Enhancing Strategies for CRISPR/Cas9 Based T cell Engineering

Candidate:

Roland Preece

Supervisors:

Prof Waseem Qasim
Dr Christos Georgiadis

Great Ormond Street
Institute of Child health
University College London

A thesis submitted for the degree of Doctor of Philosophy
Declaration:

I, Roland Preece confirm that the work presented in this thesis is my own. Where information has been derived from other sources, I confirm that this has been indicated in this thesis.

Signature........................................

Date..............................................
Abstract:

Genome editing tools are being rapidly developed, accelerating many areas of cell and gene therapy research, and are now entering clinical phase testing. Each successive genome editing technology promises increased efficacy, improved specificity, reduced manufacturing cost and design complexity; all of which are epitomised by the clustered regularly interspaced short palindromic repeats (CRISPR) and CRISPR-associated protein (Cas9) platform. Implementation of CRISPR/Cas9 in existing methodologies has been instrumental to recent progress in the treatment of cancer, primary immunodeficiency, and infectious diseases. To this end, T cell therapies attempting to redirect antigen recognition have been enhanced through CRISPR/Cas9 genome editing, endowing them with increased potency and persistence, as well as allowing the generation of allogeneic products capable of overcoming barriers to transplant. This project aims to improve CRISPR/Cas9 strategies for T cell engineering through the development of a novel self-inactivating lentiviral vector platform (terminal-CRISPR) that couples expression of CRISPR single guide RNA (sgRNA) and transgene expression, through incorporation of a sgRNA expression cassette within the ΔU3 region of the 3‘ long terminal repeat sequence. Initial investigation coupled expression of a chimeric antigen receptor and a T cell receptor alpha constant specific sgRNA, resulting in CAR+ cells devoid of alloreactive T cell receptor (TCR)αβ complexes. Moreover edited TCRαβ/CAR+ effectors demonstrated enhanced anti-leukemic outcomes in an in vivo tumour model. The terminal-CRISPR configuration further provided a platform for exploring the architecture of the U6 promoter, enabling the discovery of a minimal U6 promoter, which has been applied to the optimisation of a multiplex terminal-CRISPR vector. Moreover, concerns surrounding the generation of double strand DNA breaks prompted the application of emerging cytidine deaminase base editing technologies, showing high levels of targeted base conversions at multiple genomic loci, while also reducing large chromosomal translocation events. The terminal-CRISPR platform was further proved effective at linking delivery of a Hepatitis B virus specific recombinant TCR and sgRNA for disruption of the endogenous TCR. Discriminatory enrichment of
antigen specific T cells with removal of the endogenous TCR offers an enhanced, highly targeted T cell therapy. This strategy is currently under evaluation for application in phase 1 clinical testing targeting B cell malignancies.
Impact statement:

Immunotherapy has gained significant momentum providing a novel treatment strategy for otherwise untreatable patient populations. Particular success in the field of hematologic malignancies using anti CD19 CAR T cells has allowed the field to rapidly progress leading to a global initiation of international clinical trials. The rapid development of targeted genome editing tools has spurred development of these CAR T cell therapies, by instilling increased potency, persistence, as well as allowing the development of allogeneic off-the-shelf products devoid of alloreactive markers.

This report described steps taken to enhance the application of CRISPR/Cas9 genome editing in the development of immunotherapeutics. Results in Chapter 3 described the development of a novel lentiviral vector configuration that couples expression of CRISPR single guide RNA (sgRNA) with that of a therapeutic transgene (terminal-CRISPR) to produce more homogenous T cell products. Chapter 4 sees the development of a minimal U6 promoter used to multiplex sgRNA expression from the terminal-CRISPR vector. Further development of these configurations has shown their compatibility with cytidine deaminase base editing technology (Chapter 5) possessing an improved safety profile, as well as increasing expression and function of recombinant T cell receptor bearing T cells (Chapter 6). Beyond the results presented in this report, these lentiviral configurations have proven scalable under GMP conditions in preparation for early phase 1 clinical trial.

Although this report presents applications for the field of immunotherapeutics, the configuration described here would be suited to a wide variety of biologic research, providing a versatile tool for the expression of both a transgene and short non-coding RNA sequences. Additionally, this platform should be ready adjustable to a number of different CRISPR systems by changing the configuration of the sgRNA expression cassette, thereby expanding the potential impact of this lentiviral system.

The terminal-CRISPR platform has patented (WO/2018/115887), as well as the development of the minimal U6 promoter (N415912GB).
These results have significantly contributed to a number of publications over the course of this PhD:


With a further report that was recently submitted:

Acknowledgements:

This work would not have been possible without the supervision of Professor Waseem Qasim and Dr Christos Georgiadis, and I would like to start by thanking them for all of their time and guidance over the course of my studies at GOS ICH. Their genuine interest in the direction of this project, as well as their ever presence support has made this the best Academic experience of my career. There is really not enough space here to fulling thank them for everything they have done. Building upon this excellent supervision, the members of the Qasim group have provided me with an amazing working environment/ support network throughout my PhD. Their friendships and expertise has been invaluable, without which much of this work would not have been possible.

Moreover I would like to thank the whole of the Molecular and Cellular Immunology department, all of which have provided me with valuable advice and whose company has made my time at the institute a pleasure, and I only wish there was enough space here to detail all of these amazing individuals. I must take the time to thank Ayad Eddaoudi in the flow cytometry facility for answering countless questions, without him troubleshooting these devices would be impossible. Additionally, I would like to give a special thanks to all the people who have acted as blood donors for these experiments.

I would really like to extend a big thanks to the NIHR BTRU for doing so much more than just funding this project, but providing the opportunity to network with other NHIR BTRU funded centres, and take part in public engagement events. In this, I would be remise in not mentioning Linda Von Neree who is key for the organisation of these events.

Finally, I would like to extent a huge thanks to my friends and family who have been instrumental in supporting me over the course of this project. Of course a big thanks has to go to Klaudia Karcz who provides great advice (even if it is not always followed) and has been very supportive especially while I have been writing this report.
Contents:

Abstract: ......................................................................................................................... 3
Impact statement: ............................................................................................................ 5
Figures: ............................................................................................................................. 13
Supplementary figures ................................................................................................. 15
Tables: ............................................................................................................................... 16
Abbreviations: .................................................................................................................. 17

Chapter 1  Introduction ................................................................................................. 23

1.1  Genome editing ....................................................................................................... 23
1.2  Protein-DNA interacting genome editing tools ...................................................... 24
1.3  CRISPR/ Cas9 from bacterial immune system to genome editing tool ................. 24
1.4  CRISPR/Cas9 genome editing ............................................................................... 28
    1.4.1  Damage repair pathways post SpCas9 genome editing ................................. 29
    1.4.2  Opening up the genome by expanding the PAM recognition profile of
            CRISPR/Cas ......................................................................................................... 32
    1.4.3  Minimising off-target activity of CRISPR/Cas genome editing ................. 37
1.5  Detection of off-target CRISPR/Cas9-mediated cleavage .................................... 42
1.6  CRISPR uncut: seamless base conversion ............................................................. 42
1.7  Delivery of CRISPR reagents .................................................................................. 45
1.8  Brief description of gene therapy lentiviral vectors ............................................... 49
1.9  CRISPR/Cas9 T cell engineering for generation of enhanced immunotherapeutics .......................................................... 51
    1.9.1  Recombinant T cell receptors ......................................................................... 52
    1.9.2  CAR-T cell therapies ...................................................................................... 54
1.10  Current project ...................................................................................................... 59
## Chapter 2 Materials and Methods

### 2.1 Materials

<table>
<thead>
<tr>
<th>Subsection</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>Reagents used for DNA</td>
<td>61</td>
</tr>
<tr>
<td>Reagents used for protein</td>
<td>62</td>
</tr>
<tr>
<td>Reagents for bacterial culture</td>
<td>64</td>
</tr>
<tr>
<td>Reagents for cell culture</td>
<td>65</td>
</tr>
<tr>
<td>Flow cytometry/ western blot antibodies:</td>
<td>67</td>
</tr>
<tr>
<td>Kits</td>
<td>68</td>
</tr>
<tr>
<td>Buffers</td>
<td>70</td>
</tr>
<tr>
<td>Cell culture medium</td>
<td>71</td>
</tr>
<tr>
<td>Cell types</td>
<td>71</td>
</tr>
<tr>
<td>NSG mouse strain</td>
<td>72</td>
</tr>
<tr>
<td>Trilink biotechnologies (San Diego, USA) mRNA</td>
<td>72</td>
</tr>
<tr>
<td>Synthego (California, USA) sgRNA</td>
<td>73</td>
</tr>
<tr>
<td>Oligo’s</td>
<td>73</td>
</tr>
<tr>
<td>sgRNA sequences</td>
<td>73</td>
</tr>
<tr>
<td>Software</td>
<td>75</td>
</tr>
</tbody>
</table>

### 2.2 Methods

<table>
<thead>
<tr>
<th>Subsection</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>Bacterial transformation and plasmid preparation</td>
<td>75</td>
</tr>
<tr>
<td>Restriction digestion</td>
<td>76</td>
</tr>
<tr>
<td>Deposphorylation of DNA ends</td>
<td>77</td>
</tr>
<tr>
<td>DNA ligation reaction</td>
<td>77</td>
</tr>
<tr>
<td>Polymerase chain reaction</td>
<td>77</td>
</tr>
<tr>
<td>In-Fusion HD cloning</td>
<td>78</td>
</tr>
<tr>
<td>Site-Directed Mutagenesis</td>
<td>78</td>
</tr>
<tr>
<td>CRISPR single guide RNA design</td>
<td>80</td>
</tr>
</tbody>
</table>
2.2.9 Protoscaler cloning for sgRNA expression ........................................... 81
2.2.10 Eurofins Genomics Sanger sequencing ............................................. 81
2.2.11 Lentiviral vector production .............................................................. 82
2.2.12 Lentiviral vector titration .................................................................... 83
2.2.13 Sanger sequencing analysis of non-homologous end joining, and targeted Cytidine deamination events ......................................................... 83
2.2.14 Cell line culture .................................................................................... 84
2.2.15 Primary human lymphocyte culture and modification ....................... 84
2.2.16 Cryopreservation and recovery .......................................................... 85
2.2.17 Electroporation ................................................................................... 85
2.2.18 Flow cytometry .................................................................................. 86
2.2.19 Depletion by magnetic bead separation ............................................. 86
2.2.20 Cas9 western blot ............................................................................... 87
2.2.21 Statistics ............................................................................................. 87

Chapter 3 Development of terminal-CRISPR lentiviral vector for coupling transgene expression with CRISPR editing effects ........................................... 88
3.1 Design concept of terminal-CRISPR lentiviral vector ............................ 89
3.2 Optimisation of genome editing effects in primary T cells using TT-CAR19 lentiviral vector ..................................................................................... 93
3.3 Anti-leukaemic activity of TT-CAR19 T cells in vivo ............................... 97
3.4 Knockout of B2M, and CD52 loci using Terminal-CRISPR vectors to create immunologic stealthy/ chemotherapy resistant CAR T cells .................. 101
  3.4.1 Disruption of HLA class I by B2M knockout ....................................... 102
  3.4.2 Serotherapy resistance by CD52 knockout ....................................... 105
3.5 Chapter summary ..................................................................................... 108

Chapter 4 Generation of minimal U6 pol III promoter, supports development of multiplex terminal-CRISPR lentiviral vector ........................................... 109
4.1 Human U6 demonstrates redundancy of positional nucleosome spacer sequence ................................................................. 113

4.2 Addition of a second tandem CRISPR expression cassette allows multiplex genome editing ................................................................. 115

4.3 Incorporation of alternative sgRNA scaffold sequence reduced recombination events during reserve transcription ........................................ 119

4.4 Application of minimal pol III promoters supports efficient coupled multiplex genome editing ................................................................. 122

4.5 Chapter summary .................................................................................. 127

Chapter 5 Optimisation of coupled and uncoupled cytidine base editing for primary human T cell engineering ................................................. 128

5.1 Third generation cytidine deaminase base editor (BE3) ........................ 128

5.2 Modelling coBE3 editing in a Jurkat E6.1 T cell leukaemia line .............. 128

5.3 Efficient base editing by coBE3 mRNA in primary human T cells .......... 132

5.4 Application of coBE3, to CD52 and B2M genome editing ..................... 134

5.5 Coupled multiplexed cytidine deamination using terminal-CRISPR vector .................................................................................. 135

5.6 Reduced vector expressed sgRNA editing in the presence of chemically modified sgRNA ........................................................................ 138

5.7 Multiplexed uncoupled editing of TRBC and CD52 ............................... 141

5.8 Modification of guide scaffold sequence to improve sgRNA expression and RNP stability ............................................................. 142

5.9 Minimal U6 promoter demonstrates compatibility with coBE3 base editing .................................................................................. 152

5.10 Detection of SpCas9 and BE3 protein expression by western blot ...... 153

5.11 Screening of large translocation events with SpCas9 and BE3 .......... 156

5.12 Chapter summary: ................................................................................ 160

Chapter 6 Application of terminal-CRISPR for rTCR cellular immunotherapy .161
6.1 Recombinant TCR based immunotherapies for HBV driven HCC .......... 161
6.2 Composition of HBV specific rTCR (S183-91) ................................................. 162
6.3 TT-rTCR (S183-91) vector with transient SpCas9 delivery allowed enrichment of the eTCR+rTCR+ population ................................................................. 164
6.4 Targeted cytidine deamination of endogenous TRBC, tolerates generation of rTCR cell products ................................................................................................. 166
6.5 Molecular characterisation of base editing effects ........................................ 171
6.6 Multiplex base editing of both endogenous TRAC and TRBC loci ........... 173
6.7 Chapter summary .................................................................................................. 176

Chapter 7 Discussion ............................................................................................... 178
7.1 Coupled CRISPR/Cas9 editing and CAR19 transgene expression .............. 178
7.2 U6 RNA Pol III promoter exhibits redundancy of DNA spacer sequence 180
7.3 Multiplex terminal-CRISPR configuration for simultaneous disruption of two loci ............................................................... 181
7.4 Coupled and uncoupled multiplex cytidine base editing .......................... 184
7.5 CRISPR mediated base conversion allows discriminatory depletion of eTCRαβ, and enhanced rTCR (S183-91) potency .......................... 187
7.6 Concluding remarks ......................................................................................... 189

Chapter 8 References ............................................................................................... 190

Supplementary figures ........................................................................................... 219

Appendices ............................................................................................................. 225
Figures:

Figure 1-1 *Streptococcus pyogenes* CRISPR/Cas9 bacterial adaptive immune system: ................................................................................................................................. 27
Figure 1-2 Schematic of CRISPR-Cas9 and resulting DSB repair by C-NHEJ: ................. 31
Figure 1-3 Third generation cytidine base editor: .............................................................. 44
Figure 1-4 Delivery strategies of CRISPR/Cas9 reagents: .................................................. 48
Figure 1-5 Tumour target recognition by recombinant TCR and CAR expressing T cells: ................................................................................................................................. 53
Figure 1-6 Personalised autologous versus universal allogeneic CAR T cell therapies: ........................................................................................................................................ 58
Figure 3-1 Terminal-CRISPR configuration coupling lentiviral sgRNA and CAR19 expression: ................................................................................................................................. 90
Figure 3-2 Validation and optimisation of Terminal-TRAC-CAR19 for T cell engineering: ........................................................................................................................................ 96
Figure 3-3 In vivo functionality of TT-CAR19 effector T cells against CD19+/GFP+/Luciferase+ Daudi tumour cells: ................................................................. 100
Figure 3-4 CRISPR/Cas9 mediated disruption of HLA class I using the terminal-CRISPR configuration: ............................................................................................................................. 104
Figure 3-5 Validation of CRISPR/ Cas9 knockout of CD52 using a terminal-CRISPR expressed sgRNA: ............................................................................................................................. 107
Figure 4-1 RNA polymerase III promoter types and example transcript function: . 110
Figure 4-2 RNA polymerase III, type III promoter structure: ............................................ 112
Figure 4-3 Development of minimal U6 promoter devoid of nucleosome spacer sequence: ........................................................................................................................................ 114
Figure 4-4 Multiplex terminal-CRISPR configuration containing tandem sgRNA expression cassettes: ............................................................................................................................. 118
Figure 4-5 Multiplex terminal-CRISPR vector configuration containing alternative scaffold sequences: ............................................................................................................................. 121
Figure 4-6 Multiplex terminal-CRISPR configuration incorporating alternative scaffold sequences, and minimal RNA Pol III promoters:  .................................................................................. 123
Figure 4-7 Generation of high homogenous HLA I-/TCRαβ-/-CAR19+ populations with multiplex terminal configuration, containing minimal RNA Pol III promoters: ....... 125
Figure 5-1 Modelling coBE3 activity in a CD3+ Jurkat cell line: ............................... 131
Figure 5-2 Efficient TCRαβ knockout in primary T cells by targeted cytidine deamination: .......................................................................................................................... 133
Figure 5-3 Optimisation of CD52 and B2M knockout with coBE3: ......................... 135
Figure 5-4 Multiplex terminal-CRISPR-CAR19 vectors for simultaneous disruption of two genomic loci: ........................................................................................................ 136
Figure 5-5 Titration of coBE3 mRNA in primary human T cells transduced with multiplex terminal-CRISPR-CAR19 vector: .............................................................. 138
Figure 5-6 Multiplex coBE3 editing through coupled and uncoupled deamination: .......................................................................................................................... 140
Figure 5-7 Uncoupled multiplexed base editing: ...................................................... 142
Figure 5-8 Optimisation of sgRNA scaffold to increase expression and RNP stability:  .......................................................................................................................... 144
Figure 5-9 Genomic analysis of C>T conversion at TRBC, CD52 and B2M loci: ...... 146
Figure 5-10 Uncoupled sgRNA titration with TTRBCC+5-CAR19 vector: ............... 149
Figure 5-11 Comparison of cr1 and C+5 scaffold sequences: .................................. 151
Figure 5-12 Comparison between wild type U6 and mini-U6 promoter: ................. 153
Figure 5-13 Transient expression of BE3 protein: ................................................... 155
Figure 5-14 Reduction of large translocation events with BE3 compared to SpCas9 editing in primary human T cells: .............................................................................. 159
Figure 6-1 Terminal-CRISPR configuration coupling TRAC sgRNA and rTCR (S183-91) expression: ........................................................................................................ 165
Figure 6-2 Vector design for generation of eTCR-/rTCR+ cells using coupled base editing: .................................................................................................................... 169
Figure 6-3 Molecular analysis of on-target DNA editing: ....................................... 172
Figure 6-4 Production of rTCR (S183-91) T cells by double knockout of both TRAC and TRBC: ................................................................................................................ 175
Supplementary figures

Supplementary figure 1 DNA sequences for all tested U6 variants: .......................... 219
Supplementary figure 2 Comparison of wild type and minimal U6 promoters in HEK 293T cells: ........................................................................................................................................ 220
Supplementary figure 3 Comparison of wild type and minimal promoters across a gradient of SpCas9 mRNA concentrations: ........................................................................................................ 221
Supplementary figure 4 Multiplex terminal-CRISPR for combined CD52 and TRAC disruption: .................................................................................................................................................. 222
Supplementary figure 5 Comparison between wild type U6 and mini-U6 promoter: ........................................................................................................................................................................... 223
Supplementary figure 6 Translocation events between TRBC and CD52 loci: ....... 224
Tables:

Table 1 Cas endonucleases: .................................................................................. 35
Table 2 High-Fidelity Cas variants: ..................................................................... 41
Table 3 List of the reagents used for DNA ......................................................... 61
Table 4 List of reagents used for protein ............................................................. 62
Table 5 List of reagents used for bacterial culture ............................................. 64
Table 6 List of reagents used during cell culture ............................................... 65
Table 7 List of primary and secondary antibodies .......................................... 67
Table 8 List of manufacture designed kits ......................................................... 68
Table 9 List of Cell types .................................................................................. 71
Table 10 Protospacer sequences ....................................................................... 73
Abbreviations:

