITCh systems.* Nat Protoc, 2016. **11**(1): p. 118-33.
373. Suzuki, K., et al., *In vivo genome editing via CRISPR/Cas9 mediated homology-independent targeted integration.* Nature, 2016. **540**(7631): p. 144-149.

374. Fulcher, M.L., et al., *Novel human bronchial epithelial cell lines for cystic fibrosis research.* American Journal of Physiology - Lung Cellular and Molecular Physiology, 2009. **296**(1): p. L82-L91.
375. Manunta, M.D., et al., *Delivery of ENaC siRNA to epithelial cells mediated by a targeted nanocomplex: a therapeutic strategy for cystic fibrosis.* Sci Rep, 2017. **7**(1): p. 700.
376. Fernando, O., et al., *Development of Targeted siRNA Nanocomplexes to Prevent Fibrosis in Experimental Glaucoma Filtration Surgery.* Mol Ther, 2018. **26**(12): p. 2812-2822.
377. Tagalakis, A.D., et al., *Effective silencing of ENaC by siRNA delivered with epithelial-targeted nanocomplexes in human cystic fibrosis cells and in mouse lung.* Thorax, 2018. **73**(9): p. 847.
378. Yu-Wai-Man, C., et al., *Receptor-targeted liposome-peptide-siRNA nanoparticles represent an efficient delivery system for MRTF silencing in conjunctival fibrosis.* Scientific Reports, 2016. **6**: p. 21881.
379. Tagalakis, A.D., et al., *A Receptor-targeted Nanocomplex Vector System Optimized for Respiratory Gene Transfer.* Molecular Therapy, 2008. **16**(5): p. 907-915.
380. Munye, M.M., et al., *Role of liposome and peptide in the synergistic enhancement of transfection with a lipopolyplex vector.* Sci Rep, 2015. **5**: p. 9292.
381. Park, I.K., et al., *Bmi-1 is required for maintenance of adult self-renewing haematopoietic stem cells.* Nature, 2003. **423**(6937): p. 302-5.
382. Fasano, C.A., et al., *Bmi-1 cooperates with Foxg1 to maintain neural stem cell self-renewal in the forebrain.* Genes Dev, 2009. **23**(5): p. 561-74.
383. Munye, M.M., et al., *BMI-1 extends proliferative potential of human bronchial epithelial cells whilst retaining their mucociliary differentiation capacity.* American Journal of Physiology - Lung Cellular and Molecular Physiology, 2016.
384. Jenkins, R.G., et al., *An integrin-targeted non-viral vector for pulmonary gene therapy.* Gene Therapy, 2000. **7**: p. 393.
385. Manunta, M.D.I., et al., *Nebulisation of receptor-targeted nanocomplexes for gene delivery to the airway epithelium.* PloS one, 2011. **6**(10): p. e26768-e26768.
386. Manunta, M.D.I., et al., *Airway Deposition of Nebulized Gene Delivery Nanocomplexes Monitored by Radioimaging Agents.* American Journal of Respiratory Cell and Molecular Biology, 2013. **49**(3): p. 471-480.
387. Felgner JH, K.R., Sridhar CN, Wheeler CJ, Tsai YJ, Border R, Ramsey P, and F.P. Martin M, *Enhanced gene delivery and mechanism studies with a novel series of cationic lipid formulations. .* J Biol Chem, 1994. **269**((4)): p. 2550-61.
388. Leonhardt, C., et al., *Single-cell mRNA transfection studies: delivery, kinetics and statistics by numbers.* Nanomedicine: Nanotechnology, Biology and Medicine, 2014. **10**(4): p. 679-688.
389. Writer, M.J., et al., *Targeted gene delivery to human airway epithelial cells with synthetic vectors incorporating novel targeting peptides selected by phage display.* J Drug Target, 2004. **12**(4): p. 185-93.
390. Writer, M.J., et al., *Targeted gene delivery to human airway epithelial cells with synthetic vectors incorporating novel targeting peptides selected by phage display.* Journal of drug targeting, 2004. **12**(4): p. 185-193.
391. Tagalakis, A.D., et al., *Receptor-targeted liposome-peptide nanocomplexes for siRNA delivery.* Biomaterials, 2011. **32**(26): p. 6302-6315.

392. Tagalakis, A.D., et al., *Comparison of nanocomplexes with branched and linear peptides for siRNA delivery.* Biomacromolecules, 2013. **14**(3): p. 761-770.
393. Mashal, R.D., J. Koontz, and J. Sklar, *Detection of mutations by cleavage of DNA heteroduplexes with bacteriophage resolvases.* Nature Genetics, 1995. **9**(2): p. 177-183.
394. Brinkman, E.K., et al., *Easy quantitative assessment of genome editing by sequence trace decomposition.* Nucleic Acids Research, 2014. **42**(22): p. e168-e168.
395. Shalem, O., et al., *Genome-Scale CRISPR-Cas9 Knockout Screening in Human Cells.* Science (New York, N.Y.), 2014. **343**(6166): p. 84-87.
396. Hart, S.L., et al., *Lipid-mediated enhancement of transfection by a nonviral integrin-targeting vector.* Hum Gene Ther, 1998. **9**(4): p. 575-85.
397. Chen, Z., et al., *Enhanced DNA release from disulfide-containing layered nanocomplexes by heparin-electrostatic competition.* Journal of Materials Chemistry B, 2015. **3**(2): p. 225-237.
398. Freimann, K., et al., *Optimization of in vivo DNA delivery with NickFect peptide vectors.* Journal of Controlled Release, 2016. **241**: p. 135-143.
399. Gaj, T., C.A. Gersbach, and C.F. Barbas, 3rd, *ZFN, TALEN, and CRISPR/Cas-based methods for genome engineering.* Trends Biotechnol, 2013.
400. Rayner, R.E., et al., *Optimization of Normal Human Bronchial Epithelial (NHBE) Cell 3D Cultures for in vitro Lung Model Studies.* Scientific Reports, 2019. **9**(1): p. 500.
401. Gregory, R.J., et al., *Expression and characterization of the cystic fibrosis transmembrane conductance regulator.* Nature, 1990. **347**(6291): p. 382-6.
402. Cheng, S.H., et al., *Defective intracellular transport and processing of CFTR is the molecular basis of most cystic fibrosis.* Cell, 1990. **63**(4): p. 827-34.
403. Glaser, A., B. McColl, and J. Vadolas, *GFP to BFP Conversion: A Versatile Assay for the Quantification of CRISPR/Cas9-mediated Genome Editing.* Molecular Therapy. Nucleic acids, 2016. **5**(7): p. e334-e334.
404. Glaser, A., B. McColl, and J. Vadolas, *GFP to BFP conversion: a versatile assay for the quantification of CRISPR/Cas9-mediated genome editing.* Molecular Therapy-Nucleic Acids, 2016. **5**.
405. Tagalakis, A.D., et al., *Effective silencing of ENaC by siRNA delivered with epithelial-targeted nanocomplexes in human cystic fibrosis cells and in mouse lung.* Thorax, 2018: p. thoraxjnl-2017-210670.
406. Manunta, M.D., et al., *Delivery of ENaC siRNA to epithelial cells mediated by a targeted nanocomplex: a therapeutic strategy for cystic fibrosis.* Scientific Reports, 2017. **7**(1): p. 700.
407. Kwok, A., et al., *Systematic comparisons of formulations of linear oligolysine peptides with siRNA and plasmid DNA.* Chemical biology & drug design, 2016. **87**(5): p. 747-763.
408. Tagalakis, A.D., et al., *Multifunctional, self-assembling anionic peptide-lipid nanocomplexes for targeted siRNA delivery.* Biomaterials, 2014. **35**(29): p. 8406-8415.
409. Felgner, J.H., et al., *Enhanced gene delivery and mechanism studies with a novel series of cationic lipid formulations.* J Biol Chem, 1994. **269**(4): p. 2550-61.