*Acidaminococcus* Cas12a  
AsCas12a

Adeno-Associated Viruses  
AAVs

Adenosine deaminase base editors  
ABEs

Adoptive T cell transfer  
ACT

*Alicyclobacillus acidiphilus* Cas12b  
AaCas12b

Alternative End Joining  
Alt-EJ

Amino acid  
AA

Aprataxin-and-PNK-like factor  
APLF

Base excision repair  
BER

Beta-2-Microglobulin  
B2M

Bioluminescence imagining  
BLI

Bone marrow  
BM

Bovine serum albumin  
BSA

*Campylobacter jejuni* Cas9  
CjCas9

C-C Motif Chemokine Ligand 4  
CCL4

Chimeric antigen receptor  
CAR

Classical NHEJ  
C-NHEJ

Cluster of differentiation  
CD

Clustered-Regularly Interspaced Short Palindromic Repeats  
CRISPR
<table>
<thead>
<tr>
<th>Term</th>
<th>Abbreviation</th>
</tr>
</thead>
<tbody>
<tr>
<td>Constant</td>
<td>C</td>
</tr>
<tr>
<td>CRISPR-associated complex for antiviral defense</td>
<td>Cascade</td>
</tr>
<tr>
<td>CRISPR-Associated Protein</td>
<td>Cas</td>
</tr>
<tr>
<td>CRISPR RNA</td>
<td>crRNA</td>
</tr>
<tr>
<td>Cytidine base editors</td>
<td>CBEs</td>
</tr>
<tr>
<td>Cytomegalovirus</td>
<td>CMV</td>
</tr>
<tr>
<td>Deoxyribonucleic acid</td>
<td>DNA</td>
</tr>
<tr>
<td>Dimethyl sulfoxide</td>
<td>DMSO</td>
</tr>
<tr>
<td>Distal sequence element</td>
<td>DSE</td>
</tr>
<tr>
<td>Dithiothreitol</td>
<td>DTT</td>
</tr>
<tr>
<td>Donor leukocyte infusion</td>
<td>DLI</td>
</tr>
<tr>
<td>Double Strand Breaks</td>
<td>DSBs</td>
</tr>
<tr>
<td>Dulbecco’s Modified Eagle Medium</td>
<td>DMEM</td>
</tr>
<tr>
<td>Dulbecco’s Phosphate-Buffered Saline</td>
<td>PBS</td>
</tr>
<tr>
<td>Endogenous TCR</td>
<td>eTCR</td>
</tr>
<tr>
<td><em>Escherichia coli</em></td>
<td><em>E. coli</em></td>
</tr>
<tr>
<td>Foetal Calf Serum</td>
<td>FCS</td>
</tr>
<tr>
<td><em>Francisella novicida</em> Cas9</td>
<td>FnCas9</td>
</tr>
<tr>
<td>Genome-wide Unbiased Identification of DSBs Enabled by Sequence</td>
<td>GUIDE-seq</td>
</tr>
<tr>
<td>Graft versus host disease</td>
<td>GVHD</td>
</tr>
<tr>
<td>Term</td>
<td>Abbreviation</td>
</tr>
<tr>
<td>-------------------------------------------</td>
<td>--------------</td>
</tr>
<tr>
<td>Hepatitis B virus</td>
<td>HBV</td>
</tr>
<tr>
<td>Hepatocellular carcinoma</td>
<td>HCC</td>
</tr>
<tr>
<td>High performance liquid chromatography</td>
<td>HPLC</td>
</tr>
<tr>
<td>Homology Directed Repair</td>
<td>HDR</td>
</tr>
<tr>
<td>Human leukocyte antigens</td>
<td>HLA</td>
</tr>
<tr>
<td>human phosphoglycerate kinase</td>
<td>hPGK</td>
</tr>
<tr>
<td>Immunoglobulin</td>
<td>Ig</td>
</tr>
<tr>
<td>Insertions and Deletions</td>
<td>indels</td>
</tr>
<tr>
<td>Integration deficient lentiviral vectors</td>
<td>IDLV</td>
</tr>
<tr>
<td>Interferon-gamma</td>
<td>IFN\gamma</td>
</tr>
<tr>
<td>In Vitro Tanscribed</td>
<td>IVT</td>
</tr>
<tr>
<td>Jurkat, Clone E6-1</td>
<td>JE6.1</td>
</tr>
<tr>
<td><em>Lachnospiraceae bacterium</em> Cas12a</td>
<td>LbCas12a</td>
</tr>
<tr>
<td>Long terminal repeat</td>
<td>LTR</td>
</tr>
<tr>
<td>Major histocompatibility complex</td>
<td>MHC</td>
</tr>
<tr>
<td>Megenucleases</td>
<td>MNs</td>
</tr>
<tr>
<td>Methanol extraction residue of Bacillus Calmette-Guerin</td>
<td>MER</td>
</tr>
<tr>
<td>Minimum Essential Medium</td>
<td>Opti-MEM</td>
</tr>
<tr>
<td>Miss match</td>
<td>MM</td>
</tr>
<tr>
<td>Mre11-Rad50-Nbs1</td>
<td>MRN</td>
</tr>
<tr>
<td>Term</td>
<td>Abbreviation</td>
</tr>
<tr>
<td>---------------------------------------------------</td>
<td>--------------</td>
</tr>
<tr>
<td>Multiplicity of infection</td>
<td>MOI</td>
</tr>
<tr>
<td><em>Neisseria meningitidis Cas9</em></td>
<td>NmCas9</td>
</tr>
<tr>
<td>Next generation sequencing</td>
<td>NGS</td>
</tr>
<tr>
<td>Non-obese diabetic/ severe combined immunodeficiency/ γc⁻/⁻</td>
<td>NSG</td>
</tr>
<tr>
<td>Non-Homologous End Joining</td>
<td>NHEJ</td>
</tr>
<tr>
<td>Octamer</td>
<td>OCT</td>
</tr>
<tr>
<td>Penicillin-Streptomycin</td>
<td>Pen-Strep</td>
</tr>
<tr>
<td>peripheral blood mononuclear cells</td>
<td>PBMCs</td>
</tr>
<tr>
<td>Phenylmethylsulfonyl fluoride</td>
<td>PMSF</td>
</tr>
<tr>
<td><em>Planctomycetes CasX</em></td>
<td>PlmCasX</td>
</tr>
<tr>
<td>Polyethyleneimine</td>
<td>PEI</td>
</tr>
<tr>
<td>Polymerase</td>
<td>Pol</td>
</tr>
<tr>
<td>Pre-CRISPR RNA</td>
<td>Pre-crRNA</td>
</tr>
<tr>
<td>Programed cell death protein 1</td>
<td>PD1</td>
</tr>
<tr>
<td>Protospacer Adjacent Motif</td>
<td>PAM</td>
</tr>
<tr>
<td>proximal sequence element</td>
<td>PSE</td>
</tr>
<tr>
<td>Recombinant TCR</td>
<td>rTCR</td>
</tr>
<tr>
<td>Replication protein A</td>
<td>RPA</td>
</tr>
<tr>
<td>Ribonucleic acid</td>
<td>RNA</td>
</tr>
<tr>
<td>Ribonucleoprotein</td>
<td>RNP</td>
</tr>
<tr>
<td>Term</td>
<td>Acronym</td>
</tr>
<tr>
<td>-------------------------------------------</td>
<td>----------</td>
</tr>
<tr>
<td>Ribosomal RNA</td>
<td>rRNA</td>
</tr>
<tr>
<td>RNA interference</td>
<td>RNAi</td>
</tr>
<tr>
<td>Roswell Park Memorial Institute</td>
<td>RPMI</td>
</tr>
<tr>
<td>Self-inactivating</td>
<td>SIN</td>
</tr>
<tr>
<td>Short Regularly Spaced Repeats</td>
<td>SRSRs</td>
</tr>
<tr>
<td>Single chain variable fragment</td>
<td>scFv</td>
</tr>
<tr>
<td>Single guide RNA</td>
<td>sgRNA</td>
</tr>
<tr>
<td>single nucleotide polymorphisms</td>
<td>SNPs</td>
</tr>
<tr>
<td>Single-Strand Break Repair</td>
<td>SSBR</td>
</tr>
<tr>
<td>Single stranded RNA</td>
<td>ssRNA</td>
</tr>
<tr>
<td>Single-stranded synthetic oligodeoxynucleotide</td>
<td>ssODN</td>
</tr>
<tr>
<td>Site directed mutagenesis</td>
<td>SDM</td>
</tr>
<tr>
<td>Small nuclear RNA</td>
<td>snRNA</td>
</tr>
<tr>
<td>Small nucleolar RNAs</td>
<td>snoRNA</td>
</tr>
<tr>
<td>snRNA activating protein complex</td>
<td>SNAPc</td>
</tr>
<tr>
<td>Sodium Dodeyl Sulphate</td>
<td>SDS</td>
</tr>
<tr>
<td><em>Staphylococcus aureus</em> Cas9</td>
<td>SaCas9</td>
</tr>
<tr>
<td><em>Streptococcus canis</em> Cas9</td>
<td>ScCas9</td>
</tr>
<tr>
<td><em>Streptococcus pyogenes</em> Cas9</td>
<td>SpCas9</td>
</tr>
<tr>
<td><em>Streptococcus thermophiles</em> LMD-9</td>
<td>St1Cas9</td>
</tr>
<tr>
<td>Term</td>
<td>Abbreviation</td>
</tr>
<tr>
<td>----------------------------------------------------------------------</td>
<td>--------------</td>
</tr>
<tr>
<td>T cell receptor</td>
<td>TCR</td>
</tr>
<tr>
<td>T cell receptor alpha constant</td>
<td>TRAC</td>
</tr>
<tr>
<td>T cell receptor beta constant</td>
<td>TRBC</td>
</tr>
<tr>
<td>T cells redirected for universal cytokine-mediated killing</td>
<td>TRUCKs</td>
</tr>
<tr>
<td>Tet Methlycytosine Dioxygenase 2</td>
<td>TET2</td>
</tr>
<tr>
<td>Transactivating RNA</td>
<td>tracrRNA</td>
</tr>
<tr>
<td>Transcription-Activator-Like Effector Nucleases</td>
<td>TALENs</td>
</tr>
<tr>
<td>Transfer RNA</td>
<td>tRNA</td>
</tr>
<tr>
<td>Tris-acetate-EDTA</td>
<td>TAE</td>
</tr>
<tr>
<td>Transducing units</td>
<td>TUs</td>
</tr>
<tr>
<td>Tumour infiltrating lymphocytes</td>
<td>TILs</td>
</tr>
<tr>
<td>Tumour Necrosis Factor alpha</td>
<td>TNFα</td>
</tr>
<tr>
<td>Uracil DNA glycosylase</td>
<td>UDG</td>
</tr>
<tr>
<td>Uracil DNA glycosylase inhibitor</td>
<td>UGI</td>
</tr>
<tr>
<td>Variable</td>
<td>V</td>
</tr>
<tr>
<td>Woodchuck hepatitis virus posttranscriptional regulatory element</td>
<td>WPRE</td>
</tr>
<tr>
<td>X-ray cross complementing protein 4</td>
<td>XRCC4</td>
</tr>
<tr>
<td>XRCC4-like factor</td>
<td>XLF</td>
</tr>
<tr>
<td>Zinc Finger Nucleases</td>
<td>ZFNs</td>
</tr>
<tr>
<td>Zinc Finger Proteins</td>
<td>ZFPs</td>
</tr>
</tbody>
</table>
Chapter 1  Introduction

1.1 Genome editing

Targeted genome editing technologies have provided researchers with tools to investigate gene disruption, activation, repression, site specific integration of exogenous or corrective DNA sequences, and recently seamless base conversion. Such investigations have help decipher the genetic contribution to biological function, generation of both cell and animal models, as well as development of potential treatments for previously incurable diseases. Zinc Finger Nucleases (ZFNs) represented an early conceptualisation and realisation of a genome editing tool [1], closely followed by meganucleases (MNs) [2], transcription-activator-like effector nucleases (TALENs) [3, 4], megaTALs [5, 6], and clustered-regularly interspaced short palindromic repeats (CRISPR)/ CRISPR-associated protein (Cas9) [7-9]. These technologies have made their appearance in quick succession, with subsequent tools claiming increased activity and specificity, while reducing complexity and manufacturing costs. Such proliferative expansion of the available technologies has immensely benefited several sectors of research including the immunology field.

These genome editing platforms achieve genetic modification through the induction of site specific DNA double strand breaks (DSBs), repair of which depends upon the cell cycle phase, and the presence of a suitable repair template with homologous terminal ends. The preferred DSB repair pathway, referred to as non-homologous end joining (NHEJ) is an error-prone system frequently used to induce frame shift mutations, resulting in gene knockout or production of truncated protein products. Alternatively the cell may employ the homology directed repair (HDR) pathway in the presence of an ssDNA or dsDNA donor template flanked by homologous regions to both the 5’ and 3’ termini of the DSB. This non-error prone HDR pathway can be used to retain sequence fidelity, or to insert exogenous or corrective sequences at specific locations within the genome.
1.2 Protein-DNA interacting genome editing tools

Existing genome editing platforms are often split into two groups relying on either protein-DNA interaction for site specific binding, such as MNs, ZFNs, and TALENs or an RNA guided system through Watson: Crick base pairing as is the case for CRISPR/Cas9.

Application of MNs has been limited by the need to engineer protein residues that contact the DNA to alter the recognition site. Moreover, as DNA binding and nuclease activity share a single domain within these MNs, changing recognition can often affect cleavage activity. Furthermore, the relationship between these protein residues and DNA recognition was initially unclear, however, extensive work in these areas has helped overcome these hurdles [10, 11]. Although the zinc finger proteins (ZFPs) making up ZFNs display less uncertainty regarding the relationship between protein residues and DNA specificity, they do demonstrate context-dependent binding in large arrays. This complicates the optimisation of ZFNs for targeting desired DNA binding domains, and necessitates extensive screening of new ZFN pairs [12, 13]. TALEN DNA binding monomers show far less context dependence than ZFPs, although their highly repetitive sequences can act as a barrier for the production of novel TALEN arrays, as well as their delivery via retroviral vectors [12]. Both ZFNs and TALENs form DSBs by operating in pairs, thereby reducing the risks of off-target DNA cleavage, while complicating simultaneous disruption of multiple genomic loci due to the requirement to deliver multiple pairs of large proteins. Although significant strides have been made to overcome initial production, optimisation, screening and delivery hurdles associated with these protein-DNA interacting genome editing tools, further investigation is required. The thesis presented here will however focus on the use of the RNA guided CRISPR/Cas9 system, and associated technologies as detailed below.

1.3 CRISPR/Cas9 from bacterial immune system to genome editing tool

Ishino and associates (1987) described an unusual locus in *Escherichia coli* (*E. coli*) containing a structure of five highly homologous stretches of 29 nucleotides (nts),
separated by 32nt spacer sequences [14]. In hand with strides in DNA sequencing technologies, similar loci were observed in other prokaryotes from distinct phylogenetic groups. Mojica and colleagues (2000), through computational methods brought to light the presence of similar sequence motifs in a number of phylogenetic groups, suggesting these to be the most widespread family of repeats among prokaryotic genomes. They further characterised the layout of these loci as regularly spacer short palindromic repeats of between 20-40nts, separated by unique spacer sequences of a fixed length ranging from 20-58nts, dubbing such motifs as short regularly spaced repeats (SRSRs) [15]. Furthermore it became apparent that such SRSRs had a common origin due to the highly conserved nature of the repeat sequences with each phylogenetic group exemplified by E. coli and Salmonella typhi sharing all but one nucleotide of the 29bp repeat sequence. Interestingly, even at this early stage it was noted that the repeat sequences possessed characteristic of recognition sites on some DNA-binding proteins, such as size and inner short inverted repeats. Observations of the widespread and conserved nature of SRSR motifs supported the notion of a shared function.

Two years later four proteins were identified that were only present in prokaryotes possessing SRSR sequence motifs. Furthermore, the coding sequences for these proteins were found in close proximity to the SRSR loci suggesting that they were components of the same operon. Jansen et al (2002), therefore coined the term clustered-regularly interspaced short palindromic repeats (CRISPR) in regard to SRSRs and CRISPR-associated protein (Cas) in respect to their associated proteins (Cas1-4) [16]. Investigation of these cas genes revealed structural characteristics associated with roles in DNA metabolism/ repair or gene expression [17-19]. In 2005, multiple reports noted that new spacer sequences could be identified within CRISPR loci compared to their ancestral counterparts. Additionally these spacer DNA sequences shared homology with extra-chromosomal elements including phages, plasmids, and viruses [20-22]. It was suggested at this point that the CRISPR loci working in tandem with Cas proteins may provide an immunity against DNA from invading pathogens [21, 23]. Parallels were drawn between the function of the CRISPR loci and the eukaryotic RNA interference (RNAi) system [24]. However, it was observed that these
spacer sequences were removed or added modifying the phage resistant phenotype and suggesting that the CRISPR/Cas system represents an adaptive immune response [25]. Further supporting the idea that CRISPR/Cas represents an adaptive immune system, viruses were shown to undergo extensive recombination to evade a CRISPR mediated response [26].

The bacterial CRISPR immune system acts in three distinct stages, adaptation, expression, and interference. Adaptation consists of acquiring new spacer sequences referred to as protospacers, located next to the correct protospacer adjacent motif (PAM) from foreign DNA and inserting these sequences into CRISPR arrays (Figure 1-1 A). The Cas1 and Cas2 proteins are highly conserved across CRISPR loci containing species, which act to form a multiprotein complex (central Cas2 dimer, flanked by Cas1 dimers) that is essential for the acquisition of new spacer sequences [27]. The expression stage describes the transcription of the CRISPR array (Figure 1-1 B). Transcription of the CRISPR loci initially produces a single long non-coding RNA transcript (pre-crRNA), which interacts with a second transactivating RNA (tracrRNA) molecule, and endogenous RNase III enzyme, separating pre-crRNA into individual mature CRISPR RNAs (crRNAs) [28]. Binding of the tracrRNA at the repeat regions of the crRNA enables these hybrid RNA molecules to form ribonucleoprotein (RNP) complexes with Cas proteins, capable of recognising a specific sequence. During the interference stage of CRISPR immunity, RNP complexes exert nuclease activity at sites complementary to their protospacer sequences within invading DNA species (Figure 1-1 C). Importantly, the PAM sequence required for interference is not present within the CRISPR array to avoid self-cleavage by RNPs. For example the Cas9 endonuclease from *Streptococcus pyogenes* (SpCas9) recognises a 5’ NGG 3’ PAM, and as such protospacer sequences within the CRISPR array have a fixed 5’ GTT 3’ sequence, preventing their recognition.
**Figure 1-1 Streptococcus pyogenes CRISPR/Cas9 bacterial adaptive immune system:**

**A.** During the acquisition phase DNA from invading bacteriophages are incorporated within the *clustered-regularly interspaced short palindromic repeats* (*CRISPR*) locus as new protospacer sequences between repeat sequences. The protospacer-adjacent motif (PAM) is only present in the foreign DNA sequence, and not in the *CRISPR* loci to prevent cleavage within the bacterial genome. **B.** The *CRISPR* loci are expressed as a single long transcript (pre-crRNA). A second single stranded RNA (ssRNA) molecule is also expressed containing a region of complimentary sequence to that of the repeat.
sequence within the pre-crRNA, referred to as transactivating RNA (tracrRNA). Interactions between pre-crRNA and tracrRNA at the repeat regions allows processing by the RNase III enzyme to separate pre-crRNA into single CRISPR RNAs (crRNA). C. Mature crRNA forms ribonucleoprotein (RNP) complexes with CRISPR-associated protein 9 (Cas9) endonuclease. RNP complexes scan the foreign DNA for PAM sequences, and check for matching between the protospacer and target sequence. Successful matching of protospacer and target sequence permits the Cas9 endonuclease-mediated double strand break formation.

There are currently 2 major classes of CRISPR/Cas system that can be subdivided into 6 types and 34 subtypes. The class 1 CRISPR/Cas system includes types I, III, and IV and are characterised by the formation of multiprotein effector complexes, exemplified by the CRISPR-associated complex for antiviral defence (Cascade). On the other hand, the class 2 CRISPR/Cas system covers types II, V, and VI, and in contrast to class I relies on a single effector nuclease, such as Cas9, Cas12 or Cas13. Although both classes of CRISPR/Cas loci share the ability to encode for RNA-guided nuclease activity, distinct differences can be seen in the mode of action, targeted template (DNA/ RNA), as well as PAM sequence recognition [29, 30]. The naturally occurring class 2, type II CRISPR system is dependent on only four components; pre-crRNA, tracrRNA and RNase III enzyme for processing of pre-crRNA, as well as the Cas9 endonuclease [31, 32]. This was further simplified in 2012 by Jinek and colleagues, by fusing the crRNA and tracrRNA into a short single guide RNA (sgRNA) transcript, eliminating the need for RNase III and two separate RNA molecules, thereby reducing the required components to just sgRNA and Cas9 [8]. The relative simplicity of the system has promoted its investigation and application as a genome editing tool with primary focus on type II CRISPR/Cas9 from Streptococcus pyogenes.

1.4 CRISPR/Cas9 genome editing

This project utilises the class 2, type II, CRISPR/Cas9 system from Streptococcus pyogenes. As described above this system has been optimised to function as a precision genome editing tool that relies on two components, the SpCas9 endonuclease, and a sgRNA. Single guide RNA contains an 80 nucleotide scaffold
sequence that facilitates RNP complex formation with the SpCas9 endonuclease, while the 20 nucleotide protospacer determines target site recognition [8]. The SpCas9 component scans the DNA for suitable PAM sequences (5’ NGG 3’), at which point it unwinds the dsDNA and checks for complementary sequence between the protospacer and the DNA, forming a RNA-DNA hybrid. Upon successful site recognition, RNP complexes are able to disrupt the target sequence by inducing blunt-end DSBs between nucleotides 3-4 proximal to the PAM. Blunt-end DSBs are formed by the RuvC and HNH nuclease domains possessed by the SpCas9. The RuvC domain cleaves the single stranded DNA (non-target strand), while the HNH domain cleaves the sgRNA-DNA hybrid duplex (target strand) (Figure 1-2 A) [33].

This genome editing tool provides a means of precision genome engineering in a wide range of cell types and organisms. Additionally, the relative simplicity of retargeting this system has allowed its rapid application in gene function studies, drug target discovery, as well as production of modified crops [31, 34].

1.4.1 Damage repair pathways post SpCas9 genome editing

CRISPR/Cas9 genome editing triggers DSB formation at a specific site, determined by the sgRNA protospacer. Eukaryotic cells naturally encounter such double stranded DNA breaks by DNA damaging agents, and as part of developmental events including V(D)J recombination [35]. DNA repair mechanisms are vital, as incorrectly repaired DSBs can induce senescence, apoptosis and chromosomal instability. DSBs, will in the absence of an appropriate template be repaired by the immediate error prone non-homologous end joining (NHEJ) pathway. Processing DSBs in this manner often results in nucleotide insertions/ deletions (indels), leading to loss of function.

NHEJ can be split between two pathways with highly distinct sets of proteins. Canonical-NHEJ (C-NHEJ) is the predominant pathway mediating direct ligation of the juxtaposed broken DNA ends, whereas alternative end joining (Alt-EJ) occurs where standard repair has failed, and is associated with genomic instability [36, 37]. As NHEJ does not require the presence of a template, it can theoretically occur during any stage of the cell cycle. The C-NHEJ takes place in four sequential steps (i) NHEJ
complex formation upon DNA DSB end recognition (ii) bridging of the DNA ends, (iii) DNA end processing, and (iv) ligation followed by removal of the NHEJ complex (Figure 1-2 B). DNA end recognition and NHEJ complex formation is initiated by binding of the ring shaped Ku heterodimer (Ku70- Ku80), which plays a key role in DNA end processing by excision of abasic sites near the DSBs [38]. The Ku70 component contains a carboxyl-terminal region SAF-A/B, Acinus and PIAS (SAP) domain which is involved in DNA binding, while the Ku80 carboxyl-terminal region contains a protein scaffold for protein-protein interactions. Localisation of Ku heterodimer to DSBs occurs in seconds, binding the sugar backbone to allow it to bind in a sequence independent manner, and act as a scaffold for NHEJ complex formation [39]. The multiprotein NHEJ complex contains the DNA-dependent protein kinase catalytic subunit (DNA-PKcs), Artemis, X-ray cross complementing protein 4 (XRCC4), DNA ligase IV, polynucleotide kinase (PNK), and XRCC4-like factor (XLF). The DNA-PKcs, is the first component of the NHEJ complex to be recruited through interaction with the C-terminus of Ku80 [40]. Artemis, when bound to DNA-PKcs exhibits 5’ and 3’ endonuclease activity key to DNA end processing. The XRCC4 does not appear to possess any enzymatic activity, but has been shown to recruit PNK, XLF and DNA ligase IV. The PNK allows the addition of a 5’ phosphate group that is required for DNA ligation. XRCC4-XLF complex has been suggested to play a role in bridging of DNA ends through filament formation. Additionally ligation of the juxtaposed DNA ends is carried out by XRCC4 bound DNA ligase IV. Depending on the complexity of the DSB, it may be required to fill gaps in the DNA. This task is undertaken by the X polymerase family, which includes DNA polymerase µ and λ. Polymerase µ, in the presence of Ku and XRCC4-DNA ligase IV acts on discontinuous template, while polymerase λ can function in a template independent manner allowing removal of damaged bases. These mechanisms often results in indel formation, rarely restoring the original sequence.
Figure 1-2 Schematic of CRISPR-Cas9 and resulting DSB repair by C-NHEJ:

A. SpCas9 complexed with sgRNA. Complementary RNA-DNA duplex activates HNH, and RuvC nuclease domains resulting in blunt-end double strand break (DSB) formation (red stars). B. Canonical-non-homologous end joining (C-NHEJ) resolves DSB by initial binding by Ku heterodimer, followed by NHEJ complex formation often resulting in nucleotide insertions/deletions (indels). DNA-dependent protein kinase catalytic subunit (DNA-PKcs), X-ray cross complementing protein 4 (XRCC4), polynucleotide kinase (PNK), and XRCC4-like factor (XLF).

Genome editing also holds promise for targeted correction/insertion of endogenous and exogenous sequences. However, this requires delivery of DNA template containing the desired sequence for initiation of the homologous-directed repair (HDR) pathway which is restricted to either the S or G2 cell-cycle phases. The canonical HDR mechanism is thought to be more complex than the C-NHEJ pathway. In brief, HDR is initiated by Mre11-Rad50-Nbs1 (MRN) complex binding to the DSB site. MRN binding is closely followed by CtIP endonuclease activity, generating long 3’ ssDNA which is subsequently coated by replication protein A (RPA). Replacement of RPA by Rad51, allows nucleoprotein presynaptic filament formation permitting the search for donor template sequences with homology to the DNA break site. Upon recognition of the donor template, Rad51 dissociates, followed by synthesis and ligation [41]. There are a number of genetically distinct non-canonical HDR pathways.
which are error prone in nature. The mechanism for these alternative HDR pathways lie outside the scope of this report and have been summarised elsewhere [42].

Both integration deficient lentiviral vectors (IDLVs), and AAVs have been adapted as templates for HDR-based gene therapy. These have allowed targeted integration of large templates or entire genes with modest efficiency, predominantly focused on earlier generation editing tools with recent studies using CRISPR/Cas9 [43-45]. Both these templates require the use of viral vectors, which are costly to produce, time consuming to optimise, as well as requiring the correct infrastructure for their production. To overcome these barriers, a recent study investigated the use of an alternative, non-viral approach for delivery of dsDNA for correction of inherited autoimmune-associated mutations in T cells [46]. Contrary to expectations, they found that dsDNA templates >1kb amplified by PCR could be integrated into DSBs, with only modest impact on cell viability when co-delivered alongside CRISPR/Cas9 RNPs [46, 47]. However, the authors reported non-HDR insertions at the CRISPR/Cas9 induced DSB (~1%), as well as naturally occurring endogenous (~0.01%) DSBs. Single-stranded synthetic oligodeoxynucleotide (ssODN) represents another non-viral HDR template that has gained traction in recent years. These ssODNs have proven effective at correction of site-specific mutations or replacement of small sections, as they span approximately 200bp, 30-60bp of which comprise the 5’ and 3’ homology arms [46, 48-50]. Technical progressions such as these can broaden applicability, and importantly, have a significant clinical impact by dramatically reducing costs and manufacturing hurdles associated with viral approaches.

1.4.2 Opening up the genome by expanding the PAM recognition profile of CRISPR/Cas

In order to expand the applications of CRISPR/Cas genome editing in eukaryotic systems, attempts have been made to increase the breath of recognised PAMs (Table 1). This has involved the characterisation of other class 2, type II Cas9 orthologs recognising distinctive PAM sequences. These have included the Cas9 from *Staphylococcus aureus* (SaCas9) which scans DNA for a 5’ NNGRRT 3’ (R= A/G) PAM,
and has seen wide application due to its comparatively small size (1053 amino acid (AA) compared to the 1368AA SpCas9) [51]. While the recently characterised *Streptococcus canis* Cas9 (ScCas9), shares 89.2% homology with SpCas9, but displays a minimal 5’ NNG 3’ PAM sequence, further increasing the frequency of CRISPR/Cas editable sites [52].

Additionally, engineering of the commonly used SpCas9 endonuclease has been performed to generate variants with altered PAM specificities. Work from the Joung group presented SpCas9-VQR, EQR, and VRER variants which favoured NGAN/NGCG, NGAG and NGCG PAM sequences (5’-3’), respectively [53]. Building upon this, two reports were published in quick succession reporting the generation of SpCas9 variants with further relaxed PAM specificity. The Liu group used phage-assisted continuous evolution to evolve a SpCas9 capable of targeting NG, GAA, and GAT PAM motifs (5’-3’), dubbed xCas9(3.7) [54]. Similarly, Nishimasu et al (2018), used a rational design approach to produce a SpCas9 variant able to target 5’ NG 3’ PAMs (SpCas9-NG) [55]. Although both xCas9(3.7), and spCas9-NG are able to target 5’ NG 3’ PAMs, SpCas9-NG has been reported to have superior editing at NGC, NGT, and NGA PAMs. However, SpCas9-NG shows marked reduction in editing at traditional 5’ NGG 3’ PAMs compared to SpCas9, which is not observed with the xCas9(3.7) variant [55]. Engineering of the SpCas9 orthologs, SaCas9 and FnCas9 has also yielded variants with a more relaxed PAM sequence (KKH-SaCas9, and RHA-FnCas9) [56, 57]. These variants all have specific niches, expanding the targetable regions of the genome.

Further expanding the CRISPR/Cas genome editing toolbox; use of the class 2, type V CRISPR system has gained traction in recent years. Zetsche et al (2015) characterised the class 2, type V, endonuclease Cas12a (formally known as Cpf1) from both *Acidaminococcus* (AsCas12a) and *Lachnospiraceae bacterium* (LbCas12a), for genomic editing applications in HEK 293T cells [58]. These type V Cas12a endonuclease show serval distinctions from that of the type II Cas9 enzymes. Notably, Cas12a proteins are smaller, and possess RNase activity allowing them to process their own pre-crRNA without the need for tracrRNA. Additionally, the Cas9 endonuclease relies on both a RuvC and HNH domain to cleave the non-target and
target strands between bases 3-4 proximal to the PAM to form blunt end DSBs. In contrast, Cas12a molecules contain a single RuvC-like domain, sequentially cleaving the target and non-target strands to form DSBs with staggered ends (5 or 8 nt 5’ overhangs). Importantly, application of these as genome editing tools further expands the editable sites within the genome by substituting the G-rich PAM generally exhibited by Cas9, for the T-rich PAM recognised by Cas12a molecules [59].

Despite encouraging results with these Cas12a molecule the relatively long 5’ TTTV 3’ (V= A/G/C) PAM recognised by both AsCas12a, and LbCas12a enzymes has been predicted to reduce the editable sites in the genome by ~6 fold compared to the SpCas9 5’ NGG 3’ PAM [60]. In a similar tact to that used for SpCas9, characterisation of other type V CRISPR/Cas systems with more relaxed PAM sequences for genome editing are starting to bridge this gap. Notable for applications in mammalian genome editing, both the type V-B *Alicyclobacillus acidiphilus* Cas12b (AaCas12b), and type V-E *Planctomycetes* CasX (PlmCasX) enzymes have been described and possess distinctive T-rich PAMs [61, 62]. Furthermore, Goa and colleagues (2017) used targeted mutagenesis of the 60 AA residues in proximity to the PAM within the SaCas12a, to screen variants able to recognise non-canonical PAMs [60]. Step wise interrogation of these variants showed that the AsCas12a-RR and AsCas12a-RVR had the highest activity in 293T cells, and targeted non-canonical TYCV and TATV PAM motifs (5’-3’ Y= C/T, V= A/G/C) respectively. Recently Kleinstiver et al (2019) further expanded the targetable range of this type V system by describing an enhanced AsCas12a (enAsCas12a) variant, able to generate robust editing at the canonical 5’ TTTV 3’, and non-canonical VTTV, TTTT, TTCN, and TATV (5’-3’) PAM sequences [63]. Furthermore, they demonstrate that inclusion of the E174R mutation in their enAsCas12a enzyme increased on-target cleavage at the canonical 5’ TTTV 3’ PAM compared to the wild type AsCas12a, across a range of temperatures (25-37°C). Inclusion of this E174R mutation into the previously described AsCas12a-RR and RVR variants, demonstrated a similarly enhanced on-target activity (enRR and enRVR) [63]. These advances, increase editable regions as well as increase the activity of this type V systems for genome editing applications.
Although class 1, type I CRISPR/Cas systems represent the most widespread of these loci within bacteria and archaea (~90%) compared to class 2 (~10%), there still remains a lot of unknowns regarding their CRISPR arrays, multiprotein complexes, crRNA and PAM specificity. This gap in understanding, as well as the need to deliver the multiprotein complex Cascade with its associated Cas3 helicase-nuclease hinders the application of type I systems for the purpose of genome editing [30, 64]. However, further understanding of these loci has recently led to the first applications of the type I-E CRISPR/Cas system from *E.coli* and *Thermobifida fusca* as a genome editing tool in both human cell lines and human embryonic stem cells [64, 65]. Although these type I-E system present with relatively small A rich PAMs providing access to other niche areas of the genome, they are currently cumbersome to deliver and not fully characterised. However, advancements in the understanding of these systems may yield genome editing tools with novel applications.

Table 1 Cas endonucleases:

<table>
<thead>
<tr>
<th>Cas</th>
<th>Class / type</th>
<th>PAM (5’-3’)</th>
<th>Publication</th>
</tr>
</thead>
<tbody>
<tr>
<td><em>Streptococcus pyogenes</em> Cas9</td>
<td>2/ II</td>
<td>NGG</td>
<td>[8]</td>
</tr>
<tr>
<td><em>(SpCas9)</em></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><em>Streptococcus thermophiles</em> LMD-9 (St1Cas9)</td>
<td>2/ II</td>
<td>NNAGAAW</td>
<td>[7]</td>
</tr>
<tr>
<td><em>Neisseria meningitidis</em> Cas9</td>
<td>2/ II</td>
<td>NNNNGATT</td>
<td>[66]</td>
</tr>
<tr>
<td><em>(NmCas9)</em></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>SpCas9 A262T, R324L, S409I, E480K, E543D, M694I, E1219E (xCas9 3.7)</td>
<td>2/ II</td>
<td>NG, GAA, GAT</td>
<td>[54]</td>
</tr>
<tr>
<td>Organism</td>
<td>Cas Type</td>
<td>Protospacer</td>
<td>Reference</td>
</tr>
<tr>
<td>----------------------------------</td>
<td>------------</td>
<td>-----------------------</td>
<td>-----------</td>
</tr>
<tr>
<td>Staphylococcus aureus Cas9 (SaCas9)</td>
<td>2/ II</td>
<td>NNGRRT</td>
<td>[51]</td>
</tr>
<tr>
<td>SaCas9 E782K, N968K, R1015H (KKH-SaCas9)</td>
<td>2/ II</td>
<td>NNNRRT</td>
<td>[56]</td>
</tr>
<tr>
<td>Acidaminococcus Cas12a (AsCas12a)</td>
<td>2/ V-A</td>
<td>TTTTV</td>
<td>[58]</td>
</tr>
<tr>
<td>AsCas12a S542R, K607R (AsCas12a-RR)</td>
<td>2/ V-A</td>
<td>TYCV</td>
<td>[60]</td>
</tr>
<tr>
<td>AsCas12a E174R, S542R, K607R (enRR)</td>
<td>2/ V-A</td>
<td>TYCV</td>
<td>[63]</td>
</tr>
<tr>
<td>AsCas12a S542R, K548V, N552R (AsCas12a-RVR)</td>
<td>2/ V-A</td>
<td>TATV</td>
<td>[60]</td>
</tr>
<tr>
<td>AsCas12a E174R, S542R, K548V, N552R (enRVR)</td>
<td>2/ V-A</td>
<td>TATV</td>
<td>[63]</td>
</tr>
<tr>
<td>AsCas12a E174R, S542R, K548R (enAsCas12)</td>
<td>2/ V-A</td>
<td>TTTT, VTTV, TTTT, TTCN, TATV</td>
<td>[63]</td>
</tr>
<tr>
<td>Lachnospiraceae bacterium Cas12a (LbCas12a)</td>
<td>2/ V-A</td>
<td>TTTT</td>
<td>[58]</td>
</tr>
<tr>
<td>Alicyclobacillus acidiphilus Cas12b (AaCas12b)</td>
<td>2/ V-B</td>
<td>TTN</td>
<td>[61]</td>
</tr>
<tr>
<td>Planctomycetes CasX (PlmCasX)</td>
<td>2/ V-E</td>
<td>TTCN</td>
<td>[62]</td>
</tr>
<tr>
<td>SpCas9 D1135V, R1335Q, T1337R (SpCas9-VQR)</td>
<td>2/ II</td>
<td>NGAN, NGCG</td>
<td>[53]</td>
</tr>
<tr>
<td>Cas9 Name</td>
<td>PAM Motif</td>
<td>Ref.</td>
<td></td>
</tr>
<tr>
<td>-----------------------------------------------</td>
<td>-----------</td>
<td>------</td>
<td></td>
</tr>
<tr>
<td>SpCas9 D1135E, R1335Q, T1337R (SpCas9-EQR)</td>
<td>2/ II</td>
<td>NGAG</td>
<td>[53]</td>
</tr>
<tr>
<td>SpCas9 D1135V, G1218R, R1335E, T1337R (SpCas9-VRER)</td>
<td>2/ II</td>
<td>NGCG</td>
<td>[53]</td>
</tr>
<tr>
<td>Francisella novicida (FnCas9)</td>
<td>2/ II</td>
<td>NGG</td>
<td>[57]</td>
</tr>
<tr>
<td>FnCas9 E1369R, E1449H, R1556A (RHA-FnCas9)</td>
<td>2/ II</td>
<td>YG</td>
<td>[57]</td>
</tr>
<tr>
<td>Campylobacter jejuni Cas9 (CjCas9)</td>
<td>2/ II</td>
<td>NNNNACA</td>
<td>[67]</td>
</tr>
<tr>
<td>Streptococcus canis Cas9 (ScCas9)</td>
<td>2/ II</td>
<td>NNG</td>
<td>[52]</td>
</tr>
<tr>
<td>E.coli type I-E Cascade-Cas3</td>
<td>1/ I-E</td>
<td>ARG</td>
<td>[64]</td>
</tr>
<tr>
<td>Thermobifida fusca type I-E Cascade-Cas3</td>
<td>1/ I-E</td>
<td>AAG</td>
<td>[65]</td>
</tr>
</tbody>
</table>

**Cas endonucleases currently applied to genome editing within eukaryotic systems.** N= any base pair, W= A/T, R= A/G, V= A/G/C, and Y= C/T.

### 1.4.3 Minimising off-target activity of CRISPR/Cas genome editing

The CRISPR/ Cas9 system is highly efficient at causing DNA cleavage when directed to a specific site by a 20nt protospacer sequence. However, it has been determined that 1-2nt mismatches are tolerated in the guide sequences with varying efficiencies depending on their location resulting in off-targeted cleavage. Additionally, it has been determined that as many as 5nt mismatches from the guide sequence can be tolerated [68]. The scope of such off-target cleavage effects are wide and could be detrimental to cell survival or could potentially be oncogenic.

Fu and colleagues (2014) hypothesised that reducing the length of the sgRNA-DNA interface would minimise the potential for off-target cleavage [69]. To achieve this,
truncated protospacers between 15-18nt were generated by removal of nucleotides from the 3’ end. It appeared that comparable on-target Cas9 cleavage efficiency could be retained with truncated guides measuring 17-18nt in a 293T cell line. They further demonstrated reduction in off-target Cas9 scission of as much as 5000 fold using these truncated guide sequences. These observations were supported by Tsai et al (2015), who reported substantially decreased off-target activity with 3 truncated sgRNAs compared to their full length counterparts when measured by Genome-wide Unbiased Identification of DSBs Enabled by Sequence (GUIDE-seq) [70]. Although promising, investigations by Zhang et al (2016) revealed that use of such truncated guide sequences (17nt) in mesenchymal stem cells (MSCs) and induced pluripotent stem cells (iPSCs) reduced on-target Cas9 editing by 10-20% compared to using a 20nt protospacer [71].

Further work to minimise off-target events has taken advantage of mutations that change the SpCas9 endonuclease into a nickase [72]. As described above, the SpCas9 endonuclease contains two nuclease domains referred to as RuvC and HNH, which cleave the non-target and target strands respectively, facilitating DSB formation [33]. Mutation of the RuvC (D10A nCas9) or HNH (H840A nCas9) domain prevents the formation of DNA DSBs, allowing cleavage of only a single DNA strand. Ran and associates (2013) describe the use of both nCas9 mutants with a pair of offset sgRNAs, binding opposing DNA strands to improve Cas9 cleavage specificity [72]. Nicking of DNA at both sgRNA target sites results in DNA DSBs and NHEJ mediated insertions and deletions (indels), whereas individual nicks in the DNA are repaired via the robust single-strand break repair (SSBR) pathway. This therefore invokes a level of similarity to ZFNs and TALENs where DNA cleavage is dependent on the synergistic interaction of a pair of DNA binding molecules. Use of this strategy has been shown to reduce off-target Cas9 scission by 50-1000 fold in cell lines [72]. Moreover, recent work by Gopalappa et al (2018) established that the HNH domain of Cas9 cleaves more efficiently than the RuvC domain in mammalian cells, leading to increased gene disruption using a D10A Cas9 strategy [73]. Furthermore, use of D10A Cas9 nickase has been combined with truncated guide RNA to further reduce off-target cleavage [69]. Although results are promising, D10A Cas9 directed with a single guide has
shown the capacity to induce indels at some target sites [68, 69, 74, 75]. Additionally, the use of a second guide introduces the need to screen more predicted off-target sites, as well as precluding some methods of delivery. A variant containing both D10A and H840A mutations is referred to as catalytically dead SpCas9 (dCas9), as it is unable to cleave either DNA strand [8]. Although unable to produce DSBs, dCas9 has proven invaluable in gene activation/ inactivation studies. At the time of writing this dual nickase approach is not possible for the any of the type V Cas12 enzymes due to their general structure containing only a single RuvC-like nuclease domain; mutation of which results in a catalytically dead Cas12 [58].

Alternatively a number of high-fidelity SpCas9 variants have been described to circumvent off-target activity while retaining protospacer length and use of a single endonuclease (Table 2). Back to back reports from the Zhang and Joung labs reported development of SpCas9 variants with reduced off-target activity [76]. Slaymaker and associates (2016), hypothesised that neutralising the positive residues within the SpCas9 non-target strand groove would weaken non-target strand binding. This in turn would encourage re-hybridisation of the target and non-target strands, thereby creating an environment that favours sgRNA-DNA hybrid duplexes with strong Watson-Crick base pairing. Through the screening of 32 variants, they were able to generate a version that demonstrated enhanced specificity (eSpCas9 1.1) [77]. On the other hand, Kleinstiver and colleagues (2016) hypothesised that wild type SpCas9-sgRNA complexes binds the target strand with more energy than is required, allowing tolerance of mismatches within the protospacer sequence [76]. Based on this excess energy hypothesis, they tested 15 variants that possessed a combination of mutations in four amino acids sharing direct hydrogen bonding between the SpCas9 and the target strand (N497, R661, Q695, Q926), while still retaining on-target activity. Based on the scope of off-target activity in an EGFP disruption assay they recommend the use of a quadruple mutant variant referred to as high-fidelity SpCas9 (SpCas9-HF1). This SpCas9-HF1 variant demonstrated comparable activity to wild type SpCas9 (>85% of target sites) with greatly reduced disruption at sites containing mismatches. Interestingly, further work by Chen and colleagues (2017) established similar DNA binding affinity of wild type SpCas9 and SpCas9-HF1 suggesting an
alternative mechanism for reduced off-target cleavage distinct from the excess energy hypothesis [78]. Discovery that the HNH domain of SpCas9 is trapped in an inactive state when bound to substrate with mismatches led Chen and colleagues (2017) to hypothesise that SpCas9-HF1 displays heightened sensitivity in undergoing conformational rearrangements resulting in HNH inactivation in the presence of mismatched DNA template [78]. This mechanism was supported by investigation from the Doudna group, who observed using Forster resonance energy transfer that the HNH nuclease domains of both SpCas9-HF1 and eSpCas9 1.1 were trapped in an inactive conformation when bound to mismatched targets [78]. Informed by structural details, mutations were made within the Cas9 recognition (REC) lobe (REC3), which interacts with the RNA-DNA hybrid duplex and undergoes conformational changes when bound to the target DNA template. A quadruple mutant demonstrated equivalent on-target editing compared to both eSpCas9 1.1 and SpCas9-HF1, while demonstrating comparable levels to that of wild-type SpCas9 (>70%). Additionally GUIDE-seq analysis revealed that this quadruple variant, improved specificity compared to wild-type SpCas9, and was equivalent/better than both eSpCas9 1.1 or SpCas9-HF1, leading to the name hyper-accurate Cas9 (HypaCas9) [78]. Building upon this, Casini et al (2018), developed a yeast based assay for screening SpCas9 variants with mutations in REC3 regions for improved on- and off-target activity. This yeast-based assay, with further validation in a 293T cell-line generated a variant referred to as evolved Cas9 (evoCas9) boasting on-target activities similar to eSpCas91.1, and improved efficiency compared to SpCas9-HF1 while demonstrating increased fidelity compared to both [79]. Despite advancements, eSpCas9 1.1, SpCas9-HF1, and HypaCas9 variants presented highly reduced on-target activity when delivered in a more transient, clinically applicable, RNP complex format. However, using an unbiased high-throughput bacterial selection assay, Vakulskas et al (2018) presented a single R691A mutation in the SpCas9 REC3 domain retained on-target activity compared to wild-type SpCas9, while substantially reducing off-target activity when delivered as RNP complexes (HiFi Cas9) [80]. With the rapidly expanding nature of this field high-fidelity variants are becoming available for Cas9 orthologs including SaCas9, as well as for the type V AsCas12a endonuclease [63, 77, 81].
Table 2 High-Fidelity Cas variants:

<table>
<thead>
<tr>
<th>Cas</th>
<th>Class / type</th>
<th>PAM (5’-3’)</th>
<th>Publication</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>SpCas9 K848A, K1003A, R1060A</strong> (eSpCas9 1.1)</td>
<td>2/ II</td>
<td>NGG</td>
<td>[77]</td>
</tr>
<tr>
<td><strong>SpCas9 N497A, R661A, Q695A, Q962A (SpCas9-HF1)</strong></td>
<td>2/ II</td>
<td>NGG</td>
<td>[76]</td>
</tr>
<tr>
<td><strong>SpCas9 N69A, M694A, Q695A, H698A (HypaCas9)</strong></td>
<td>2/ II</td>
<td>NGG</td>
<td>[78]</td>
</tr>
<tr>
<td><strong>SpCas9 M495V, Y515N, K526E, R661Q (evoCas9)</strong></td>
<td>2/ II</td>
<td>NGG</td>
<td>[79]</td>
</tr>
<tr>
<td><strong>SpCas9 R691A (HiFi Cas9)</strong></td>
<td>2/ II</td>
<td>NGG</td>
<td>[80]</td>
</tr>
<tr>
<td><strong>SaCas9 R499A, Q500A, R654A, G655A (S-HF)</strong></td>
<td>2/ II</td>
<td>NNGRRT</td>
<td>[77]</td>
</tr>
<tr>
<td><strong>SaCas9 R245A, N413A, N419A, R654A (SaCas9-HF)</strong></td>
<td>2/ II</td>
<td>NNGRRT</td>
<td>[81]</td>
</tr>
<tr>
<td><strong>AsCas12a E174R, N282A, S542R, K548R (enAsCas12-HF)</strong></td>
<td>2/ V-A</td>
<td>TTV, VTTV, TTTT, TTCN, TATV</td>
<td>[63]</td>
</tr>
</tbody>
</table>

*High-fidelity Cas variants applied to eukaryotic genome editing. N= any base pair, R= A/G, and V= A/G/C.*
1.5 Detection of off-target CRISPR/Cas9-mediated cleavage

As detailed above, the scope of off-target CRISPR/Cas9 activity, has been narrowed by using truncated protospacers [82], paired nCas9 [83], or high-fidelity Cas9 variants [80]. These advancements have improved the safety profile of many CRISPR/Cas genome editing systems, although the potential remains for off-target indel formation, large deletions extending over kilobases, and chromosomal translocations [84-86]. To assist in locating potential CRISPR/Cas9 off-target sites, many protospacer design tools are available providing an in silico prediction of the on- and off-target potential [87, 88]. In silico predicted off-target scores are generally based on sequence alignments, searching for genomic locations possessing mismatches to the protospacer sequence, however, these have shown limitations [89, 90]. Development of predictive tools using deep-learning models, such as DeepCRISPR (http://www.deepcrispr.net/), and DeepSpCas9 (http://deepcrispr.info/DeepSpCas9) aims to increase their accuracy [91, 92]. Rather than computational based predictions, a number of in vitro techniques that rely on high throughput sequencing of genomic DNA treated with RNP complexes have been developed [51, 70, 93, 94]. Nevertheless, at the time of writing there remains no industry standard for prediction and detection of off-target events, and relatively little application outside of SpCas9 based editing. Furthermore, it has become apparent that indel formation by the NHEJ pathway may not be as unpredictable as previously thought, with Shen et al (2018) developing a machine learning model that allows researchers to choose sgRNA sequences that are predicted to preferentially resolve in specific indels (inDelphi: https://www.crisprindelphi.design) [95]. However, this indel prediction tool is currently only optimised for specific cell types with SpCas9 endonuclease based editing, and will require further development.

1.6 CRISPR uncut: seamless base conversion

In order to avoid the issues described above associated with SpCas9 induced DSB formation both the Liu and Kondo groups, in 2016 described CRISPR/Cas9 cytidine base editors (CBEs) capable of generating precise base conversions without DNA DSBs, or exogenous template DNA [74, 96]. Komor and associates (2016), described
the development of three generations of CBE, the first of which fused a rat apolipoprotein B mRNA editing enzyme catalytic polypeptide 1 (rAPOBEC1) cytidine deaminase to a dCas9 (D10A, H840A) allowing for C>U conversions which were corrected to T bases upon DNA replication (BE1). Cytidine deamination events can occur naturally within eukaryotic cells, and as such they possess uracil DNA glycosylase (UDG) activity that resolves U:G heteroduplexes by removing U’s in the DNA through the base excision repair (BER) pathway. With this in mind, further refinements saw the addition of a uracil DNA glycosylase inhibitor (UGI) to prevent early removal of edited Cs (BE2). Additionally, it was theorised that nicking the unedited strand may promote correction from the edited uncut strand, promoting the replacement of the dCas9 with the D10A nCas9 forming (BE3) [74] (Figure 1-3).

Chapter 5 of this report details the application of BE3 for T cell genome editing. However, CBEs have been rapidly optimised for increased activity, reduced indel formation, and non-C>T changes [97-99]. Moreover, changes to the deaminase activity window within the protospacer have increased both precision and utility [100, 101]. Similarly, a number of groups have characterised alternative deaminases, which display unique editing efficacies, windows, and preferences including temperature, sequence context, and methylation status [102-104]. Expansion of the number of targetable bases has also been achieved by replacing the SpCas9 nickase with other endonucleases such as xCas9, SaCas9, or dAsCas12a [54, 97, 101, 105]. The Liu group have additionally developed an adenosine deaminase base editor (ABE), able to generate A>G modifications by replacing rAPOBEC1 with a modified E.coli TadA enzyme, which also negates the requirement for the UGI subunit [106].

The combination of both CBE and ABE technology in combination with the rapid refinement of these tools, has huge potential for the correction of pathogenic single nucleotide polymorphisms (SNPs) [97].
Third generation cytidine base editor:

Figure 1-3 Third generation cytidine base editor (BE3), comprising an N terminal rat APOBEC1 (rAPOBEC1, yellow), D10A nickase SpCas9 (nCas9, green), followed by a uracil glycosylases inhibitor (UGI, red), and C terminal SV40 nuclear localisation signal (NLS, blue). Upon nCas9 complexing with single guide RNA (sgRNA) and subsequent target site recognition, the rAPOBEC1 deaminates cytidines (Cs) to uracils (Us) between protospacer positions 4-8 (red box) distal to the protospacer adjacent motif (PAM), on the single stranded DNA (ssDNA). The UGI prevents the endogenous uracil DNA glycosylases (UDG) from recognising the U:G heteroduplexes, thereby blocking their removal by base excision repair. Additionally, nCas9 cleaves the unedited sgRNA bound DNA strand between positions 17 and 18 distal to the PAM (red star), promoting correction from the uncut, edited strand. CBES can be used to disrupt gene expression by removal of the start codon (silencing), mRNA splice donor site inference (retention of intronic sequence), mRNA splice
acceptor inference (exon skipping), and introduction of a premature stop codon (truncated protein product).

Since its conceptualisation and realisation, base editing technology has seen applications in a number of fields including crop improvement [107-110], generation of animal models [108, 111, 112], development of gene disruption tools [101, 113, 114], exon skipping strategies [115, 116], and correction of pathogenic single nucleotide polymorphisms (SNPs) [74, 117]. As such, base editing tools have been tested in a variety of mammalian cells types, including fibroblasts [98, 99], post mitotic sensory cells [118], cancer cell lines [99], blastocysts and embryonic cells [109, 111, 119-122], with Webber and associated (2019) exploring their use in T lymphocytes for seamless knockout while minimising concerns associated with current nuclease platforms [123].