410. Lopalco, A., et al., *Oxcarbazepine-loaded polymeric nanoparticles: development and permeability studies across in vitro models of the blood-brain barrier and human placental trophoblast.* Int J Nanomedicine, 2015. **10**: p. 1985-96.
411. Gaumet, M., et al., *Nanoparticles for drug delivery: the need for precision in reporting particle size parameters.* Eur J Pharm Biopharm, 2008. **69**(1): p. 1-9.

412. Kou, L., et al., *The endocytosis and intracellular fate of nanomedicines: Implication for rational design.* Asian Journal of Pharmaceutical Sciences, 2013. **8**(1): p. 1-10.
413. Champion, J.A., A. Walker, and S. Mitragotri, *Role of particle size in phagocytosis of polymeric microspheres.* Pharm Res, 2008. **25**(8): p. 1815-21.
414. Byrne, A.J., et al., *Pulmonary macrophages: key players in the innate defence of the airways.* Thorax, 2015. **70**(12): p. 1189.
415. Suk, J.S., et al., *The penetration of fresh undiluted sputum expectorated by cystic fibrosis patients by non-adhesive polymer nanoparticles.* Biomaterials, 2009. **30**(13): p. 2591-2597.
416. Elouahabi, A. and J.M. Ruysschaert, *Formation and intracellular trafficking of lipoplexes and polyplexes.* Mol Ther, 2005. **11**(3): p. 336-47.
417. Tagalakis, A.D., et al., *PEGylation improves the receptor-mediated transfection efficiency of peptide-targeted, self-assembling, anionic nanocomplexes.* Journal of Controlled Release, 2014. **174**: p. 177-187.
418. Uzgun, S., et al., *PEGylation improves nanoparticle formation and transfection efficiency of messenger RNA.* Pharm Res, 2011. **28**(9): p. 2223-32.
419. Briuglia, M.L., et al., *Influence of cholesterol on liposome stability and on in vitro drug release.* Drug Deliv Transl Res, 2015. **5**(3): p. 231-42.
420. Zielenski, J., et al., *Analysis of CFTR transcripts in nasal epithelial cells and lymphoblasts of a cystic fibrosis patient with 621 + 1G-->T and 711 + 1G-->T mutations.* Hum Mol Genet, 1993. **2**(6): p. 683-7.
421. Kerem, E., et al., *A cystic fibrosis transmembrane conductance regulator splice variant with partial penetrance associated with variable cystic fibrosis presentations.* American Journal of Respiratory and Critical Care Medicine, 1997. **155**(6): p. 1914-1920.
422. Nissim-Rafinia, M., et al., *Restoration of the cystic fibrosis transmembrane conductance regulator function by splicing modulation.* EMBO reports, 2004. **5**(11): p. 1071-1077.
423. Sanz, D.J., et al., *Cas9/gRNA targeted excision of cystic fibrosis-causing deep-intronic splicing mutations restores normal splicing of CFTR mRNA.* PLoS One, 2017. **12**(9): p. e0184009.
424. Pan, C., et al., *Comparative proteomic phenotyping of cell lines and primary cells to assess preservation of cell type-specific functions.* Mol Cell Proteomics, 2009. **8**(3): p. 443-50.
425. Uhlen, M., et al., *Proteomics. Tissue-based map of the human proteome.* Science, 2015. **347**(6220): p. 1260419.
426. Zucco, F., et al., *An inter-laboratory study to evaluate the effects of medium composition on the differentiation and barrier function of Caco-2 cell lines.* Altern Lab Anim, 2005. **33**(6): p. 603-18.
427. Ott, C.J., et al., *Novel regulatory mechanisms for the CFTR gene.* Biochem Soc Trans, 2009. **37**(Pt 4): p. 843-8.
428. Papapetrou, E.P., et al., *Genomic safe harbors permit high beta-globin transgene expression in thalassemia induced pluripotent stem cells.* Nat Biotechnol, 2011. **29**(1): p. 73-8.
429. Henckaerts, E. and R.M. Linden, *Adeno-associated virus: a key to the human genome?* Future Virol, 2010. **5**(5): p. 555-574.