1.7 Delivery of CRISPR reagents

Wide spread application of the CRISPR/Cas9 system has promoted the development of a variety of delivery methods to target cells (Figure 1-4). Delivery strategy is often tailored to cell type and experimental design. Development of novel techniques, characterising Cas endonucleases, as well as screening Cas variants often rely on plasmid DNA based of entire CRISPR/Cas9 cassettes [72, 124-126] (Figure 1-4 A, and B). The reasoning behind this choice is that transfection with a DNA expression plasmid such as PX458 (addgene: #48138) provides an inexpensive and relatively simple method of expressing both SpCas9 endonuclease and sgRNA without the need for viral vector production, or transfection of multiple components in vitro. However, this approach is not suitable for all cells types due to the innate immune DNA-sensing pathway damaging the host cell, leading to senescence [127]. Alternatively, lentiviral vectors have been adapted for the integrated or non-integrated delivery of CRISPR/Cas9 due to their high infection efficiency, mild immunogenicity, large packaging capacity and stable expression following integration of viral DNA into the host genome. Vectors such as lentiCRISPR v2 (addgene: #52961) can be designed to express both the SpCas9 endonuclease and the sgRNA. Adeno-associated viruses (AAVs) have also been applied to CRISPR/Cas delivery due to the availability of a wide
arrange of serotypes, their non-pathogenic nature, low immunogenicity, and avoidance of DNA sensing. Additionally, most serotypes of AAV are expressed transiently, not integrating into the DNA thereby minimising the risk of insertional mutagenesis. However, the application of many CRISPR/Cas9 systems in AAV is complicated by a relatively small packaging capacity (~4.5kb). Although it has proven feasible to package SpCas9 and sgRNA expression cassette (~4.2kb) into a single AAV, it leaves very little room for regulatory elements or other transgenes. To overcome limited AAV capacity, truncated SpCas9, smaller Cas9 orthologs, or application of the more compact type V CRISPR/Cas system has been used [33]. Moreover, expressing the Cas9 and sgRNA on separate AAVs, has been shown to mitigate this issue, however it necessitates the purification of two AAVs, as well as co-infection of target cells [33]. Furthermore, these delivery strategies are hampered by concerns surrounding off-target cleavage and activation of pre-existing immunity in the face of the prolonged, or constitutive expression of prokaryote derived SpCas9, deeming such strategies unfit for clinical translation [128-130].

Transient delivery of individual CRISPR/Cas9 components through co-electroporation of sgRNA with SpCas9 supplied as mRNA or protein, circumvents these concerns and represents a more clinically applicable strategy [131-133] (Figure 1-4 C, and D). Delivery of SpCas9 mRNA to the cytoplasm can be readily translated and exert DNA cleavage effects, which is well tolerated by both primary cells and cell lines. Pre-complexed Cas9 protein with sgRNA is the most common strategy as there is no requirement for transcription, codon optimisation, promoter choice, or vector design. Additionally, as no translation is required, RNPs present the most immediate and highly transient strategy. RNPs are also associated with high editing efficiency using both wild-type and HiFi-Cas9, reduced off-target, as well as low toxicity, and undetectable adaptive immune responses [134].

CRISPR/Cas9 genome editing also relies on the appropriate delivery of sgRNA. However, initial reports using enzymatically in vitro transcribed (IVT) sgRNAs were riddled by low editing efficiencies, attributed to their rapid degradation [135]. Furthermore, incomplete removal of triphosphates at the 5’ end of IVT sgRNAs have been shown to trigger potent innate immune responses in mammalian cells, akin to
those against RNA viruses through upregulation of retinoic acid-inducible gene I (RIG-1) [136]. The development of automated solid-phase RNA synthesis of sgRNAs not only increased levels of editing, but allowed addition of chemically modifications that act to evading innate immune sensing [135, 137]. The addition of 2'-O-methyl 3’phosphorothioate modifications to the first and last three nt’s in the sgRNA, has immensely improved stability and editing efficiencies by more than 30-fold over non-modified sgRNAs when transiently co-delivered with Cas9 mRNA or as ribonucleoprotein (RNP) complexes in primary T cells and are now being implemented in the vast majority of studies [50, 135, 138].

Delivery of Cas9 mRNA with sgRNA or as RNPs is most often achieved by physical non-viral methods, including nanoparticles, microinjection, lentiviral-like particles, and electroporation [50, 133, 139-142]. The method chosen is suited to the experimental design as each has advantages and disadvantages. Of these methods, electroporation is compatible with all CRISPR/Cas systems and is widely used in mammalian cells. Upon electroporation the cell membrane is temporarily made more permeable, allowing proteins or nucleic acids to enter the cytoplasm. However electroporation protocols have to be carefully optimised for each target cell type, as this process can result in high levels of cell death. The CRISPR/ Cas9 system can be multiplexed to allow simultaneous disruption of multiple loci by a variety of methods. Delivery of RNPs or Cas9 mRNA with sgRNA presents a relatively straightforward route to multiplexing by the addition of multiple sgRNA sequences within a single reaction [143]. Other delivery methods often require the addition of multiple sgRNA expression cassettes requiring further vector design.

The delivery strategies described above are also applicable to SpCas9 orthologs, as well as other CRISPR systems. However, many of the recently characterised systems are currently available only as DNA expression plasmids, with few including the SaCas9 and AsCas12a commercially available as protein.
Figure 1-4 Delivery strategies of CRISPR/Cas9 reagents:

A. Integrating viral vectors encoding sgRNA and SpCas9 or a combination of the two can be delivered to a variety of cell types by viral transduction. Stable integration into the target cell genome permits long term constitutive expression of CRISPR/Cas9 components. Once integrated, sgRNA is transcribed by RNA polymerase III promoters, and mRNA SpCas9 transcribed by an RNA Pol II promoter before undergoing translation. Ribonucleoprotein (RNP) complexes are then trafficked into the nucleus for targeted genomic cleavage. B. Plasmid DNA expressing either sgRNA, SpCas9 or both can be delivered by transfection or electroporation protocols. Plasmid DNA remains episomal, and will be diluted out over the course of cell divisions providing a route for transient delivery. Transcription of both sgRNA and SpCas9, and translation of SpCas9 protein, is followed by formation of RNPs that will enter the nucleus and cleave the target genome. C. Synthesised sgRNA and SpCas9 mRNA can be transiently delivered by transfection or electroporation. SpCas9 mRNA is translated into protein, which forms RNP complexes that are able to enter the nucleus and edit the target sequence. D. Pre-complexing of SpCas9 protein and synthesized sgRNA forms RNPs that can be delivered by transfection or electroporation. RNPs will traffic to the nucleus where they cleave target sequence.
In this project the delivery method chosen has to be clinically applicable for primary human T cells, allowing the expression of a immunotherapeutic transgene, as well as providing efficiency CRISPR/Cas9 genome editing, where SpCas9 endonucleases expression is transient. To achieve these criteria, transgene expression is provided by lentiviral vector, due to their ability to efficiently transduce a diverse range of dividing and quiescent cell types, semi-randomly integrate into the host genome, and permit long term transgene expression [144]. Furthermore, the lentiviral vector used here also allows for constitutive sgRNA expression from a RNA polymerase III promoter, while also restricting knockout to the transduced cell population. SpCas9 endonuclease can subsequently be delivered to vector transduced cells by electroporation in a transient format, such as mRNA or protein.

1.8 Brief description of gene therapy lentiviral vectors

Lentiviral vectors used for gene therapy applications are based on the retroviridae family member HIV-1 and are phenotypically spherical, ranging from 80-120nm. Lentiviral vectors contain a diploid positive-sense ssRNA genome, complexed with nucleocapsid protein, reserve transcriptase, integrase, and protease. The ssRNA viral genome is further surrounded by a capsid protein shell making up the viral core. The viral core is in turn surrounded by matrix proteins which interact with the virion lipid envelope incorporating viral envelope glycoproteins. Envelope glycoproteins consist of a transmembrane GP41 and surface GP120 domain, playing critical roles in virion entry into the cells [145].

The wild type viral genome is organised in a 5’ to 3’ orientation containing GAG, POL, ENV and REV gene sequences flanked by long terminal repeats (LTRs) containing a strong viral promoter. Generally, GAG encodes the structural proteins, POL encodes the reverse transcriptase, integrase, and protease, ENV encodes the envelope glycoprotein necessary for target cell entry, and REV facilitates nuclear export of the viral genome. Complex retroviruses such as Lentiviruses have additional cis-acting sequences, including the rev response element (RRE) allowing nuclear export of the viral genome, the RNA packaging signal (ψ) for incorporation of the viral genome into virions, and the central polyuridine tract (cPPT) employed during reverse
transcription. Additionally, the LTRs contain a trans-activating response (TAR) element embedded in their repeat (R) regions that functions as a binding site for the trans-activator (Tat) protein, facilitating transcription of the viral genome [144].

Lentiviral gene therapy vectors were engineered by splitting viral genome components GAG, POL, ENV, and REV from the lentiviral genome as a way of reducing the likelihood of recombinant lentiviral particles forming [144, 146]. Second generation lentiviral vectors expressed GAG, POL, and REV from the same plasmid DNA, with late third generation vectors further splitting REV expression into a separate plasmid. Removal of these genes opened up space within the viral genome for the incorporation and expression of transgenes from the viral promoter in the 5’LTR. Further advancement led to safer configurations devoid of promoter elements in the U3 region of the 3’LTR which are in turn inherited by the 5’LTR upon reverse transcription. This deletion effectively inactivates the LTR promoter perturbing their ability to cause insertional activation of oncogenes, as well as preventing continued expression viral genome from the proviral DNA. These self-inactivating (SIN) lentiviral vectors instead support transgene expression through an internal expression cassette, containing an RNA polymerase II promoter [147]. Use of different internal promoters allows tuning of transgene expression level, as well as tissue specific expression [148]. Furthermore, increased performance from these lentiviral vectors was achieved through the addition of posttranscriptional elements, such as the woodchuck post-transcriptional regulatory element (WPRE) [149]. Furthermore, in a process known as pseudotyping the glycoproteins in the lentiviral vector enveloped can be replaced with those derived from other viruses, such as vesicular stomatitis virus glycoprotein (VSV-G), murine leukaemia virus (MLV), gibbon ape leukaemia virus (GALV), or baboon endogenous retrovirus (BaEV). This process has been used to expand the tropism of lentiviral vectors to permit gene transfer to almost any human cell type [150]. Alternatively, receptor-targeted lentiviral vectors, devoid of natural envelop glycoprotein receptor binding, instead possessing a ligand specific for a cell surface antigen, have been used to limit vector tropism to specific cell subsets [151]. This report utilises the VSV-G pseudotyped envelop, due to its wide tropism, and viral particle stability [152]. Viral particles containing replication
deficient lentiviral genomes are produced by co-transfection of DNA plasmids allowing transient expression of viral proteins GAG, POL, ENV, and REV in a HEK293T packaging cell lines. These viral particles retain the ability to enter the target cells, uncoat, undergo reserve transcription, and proviral integration. However, as they no longer contain the GAG, POL, ENV or REV gene, and often process no promoter activity in the 5’LTR they are unable to undergo further transcription/translation of viral genes, virion assembly, and release of new viral particles by virion budding

1.9 CRISPR/Cas9 T cell engineering for generation of enhanced immunotherapeutics

T cells are an essential component of the adaptive immune system, performing roles in tumour surveillance, as well as the recognition and clearance of pathogenic organisms. T cell immunotherapies aim to redirect these adaptive immune responses toward the recognition and lysis of desired antigens while minimising detrimental ‘off-target’ effects on healthy tissue. Such therapeutic strategies initially relied on stimulation of the host immune system in a non-specific manner with hopes of increasing reactivity to tumour antigens [153]. Initial studies indicated the need for immunotherapies with more specific responses, and as such, non-specific immune activation was widely abandoned in favour of adoptive T cell transfer (ACT) [153]. These ACT treatments were given in an autologous manner, relying on ex vivo isolation, activation, and expansion prior to returning these antigen specific cells to the patient [154, 155]. Pioneering work by Rosenberg et al (1988), demonstrated that tumour infiltrating lymphocytes (TILs) could be extracted from resected melanomas, and subsequently expanded with IL-2. Autologous delivery of such products resulted in objective regression in 9 of 15 patients [156]. However, TIL based therapies are reliant on the pre-existence of TILs in the resected tumour, and are further limited by the laborious procedure of isolating and expanding these cells within a suitable time course for the patients.

Building upon these observations, allogeneic donor leukocyte infusion (DLI) protocols were developed that aimed to balance beneficial graft versus leukaemia (GVL) effects with potentially fatal graft versus host disease (GVHD).
In this setting, patients who present with either recurrent haematological malignancies or low bone marrow donor chimerism after an allogeneic haematopoietic stem cell transplant would receive T cells from the allogeneic donor. Despite modest efficacy in combating chronic myeloid leukaemia (CML), there has been little benefit seen in patients with acute leukaemia, due to low GVL response, and development of GVHD. In a seminal study conducted by Kolb et al (1990), sustained remission was achieved following DLI in 3/3 patients who had relapsed with CML patient after receiving an allogeneic haematopoietic stem cell transplant highlighting the potential of this strategy [157]. However, it also made clear the need for more refined T cell products as 2/3 patients presented with clinically significant GVHD [157, 158]. The advent of T cell engineering has seen marked improvement of both autologous and allogeneic cell therapies, by allowing ex vivo gene modification of peripheral blood T cells to express receptors of desired specificity. Outlined below are the two most widely applied methods of redirecting T lymphocytes.

1.9.1 Recombinant T cell receptors

The T cell receptor (TCR) on the surface of T lymphocytes represent the means by which these cells physiologically recognise and engage specific short linear peptide epitopes, including tumour associated antigens in the context of major histocompatibility complex (MHC) molecules. The majority of these (~95%) are composed of α and β chains (TCRαβ), with the remainder consisting of α γ and δ chain heterodimer (TCRγδ). Recombinant TCRs (rTCRs) can be used to redirect T cells towards specific antigens, requiring the expression of full-length TCRα and β chains. Complexes formed between rTCR chains and the endogenous multiprotein cluster of differentiation (CD) 3 complex (CD3ε, CD3δ, CD3γ, and CD3ζ chains) facilitate signal transduction and cytotoxic response upon MHC-TCR-CD3 engagement (Figure 1-5 A).

Encouraging results from their application in a variety of tumour types including Hepatocellular Carcinoma [159], B-cell malignancies [160], WT1 expressing tumours [161, 162], sarcoma and melanoma [163] have been reported.
Tumour target recognition by recombinant TCR and CAR expressing T cells:

**A.** The recombinant T cell receptors (rTCRs) are delivered as an α and β chain which form complexes with the endogenous multiprotein CD3 complex (γ, δ, ε, and ζ chains). Tumour associated peptides presented by major histocompatibility complex class I (MHC class I) molecules on the target cell are recognised by the rTCR. Engagement of the rTCR results in effector T cell activation and degranulation leading to target cell destruction. **B.** Chimeric antigen receptors (CARs) undergo tumour cell recognition through a single chain variable fragment (scFv) of a monoclonal antibody. In a second generation CAR the extracellular scFv domain is linked to a hinge region, transmembrane stalk, co-stimulatory domain (4-1BB or CD28) and a CD3ζ signalling domain. Tumour antigen recognition by the scFv, with subsequent signalling through co-stimulatory and CD3ζ domains activates the CAR T cell to lyse the target cell. Unlike rTCRs, recognition through the CAR scFv is MHC class I independent.

Although the above rTCR based ACT immunotherapies have proven effective, their application is limited by unpredictable ‘off-target effects’ due to TCR cross-reactivity (for example cardiac toxicity following therapy with MAGE-A3 rTCR) [164-167]. Development of predictive models to help characterise the potential cross-reactivity events of a rTCR prior to clinical application have reduced concerns in this area, although these still remain imperfect [168]. Furthermore, application of rTCRs have been limited by concerns surrounding the potential for mispairing between endogenous and recombinant TCR chains, resulting in TCR-CD3 complexes with unknown specificity, displaying potential ‘off-target-off-tumour’ responses in healthy tissue [161, 166, 169]. However, TCR mispairing has been addressed by promoting...
exclusive rTCR pairing via additional disulfide bonds, use of high-affinity TCRs, murinisation of rTCR constant regions, hybrid TCR-Chimeric antigen receptor molecules, as well as development of single chain TCRs through covalently linking the variable domains [170-177]. Legut et al (2018), took advantage of the restricted nature of pairing between αβ and γδ chains, by expressing a γδ transgenic TCR along with CRISPR/Cas9-based knockout of the endogenous β chain [178]. Furthermore, expression of rTCRs is limited by competition for shared components of the CD3 complex between recombinant and endogenous TCR chains [161, 179, 180]. To increase the expression and function of less competitive rTCRs, genome editing tools have been used to disrupt the endogenous TCR chains using ZFNs [179, 180], TALENs [169, 180], megaTALs [181], and CRISPR/Cas9 [46, 166, 180, 181]. Roth and associates (2018), showed proof of principle for a viable CRISPR/Cas9-based targeted insertion of an NY-ESO-1-directed rTCR into the endogenous T cell receptor alpha constant (TRAC) locus. Using a non-viral PCR-amplified dsDNA template in conjugation with RNPs, they achieved rTCR expression of ~12%, as well as alleviating competition for CD3, while regulating expression of the rTCR from the endogenous TRAC promoter [46].

Successful reports of CRISPR/Cas9 genome editing in conjunction with rTCR expression has fuelled a phase 1 clinical trial in multiple myeloma patients (NCT03399448), where patients have received autologous T cells transduced with an NY-ESO-1 specific TCR, edited for removal of programmed cell death protein 1 (PD1) as well as, TCRα and β chains [133]. Although tumour clearance was not achieved, Stadtmauer et al (2020), observed trafficking of edited T cells to the tumour sites, and a reduction in the presence of targeted NY-ESO-1 antigen [133]. These preliminary clinical results provide pivotal data indicating the safety and feasibility of CRISPR edited T cells, as well as their ability to engraft for a sustained period of time (~9 months).

1.9.2 CAR-T cell therapies

Unlike rTCRs, chimeric antigen receptors (CARs) recognise their desired target through a highly specific, single-chain variable fragment (scFv) from a monoclonal
antibody. This extracellular scFv domain is generally fused to, a spacer region allowing for optimal antigen engagement, followed by a transmembrane stalk for signal transduction, and intracellular activation and costimulatory domains of a TCR. This structure endows the T cells with MHC-independent target recognition and active T cell signalling responses, similar to those elected by the endogenous TCR. Modification of these structural features allows redirection of specificity and antigen binding, modulation of signal transduction, and effector T cell metabolism [182-184] (Figure 1-5 B). Since their conception, a number of CAR generations claiming improved efficacy have been developed. The first generation of CARs, initially termed T bodies, or chimeric TCRs relied solely on a CD3ζ chain, or Fc(epsilon)RI gamma-chain, for the initiation of T cell activation, fused to an extracellular scFv ligand binding domain [184-188]. The structure of these first generation CARs was reminiscent of TCRs. However, they proved inefficient at promoting IL-2 production, resulting in low T cell activation and low persistence and efficacy against target cells in vitro and in vivo [185, 186]. These observations led to the development of a more synthetic configuration, containing a dual signalling domain, incorporating a co-stimulatory domain more typical of T cell activation. Inclusion of co-stimulatory domains in the intracellular cytoplasmic tail of the CAR, increased IL-2 secretion, CAR T cell persistence, and expansion upon repeated antigen recognition [185, 189, 190]. This structure is often referred to as a second generation CAR, with a number of co-stimulatory domains being utilised to provide different degrees of activation, proliferation, and persistence. Of these, second generation CARs containing either Ig superfamily member CD28ζ, or tumour necrosis factor receptor (TNFR) 4-1BB (CD137) co-stimulatory domains have undergone rigorous clinical testing, and have both proven effective against refractory relapsed B cell malignancies [191]. Notably these distinct co-stimulatory domains endow CAR T cells with distinct functional and metabolic profiles.

CAR molecules containing the CD28ζ chain present with rapid expansion and greater effector T cell activity, although they appear more transient compared to 4-1BB based CARs, which instead allow persistent anti-tumour effects, and prolonged T cell division [192, 193]. Interestingly, Kawaleker and colleagues (2016) helped elucidate
the differences in CAR T cells dynamics by demonstrating that CD28-based CARs favour an effector memory T cell phenotype, with a genetic signature that suggests a dependence on glucose based metabolism, whereas 4-1BB-CARs elicit a more naive CD8+ central memory phenotype with lipid oxidation based metabolism [194]. Co-stimulation is a dynamic process involving numerous molecules which are now being characterised for incorporation into CAR molecules, including ICOS, OX40, and CD27, affording CAR T cells with different phenotypes and cytokine profiles, as a means to tune activation and cell fate [195-197].

Updated designs have included additional co-stimulatory sequences to generate stronger cytokine production and increased target killing potential [184, 185]. Preclinical models suggested that third generation CAR molecules, incorporating both CD28 and 4-1BB co-stimulatory signals retained the rapid tumour clearance of CD28-based CARs, with the persistence of 4-1BB-based CARs [193]. Ramos and associates (2018), reported clinical phase testing of CD28-based second generation CAR T cells compared to a CD28/4-1BB third generation CAR against B cell malignancies [198]. Increased expansion and persistence was observed in CAR T cells expressing the third generation configuration, which appeared most apparent in a patients with a low tumour burden. Despite preclinical and early phase clinical data there is not conclusive evidence indicating an improved outcome when using third generation CARs over their second generation counterparts, necessitating their further testing. Furthermore, the optimal dose to achieve tumour clearance, while avoiding cytokine release syndrome and neurotoxicity, has yet to be established for third generation CAR T cells. Moreover, the addition of extra co-stimulatory domains may exacerbate antigen-dependent tonic signalling leading to increased exhaustion [199]. However, if these co-stimulatory domains prove to be problematic when combined, use of multiple second generation CAR populations with distinct co-stimulatory domains, expression of full length co-stimulatory molecules in trans, or inducible co-stimulatory molecules instead could help mitigate this [191, 200].

Additionally, fourth generation CAR molecules have been engineered, with the ability to recruit other immune cells to the tumour site. This fourth generation configuration is often referred to as T cells redirected for universal cytokine-mediated killing
TRUCKs, and recruits other immune cells through expression of transgenic cytokines, such as IL-2, upon target recognition also making them resilient to immunosuppressive tumour environments [201]. Although TRUCKs could have wider applications than previous CAR generations, including viral infection, metabolic disorders and tumours with a heterogeneous microenvironment; concerns including delivery method and leakage of the system leading to off-target, off-tumour immune responses have to be addressed [185, 201].

As well as structural design, the method of expressing these CARs in T cells has also been investigated. Use of semi-randomly integrating viral vectors such as γ-retroviruses and lentiviruses has been the most commonly used delivery strategy. However, γ-retroviruses have shown the potential of insertional mutagenesis leading to clonal expansion and oncogenic transformation. Additionally, both γ-retroviruses and lentiviruses display inconsistent intensity of transgene expression, promoter silencing, and tonic CAR signalling [199, 202, 203]. Genome editing technologies have allowed sequence specific integration within the host genome mitigating concerns surrounding vector insertion site. Eyquem et al (2017), demonstrated the ability to place a second generation CAR under the transcriptional control of the endogenous TRAC promoter, providing a number of key advantages, including uniform CAR expression leading to increased CAR T cell potency, avoidance of tonic CAR signalling, and reduced exhaustion upon repeat antigen exposure. Importantly, CRISPR-Cas9 disruption of the TRAC locus removes alloreactive cell surface TCR, allowing for allogeneic T cell application [44]. These promising results have indicated the need for both suitable CAR design and method of expression.

These CAR T cell-based immunotherapeutics have been predominantly applied clinically in the field of haematopoietic malignancies using autologous anti-CD19 CAR T cells for the treatment of adult and paediatric B cell acute lymphoblastic leukaemia (B-ALL) [204-207]. The principal target for these CARs has been CD19 due to it ubiquitous expression on B cells, as well as its lack of homology with other proteins. Despite highly encouraging results using autologous products, collection of material from geriatric and paediatric patients presents logistical roadblocks, batch variability, and demanding manufacture infrastructure greatly limiting access to CAR T cell
therapies (*Figure 1-6 A*). Genome editing has facilitated the removal of alloreactive surface antigens, allowing the generation of ‘off-the-shelf’ ‘universal’ CAR T cells that overcome HLA barriers (*Figure 1-6 B*).

**Figure 1-6 Personalised autologous versus universal allogeneic CAR T cell therapies:**

**A.** Autologous CAR T cell therapies are personalised cell therapeutics generated from the patient’s own cells. Procedure requires harvesting of the patient’s peripheral blood T cells, prior to ex vivo modification with an integrating viral vector encoding the desired CAR. These CAR T cells are expanded, before being re-infused into the patient and monitored for tumour clearance. **B.** Off-the-shelf allogeneic CAR T cell therapies use healthy donor material from an unmatched recipient. Manufacturing involves isolation of healthy peripheral blood T lymphocytes, before ex vivo genome editing to remove alloreactive molecules or chemotherapeutic antibody targets. Moreover, gene addition is used to express the desired CAR, before expansion of
universal CAR T cells and cryopreservation. The therapeutic can be administered to multiple patients in an ‘off-the-shelf’ manner without the requirement for donor matching.

Early applications led by Qasim et al (2017), illustrated the first clinical outcome of the allogeneic ‘off-the-shelf’ ‘universal’ CAR T cell (UCART19) therapies. In this setting TALEN-mediated genome editing was used to disrupt both the TRAC and CD52 loci [208, 209]. Disruption of TRAC, removes the alloreactive TCRαβ expression from the T cells’ surface, while CD52 editing permits resistance to the CD52-targeting chemotherapeutic antibody Alemtuzumab. The UCART19 cells exhibited highly potent anti-leukaemic effects in two paediatric lymphoblastic leukaemic relapse patients, entering molecular remission by day 28 prior allogeneic stem cell transplantation [210]. Further clinical trials of allogeneic ‘universal’ immunotherapeutics are currently under investigation in both Europe and the US (NCT04093596, NCT04230265).

1.10 Current project

As detailed above, understanding of both CRISPR/Cas9 genome editing, and T cell engineering for development of immunotherapies has increased significantly over the last decade. Broadly, the current project develops methods for efficient T cells engineering, combining both lentiviral vector mediated transgene expression, with CRISPR/Cas9-based genome editing, relevant to the allogeneic immunotherapeutic field. Allogeneic ‘off-the-shelf’ T lymphocytes expressing either rTCRs, or CARs, using both TALEN and CRISPR/Cas9 genome editing for disruption of alloreactive markers has recently been reported with successful entry into the clinic [133, 210]. These protocols rely on removal of residual alloreactive TCRαβ expression, which is normally facilitated by magnetic bead mediated depletion. Despite this being highly effective, there is often ~1% carriage of TCRαβ+ cells, which constitutes a risk of GVHD. Taking this risk into account, clinical doses are often capped at 5x10⁴ TCRαβ+ T cells/kg, restricting the total cell number delivered [211]. The proportion of the cell inoculum expressing either the rTCR, or CAR, often depends upon variable transduction/ HDR efficiencies, depending upon the route of delivery. Therefore, the
restrictions on delivery of these TCRαβ+ cells can limit the therapeutic rTCR/ CAR T cell dose.

This problem exists where delivery of genome editing components and CAR expression act independently. To this end, I aim to:

1. Develop a lentiviral vector platform that would facilitate coupling of a CAR19 transgene, with CRISPR/Cas9 genome editing of the TRAC locus, to produce CAR19 T cells devoid of alloreactive TCRαβ expression. By virtue of the RNA guided nature of CRISPR/Cas9, optimisation of a third generation lentiviral vector configuration, incorporating a sgRNA expression cassette within the ΔU3 region of the 3’ long terminal repeat (LTR) was developed, referred to as ‘terminal-CRISPR’. Expression of the sgRNA from the lentiviral vector, would allow CRISPR/Cas9 knockout restricted to the transduced cell population upon delivery of the SpCas9 endonuclease. Furthermore, placement of the sgRNA expression cassette within the 3’LTR allows duplication of the sgRNA expression cassette during reverse transcription [212], while also minimising interference with the internal promoter used to express the transgene [213, 214].

2. Multiplex the expression of sgRNA from this terminal-CRISPR configuration. This will involve testing sgRNA sequences for the knockout of both CD52 and Beta-2-Microglobulin (B2M) loci, which will be expressed in tandem with a TRAC specific sgRNA.

3. Refine of the sgRNA expression cassettes and exploration of RNA polymerase III promoter sequence and architecture, by screening a number of mutant sequences, devoid of promoter elements

4. Validate the recently described CBE technology in the context of the coupled terminal-CRISPR vector configuration. Moreover, I will endeavour to optimise multiplexed editing using CBE technology with an assessment of predicted translocation events by targeted PCR.

5. Finally, efforts will be made to apply these developed methods to the expression of a rTCR with coupled targeted knockout of both the endogenous T cell receptor β constant (TRBC), and TRAC loci.
Chapter 2  Materials and Methods

2.1  Materials

2.1.1  Reagents used for DNA

Table 3 List of the reagents used for DNA

<table>
<thead>
<tr>
<th>Reagent:</th>
<th>Manufacturer:</th>
<th>Catalogue number:</th>
</tr>
</thead>
<tbody>
<tr>
<td>50X Tris-acetate-EDTA (TAE) buffer</td>
<td>ThermoFisher Scientific, Massachusetts, USA</td>
<td>B49</td>
</tr>
<tr>
<td>Nuclease-Free Water (not DEPC-Treated)</td>
<td>ThermoFisher Scientific, Massachusetts, USA</td>
<td>AM9939</td>
</tr>
<tr>
<td>1Kb plus DNA ladder</td>
<td>ThermoFisher Scientific, Massachusetts, USA</td>
<td>10787018</td>
</tr>
<tr>
<td>dNTP Set 100mM Solutions</td>
<td>ThermoFisher Scientific, Massachusetts, USA</td>
<td>R0181</td>
</tr>
<tr>
<td>Gel loading dye: Orange G 6X</td>
<td>ThermoFisher Scientific, Massachusetts, USA</td>
<td>R0631</td>
</tr>
<tr>
<td>UltraPure™ Agarose</td>
<td>ThermoFisher Scientific, Massachusetts, USA</td>
<td>16500500</td>
</tr>
<tr>
<td>SYBR™ Safe DNA Gel Stain</td>
<td>ThermoFisher Scientific, Massachusetts, USA</td>
<td>S33102</td>
</tr>
<tr>
<td>Branched Polyethyleneimine (PEI)</td>
<td>Sigma-Aldrich, Dorset, UK</td>
<td>408727</td>
</tr>
<tr>
<td>T4 DNA ligase (5U/µl)</td>
<td>ThermoFisher Scientific, Massachusetts, USA</td>
<td>EL0014</td>
</tr>
<tr>
<td>Reagent:</td>
<td>Manufacturer:</td>
<td>Catalogue number:</td>
</tr>
<tr>
<td>------------------------------</td>
<td>-------------------------------------------------------</td>
<td>-------------------</td>
</tr>
<tr>
<td>T4 Polynucleotide Kinase</td>
<td>NEW ENGLAND BioLabs, Massachusetts, USA</td>
<td>M0201S</td>
</tr>
<tr>
<td>FastAP Thermosensitive</td>
<td>ThermoFisher Scientific, Massachusetts, USA</td>
<td>EF0651</td>
</tr>
<tr>
<td>Alkaline Phosphatase (1U/µl)</td>
<td>ThermoFisher Scientific, Massachusetts, USA</td>
<td></td>
</tr>
<tr>
<td>FastDigest XbaI</td>
<td>ThermoFisher Scientific, Massachusetts, USA</td>
<td>FD0684</td>
</tr>
<tr>
<td>FastDigest BpiI (IIs class)</td>
<td>ThermoFisher Scientific, Massachusetts, USA</td>
<td>FD1014</td>
</tr>
<tr>
<td>FastDigest Esp3I (IIs class)</td>
<td>ThermoFisher Scientific, Massachusetts, USA</td>
<td>FD0454</td>
</tr>
<tr>
<td>FastDigest NheI</td>
<td>ThermoFisher Scientific, Massachusetts, USA</td>
<td>FD0974</td>
</tr>
</tbody>
</table>

### 2.1.2 Reagents used for protein

Table 4 List of reagents used for protein

<table>
<thead>
<tr>
<th>Reagent:</th>
<th>Manufacturer:</th>
<th>Catalogue number:</th>
</tr>
</thead>
<tbody>
<tr>
<td>Amersham™ ECL™ Rainbow™</td>
<td>GE LifeSciences, Buckinghamshire, UK</td>
<td>GERPN800N</td>
</tr>
<tr>
<td>Markers- Full Range</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Amersham Hybond P 0.22μm</td>
<td>GE LifeSciences, Buckinghamshire, UK</td>
<td>10600021</td>
</tr>
<tr>
<td>PVDF blotting membrane</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Product</td>
<td>Supplier</td>
<td>Code</td>
</tr>
<tr>
<td>----------------------------------------------</td>
<td>---------------------------------</td>
<td>--------</td>
</tr>
<tr>
<td>Tran-Blot® Turbo™ Mini PVDF Transfer Packs (0.22µm)</td>
<td>Bio-Rad, California, USA</td>
<td>1704156</td>
</tr>
<tr>
<td>NuPAGE™ 4-12% Bis-Tris Protein Gels, 1.0 mm, 10-well</td>
<td>ThermoFisher Scientific, Massachusetts, USA</td>
<td>NP0321BOX</td>
</tr>
<tr>
<td>4-15% Mini-PROTEAN® TGX™ Precast Protein Gels, 10-well, 30µl</td>
<td>Bio-Rad, California, USA</td>
<td>456-1083</td>
</tr>
<tr>
<td>Complete protease Inhibitor cocktail tables</td>
<td>Roche, Basel, Switzerland</td>
<td>11697498001</td>
</tr>
<tr>
<td>Dithiothreitol (DTT)</td>
<td>ThermoFisher Scientific, Massachusetts, USA</td>
<td>R0861</td>
</tr>
<tr>
<td>Phenylmethylsulfonyl fluoride (PMSF)</td>
<td>Sigma-Aldrich, Dorset, UK</td>
<td>10837091001</td>
</tr>
<tr>
<td>Triton™ X-100</td>
<td>Sigma-Aldrich, Dorset, UK</td>
<td>X100-5ML</td>
</tr>
<tr>
<td>Sodium Dodecyl Sulphate (SDS)</td>
<td>Sigma-Aldrich, Dorset, UK</td>
<td>71725-50G</td>
</tr>
<tr>
<td>TrueCut™ Cas9 Protein v2</td>
<td>ThermoFisher Scientific, Massachusetts, USA</td>
<td>A36496</td>
</tr>
<tr>
<td>10x Tris/Tricine/SDS Running Buffer, 1L</td>
<td>Bio-Rad, California, USA</td>
<td>1610744</td>
</tr>
<tr>
<td>Glycerol</td>
<td>Sigma-Aldrich, Dorset, UK</td>
<td>G5516-100ML</td>
</tr>
<tr>
<td>β-mercaptothanol</td>
<td>Sigma-Aldrich, Dorset, UK</td>
<td>M6250</td>
</tr>
<tr>
<td>Reagent:</td>
<td>Manufacturer:</td>
<td>Catalogue number:</td>
</tr>
<tr>
<td>--------------------------------</td>
<td>----------------------------------------</td>
<td>-------------------</td>
</tr>
<tr>
<td>Dried skimmed milk power</td>
<td>Marvel</td>
<td>n/a</td>
</tr>
<tr>
<td>Trizma® base</td>
<td>Sigma-Aldrich, Dorset, UK</td>
<td>T1503-25G</td>
</tr>
<tr>
<td>Bromophenol blue</td>
<td>Sigma-Aldrich, Dorset, UK</td>
<td>114391-5G</td>
</tr>
<tr>
<td>TWEEN® 20</td>
<td>Sigma-Aldrich, Dorset, UK</td>
<td>P9416-50ML</td>
</tr>
<tr>
<td>UltraPure™ 0.5M EDTA pH 8.0</td>
<td>ThermoFisher Scientific, Massachusetts, USA</td>
<td>15575-038</td>
</tr>
<tr>
<td>NaCL</td>
<td>Sigma-Aldrich, Dorset, UK</td>
<td>S9888</td>
</tr>
</tbody>
</table>

### 2.1.3 Reagents for bacterial culture

Table 5 List of reagents used for bacterial culture

<table>
<thead>
<tr>
<th>Reagent:</th>
<th>Manufacturer:</th>
<th>Catalogue number:</th>
</tr>
</thead>
<tbody>
<tr>
<td>Ampicillin Sodium salt</td>
<td>Sigma-Aldrich, Dorset, UK</td>
<td>A0166-25G</td>
</tr>
<tr>
<td>Kanamycin sulfate from Streptomyces kanamyceticus</td>
<td>Sigma-Aldrich, Dorset, UK</td>
<td>K1377-25G</td>
</tr>
<tr>
<td>LB broth</td>
<td>Sigma-Aldrich, Dorset, UK</td>
<td>L3022-250G</td>
</tr>
<tr>
<td>LB agar</td>
<td>Sigma-Aldrich, Dorset, UK</td>
<td>L3147-1KG</td>
</tr>
<tr>
<td>S.O.C. Medium</td>
<td>ThermoFisher Scientific, Massachusetts, USA</td>
<td>15544034</td>
</tr>
<tr>
<td>One Shot™ Stbl3™ Chemically Competent E. coli</td>
<td>ThermoFisher Scientific, Massachusetts, USA</td>
<td>C737303</td>
</tr>
</tbody>
</table>
### Stellar™ Competent Cells

<table>
<thead>
<tr>
<th>Manufacturer:</th>
<th>Takara Bio Europe, Saint-Germain-en-Laye, France</th>
</tr>
</thead>
<tbody>
<tr>
<td>Catalogue number:</td>
<td>636766</td>
</tr>
<tr>
<td>Product Name</td>
<td>Supplier Information</td>
</tr>
<tr>
<td>--------------------------------------------------</td>
<td>--------------------------------------------------------</td>
</tr>
<tr>
<td>Penicillin-Streptomycin (Pen Strep), 10,000 U/ml, 100ml</td>
<td>ThermoFisher Scientific, Massachusetts, USA</td>
</tr>
<tr>
<td>Dulbecco’s Phosphate-Buffered Saline (DPBS),</td>
<td>ThermoFisher Scientific, Massachusetts, USA</td>
</tr>
<tr>
<td>TexMACS™ Medium</td>
<td>Miltenyi Biotech, Surrey, UK</td>
</tr>
<tr>
<td>Gemcell Human Serum Ab U.s. Origin Male Donors Only</td>
<td>Seralabs, Brussels, Belgium</td>
</tr>
<tr>
<td>Human IL-2 IS, premium grade</td>
<td>Miltenyi Biotech, Surrey, UK</td>
</tr>
<tr>
<td>T Cell TransACT™, human</td>
<td>Miltenyi Biotech, Surrey, UK</td>
</tr>
<tr>
<td>OneComp eBeads™</td>
<td>ThermoFisher Scientific, Massachusetts, USA</td>
</tr>
<tr>
<td>Anti-Biotin MicroBeads UltraPure</td>
<td>Miltenyi Biotech, Surrey, UK</td>
</tr>
<tr>
<td>LD Columns</td>
<td>Miltenyi Biotech, Surrey, UK</td>
</tr>
</tbody>
</table>
# 2.1.5 Flow cytometry/ western blot antibodies:

Table 7 List of primary and secondary antibodies

<table>
<thead>
<tr>
<th>Target:</th>
<th>Fluorochrome:</th>
<th>Manufacturer:</th>
<th>Catalogue number:</th>
</tr>
</thead>
<tbody>
<tr>
<td>CD3</td>
<td>VioBlue</td>
<td>Miltenyi Biotech, Surrey, UK</td>
<td>130-114-519</td>
</tr>
<tr>
<td>CD3</td>
<td>APC</td>
<td>Miltenyi Biotech, Surrey, UK</td>
<td>130-113-135</td>
</tr>
<tr>
<td>CD45</td>
<td>VioGreen</td>
<td>Miltenyi Biotech, Surrey, UK</td>
<td>130-110-638</td>
</tr>
<tr>
<td>TCRαβ</td>
<td>APC</td>
<td>Miltenyi Biotech, Surrey, UK</td>
<td>130-091-237</td>
</tr>
<tr>
<td>TCRαβ</td>
<td>PerCP-vio700</td>
<td>Miltenyi Biotech, Surrey, UK</td>
<td>130-109-924</td>
</tr>
<tr>
<td>TCRαβ</td>
<td>biotin</td>
<td>Miltenyi Biotech, Surrey, UK</td>
<td>130-098-219</td>
</tr>
<tr>
<td>CD52</td>
<td>APC</td>
<td>Miltenyi Biotech, Surrey, UK</td>
<td>130-099-632</td>
</tr>
<tr>
<td>CD52</td>
<td>VioBlue</td>
<td>Miltenyi Biotech, Surrey, UK</td>
<td>130-099-628</td>
</tr>
<tr>
<td>CD279(PD1)</td>
<td>PE</td>
<td>Miltenyi Biotech, Surrey, UK</td>
<td>130-096-164</td>
</tr>
<tr>
<td>HLA-ABC</td>
<td>APC</td>
<td>BioLegend, London, UK</td>
<td>311410</td>
</tr>
<tr>
<td>Reagent:</td>
<td>Manufacturer:</td>
<td>Catalogue number:</td>
<td></td>
</tr>
<tr>
<td>--------</td>
<td>---------------</td>
<td>------------------</td>
<td></td>
</tr>
<tr>
<td>HLA-ABC</td>
<td>Miltenyi Biotech, Surrey, UK</td>
<td>130-101-463</td>
<td></td>
</tr>
<tr>
<td>Biotin-SP (long spacer) AffiniPure F(ab') Fragment Goat Anti-Mouse IgG, F(ab') Fragment Specific antibody</td>
<td>Stratech Scientific Limited, Suffolk, UK</td>
<td>115-066-072-JIR</td>
<td></td>
</tr>
<tr>
<td>Streptavidin</td>
<td>BioLegend, London, UK</td>
<td>405207</td>
<td></td>
</tr>
<tr>
<td>Streptavidin</td>
<td>Miltenyi Biotech, Surrey, UK</td>
<td>130-106-789</td>
<td></td>
</tr>
<tr>
<td>Anti-CRISPR-Cas9</td>
<td>Abcam, Cambridge, UK</td>
<td>ab191468</td>
<td></td>
</tr>
<tr>
<td>Anti-β-Actin</td>
<td>Abcam, Cambridge, UK</td>
<td>Ab115777</td>
<td></td>
</tr>
<tr>
<td>Amersham ECL HRP-linked whole Ab (from sheep)</td>
<td>GE LifeSciences, Buckinghamshire, UK</td>
<td>NXA931</td>
<td></td>
</tr>
</tbody>
</table>

### 2.1.6 Kits

**Table 8 List of manufacture designed kits**

<table>
<thead>
<tr>
<th>Reagent:</th>
<th>Manufacturer:</th>
<th>Catalogue number:</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

68
<table>
<thead>
<tr>
<th>Product Name</th>
<th>Company</th>
<th>Location</th>
<th>Catalog Number</th>
</tr>
</thead>
<tbody>
<tr>
<td>Bio-Rad Protein assay kit II</td>
<td>Bio-Rad, California, USA</td>
<td>5000002</td>
<td></td>
</tr>
<tr>
<td>Clarity™ Western ECL Substrates, 200ml</td>
<td>Bio-Rad, California, USA</td>
<td>1705060</td>
<td></td>
</tr>
<tr>
<td>Q5® High-Fidelity DNA Polymerase</td>
<td>NEW ENGLAND BioLabs, Massachusetts, USA</td>
<td>M0491L</td>
<td></td>
</tr>
<tr>
<td>Monarch® Plasmid Miniprep Kit</td>
<td>NEW ENGLAND BioLabs, Massachusetts, USA</td>
<td>T1010L</td>
<td></td>
</tr>
<tr>
<td>Plasmid Maxi Kit</td>
<td>QIAGEN, Hilden, Germany</td>
<td>12165</td>
<td></td>
</tr>
<tr>
<td>Monarch® PCR &amp; DNA Cleanup Kit (5µg)</td>
<td>NEW ENGLAND BioLabs, Massachusetts, USA</td>
<td>T1030L</td>
<td></td>
</tr>
<tr>
<td>Monarch® DNA Gel extraction Kit</td>
<td>NEW ENGLAND BioLabs, Massachusetts, USA</td>
<td>T1020L</td>
<td></td>
</tr>
<tr>
<td>DNeasy Blood &amp; Tissue Kit (250)</td>
<td>QIAGEN, Hilden, Germany</td>
<td>69506</td>
<td></td>
</tr>
<tr>
<td>In-Fusion® HD Cloning Plus</td>
<td>Takara Bio Europe, Saint-Germain-en-Laye, France</td>
<td>638910</td>
<td></td>
</tr>
<tr>
<td>QuikChange Lighting Multi Site-Directed Mutagenesis Kit</td>
<td>Agilent, California, USA</td>
<td>210513</td>
<td></td>
</tr>
<tr>
<td>Neon™ Transfection System 100 µl Kit</td>
<td>ThermoFisher Scientific, Massachusetts, USA</td>
<td>MPK10096</td>
<td></td>
</tr>
<tr>
<td>P3 Primary Cell 4D-Nucleofector™ X Kit L</td>
<td>Lonza, Basel, Switzerland</td>
<td>V4XP-3024</td>
<td></td>
</tr>
</tbody>
</table>
2.1.7 Buffers

dNTPs for PCR (10mM): dNTPs from ThermoFisher Scientific come as 250μl aliquots of dATP, dCTP, dGTP, and dTTP at 100mM. These were mixed together to make a 1ml solution at 25mM of each dNTP. Nuclease free water was used to further dilute this to 10mM of each dNTP (total volume 2.5ml). Aliquots were made at both 10 and 100μl volumes and stored at -20°C.

Lysis buffer base: 50mM Tris/HCL pH 8.0, 150mM NaCL, and 5mM EDTA.

Lysis buffer: 84.5% lysis buffer base, 10% Triton-X (10 fold diluted in PBS), 5% protease inhibitor cocktail (cOmplete), 0.5% PMSF.

Assay buffer: 98% lysis buffer base, and 2% lysis buffer.

10X Wash buffer: 0.5% Tween-20 in PBS. Stored at RT.

5X sample buffer: 10% SDS, 30% Glycerol, 10% β – mercaptoethanol, 0.001% bromophenol blue, and ~60% 0.5M Tris/HCL, pH6.8. Stored at -20°C.

Blocking solution: 5% dried skimmed milk powder in 1X Wash buffer.

Western blot primary antibody: 3% BSA in 1X wash buffer.

Western blot secondary antibody: 3% dried skimmed milk powder in 1X Wash buffer. Stored at 4°C.

Flow cytometry wash and staining buffer: 2% FCS in PBS. Stored at 4°C.

Column depletion MACS buffer: 0.5% bovine serum albumin (BSA), 2mM EDTA, in PBS. Stored at 4°C.

LB broth: 20g LB broth powder per L of H₂O. Autoclaved at 121°C for 15min. Once at RT, either kanamycin (50μg/ml) or ampicillin (100μg/ml) was added. Stored RT till antibiotic was added and then moved to 4°C.

LB agar: 36g LB agar powder per L of H₂O. Autoclaved at 121°C for 15min. Allow LB agar to cool to ~60°C before adding either kanamycin (50μg/ml) or ampicillin
(100µg/ml). Once selective antibiotic was added, 18-20ml LB was added per plate. Stored at 4°C.

**Bacterial glycerol stock:** 100% glycerol was mixed in a 1:1 ratio with H₂O to make a 50% glycerol solution. This 50% glycerol solution was then mixed in a 1:1 ratio with overnight bacterial culture (500µl: 500µl) in a screw cap cryopreservation tube, and stored at -80°C.

**Primary T cell freezing mix:** 10% DMSO, 45% TexMACS, 45% human serum Ab. Stored at 4°C.

**Cell line freezing mix:** 10% DMSO, and 90% FCS. Stored at 4°C.

### 2.1.8 Cell culture medium

**Complete DMEM:** DMEM with 10% FCS, and 1% Pen/Strep. Stored at 4°C.

**Complete RPMI:** RPMI with 10% FCS, and 1% Pen/Strep. Stored at 4°C.

**T cell Medium:** TexMACS with 3% human serum Ab, and 100U/ml (20ng/ml) IL-2. Stored at 4°C.

### 2.1.9 Cell types

**Table 9 List of Cell types**

<table>
<thead>
<tr>
<th>Cell ID</th>
<th>Tissue type</th>
<th>Immortalisation</th>
<th>Medium</th>
</tr>
</thead>
<tbody>
<tr>
<td>HEK-293T (293T)</td>
<td>Human Embryonic Kidney Cell Line (Adherent)</td>
<td>SV40 large T antigen</td>
<td>Complete DMEM</td>
</tr>
<tr>
<td>Jurkat, Clone E6-1 (JE6.1)</td>
<td>Peripheral Blood Cell Line, T Lymphoblast (Suspension)</td>
<td>Acute T cell Leukaemia</td>
<td>Complete RPMI</td>
</tr>
<tr>
<td></td>
<td>Peripheral Blood Cell Line, B lymphoblast (Suspension)</td>
<td>Burkitt’s Lymphoma</td>
<td>Complete RPMI</td>
</tr>
<tr>
<td>--------------------------</td>
<td>-------------------------------------------------------</td>
<td>--------------------</td>
<td>---------------</td>
</tr>
<tr>
<td>Daudi (lentiviral vector transduced to express GFP and Luciferase)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Primary human T cells from healthy donors</td>
<td>Peripheral Blood, T lymphocyte (Suspension)</td>
<td>n/a</td>
<td>T cell medium</td>
</tr>
</tbody>
</table>

2.1.10 **NSG mouse strain**

Non-obese diabetic (NOD)/ severe combined immunodeficiency (SCID)/γc−/− (NSG) mice (*NOD.Cg-Prkdc<sup>scid</sup> Il2rg<sup>tm1wjI</sup>/SzJ*) are an immunodeficient, inbred line. These were obtained from Charles River Laboratories.