430. Sadelain, M., E.P. Papapetrou, and F.D. Bushman, *Safe harbours for the integration of new DNA in the human genome.* Nat Rev Cancer, 2011. **12**(1): p. 51-8.
431. Sentmanat, M.F., et al. *A Survey of Validation Strategies for CRISPR-Cas9 Editing*. Scientific Reports, 2018. **8**, 888 DOI: 10.1038/s41598-018-19441-8.
432. Maule, G., et al., *Allele specific repair of splicing mutations in cystic fibrosis through AsCas12a genome editing.* Nature Communications, 2019. **10**(1): p. 3556.
433. McCarron, A., M. Donnelley, and D. Parsons, *Airway disease phenotypes in animal models of cystic fibrosis.* Respiratory research, 2018. **19**(1): p. 54-54.
434. Duguépéroux, I. and M. De Braekeleer, *The CFTR 3849+10kbC->T and 2789+5G->A alleles are associated with a mild CF phenotype.* European Respiratory Journal, 2005. **25**(3): p. 468.
435. Chillón, M., et al., *A novel donor splice site in intron 11 of the CFTR gene, created by mutation 1811+1.6kbA-->G, produces a new exon: high frequency in Spanish cystic fibrosis chromosomes and association with severe phenotype.* American journal of human genetics, 1995. **56**(3): p. 623-629.
436. Beck, S., et al., *Cystic fibrosis patients with the 3272-26A-->G mutation have mild disease, leaky alternative mRNA splicing, and CFTR protein at the cell membrane.* Hum Mutat, 1999. **14**(2): p. 133-44.
437. Costantino, L., et al., *Fine Characterization of the Recurrent c.1584+18672A>G Deep-Intronic Mutation in the Cystic Fibrosis Transmembrane Conductance Regulator Gene.* American Journal of Respiratory Cell and Molecular Biology, 2013. **48**(5): p. 619-625.
438. Lee, M., et al., *Systematic Computational Identification of Variants That Activate Exonic and Intronic Cryptic Splice Sites.* Am J Hum Genet, 2017. **100**(5): p. 751-765.
439. Vaz-Drago, R., N. Custódio, and M. Carmo-Fonseca, *Deep intronic mutations and human disease.* Human Genetics, 2017. **136**(9): p. 1093-1111.
440. Cheng, T.C., et al., *beta-Thalassemia in Chinese: use of in vivo RNA analysis and oligonucleotide hybridization in systematic characterization of molecular defects.* Proc Natl Acad Sci U S A, 1984. **81**(9): p. 2821-5.
441. den Hollander, A.I., et al., *Mutations in the CEP290 (NPHP6) Gene Are a Frequent Cause of Leber Congenital Amaurosis.* The American Journal of Human Genetics, 2006. **79**(3): p. 556-561.
442. Valdmanis, P.N., et al., *A mutation that creates a pseudoexon in SOD1 causes familial ALS.* Ann Hum Genet, 2009. **73**(Pt 6): p. 652-7.
443. Darquet, A., et al., *A new DNA vehicle for nonviral gene delivery: supercoiled minicircle.* Gene therapy, 1997. **4**(12): p. 1341.
444. Suzuki, K., et al., *In vivo genome editing via CRISPR/Cas9 mediated homology-independent targeted integration.* Nature, 2016. **540**(7631): p. 144.
445. Chen, Z.-Y., et al., *Minicircle DNA vectors devoid of bacterial DNA result in persistent and high-level transgene expression in vivo.* Molecular Therapy, 2003. **8**(3): p. 495-500.
446. Chen, Z.-Y., et al., *Silencing of episomal transgene expression in liver by plasmid bacterial backbone DNA is independent of CpG methylation.* Molecular Therapy, 2008. **16**(3): p. 548-556.
447. Munye, M.M., et al., *Minicircle DNA Provides Enhanced and Prolonged Transgene Expression Following Airway Gene Transfer.* Scientific Reports, 2016. **6**: p. 23125.
448. Lu, J., F. Zhang, and M.A. Kay, *A mini-intronic plasmid (MIP): a novel robust transgene expression vector in vivo and in vitro.* Molecular Therapy, 2013. **21**(5): p. 954-963.

449. Diecke, S., et al., *Novel codon-optimized mini-intronic plasmid for efficient, inexpensive, and xeno-free induction of pluripotency.* Scientific Reports, 2015. **5**: p. 8081.
450. Hollywood, J.A., et al., *Analysis of gene repair tracts from Cas9/gRNA double-stranded breaks in the human CFTR gene.* Sci Rep, 2016. **6**: p. 32230.
451. Madisen, L., et al., *A robust and high-throughput Cre reporting and characterization system for the whole mouse brain.* Nature neuroscience, 2010. **13**(1): p. 133.
452. Bantikassegn, A., X. Song, and K. Politi, *Isolation of epithelial, endothelial, and immune cells from lungs of transgenic mice with oncogene-induced lung adenocarcinomas.* American journal of respiratory cell and molecular biology, 2015. **52**(4): p. 409-417.
453. Ahmad, A., *Correction of the ΔF508 mutation in the CFTR Gene by CRISPR/Cas9 system*, in *Genetics and Genomic Medicine*. 2018, University College London: Great Ormond Street Institute of Child Health p. 287.
454. Zhanmu, O., et al., *Maintenance of Fluorescence During Paraffin Embedding of Fluorescent Protein-Labeled Specimens.* Frontiers in Neuroscience, 2019. **13**(752).
455. Munye, M., et al., *Minicircle DNA Provides Enhanced and Prolonged Transgene Expression Following Airway Gene Transfer.* Scientific Reports, 2016. **6**: p. 23125.
456. Du, Z., et al., *The Role of the Helper Lipid on the DNA Transfection Efficiency of Lipopolyplex Formulations.* Scientific Reports, 2014. **4**: p. 1-6.
457. Koenig, M., et al., *Complete cloning of the Duchenne muscular dystrophy (DMD) cDNA and preliminary genomic organization of the DMD gene in normal and affected individuals.* Cell, 1987. **50**(3): p. 509-17.
458. Nakamura, A., et al., *Follow-up of three patients with a large in-frame deletion of exons 45-55 in the Duchenne muscular dystrophy (DMD) gene.* J Clin Neurosci, 2008. **15**(7): p. 757-63.
459. Taglia, A., et al., *Clinical features of patients with dystrophinopathy sharing the 45-55 exon deletion of DMD gene.* Acta myologica : myopathies and cardiomyopathies : official journal of the Mediterranean Society of Myology, 2015. **34**(1): p. 9-13.
460. Tabebordbar, M., et al., *In vivo gene editing in dystrophic mouse muscle and muscle stem cells.* Science, 2016. **351**(6271): p. 407-411.
461. Lee, K., et al., *Nanoparticle delivery of Cas9 ribonucleoprotein and donor DNA in vivo induces homology-directed DNA repair.* Nature Biomedical Engineering, 2017. **1**(11): p. 889-901.
462. Staahl, B.T., et al., *Efficient genome editing in the mouse brain by local delivery of engineered Cas9 ribonucleoprotein complexes.* Nat Biotechnol, 2017. **35**(5): p. 431-434.
463. Leitch, A.E.a.R.H.C., *Cystic Fibrosis.* J R Coll Physicians Edinb. **43(2)**: p. 144-50.
464. Mirtajani, S., et al., *Geographical distribution of cystic fibrosis; The past 70 years of data analyzis.* Biomedical and Biotechnology Research Journal (BBRJ), 2017. **1**(2): p. 105-112.
465. Lu, J., F. Zhang, and M.A. Kay, *A mini-intronic plasmid (MIP): a novel robust transgene expression vector in vivo and in vitro.* Mol Ther, 2013. **21**(5): p. 954-63.
466. Liang, X., et al., *Enhanced CRISPR/Cas9-mediated precise genome editing by improved design and delivery of gRNA, Cas9 nuclease, and donor DNA.* Journal of Biotechnology, 2017. **241**(Supplement C): p. 136-146.