2.1.11 **Trilink biotechnologies (San Diego, USA) mRNA**

**CleanCap® Cas9 mRNA (Cat. No. L-7606):**

This off-the-shelf mRNA encodes the SpCas9 endonuclease (transcript length: 4521bp). Trilink have incorporated two nuclear localisation signals, one at either terminus of the protein to increase trafficking to the nucleus. Additionally co-transcriptional capping supported a naturally occurring Cap 1 structure which in conjunction with polyadenylation optimises mRNA expression and stability.

**Custom made CleanCap® coBE3 mRNA (Cat. No. L-7007):**

This mRNA is a custom made product from Trilink, encoding a human codon optimised, third generation CBE (coBE3) (transcript length: 5664bp). The plasmid DNA used for mRNA synthesis containing the coBE3 sequence was sent to Trilink for mRNA production (180µg). This involved template linearization by BtgZI restriction digest, and mRNA purification by high performance liquid chromatography (HPLC) to help remove truncated mRNA by-products, increasing product purity. Unlike the SpCas9 mRNA, the coBE3 only contains a single NLS at the C terminus. Trilink’s
CleanCap technology was used to add a co-transcriptional Cap 1 structure, and this mRNA was further polyadenylated to increase expression and stability.

### 2.1.12 Synthego (California, USA) sgRNA

Synthetic sgRNA were manufactured by Synthego using automated solid-phase synthesis with 2’-O-methyl 3’ phosphorothioate modifications in the first and last 3nt’s. Single guide RNA containing a 20 nucleotide protospacer with an 80 nucleotide CRISPR scaffold were produced at either a 1.5 nmol (~50µg), or 3 nmol (~100µg) production scale. These were eluted in nuclease-free Tris-EDTA buffer provide by Synthego at 2µg/µl.

### 2.1.13 Oligo’s

All primers were ordered in the 5’-3’ orientation from ThermoFisher Scientific, Massachusetts, USA. Primers were resuspended at 10µM.

### 2.1.14 sgRNA sequences

**Table 10 Protospacer sequences**

<table>
<thead>
<tr>
<th>sgRNA name:</th>
<th>Target/ Exon</th>
<th>Sequence 5’ – 3’</th>
<th>Nuclease</th>
<th>Benchling on-target (off-target) score</th>
</tr>
</thead>
<tbody>
<tr>
<td>TRAC</td>
<td>TRAC/ Exon 1</td>
<td>TCTCTCAGCTGGTACACGGCC</td>
<td>SpCas9</td>
<td>51.8 (85.6)</td>
</tr>
<tr>
<td>TRAC ex1 SD</td>
<td>TRAC/ Exon 1</td>
<td>CTTACCTGGGCTGGGGAAGA</td>
<td>BE3</td>
<td>C1 0.5, C5 21.9, C6 21.4, (27.3)</td>
</tr>
<tr>
<td>TRAC ex3</td>
<td>TRAC/ Exon 3</td>
<td>TTTCAAAACCTGTCAGTGAT</td>
<td>BE3</td>
<td>C4 21.7 (56.9)</td>
</tr>
<tr>
<td>TRAC ex3 SA</td>
<td>TRAC/ Exon 3</td>
<td>TTCGTATCTGTAACCAAG</td>
<td>BE3</td>
<td>C3 16.0, C8 25.4, (52.0)</td>
</tr>
<tr>
<td><strong>TRBC ex 1-1</strong></td>
<td>TRBC 1 and 2/ Exon 1</td>
<td>CCACACCCAAAAGGCCACAC</td>
<td>BE3</td>
<td>C1 0.5, C2 11.0, C4 11.6, C6 13.5, C7 22.4, C8 8.7, (26.1)</td>
</tr>
<tr>
<td>---</td>
<td>---</td>
<td>---</td>
<td>---</td>
<td>---</td>
</tr>
<tr>
<td><strong>TRBC ex 1-2</strong></td>
<td>TRBC 1 and 2/ Exon 1</td>
<td>CCCACCCAGCTCAGCTCCACG</td>
<td>BE3</td>
<td>C1 0.8, C2 11.0, C3 5.7, C5 21.9, C6 21.4, (25.7)</td>
</tr>
<tr>
<td><strong>TRBC ex 1-3</strong></td>
<td>TRBC 1 and 2/ Exon 1</td>
<td>CGCTGCAGTCGGCTGTTCTA</td>
<td>BE3</td>
<td>C1 0.5, C3 0.5, C7 25.2, (31.0)</td>
</tr>
<tr>
<td><strong>TRBC ex 1-4</strong></td>
<td>TRBC 1 and 2/ Exon 1</td>
<td>GCTGCAAGTCCAGTTCTAC</td>
<td>BE3</td>
<td>C2 0.4, C6 26.6, (29.7)</td>
</tr>
<tr>
<td><strong>CD52 ex1-1</strong></td>
<td>CD52/ Exon 1</td>
<td>CACCATCGCCTCCTGGTTA</td>
<td>SpCas9</td>
<td>37.3 (64.7)</td>
</tr>
<tr>
<td><strong>CD52 ex1-2</strong></td>
<td>CD52/ Exon 1</td>
<td>TACCCATAACCAGGAGGCTGA</td>
<td>SpCas9</td>
<td>60.0 (57.6)</td>
</tr>
<tr>
<td><strong>CD52 ex2-1</strong></td>
<td>CD52/ Exon 2</td>
<td>AATGCCTCCGTTATTTGC</td>
<td>SpCas9</td>
<td>40.5 (83.5)</td>
</tr>
<tr>
<td><strong>CD52 ex2-2</strong></td>
<td>CD52/ Exon 2</td>
<td>TCAGCATCCAGCAACATAAG</td>
<td>SpCas9</td>
<td>58.7 (71.9)</td>
</tr>
<tr>
<td><strong>CD52 ex1 ISTOP</strong></td>
<td>CD52/ Exon 1</td>
<td>GTAAGATACAGACACGCC</td>
<td>BE3</td>
<td>C4 11.6, (86.3)</td>
</tr>
<tr>
<td><strong>CD52 ex1 SD</strong></td>
<td>CD52/ Exon 1</td>
<td>CTCTTACCTTGACTCATAACC</td>
<td>BE3</td>
<td>C1 0.2, C3 16.0, C7 19.0, C8 8.7, (82.4)</td>
</tr>
<tr>
<td><strong>CD52 ex2 SA</strong></td>
<td>CD52/ Exon 2</td>
<td>CAGTACAACTGACTCT</td>
<td>BE3</td>
<td>C1 0.8, C7 19.0, (69.8)</td>
</tr>
<tr>
<td><strong>CD52 ex2 S-P S</strong></td>
<td>CD52/ Exon 2</td>
<td>TGAGGGGCTGCTGGTTTTGCG</td>
<td>BE3</td>
<td>C8 2.8, (58.4)</td>
</tr>
<tr>
<td><strong>B2M ex1-1</strong></td>
<td>B2M/ Exon 1</td>
<td>GAGTAGCGCGAGCACAGCTA</td>
<td>SpCas9</td>
<td>56.3, (88.6)</td>
</tr>
</tbody>
</table>
Protospacer sequences used throughout this project with either SpCas9 or BE3. All protospacer sequences are presented in the 5’–3’ orientation, with cytidines within the BE3 editing window (positions 4-8 distil to the PAM) shown in Red. Benchling online tool (https://benchling.com) was used to score both on-target and off-target potential.

2.1.15 Software

FlowJo v10: Allowed import of FCS files used for all flow cytometry analysis shown in this report.

Graphpad Prism v8.0.0: Arrangement of data into appropriately formatted graphs, with subsequent statistical analysis.

SnapGene® v3.1.4: Instrumental in plasmid and gene map production, as well as restriction digest design. Additionally, this software was used to design sequencing primers for plasmid DNA and In-fusion cloning primers. Furthermore, this software was used to align Sanger sequencing to a reference sequence.

2.2 Methods

2.2.1 Bacterial transformation and plasmid preparation

All transformations were performed with 25µl of chemically competent E. coli cells. Either Stellar (Takara Bio Europe), or one-shot Stbl3 (ThermoFisher Scientific) cells were used for plasmid transformations. Competent cells were thawed on ice for 10 minutes before mixing with T4 DNA ligation reaction (2µl), or plasmids DNA (20-
50ng). This was allowed to rest on ice for 30 minutes, after which the cells were heat shocked at 42°C for 45 seconds. Transformations were then rested on ice for 2 minutes, before adding 250µl of S.O.C. medium (ThermoFisher Scientific) at RT. These were then recovered at 37°C for 1 hour, shaking at 250rpm. Once complete 150µl of the bacterial transformation was streaked on a pre-warmed selective LB agar plate containing an appropriate antibiotic and incubated at 37°C overnight. Individual colonies were picked and grown in 4ml of LB broth with antibiotic at 37°C, shaking at 250rpm overnight.

These bacterial starter cultures were used for plasmid DNA isolation (1.5ml) by a Miniprep kit (NEW ENGLAND BioLabs), and generation of glycerol stocks for long-term storage. Where appropriate, bacterial cultures were scaled up by mixing with fresh LB broth containing selective antibiotic at a 1:1000 ratio (500µl: 500ml), before incubation at 37°C, shaking at 250rpm overnight. Plasmid DNA was isolated from these large scale bacteria cultures using the Plasmid Maxi Kit (QIAGEN).

2.2.2 Restriction digestion

All Restriction enzymes used in this project are sourced from ThermoFisher Scientific in their FastDigest format.

Reaction setup:

- Plasmids DNA (0.5-1µg)
- Restriction enzyme 1µl
- 10X FastDigest Green Buffer 2.5µl
- ddH₂O up to 25µl

Digestions with BsmBI (Esp3I), required the addition of 1.25µl DTT (20mM). Reactions were incubated at 37°C for 1hr, before band discrimination by gel electrophoresis. Uncut, and single cut controls were setup in parallel to verify results.
2.2.3 Dephosphorylation of DNA ends

FastAP enzyme (ThermoFisher Scientific) was added to restriction digests (1U) to release 5’ and 3’ phosphate groups from the DNA ends, thereby preventing re-annealing of the plasmid backbone during ligation reactions, reducing the presence of background colonies.

2.2.4 DNA ligation reaction

T4 DNA ligase (ThermoFisher Scientific) was used to join dsDNA with complementary single stranded overhangs, or with blunt ends. Reactions were setup per manufacture instructions at either 10 or 20µl, depended upon DNA concentration. Linearized plasmid backbone was added at 50ng/reaction, with the desired insert added at a 1:3 molar ratio.

Where both the plasmid backbone and insert have complimentary overhangs, reactions were incubated at RT for 15 minutes. Blunt end ligations were allowed to incubate for 1hr at RT or 16°C overnight.

2.2.5 Polymerase chain reaction

PCR was used to amplify specific sequences of DNA for both cloning of new plasmids, and assessment of genomic DNA at edit loci. Q5® High-Fidelity DNA Polymerase (New England BioLabs) was used for all reactions, and setup in accordance with manufactures instructions. The dNTP were sourced from ThermoFisher Scientific and diluted to 10mM of each dNTP. A Mastercycler® nexus X2 (Eppendorf, Hamburg, Germany) thermocycler was used for these reactions. The annealing temperature was optimised for new primer pairs by setting gradient PCRs ranging between 58-68°C. Extension time was set at 30 seconds/ kilobase (kb) for the first 5kb, and 1min for every additional kb.

SnapGene® v3.1.4 was used to design primers for amplification of plasmids DNA.
To help avoid background amplification when using genomic DNA as a template, the NCBI Primer-Blast (https://www.ncbi.nlm.nih.gov/tools/primer-blast/) tool was used to design primer pairs.

### 2.2.6 In-Fusion HD cloning

A PCR based method for seamless insertion of a desired sequence within plasmid DNA without the need for restriction digestion. Both the plasmid backbone and the insert were linearized by PCR using ~1ng of template DNA. Furthermore, SnapGene® v3.1.4 was used to design primers to flank the insert with complimentary sequence to that of the backbone. As detailed above, PCR reactions were setup with Q5® High-Fidelity DNA Polymerase (New England BioLabs). Gel electrophoresis was used to identify these products, before subsequent DNA clean up using Monarch® DNA Gel extraction Kit (NEW ENGLAND Biolabs), followed by Monarch® PCR & DNA Cleanup Kit (NEW ENGLAND Biolabs).

**Reaction set up:**

- 2µl In-Fusion enzyme mix (5X)
- 50ng of plasmids backbone DNA
- 50ng of Insert DNA
- Up to 10µl ddH₂O

As per manufactures instructions, reactions were carried out in a thermocycler (Mastercycler® nexus X2) at 50°C for 15min before being rested on ice for 2min prior to transformation (2µl).

### 2.2.7 Site-Directed Mutagenesis

**QuikChange Lightning Multi Site-Directed Mutagenesis kit (Agilent):**

Used to make 1-2bp changes at multiple sites simultaneously (≤7 sites). Primers compatible with this kit were designed using the QuikChange primer design online tool (https://www.agilent.com/store/primerDesignProgram.jsp).
Reactions were setup as described in the manufactures protocol. All primers were used in the forward orientation, with 100ng of plasmid DNA. Reactions were carried out in the Mastercycler® nexus X2 (Eppendorf), using the appropriate cycling parameters.

Once the reactions had cooled to RT, 1µl of DpnI enzyme (from kit), was added to each reaction before digestion at 37°C for 15min. Transformation was carried out in chemically competent cells using 2µl of this reaction.

**Large insertion/ deletion:**

The NEB base changer online tool ([http://nebbasechanger.neb.com/](http://nebbasechanger.neb.com/)) was used to design sets of primers to either exclude sections of DNA or include/ replace appropriate sequences. Blunt ended linear DNA containing the desired changes were produced by PCR amplification using Q5® High-Fidelity DNA Polymerase (NEW ENGLAND Biolabs). Reactions setup as described above with ~1ng of template plasmid DNA. PCR reactions were run on a 1% agarose gel with SYBR™ safe DNA stain (ThermoFisher Scientific). DNA products were excised from the gel, before purification by Monarch® DNA Gel extraction Kit (NEW ENGLAND Biolabs), followed by Monarch® PCR & DNA Cleanup Kit (NEW ENGLAND Biolabs). Purified DNA, was placed in a 10µl T4 Polynucleotide Kinase reaction (NEW ENGLAND Biolabs) to achieve the 5’ phosphorylation needed for subsequent DNA ligation.

**T4 Polynucleotide Kinase reaction:**

- 50ng Purified DNA
- 0.5µl T4 Polynucleotide kinase enzyme (NEW ENGLAND Biolabs)
- 1µl T4 DNA Ligase Buffer (10X, ThermoFisher Scientific)
- Up to 10µl with dH₂O

Reactions were carried out in the Mastercycler® nexus X2 (Eppendorf), at 37°C for 30min, before inactivation of the enzyme at 65°C for 20min.

**T4 Polynucleotide Kinase reaction was allowed to reach RT before setting up a T4 DNA Ligation reaction:**
- 10µl T4 Polynucleotide Kinase reaction
- 2µl T4 DNA Ligase Buffer (10X, ThermoFisher Scientific)
- 2µl T4 DNA Ligase enzyme (ThermoFisher Scientific)
- Up to 20µl with dH₂O

These reactions were incubated for 1hr at RT, before transformation with either Stellar or Stbl3 chemically competent cells.

### 2.2.8 CRISPR single guide RNA design

Guide sequences compatible with wild type SpCas9 targeting TRAC, CD52, and B2M were originally designed using both the Massachusetts Institute of Technology (MIT) online CRISPR design tool (http://crispr.mit.edu/) and then later using Benchling (https://benchling.com). Both the MIT and Benchling design tools detect potential sgRNA sequences within a given sequence compatible with an 5’ NGG 3’ PAM. Benchling assigns each potential guide a score representing an in silico prediction of on and off-target CRISPR cutting. The on-target score is based on work by Doench and associates (2016), which takes into account the position of the predicted cut site within the translated gene sequence and provides a score from 0 – 100 (higher scores indicate better predicted on-target activity) [87]. On the other hand the off-target score is calculated based on a method proposed by Hsu and colleagues (2013) where in silico predict off-target scores are deducted from an initial score of 100; making high scoring guides preferable [88]. Both platforms further calculate the minimal number of mismatches (MMs) necessary for these guides to bind other intronic and exonic locations within the genome of the selected species. Number of MMs necessary to target an exonic location was kept to a minimum of three.

Single guide RNA sequences compatible with BE3 targeting TRAC, TRBC 1/2, CD52, and B2M loci were designed using a combination of tools. Guides predicted to generate a pre-mature stop codon by utilising C>T changes compatible with BE3 were designed using the iSTOP online tool (http://www.ciccialab-database.com/istop). Although, the CRISPR-SKIP platform (http://song.igb.illinois.edu/crispr-skip/submit) described by Gapinske et al (2018), can be used to design sgRNA sequences compatible with BE3 for the targeted disruption of splice acceptor sites, there is
currently no similar tool for the removal of mRNA splice donor sequences [115]. Therefore, the Human Splice Finder software version 3.1 (HSF v3.1, http://www.umd.be/HSF/HSF.shtml) was used to locate and score the naturally occurring mRNA splice sites within the targeted gene. Benchling was then used to look for protospacer sequences predicted to cause C>T conversions within the splice donor and acceptor sequences that would reduce the predicted score calculated by the HSF v3.1 software. All BE3 compatible guides were assigned an in silico predicted on- and off-target base editing score by Benchling, where on-target score is based on in vitro analysis using a first generation base editor [74], while off-target score is based on work by Hsu and colleagues (2013) [215]. In both instances, higher scoring guides were favoured.

2.2.9 Protospacer cloning for sgRNA expression

Plasmid DNA (1µg) containing a sgRNA expression cassette was digested with either BbsI or BsmBI type IIS restriction enzymes to remove a short DNA stuffer sequence (20bp) between the RNA Pol III promoter (U6/ H1), and sgRNA scaffold sequence. FastAP (1U) was added to this reaction to remove 5’ and 3’ phosphate groups from the DNA ends. The reaction was purified using the Monarch® PCR & DNA Cleanup Kit (NEW ENGLAND Biolabs) to removal the DNA stuffer sequence, as well as residual enzymes and buffers. Protospacer sequences were ordered as complimentary DNA oligos which when annealed formed dsDNA flanked by ssDNA overhangs complimentary to BbsI and BsmBI overhangs in the digested plasmid. Oligo annealing and phosphorylation by T4 Polynucleotide Kinase (NEW ENGLAND Biolabs) was performed as detailed by the Zhang lab GeCKO protocol (http://genome-engineering.org/gecko/wp-content/uploads/2013/12/lentiCRISPRv2-and-lentiGuide-oligo-cloning-protocol.pdf). A 10µl ligation reaction containing 50ng linearized plasmid DNA and 1µl (1:200 dilution in dH2O) annealed oligos was setup at RT for 15 minutes using T4 DNA ligase (ThermoFisher Scientific) prior to bacterial transformation.

2.2.10 Eurofins Genomics Sanger sequencing

DNA was prepared for Sanger sequencing as requested by eurofins Genomics.
- Purified PCR product: 15µl, at 5ng DNA/µl
- Gel extracted PCR product: 15µl, at 10ng DNA/µl
- Plasmid DNA: 15µl, at 50ng DNA/µl

These were all sent with a premixed sequencing primer (2µl at 10µM).

### 2.2.11 Lentiviral vector production

Lentiviral vector stocks were produced by transient transfection of 293T cells with either second or third generation packaging and transfer plasmids. All packaging plasmids were developed by the Trono lab and manufactured from PlasmidFactory (Bielefeld, Germany) at 1µg/µl, in 0.1 x TE buffer [216]. Second generation lentiviral vector production relied on two packaging plasmids, pCMVR8.74 (Addgene #22036) comprising the viral gag-pol, tat, and rev, as well as pMDG2 (Addgene #12259) expressing the vesicular stomatitis virus-G (VSV-G) envelope. In a third generation packing system, the nuclear exporter rev signal is expressed from a separate plasmid (pRSV-Rev, Addgene #12253). Additionally, tat expression has been removed entirely from the gag-pol expressing plasmid (pMDLg/pRRE, Addgene #12251). The same VSV-G envelope was used for both second and third generation vectors (pMDG2, Addgene #12259).

Both second and third generation transfer plasmids used in this project contained the HIV-1 cPPT, ψ, and RRE elements, as well as a woodchuck post-transcriptional regulatory element (WPRE). Additionally, both used internal RNA Pol II promoters for transgene expression. However, the second generation transfer plasmid relies of Tat dependent expression from the U3 promoter in the 5’LTR for viral genome production. Whereas, the U3 promoter has been replaced with a tat independent cytomegalovirus (CMV) promoter in the third generation transfer plasmid.

Viral stocks were produced by seeding 12 T175 flask at 22-25x10^6 293T cell/ flask, 1 day prior to plasmid transfection. A transfection mix was made containing the relevant plasmids, and 1x10^{-7} mol/L PEI, in reduced serum OPTI-MEM (ThermoFisher Scientific). Medium change were performed with complete DMEM at both 4hrs and 24hrs post transfection. Culture media containing lentiviral particles was harvested
at 48hrs post transfection, and passed through a 0.45μm pore cellulose acetate filter. This was then concentrated by ultracentrifugation for 2 hours at 100,000g. Virus was resuspended in reduced serum OPTI-MEM (ThermoFisher Scientific) and incubated on ice for 1 hour before storage at -80°C for further use.

2.2.12 Lentiviral vector titration

All viral titers were measured by serial dilution of the lentiviral vector (10, 2, 0.4, 0.08, 0.016, and 0.0032µl), for transduction of 0.1x10^6 293T cells. At day 3 post transduction, 1/10 of the 293T cells were harvested from the wells and stained with for the relevant transgene.

Calculation of transducing units (TUs) per ml was performed as below:

TUs/ per ml= \[\left(\frac{\text{\% transgene positive} - \text{background}}{1000}\right) \times \left(\frac{1000}{\text{lentiviral vector volume}}\right)\]

Therefore if the % transgene positive= 2.6 with 1% background at 0.016µl of lentiviral vector:

TUs/ per ml = \[(2.6-1) \times 1000\] x (1000/0.016)

TUs/ per ml = 1600 x 62500

TUs/ per ml = 1x10^8

Titration of lentiviral vectors expressing an rTCR transgene were carried out in a TRAC-JE6.1 cell line. This method provides a measure of TUs/ml which is most accurate in the cell type used during the viral titration. It is therefore acknowledged that the resulting viral titers include a margin of inaccuracy when applied to other cell types.

2.2.13 Sanger sequencing analysis of non-homologous end joining, and targeted Cytidine deamination events

Genomic DNA extraction was performed using DNeasy Blood and Tissue Kit (69504, QIAGEN) and a PCR reaction designed to amplify 400-800bp over the protospacer binding site. PCR products were discriminated by 1% agarose gel electrophoresis,
before purification by Monarch® DNA Gel extraction Kit (NEW ENGLAND Biolabs), followed by Monarch® PCR & DNA Cleanup Kit (NEW ENGLAND Biolabs). DNA was sent to eurofins Genomics for Sanger sequencing. Resulting Sanger sequencing data was analysed using either Tide (https://tide.nki.nl/) or Synthego ICE (https://ice.synthego.com/#/) online tools, to measure the frequency of indels, at the predicted SpCas9 scission site. When analysing C>T conversion rates produced by cytidine deaminase base editing technologies, EDITR software was used (https://moriaritylab.shinyapps.io/editr_v10/).

### 2.2.14 Cell line culture

**293T cells:** Adherent cells cultured in either T75 or T175 flasks depending on the scale, using complete DMEM medium. Cell handling was performed as detailed by the ATCC, and were maintained by splitting once every 2-3 days, 1 flask in 4.

**JE6.1 and Daudi cells:** Suspension cells grown in upright T75 flasks. Cultured as described by the ATCC in complete RPMI medium. Maintained in 20ml/flask cultures, by splitting 1 flask in 4 every 2-3 days.

### 2.2.15 Primary human lymphocyte culture and modification

Between 10-50ml of blood was received from a consenting healthy donor, and peripheral blood mononuclear cells (PBMCs) were isolated by ficoll density gradient. These cells were either cryopreserved for later use or used immediately. Cryopreserved cells were thawed and rested overnight before activation, whereas fresh cells were activated with TransACT reagent (Miltenyi Biotech, Surrey, UK) immediately following isolation. Lymphocytes were cultured in TexMACS medium (Miltenyi Biotech) with 3% human AB serum (Seralabs, Brussels, Belgium) and 100U/ml human IL-2 (Miltenyi Biotech). These were generally split 1:2 and given a 75% medium changes every 2 day.

Transduction with lentiviral vector was performed at day 1 post activation using a multiplicity of infection (MOI) of 5 and where applicable mRNA electroporation performed at day 4 post activation. Lymphocytes were cultured till day 11 post
activation, by which time cells might be TCRαβ/ HLA class I depleted by magnetic bead separation.