467. Renaud, J.-B., et al., *Improved Genome Editing Efficiency and Flexibility Using Modified Oligonucleotides with TALEN and CRISPR-Cas9 Nucleases.* Cell Reports. **14**(9): p. 2263-2272.
468. Richardson, C.D., et al., *Enhancing homology-directed genome editing by catalytically active and inactive CRISPR-Cas9 using asymmetric donor DNA.* Nature Biotechnology, 2016. **34**: p. 339.
469. Chen, S., et al., *CRISPR-READI: Efficient Generation of Knockin Mice by CRISPR RNP Electroporation and AAV Donor Infection.* Cell Reports, 2019. **27**(13): p. 3780-3789.e4.
470. Suzuki, K., et al., *Precise in vivo genome editing via single homology arm donor mediated intron-targeting gene integration for genetic disease correction.* Cell Research, 2019. **29**(10): p. 804-819.
471. *Generation and initial analysis of more than 15,000 full-length human and mouse cDNA sequences.* Proceedings of the National Academy of Sciences, 2002. **99**(26): p. 16899.
472. Gaudelli, N.M., et al., *Programmable base editing of A•T to G•C in genomic DNA without DNA cleavage.* Nature, 2017. **551**: p. 464.
473. Komor, A.C., et al., *Programmable editing of a target base in genomic DNA without double-stranded DNA cleavage.* Nature, 2016. **533**: p. 420.
474. Plasschaert, L.W., et al., *A single-cell atlas of the airway epithelium reveals the CFTR-rich pulmonary ionocyte.* Nature, 2018. **560**(7718): p. 377-381.
475. Montoro, D.T., et al., *A revised airway epithelial hierarchy includes CFTR-expressing ionocytes.* Nature, 2018. **560**(7718): p. 319-324.
476. Leigh, M.W., et al., *Cell proliferation in bronchial epithelium and submucosal glands of cystic fibrosis patients.* American Journal of Respiratory Cell and Molecular Biology, 1995. **12**(6): p. 605-612.
477. Labiris, N.R. and M.B. Dolovich, *Pulmonary drug delivery. Part I: physiological factors affecting therapeutic effectiveness of aerosolized medications.* British journal of clinical pharmacology, 2003. **56**(6): p. 588-599.
478. Alton, E.W., et al., *Repeated nebulisation of non-viral CFTR gene therapy in patients with cystic fibrosis: a randomised, double-blind, placebo-controlled, phase 2b trial.* Lancet Respir Med, 2015. **3**(9): p. 684-691.
479. Patel, A.K., et al., *Inhaled Nanoformulated mRNA Polyplexes for Protein Production in Lung Epithelium.* Advanced Materials, 2019. **31**(8): p. 1805116.
480. Koshkina, N.V., et al., *Biodistribution and pharmacokinetics of aerosol and intravenously administered DNA-polyethyleneimine complexes: optimization of pulmonary delivery and retention.* Mol Ther, 2003. **8**(2): p. 249-54.
481. Beck, S.E., et al., *Deposition and expression of aerosolized rAAV vectors in the lungs of Rhesus macaques.* Mol Ther, 2002. **6**(4): p. 546-54.
482. Sage, E.K., et al., *Systemic but not topical TRAIL-expressing mesenchymal stem cells reduce tumour growth in malignant mesothelioma.* Thorax, 2014. **69**(7): p. 638-47.
483. Wu, N.-H., et al., *The differentiated airway epithelium infected by influenza viruses maintains the barrier function despite a dramatic loss of ciliated cells.* Scientific Reports, 2016. **6**: p. 39668.
484. Csanády, L. and B. Töröcsik, *Cystic fibrosis drug ivacaftor stimulates CFTR channels at picomolar concentrations.* eLife, 2019. **8**: p. e46450.

485. Quinn, C., et al., *Estimating the Clinical Pipeline of Cell and Gene Therapies and Their Potential Economic Impact on the US Healthcare System.* Value in Health, 2019. **22**(6): p. 621-626.