### 2.2.16 Cryopreservation and recovery

Cell were cryopreserved in 1ml aliquots (cell lines: 20-50 x10⁶, primary cells: 10-20 x10⁶) in either cell line or primary cell freezing mix. These aliquots were first placed into a Mr. Frosty™ freezing container (ThermoFisher Scientific) filled with 100% isopropanol, and stored level in a -80°C freezer. After 24 hours, aliquots were moved to liquid N₂ for long term storage. Recovery of cells was performed by putting the tubes in a sealed bag and placing them in a 37°C water bath till only a small ice crystal remains (3-4 minutes). Tube were then removed from the bag and cells were pipetted into a falcon tube containing 10ml of appropriate pre-warmed medium. Cells were washed twice in PBS (400g for 10min), before being counted and seeded for subsequent culture.

### 2.2.17 Electroporation

**Neon transfection system (ThermoFisher Scientific):** The majority of electroporation reactions carried out as a part of this project have used the Neon transfection system, with the 100µl tip kit. Cells were electroporated at 10x10⁶ cells/ ml in buffer T, using protocol 24 (1600V, 10ms, 3pulses).

**4D-Nucleofector™ X Unit (Lonza):** This device was used as an alternative system using the 100µl cuvettes, at cell concentrations ranging between 10-50 x10⁶ cells/ ml in buffer P3, using program EH115. The application of this device has been indicated in the results section.

Regardless of the electroporation device used cells were placed a pre-warmed plate and incubated in a state of hypothermia (30°C, 5% CO₂) overnight. Cells were then restored to normal culture conditions (37°C, 5% CO₂), the next day.
2.2.18 Flow cytometry

Generally between 0.1 - 1x10⁶ cells were used for flow based phenotyping. All wash and staining steps occurred in FACS buffer (PBS with 2% FCS). Cell acquisition was carried out on a 4-laser BD LSR II (BD Biosciences, Oxford, UK). Compensation between fluorochromes was performed by using OneComp eBeads™ (ThermoFisher Scientific). Half these beads are conjugated to anti-lambda/ Kappa chain antibodies making them able to recognise hamster, mouse, and rat antibodies, providing an ideal single colour control. FACS analysis performed using FlowJo v10.

2.2.19 Depletion by magnetic bead separation

Depletions were all performed by initial staining with either anti-HLA-ABC-biotin (Miltenyi Biotech), anti-TCRαβ-biotin (Miltenyi Biotech) or a combination of both. Subsequent staining with anti-biotin microbeads ultrapure (Miltenyi Biotech) allows indirect magnetic labelling which can be passed through an LD column (Miltenyi Biotech) in a magnetic field to deplete the culture of positively stained cells. All staining and wash steps were performed in MACS buffer consisting of PBS, 0.5% BSA and 2mM EDTA. Where the number of labelled cells is >10⁸ or total cells is >5x10⁸ the following protocol was used. Cells were re-suspended in 90μl MACS buffer with the addition of 10μl of biotin labelled antibody. This was incubated at 4°C for 10 minutes before being washed twice with 2mL of MACS buffer (spun at 300g for 10 minutes). This was then re-suspended in 80μl MACS buffer with the addition of 20μl anti-biotin microbeads ultrapure (Miltenyi Biotech). This was incubated for 15 minutes at 4°C before being washed with MACS buffer another two times. During these wash steps an LD column is placed in a magnetic field and is equilibrated with 2ml of MACS buffer (this flow through is discarded). Labelled cells are now re-suspended in 500μl of MACS buffer before being passed through the LD column. Column is then washed twice with 1ml of MACS buffer. The flow through containing depleted cells is collected and subsequently counted and cultured.
2.2.20 Cas9 western blot

Total protein was extracted from cell pellets stored at -80°C, by incubation with 40μl lysis buffer at 4°C for 15min. Samples were spun at full speed (~20,000g) for 15min at 4°C. Protein concentration was quantified using the Bio-Rad Protein assay kit II (Bio-Rad), at 1:50, 1:100, and 1:500 dilutions. Of this 20μg total protein per sample was placed in a new Eppendorf, toped up to 21μl with lysis buffer base, with the addition of 6μl 5x sample dye and 3μl DTT (total 30μl). Samples were denatured before gel electrophoresis by heating at 95°C for 3min, and Amersham™ ECL™ Rainbow™ Markers- Full Range ladder (GE LifeSciences) was used to determine protein size. Pre-cast, 10well polyacrylamide gels (4-15%, well capacity 30μl) were run in a Mini-PROTEAN® Tetra cell system (Bio-Rad) using 1X running buffer (Bio-Rad). Gel was run at 100V for 10min, followed by 120V for 1hr. Transfer of protein to a PVDF membrane (Bio-Rad) was performed using the Trans-Blot Turbo Transfer System (Bio-Rad), using the rapid blot transfer (3min) program. Membranes were placed in 1X wash buffer, before placing in blocking solution for 1hr. Primary antibody solutions were prepared in PBS containing 3% BSA (10ml anti-CRISPR-Cas9 antibody abcam 1:1000, and 10ml anti-β-Actin 1:200), and membranes were cut in two at the orange line (70kda). These were washed, before the top half was stained for Cas9 (5ml) and the bottom half was stained for β-Actin (5ml) overnight. Primary antibody was removed and membranes washed, prior to staining for 1hr with secondary antibody (5ml/ membrane, 1:3000 Amersham ECL HRP-linked whole Ab, in secondary stain buffer). Membranes were washed, and antibody staining was visualisation using 3ml Clarity ECL (Bio-Rad), in a Syngene GeneGnome XRQ western blot imager.

2.2.21 Statistics

All statistical analysis was performed using GraphPad Prism software version 8.0.0. Details of analysis preformed can be found in the relevant figure legends.
Chapter 3  Development of terminal-CRISPR lentiviral vector for coupling transgene expression with CRISPR editing effects

Cellular immunotherapies expressing either rTCRs or CARs are being applied clinically in multi-phase clinical trials, with promising results in the field of haematopoietic malignancies [217]. To date, the majority of these trials have relied on autologous T cell products, in which cells are recovered from the patient, modified and expanded \emph{ex vivo}, before delivery back to the patient. However, this approach limits the widespread availability of these treatments, while also being relatively expensive and time consuming to produce. The expanding range of genome editing tools has provided the means to overcome HLA barriers to produce ‘off-the-shelf’, ‘universal’ T cell immunotherapies [85, 179, 218]. The first clinical demonstration of such universal therapies utilised TALEN pairs targeting both TRAC and CD52, while lentiviral transduction was used to express a CAR against CD19 for the treatment of B-ALL [210]. In this setting TRAC knockout mitigated the risk of donor T cells causing GVHD in the non-HLA matched donor, while CD52 knockout protected the donor T cells from lymphodepleting chemotherapeutic Alemtuzumab [159, 210].

Although TCRαβ magnetic bead mediated depletion protocols are highly effective in removing \textasciitilde99\% of residual TCRαβ expressing T cells, the remaining \textasciitilde1\% constitutes a risk for GVHD. To this end clinical doses are capped at $5 \times 10^4$ TCRαβ$^+$ T cells/kg [211], limiting the CAR T cell dose which only accounts for a proportion of the cell inoculum depending on transduction efficiency. This caveat exists for any genome editing strategy where DNA nuclease activity and transgene expression have been performed independently. This chapter therefore describes the design and validation of a SIN 3$^{rd}$ generation lentiviral vector platform referred to as terminal-CRISPR that couples CRISPR sgRNA expression to CAR19 transgene expression, thereby restricting knockout to the transduced cell population. Initial validation of this platform expressing a TRAC specific sgRNA, demonstrated efficient removal of the allogeneic TCR in primary human T cells, while showing superior functionality in an in vivo tumour model.
3.1 Design concept of terminal-CRISPR lentiviral vector

The configuration detailed in this chapter places a second generation CAR19 transgene under the control of an internal human phosphoglycerate kinase (hPGK) promoter, while CRISPR sgRNA expression is driven from a human U6 promoter (Figure 3-1 A). This vector configuration is hereon referred to as “terminal-CRISPR” as the sgRNA expression cassette is incorporated within the ΔU3, immediately proximal to the repeat (R) region in the 3’LTR [219] (Appendix A). Although, Ren et al (2017) also demonstrated coupling of CRISPR sgRNA and CAR19 transgene expression in a lentiviral vector for T cell engineering, their sgRNA expression cassette was placed outside the LTRs between the cPPT and the internal Pol II promoter [220]. Placement of the sgRNA expression cassette within the ΔU3 region of the 3’LTR effectively takes advantage of the lentiviral lifecycle, in which this region is duplicated in the 5’LTR by reverse transcription of the ssRNA viral genome (Figure 3-1 B). Duplication of the sgRNA expression cassette theoretically increases the expression of sgRNA from the vector. Past attempts to use these flanking LTR’s have including insertion of cDNA expression cassettes [40], shRNA expression cassettes [212], and chromatin insulator elements [221]. Additionally, placement of promoter elements within the LTR also avoids transcriptional interference effects between Pol III and Pol II promoters [213, 214].
Figure 3-1 Terminal-CRISPR configuration coupling lentiviral sgRNA and CAR19 expression:

A. Third generation self-inactivating (SIN) lentiviral plasmid DNA configuration, coupling CAR19 transgene expression from an internal RNA Pol II promoter, and T cell

B. Figure 3.1: A. ΨhPGK WPRE cPPT CAR19 5’LTR 3’LTR PolIII TRAC scaffold RΔU3 U5 ~350bp ~1500bp PolIII TRAC scaffold

C. i. ACC-I-PGK-CAR19 TT-I-PGK-CAR19

D. TT-CAR19 SpCas9 mRNA electroporation Cytotoxic T cell

scFv-CD19

CD8 41BB CD3 ζ 41ΒΒ vL

TCR

90
receptor alpha constant (TRAC) specific single guide RNA (sgRNA) from an RNA Pol III promoter (U6/H1) embedded in the deleted unique 3’ (ΔU3) region of the 3’ long terminal repeat (LTR). This configuration is referred to as terminal-TRAC-CAR19 (TT-CAR19). Single stranded RNA (ssRNA) lentiviral genome expressed from the cytomegalovirus (CMV) promoter encompasses everything from the repeat (R) region in the 5’LTR to the equivalent R sequences 3’LTR (ssRNA = grey dotted line, transcription start = green arrow, transcription stop red hexagon). B. Proviral DNA of TT-CAR19 vector after reverse transcription of the ssRNA lentiviral genome shows duplication of the 3’ ΔU3 region carrying the sgRNA expression machinery into the proviral 5’LTR C. Amplification of proviral LTR sequences by PCR from both a pCCL-CAR19 control without the addition of RNA Pol III/sgRNA sequence, compared to the TT-CAR19 configuration. i. 5’ proviral LTR spanning the ΔU3 and Psi (ψ) regions, indicate successful duplication of the sgRNA expression machinery (~400bp in the pCCL-CAR19 control, ~750bp in the TT-CAR19 configuration). ii. Amplification of proviral 3’LTR demonstrating the continued presence of the CRISPR sgRNA expression cassette (~400bp in the pCCL-CAR19 control, ~750bp in the TT-CAR19 configuration). D. Both CAR19 and sgRNA are expressed constitutively from the integrated TT-CAR19 configuration. Transient delivery of Streptococcus pyogenes Cas9 (SpCas9) mRNA by electroporation, is translated to protein to form ribonucleoprotein (RNP) complexes with expressed sgRNA. This strategy couples CRISPR activity with CAR19 transgene expression to produce cytotoxic CAR19 T cells with CRISPR-mediated disruption of the T cell receptor (TCR) αβ-CD3 complex. hPGK: human phosphoglycerate kinase promoter, RRE: rev response element, cPPT: central polypurine tract, WPRE: woodchuck post-transcriptional regulatory element, scFv: single chain variable fragment, vH: variable heavy, vL: variable light.

In order to construct the terminal-CRISPR lentiviral plasmid, an sgRNA expression cassette flanked by XbaI sites containing an RNA Pol III U6 promoter, a mini-stuffer sequence containing BbsI restriction enzyme sites, a guide scaffold, and a Pol III termination signal (stretch of 7 T nts) were designed and synthesised by GeneArt. Infusion cloning was used to place the sgRNA expression cassette within the 3’LTR of a pCCL-CAR19 plasmid. Colonies were screened by visualising DNA on an agarose gel
following XbaI restriction digest, and then confirmed by Sanger sequencing. As inserting sgRNA protospacer sequences within this configuration relies upon linearisation by BbsI restriction digest, SDM was used to remove all other BbsI sites within the plasmid, as well as sites for two other type II restriction enzymes (BsmBI and SapI).

In the first instance a protospacer sequence specific for the TRAC locus was designed using the MIT CRISPR design tool, whose cleavage site shares close proximity to the TALEN pair initially characterised by Poirot et al (2015), and further used in the clinical study led by Qasim and colleagues (2017) [85, 210]. To enhance expression of protospacer sequences from the hU6 promoter, a +1G transcription start site was added [222, 223]. An oligo annealing protocol was then used to clone protospacer sequences within the terminal-CRISPR plasmid. The resulting configuration expressing both TRAC specific sgRNA, and CAR19 transgene will here be referred to as terminal-TRAC-CAR19 (TT-CAR19).

It has been noted by several groups that modification of the lentiviral LTRs can reduce titers [213, 221, 224]. However, insertion of the ~350bp sgRNA expression cassette within the ΔU3 region of the 3’LTR did not appear to effect the concentrated viral titer of the TT-CAR19 vector compared to its pCCL counterpart (1.6x10^8 versus 1.5x10^8 TU/ml), as measure by flow cytometry. This observation was supported by similar transduction efficiencies at MOI 5 between pCCL-CAR19 and TT-CAR19, in both a JE6.1 cell line (96% and 94.3%), and in primary human T cells (42.1% and 42.1%).

To confirm self-duplication of the sgRNA cassette in the primary T cells post reverse transcription, unique primer pairs were designed to amplify the proviral 5’ and 3’ LTRs. The 5’LTR was amplified using a forward primer binding the U3 region and a reverse primer specific to Psi (ψ) (Figure 3-1 C i), while amplification of the 3’LTR relied on a primer binding WRPE and U5 (Figure 3-1 C ii). Genomic DNA was extracted at day 11 post transduction to reduce the risk of amplifying the LTRs from non-integrated viral genomes. Visualisation of the proviral 3’LTR from pCCL-CAR19, and TT-CAR19 transduced cells showed expected DNA bands of ~400bp and ~750bp respectively confirming the presence of the sgRNA expression cassette. Similarly, the
proviral 5’LTR presented an increased product size in TT-CAR19 transduced cells (~750bp) compared to pCCL-CAR19 (~400bp), indicating self-duplication of the sgRNA expression cassette. Sanger sequencing of the gel this bands was carried out to confirm the integrity of the sgRNA expression cassettes after reverse transcription and integration [219](Appendix A).

As detailed above in section 1.7, there are a number of delivery strategies that can be applied to the SpCas9 endonuclease. Delivery of plasmid DNA or integrating / non-integrating viral vectors expressing SpCas9 have shown to be effective in pre-clinical applications, as they promoter effective long lasting expression. However, their use in human trials could be problematic, firstly as extended exposure to SpCas9 has proven immunogenic [225] and furthermore as this would increase the likelihood of off-target cleavage [226]. To help mitigate these issues, transient delivery of the SpCas9 endonuclease can be achieved by electroporation with either mRNA or protein SpCas9. Preliminary investigation suggested that both mRNA and protein SpCas9 formats were compatible with the TT-CAR19 configuration. However, large amounts of SpCas9 protein were required to achieve efficient disruption, making use of the SpCas9 mRNA more practical. Following translation of the SpCas9 mRNA into protein TRAC specific sgRNA expressed from the TT-CAR19 vector will complex to form RNPs which will execute precise genomic cleavage to give rise to a TCRαβ-/CAR19+ population (Figure 3-1 D).

### 3.2 Optimisation of genome editing effects in primary T cells using TT-CAR19 lentiviral vector

As TRAC sgRNA is expressed from the lentiviral vector, timing of SpCas9 mRNA electroporation post transduction may impact the knockout efficiency. In order to assay if there is an optimal time point post transduction for SpCas9 delivery, primary human PBMCs were isolated and activated one day prior to transduction with TT-CAR19 vector at MOI 5 (transduction 38.6%). Cells were subsequently electroporated with SpCas9 mRNA (100µg/ml), at day 1, 2, 3, 4, 7 and 11 post transduction, before assessment of TRAC knockout by flow cytometry 7 days post electroporation (Figure 3-2 A). These results indicated that optimal SpCas9 delivery was either on day 3 or 4
post transduction (52.6% and 63.4%, respectively, gated on CAR19+), although further repeats of this assay would have to be conducted to achieve a definitive answer.

In addition to the timing of SpCas9 delivery, the concentration of the mRNA was next investigated as it was hypothesised to directly influence CRISPR knockout efficiency. To test this, primary T cells from a healthy blood donor were activated and transduced the following day with TT-CAR19 vector at MOI 5. At day 4 post activation, transduction efficiency was assayed by flow cytometry (74.3% CAR19+), and cells were next electroporated with a range of SpCas9 mRNA (10, 25, 50, and 100µg/ml) (Figure 3-2 B). TRAC disruption measured by flow cytometry-based cell surface staining of TCRαβ expression showed CRISPR/Cas9 knockout plateauing at 25µg/ml SpCas9 (67.7% gated on CAR19+), reaching 85.7% TCRαβ disruption at 100µg/ml (gated on CAR19+). Next, to quantify editing at the genomic level the TRAC locus was amplified and sequenced. Sanger traces were then subjected to TIDE analysis which revealed a similar trend in genomic disruption to that seen phenotypically, reaching 44.8% at 100µg/ml SpCas9 mRNA (Figure 3-2 C).

As previously eluded to in section 3.1, transient expression of SpCas9 poses a safer and more clinically applicable delivery strategy, compared to constitutive expression from vector configurations such as lentiCRISPR v2 (Addgene: #52961). Thus far use of stabilised SpCas9 mRNA (capped, and polyadenylated), in combination with the TT-CAR19 vector has shown effective CRISPR/Cas9 knockout both phenotypically and genomically. However, visualisation of the SpCas9 protein produced from this mRNA overtime could provide useful information relating to its persistence when using this delivery strategy.

To confirm that SpCas9 protein could be efficiently detected, a titration of the purified protein using four fold serial dilutions from 1000 – 0.24ng was run on a western blot. The membrane was probed with an anti-SpCas9 antibody which was able to detect a band of the correct size (160kDa) at all dilutions (Figure 3-2 D).

With this in mind PBMCs were isolated from a healthy donor, and activated and transduced with TT-CAR19 as previously described. At day 4 post activation
transduction efficiency was measured (33.4%) and cells were electroporated with 100µg/ml SpCas9 mRNA. Cell pellets were collected at 0.25, 0.5, 1, 2, 3, and 7 days post SpCas9 mRNA electroporation for western blot analysis. Having confirmed on-target editing by TCRαβ knockout (52.6% gated on CAR19+), total protein was extracted from the collected cell pellets, before quantification by BCA assay. Normalised amounts of protein lysate were subjected to SDS-page and the membrane was stained for SpCas9 protein. Staining for β-actin was used as a protein loading control confirming consistent loading between samples. Unmodified cells used as a negative control showed no SpCas9 staining. Conversely, SpCas9 protein could be detected in SpCas9 mRNA treated samples as early as 0.25 days post electroporation. SpCas9 protein appeared to peak at 0.5 days, and had mostly dissipated by day 3 post electroporation. No band corresponding to SpCas9 protein could be detected by 7 days post electroporation, confirming the transient nature of this delivery method (Figure 3-2 E).
Figure 3.2 Validation and optimisation of Terminal-TRAC-CAR19 for T cell engineering:

A. Quantification of TCRαβ protein knockout by flow cytometry, in primary T cells electroporated with 100µg/ml SpCas9 mRNA at days 1, 2, 3, 4, 7, and 11 post TT-CAR19 transduction. Optimal TCRαβ disruption was observed between days 3 and 4. Calculation is based on CD45+ CAR19+ flow cytometry gating.

B. Titration of SpCas9 mRNA at 10, 25, 50, and 100µg/ml in TT-CAR19 transduced primary T cells. Quantification of TCRαβ knockout by flow cytometry displays effective CRISPR/Cas9 mediated disruption at concentrations ≥25µg/ml, peaking at 100µg/ml. All values are based on CD45+CAR19+ gating.

C. Comparable TRAC knockout at both protein (flow cytometry staining for TCRαβ) and genomic (Tide analysis of Sanger sequencing across the TRAC target site) levels.

D. Titration of purified SpCas9 protein with 1 in 4 serial dilutions from 1000-0.24ng for detection by western blot. Purified SpCas9...
protein detected at 160kDa. E. Western blot measuring transient nature of SpCas9 protein in primary human T cells transduced with TT-CAR19 vector and electroporated with 100µg/ml SpCas9 mRNA. Peak expression detected at 12 hours post electroporation, which appears to be mostly dissipated by day 3 (72hr). In this assay β-actin was used as a loading control (42kDa) and appeared consistent across the samples.

3.3 Anti-leukaemic activity of TT-CAR19 T cells in vivo

Use of the TT-CAR19 vector with mRNA SpCas9 has proven to be scalable and compliant with therapeutic manufacture using a CliniMACS Prodigy device. Following a 12 day semi-automated culturing procedure a highly homogenous population of T cells post TCRαβ depletion which were ~97% CAR19+, with <1% TCRαβ carriage could be generated [219] (Appendix A). To investigate the molecular consequences of CRISPR/Cas9 genome editing three independent methods were used to characterise both unmodified and SpCas9 treated cells; whole genome sequencing (WGS), targeted next generation sequencing (NGS), and Digenome-seq. In summary high levels of on-target editing were confirmed with undetected off-target cleavage events detectable in SpCas9 treated samples. More detail can be found in Appendix A [219].

In order to assay whether CRISPR/Cas9 editing had an impact on the cytolytic potential of the TCRαβ+/CAR19+ T cells, a $^{51}$Cr release assay was performed, demonstrating comparable killing of target cells (CD19+ SupT1 cells, and CD19+ Daudi cells), between TCRαβ+/CAR19+ and TCRαβ-/CAR19+ cells [219] (Appendix A). Once their ability to generate a robust cytotoxic response in vitro was validated, the CRISPR modified T cells were tested in an in vivo humanised murine model. Non-obese diabetic (NOD)/severe combined immunodeficiency (SCID)/γc−/− (NSG) mice were inoculated with 5x10⁵ CD19+/EGFP+/Luciferase+ Daudi tumour cells by tail vein injection. Engraftment of tumour cells was confirmed after 3 days by bioluminescence imagining (BLI) using the IVIS Lumina III. On day 4, mice received either 5x10⁶ CRISPR edited TCRαβ+/CAR19+ (n=8), or unedited TCRαβ+/CAR19+ (n=8) effector T cells by tail vein injection. Unmodified T cells (TCRαβ+/CAR19−) (n=8), and
PBS (n=3) were delivered as negative controls. Serial bioluminescence imaging was then carried out to track tumour progression overtime (Figure 3-3 A). Half the animals were culled at day 14 for bone marrow analysis, while the other half were tracked until either day 28 or 34. The timeline presented in Figure 3-3 B, represents a more detailed experimental overview.

The mice receiving CAR19+ effectors exhibited a rapid clearance of Daudi tumour cells showing little signal by day 14. This was in direct contrast with mice receiving either unmodified TCRαβ+/CAR19- effectors or PBS, which exhibited a steady increased in bioluminescent signal (p<0.001). Furthermore, the CRISPR modified TCRαβ+/CAR19+ T cells appeared to have superior tumour clearance by day 14, compared to their unedited TCRαβ+/CAR19+ counterparts (p<0.05) (Figure 3-3 C). Total bone marrow was harvested at day 14 from the two long leg bones of 4/8 TCRαβ+/CAR19+, 5/8 TCRαβ+CAR19+, 5/8 TCRαβ+/CAR19+, and 3/3 PBS receiving animals. Flow cytometry based analysis of the bone marrow supports the bioluminescence data, showing >45 fold reduction of GFP+ Daudi tumour cells in animals receiving CAR19+ effector groups compared to controls (Figure 3-3 D, and E). At day 28, the remaining mice inoculated with TCRαβ+/CAR19+ (3/8), showed a further 125 fold increase in bioluminescence signal and were culled. Mice receiving CAR19+ effectors (4/8 TCRαβ+/CAR19+, 3/8 TCRαβ+CAR19+), were monitored until day 34. Both CAR19+ groups presented with an increased signal, which was significantly higher in the unedited TCRαβ+/CAR19+ group (p<0.01), and appeared to be localised in the bone marrow (Figure 3-3 C).

Analysis of the bone marrow revealed high levels of GFP+ Daudi tumour burden in the unmodified TCRαβ+/CAR19- control group, despite an increased expansion of CD45+CD2+ T cells (Figure 3-3 D, and E). Additionally, consistent with bioluminescent imaging the unedited TCRαβ+/CAR19+ effector group showed a 6 fold increase in the presence of GFP+ Daudi tumour cells compared to day 14 (p<0.05). This did not appear to be the case in the CRISPR edited TCRαβ+/CAR19+ effector group exhibiting a decrease in GFP+ Daudi tumour burden at day 34 (p<0.05) (Figure 3-3 D, and E).

Interestingly further analysis of the GFP+ Daudi tumour cells present in the TCRαβ+/CAR19+ group revealed a subpopulation of CD20+/CD19- cells population,
indicating potential antigen escape. This was not present in groups receiving unmodified TCRαβ+/CAR19−, or CRISPR edited TCRαβ−/CAR19+ effectors (Figure 3-3F). Additionally, both TCRαβ+ effector groups presented with increased levels of programmed cell death protein 1 (PD1)+ expression (TCRαβ+/CAR19− = 48.6% day 14, and 83.5% day 28, and TCRαβ+/CAR19+ = 15.1% day 14, and 84.5% day 34), whereas the TCRαβ−/CAR19+ group presented with considerably lower PD1 expression from an average of 6.8% at day 14 to 24.8% at day 34.
Figure 3-3 In vivo functionality of TT-CAR19 effector T cells against CD19+/GFP+/Luciferase+ Daudi tumour cells:

A. Tumour progression was tracked in Non-obese diabetic (NOD)/ severe combined immunodeficiency (SCID)/γc−/− (NSG) mice by serial bioluminescence imagining (BLI) using the IVIS Lumina III, 15 minutes post intraperitoneal injection with D-luciferin,
which act as the substrate for luciferase enzyme. **B.** Experimental time line detailing tumour and effector T cell intravenous injections, BLI, and bone marrow (BM) harvesting. **C.** Tumour progression in experimental groups as measured by BLI, quantified in radiance (p/s/cm²/sr). Increased tumour burden was apparent in non-treated and control TCRαβ+/CAR19− groups. Mice receiving CAR19+ effector groups exhibited a significantly delayed tumour progression at day 14 post Daudi injection compared to the TCRαβ+/CAR19− control group. At termination, CRISPR edited TCRαβ−/CAR19+ effector group showed the lowest disease burden compared to their unedited TCRαβ−/CAR19+ counterparts. The error bars represent standard error of the mean (SEM), dotted lines represent significance a specific times, analysis by Mann-Whitney U tests, and solid line show linear regression analysis used to investigate the tumour burden over time between each group (* p <0.05, ** p < 0.01, *** p < 0.001, **** p < 0.0001). **D.** Disease burden in cells isolated from the BM at day 14, or at termination (day 28/34). Control TCRαβ+/CAR19− effector group presented with the highest presence of tumour cells in the BM at day 14, which significantly increased by termination at day 28. Tumour burden appeared comparable in the CAR19+ effector groups at day 14, however, the unedited TCRαβ+/CAR19+ effector group displayed significantly increased presence of tumour cells at day 34, whereas the edited TCRαβ−/CAR19+ group showed a significant decrease. Error bars represent SEM, with significance measured by Mann-Whitney U test. **E.** Representative flow cytometry data, of human CD45+ cells in the BM at day 14 and day 28/34. Demonstrates greatly reduced GFP+ tumour burden in CAR19+ treatment groups. **F.** Representative flow cytometry plots of GFP+ Daudi tumour cells isolated from the BM of respective effector groups. Emergence of a CD19+/CD20+ population present in animals receiving TCRαβ+/CAR19+ effectors. A CD19+/CD20+ population was not presents in the TCRαβ−/CAR19− control or TCRαβ+/CAR19+ effector group.

### 3.4 Knockout of B2M, and CD52 loci using Terminal-CRISPR vectors to create immunologic stealthy/ chemotherapy resistant CAR T cells

As detailed above the terminal-CRISPR platform efficiently removes the TCRαβ-CD3 complex from the cell surface when expressing a TRAC specific sgRNA. However, in
order to produce truly ‘universal’ cell therapies, disruption of additionally loci would have clinical benefit and expand the application of the terminal-CRISPR platform. This section will detail the knockout of other potential targets as well as the rationale for their selection.

3.4.1 Disruption of HLA class I by B2M knockout

The MHC locus encodes human leukocyte antigens (HLA), cell surface glycoproteins that are part of the immunoglobulin (Ig) superfamily. The MHC locus is highly polymorphic and contains regions called Class I, II, and III. HLA Class I molecules are membrane bound glycoproteins containing three Ig-like domains, which are complexed to B2M during its maturation within the cell. HLA class II molecules, instead, consist of a heterodimer of membrane-anchored α and β chains, each with two Ig-like domains. Class III encodes components of the complement cascade. Both HLA class I and II present peptide antigens on the cells surface using an N-terminal domain consisting of a platform of β-sheets with two parallel α-helices, effectively making a groove within which the peptide antigen is bound and displayed. HLA Class I molecules display self-molecules, and intracellular pathogens, whereas Class II complexes display products resulting from proteolytic degradation of extracellular proteins. The MHC locus contains hundreds of HLA molecules from Class I and II, with a single person expressing only a limited repertoire of HLA molecules. These polymorphic HLA molecules are grouped into A, B, and C for Class I and DR, DP, and DQ for Class II [227].

Disruption of HLA class I molecules from the cell surface, prevents the display of self-antigens. This would be advantageous in an allogeneic transplant setting, where disparate HLA class I molecules on the graft would otherwise lead to rejection by a host versus graft immune response. The highly polymorphic nature of HLA class I molecules complicates genome editing strategies, however, disruption of B2M prevents their maturation and cell surface expression (Figure 3-4 A). To this end, several groups have used CRISPR mediated disruption of the B2M gene for HLA class I removal [143].
Three sgRNA sequences targeting the *B2M* gene were designed, and previously validated in our lab by using the LentiCRISPR v2 (addgene: #52961) configuration in a JurkatE6.1 cell line. The LentiCRISPR v2 vector allows simultaneous expression of both sgRNA and SpCas9 from a single lentiviral vector. Of the three guides tested, one targeted exon 1 (*B2M* ex1-1), and two were designed against exon 2 (*B2M* ex2-1 and ex2-2) (Figure 3-4 B). Only guides B2M ex1-1 and ex2-2 showed a high level of HLA class I protein disruption in the Jurkat E6.1 cell line measured by flow cytometry (74% and 52.5% respectively).

The B2M ex1-1 protospacer sequence exhibiting the highest level of knockout was used for all subsequent experiments. B2M ex1-1 guide was cloned into a terminal-CRISPR configuration expressing the CAR19 transgene (TB2M-CAR19). Primary human T cells were transduced with TB2M-CAR19 at MOI 5, followed by electroporation with SpCas9 mRNA (100µg/ml), 3 days post transduction. HLA class I knockout as measure by flow cytometry was 60.2% (84.6% gated on CAR19+) at day 3 post electroporation (Figure 3-4 C). Residual B2M+ cells can be further depleted by antibody coated magnetic beads thereby enriching the HLA I+/CAR19+ population. Enrichment of the HLA I+/CAR19+ population resulted in a highly homogenous cell product which was 93.8% CAR19+, with <1% HLA class I carriage (Figure 3-4 D). Genomic DNA was extracted from the depleted HLA I+/CAR19+ population and next generation sequences (NGS) ran by Dr Petrova to confirm on-target CRISPR scission (Figure 3-4 E). NGS output was analysed using online tool CRISPResso and showed 100% *B2M* gene disruption, consistent with results seen at the protein level.
**Figure 3-4 CRISPR/Cas9 mediated disruption of HLA class I using the terminal-CRISPR configuration:**

**A.** Schematic representation of the HLA class I complex on the cell surface presenting peptides of self or intracellular pathogens (red line). HLA class I molecules are highly polymorphic, but contain a conserved B2M subunit. Disruption of this B2M subunit, prevents maturation of the HLA class I complex, preventing surface expression. **B.** Diagram of the B2M locus, containing 3 coding exons (purple), separated by introns (black lines), with the translation stop codon shown in red. Grey triangles indicate the location of tested sgRNA binding sites. **C.** Primary T cells transduced with a terminal-CRISPR-CAR19 vector expressing the B2M ex 1-1 guide (TB2M-CAR19) 3 days post electroporation with 100µg/ml SpCas9. Flow cytometry confirms protein knockout is isolated to the vector transduced CAR19^+ population. **D.** Magnetic bead-mediated depletion of residual HLA class I expressing cells with <1% carriage. The coupled nature of the knockout from the terminal-CRISPR vector, simultaneously enriches for CAR19 expression resulting in a homogenous 93.8% HLA I/CAR19^+. **E.** Next generation sequencing (NGS) across the B2M locus in the HLA class I depleted cell population. Output was analysed by the CRISPResso online tool confirming high levels of genome disruption, consistent with the protein knockout.
3.4.2 Serotherapy resistance by CD52 knockout

The CD52 glycoprotein consists of 12 amino acids (AA) (21-28 kDa) bound to a glycosylphosphatidylinositol cell surface anchor. Expression of CD52 is mostly limited to immune cell subsets, including mature lymphocytes, and natural killer (NK) cells; but is also expressed on other tissues including the male genital tract and mature sperm cells [228]. Due to its negative charge and high cell surface density, CD52 is suspected to play a role in anti-adhesion allowing cell motility. High expression of CD52 on immune cells makes anti-CD52 monoclonal antibody (mAb) alemtuzumab a potent lymphodepletion/ immunosuppressive agent.

Poirot and colleagues (2015) used TALEN based editing to disrupt CD52 expression, generating a cellular therapeutic that can be administered to patients undergoing alemtuzumab treatment. They further disrupted the TRAC locus, in order to generate a TCRαβ/CD52 deficient universal CAR19 T cell product (UCART19) [208]. The immunosuppressive environment, created by alemtuzumab effectively clears space and promotes engraftment for the edited T cells. Clinical results from two paediatric patients with relapsed refractory CD19+ B-ALL treated with UCART19 achieved molecular remission within 28 days in the presence of lymphodepleting alemtuzumab [209] (clinical trial identifier: NCT02808442).

To investigate the applicability/ potential of CRISPR/Cas9 for CD52 gene knockout, four sgRNAs were designed targeting either exon 1 (CD52 ex1-1 and 1-2) or exon 2 (CD52 ex2-1 and 2-2) (Figure 3-5 A). Single guide RNAs in the form of annealed oligos were separately cloned into the terminal-CRISPR configuration expressing the CAR19 transgene, by type IIS restriction enzyme Bbs I digestion. The levels of CD52 disruption for each sgRNA was tested in primary human T cells from a healthy donor, by transduction with each of the four terminal-CD52-CAR19 vectors at MOI 5. Electroporation with 25µg/ml SpCas9 mRNA was used for initial sgRNA testing based on previous SpCas9 mRNA titration data shown in Figure 3-2 B. This revealed comparable, albeit low, protein knockout of CD52 measuring at 3.4% and 4.6% for CD52 ex1-1 and ex1-2 guides, respectively (6.2% and 7.2% gated on CAR19+) (Figure 3-5 B). Both exon 2 targeting guides failed to show appreciable knockout by FACS
analysis above that of background (1.2% for ex2-1 and 0.6% for ex2-2) (data not shown). In order to improve CD52 knockout, exon 1 guides were retested in primary human T cells at 100µg/ml SpCas9 mRNA which had proven to be optimal with the TRAC guide. Increasing SpCas9 mRNA concentration promoted a marked increase in CD52 knockout with both guides, attaining 7.8% and 35.5% disruption with CD52 ex1-1 and ex1-2, respectively (13.1% and 56.6% gated on CAR19+) (Figure 3-5 C). As the CD52 ex1-2 guide exhibited higher activity, further characterisation of CD52 disruption at the genomic level was performed on, DNA extracted from the cells. PCR amplicons of the target region were further interrogated by NGS performed by Dr Petrova with subsequent analysis using online software CRISPResso (http://crispresso.rocks/). This analysis showed 29.6% NHEJ (1.9% background) corroborating protein knockout measured by flow cytometry (Figure 3-5 D).
Figure 3-5 Validation of CRISPR/ Cas9 knockout of CD52 using a terminal-CRISPR expressed sgRNA:

A. Diagram of the CD52 locus, containing 2 coding exons (green), separated by an intron (black lines), with the translation stop codon shown in red. Grey triangles indicate the location of tested sgRNA binding sites. B. Primary T cells transduced with a terminal-CRISPR-CAR19 vector expressing CD52 specific sgRNA, and electroporated with 25µg/ml SpCas9 mRNA. Only guide sequences ex1-1 and ex1-2 showed CD52 knockout in the CAR19+ population (3.4% and 4.6% respectively). C. Guides ex1-1 and ex1-2 were retested in primary human T cells electroporated with 100µg/ml SpCas9 mRNA. Analysis of protein knockout revealed superior knockout using the ex1-2 guide compared to the ex1-1 guide sequence. Terminal-CRISPR-CAR19 vector expressing the CD52 ex1-2 guide is referred to as TCD52-CAR19. D. Next generation sequencing (NGS) across the CD52 locus in cells transduced with TCD52-CAR19 vector and electroporated with 100µg/ml SpCas9 mRNA. NGS output was analysed by the CRISPResso online tool. Genomic disruption at the CD52 ex1-2 sgRNA target site was consistent with protein knockout (protein knockout= 35.5%, reads containing NHEJ = 29.6%).
3.5 Chapter summary

This chapter of work detailed the development and optimisation of a novel lentiviral vector platform referred to as terminal-CRISPR, coupling CRISPR sgRNA expression with and transgene expression. This configuration contains a sgRNA expression cassette embedded within the ΔU3 region of the 3’LTR. PCR amplification of the proviral LTRs confirmed duplication of the sgRNA expression cassette to the 5’LTR post reverse transcription. Additionally, insertion with the LTR did not appear to impact viral titers or the ability of these vectors to effectively transduce primary T cells. Furthermore, delivery of SpCas9 mRNA was optimised appearing most effective between days 3 and 4 post transduction at 100µg/ml. Additionally, the transient nature of this delivery strategy was confirmed by western blot, where SpCas9 protein fell below the level of detection by day 7 post electroporation.

Work performed by other researcher’s demonstrated compatibility of the TT-CAR19 vector with therapeutic manufacturing protocols using a CliniMACS Prodigy device. Further molecular characterisation of these cells confirmed high levels of on-target TRAC disruption, with undetectable levels of off-target SpCas9 scission effects.

These cell products were then validated in an in vivo humanised murine NOD/SCID model. In this setting CRISPR modified T cells demonstrated a superior anti-tumour effect compared to their unedited counter parts, indicating that CRISPR editing of the TRAC locus has not impeded the ability of these cells to generate a cytotoxic effect.

Additionally, protospacer sequences were validated for both B2M and CD52 loci to generate either immunologically stealthy cells (B2M⁻), or chemotherapy resistance cells (CD52⁻).
Chapter 4  Generation of minimal U6 pol III promoter, supports development of multiplex terminal-CRISPR lentiviral vector

In the eukaryotic system, three RNA polymerase (Pol) enzymes are responsible for transcription. These three enzymes (Pol I, II, and III), have distinct functions, generating altogether different classes of transcript. In general Pol I transcripts are large non-coding ribosomal RNA (rRNA), while Pol II transcripts are coding mRNA, and Pol III transcripts are essential short non-coding RNA molecules. The terminal-CRISPR configuration described in Chapter 3, incorporates promoters for both Pol II and Pol III, permitting transgene and sgRNA expression respectively.

Pol III promoters can be further subdivided into three unique types. Promoter types 1 and 2 are referred to as gene internal, as the transcribed sequence, rests between the promoter elements. Transcripts from type 1 and 2 Pol III promoters are involved in essential cell functions, including 5S ribosomal RNA (rRNA) required for ribosome structural stability (type 1), and tRNA for AA transfer during translation (type 2) (Figure 4-1). Type 3 promoters are referred to as gene external, where transcribed sequences are downstream of the promoter elements. A variety of essential RNA molecules are transcribed from these promoters, including the U6 small nuclear RNA (snRNA) component of the spliceosome (Figure 4-1) [229].
Figure 4.1 RNA polymerase III promoter types and example transcript function:

There are 3 types of RNA polymerase III (RNA Pol III) promoters whose transcripts are all involved in protein translation. These transcripts are generally short non-coding RNA (<200bp). Type 1 transcripts include small nucleolar RNAs (snoRNAs) that guide chemical modification of other such transcripts called 5S ribosomal RNA (5S rRNA), which in turn plays a functional and structural role in the 60S subunit of the ribosome. Type II RNA Pol III promoters transcribe the transfer RNA (tRNA), required for amino acid transfer during mRNA translation. Type III RNA Pol III promoters transcribe the small nuclear RNAs (snRNAs), which play a vital role in RNA processing reactions, including mRNA splicing.

Type III, Pol III promoters have been utilised for the expression of microRNAs, interference RNAs or CRISPR sgRNA in biomedical applications [7, 219, 230-232] (Appendix A). These promoters show several architectural hallmarks, including a proximal sequence element (PSE), and a TATA box upstream of a distal sequence element (DSE) comprising an octamer (OCT) and SPH sequence [229, 233, 234]. This structure is similar to that of Pol II promoters, with recent evidence demonstrating
that three of the commonly used Pol III promoters (7SK, H1 and U6) are able to recruit Pol II enzyme [235].

Functionally, the PSE attracts snRNA activating protein complex (SNAPc) binding, while the transcription factor Oct-1 binds the highly conserved OCT sequence in the DSE [234]. Protein-protein interactions between DSE bound Oct-1 and PSE bound SNAPc are required for efficient transcriptional activity. Despite this need for interaction between DSE and PSE bound transcription factors, the elements are normally separated by ~150bp spacer sequence [234] (Figure 4-2 A). Although nucleosomes are normally associated with transcriptional repression, the current model for transcriptional activity suggests that the conserved ~150bp spacer sequence is looped around a positional nucleosome, bringing DSE and PSE associated transcription factors into juxtaposition. This DNA loop model would thereby allow cooperative binding and transcriptional activity [236] (Figure 4-2 B).

In contrast to this model, Myslinki and colleagues (2001) observed that the H1 promoter has a particularly compact structure where the DSE and PSE are adjacent to one another (Figure 4-2 C i). They speculated that this compact structure allowed for cooperative binding between DSE and PSE bound transcription factors, without the need for an intervening nucleosome. They further demonstrated the functionality of a minimal H1 promoter (mini-H1) reducing the size by >60% with no notable effect on transcriptional activity [237] (Figure 4-2 C ii).
Figure 4-2 RNA polymerase III, type III promoter structure:

A. Schematic representation of the DNA layout from a generic type III, RNA Pol III promoter. These gene external promoters consist of a distal sequence element (DSE, red), separated via a ~150bp spacer sequence from a proximal sequence element (PSE, blue), which is in proximity to a TATA box (black). B. DNA loop model of transcriptional activity. A spacer DNA sequence is wrapped around a positional nucleosome (grey) bringing DSE and PSE bound transcription factors into juxtaposition and allowing cooperative binding (red) between DSE bound Oct-1 (purple) and PSE bound snRNA activating protein complex (SNAPc). C. The type III H1 RNA Pol III promoter has a naturally compact DNA structure. i. The wild type H1 promoter layout with DSE and PSE in direct proximity, followed by a 176bp spacer sequence. ii. Minimal H1 promoter, devoid of spacer DNA sequence.

Of all type III promoters, U6 has been most widely characterised [234, 238, 239]. The human U6 promoter is reported to have a more classical structure than the H1 promoter, where the DSE and PSE are separated by 147bp spacer sequence looped around a positional nucleosome [238, 239]. Mutational analysis has demonstrated
increased transcriptional activity from both the human 7SK and U6 promoters after deletion of this spacer region in the context of naked DNA. However, the spacer sequence between DSE and PSE has been considered critical for co-localisation and interaction of associated transcription factors in the presence of chromatin [238, 239].

In this chapter I utilise the TT-CAR19 vector described in the previous chapter [219](Appendix A), to explore the sequence and architectural anatomy of the hU6 promoter in more detail. These observations were then applied to the optimisation of a multiplexed terminal-CRISPR vector, allowing simultaneous KO of either B2M/TRAC or CD52/TRAC.

4.1 Human U6 demonstrates redundancy of positional nucleosome spacer sequence

Deletions were made in the wild type U6 promoter sequence (Figure 4-3 A i, Supplementary figure 1 A) by site directed mutagenesis (SDM), to produce ΔU6 variants devoid of nucleosomal spacer sequence (Figure 4-3 A ii, Supplementary figure 1 B) or further devoid of PSE and/or DSE promoter elements (Figure 4-3 A iii - v, Supplementary figure 1 C - E). Functionality of the ΔU6 promoter mutants was assessed through expression of a TRAC targeting sgRNA sequence within the TT-CAR19 vector, validated in the previous chapter. Lentiviral vectors were produced for TT-CAR19 constructs incorporating each of the ΔU6 mutant sequences. High viral titers of >1x10⁸ transducing units/ml were attained, which were comparable amongst constructs. The lentiviral vectors generated were next used to transduced frozen PBMCs from healthy blood donors at MOI 5, 24hrs post activation (biological replicates, n=4).

Transduction efficiency was confirmed 4 days post activation, prior to electroporation with SpCas9 mRNA (50μg/ml). Disruption of TCRαβ by flow cytometry was compared between wtU6 and ΔU6 variants at day 11 post activation (Figure 4-3 B). Removal of either PSE, DSE or both resulted in a significant loss of CRISPR activity compared to the wtU6 promoter (p< 0.0001) suggesting lack of TRAC
sgRNA transcription, whereas the ΔU6 variant devoid of the nucleosomal spacer region showed no significant difference compared to wtU6 promoter (p= 0.8937). The ΔU6 mutant devoid of nucleosomal spacer sequence, reduced the size of the promoter by >50% (249bp to 111bp) and is hereon in referred to as a minimal U6 promoter (mini-U6). Similar to the mini-H1 presented by Myslinki and colleagues (2001), removal of this spacer sequence from the U6 promoter brings the DSE and PSE into direct contact and could allow protein-protein interactions between their bound transcription factors (Figure 4-3 C) [237]. These effects were also confirmed in HEK 293T cells, where the mini-U6 B2M sgRNA expressing cassette was delivered as integrating lentiviral vector (Supplementary figure 2) indicating activity of this mini-U6 in an alternative cell lineage.

Figure 4-3 Development of minimal U6 promoter devoid of nucleosome spacer sequence:

A. Illustration of wild type U6 (i), and minimal U6 devoid of spacer sequence (ii). Control U6 variants were generated devoid of both essential DSE and PSE promoter
elements (iii), or deletion of just the DSE (iv), or deletion of just the PSE (v). B. TCRαβ knockout across four primary T cell donors transduced with TT-CAR19 vectors containing U6 variants and electroporated with 50µg/ml SpCas9 mRNA. No significant difference was observed in knockout efficiency when using wild type or minimal U6 promoters (p = 0.8937), while variants devoid of either DSE or PSE promoter elements showed significantly reduced knockout compared to the wild type control (p < 0.0001). Line represents mean of the group, with errors delineating 1x SEM. Comparison between groups was carried out using a one-way ANOVA with Tukey’s post hoc test. C. Schematic representation of minimal U6 promoter, with the DSE and PSE placed in close proximity. Promoter size is reduced by >50% (111bp).

4.2 Addition of a second tandem CRISPR expression cassette allow multiplex genome editing

Section 3.4 provides proof of principle that the terminal-CRISPR platform can couple CAR19 transgene expression and CRISPR mediated knockout of both B2M, and CD52 independently. Simultaneous disruption of TRAC/CD52, or TRAC/B2M have proven advantageous [143, 208]. Both studies show prevention of GVHD effects from allogeneic CAR19 T cells through knockout of the TRAC locus, while additionally cells are protected from host versus graft effects either via removal of surface antigen CD52, target of lymphodepleting chemotherapy Alemtuzumab, or by removal of HLA class I cell surface expression by B2M knockout. This section will therefore detail steps taken to multiplex the terminal-CRISPR configuration, to allow for disruption of multiple genes.

In the first instance a second sgRNA expression cassette was cloned in tandem to the first within the ΔU3 region of the 3’LTR. Due to the potential for molecular interference caused by placing two U6 promoters in tandem [240], the commonly used human H1 pol III promoter was instead incorporated into the second cassette. Here the H1 promoter is used to produce short RNA transcripts in the form of sgRNA similar to the U6 promoter. However, unlike the U6 which has +1G nucleotide as its favoured putative transcription start site, the H1 pol III promoter utilises a +1A
nucleotide. Therefore guide sequences expressed from this CRISPR expression cassettes were designed to include a +1A nucleotide at their 5’ end [223].

This multiplexed terminal-CRISPR vector was used to generate immunologically stealthy CAR19 T cells with reduced alloreactive potential. To this end the B2M ex1-1 guide sequence was placed under the transcriptional activity of the U6 promoter, while the TRAC guide sequence was expressed by the H1 promoter, coupled to expression of the CAR19 transgene (U6B2M<sub>cr1</sub>&gt;H1TRAC<sub>cr1</sub>&gt;CAR19) (Figure 4-4 A). Primary human T cells were transduced at MOI 5 with the multiplexed terminal-CRISPR vector, and electroporated after 72hrs with 100µg/ml SpCas9 mRNA. Cells were analysed by flow cytometry seven days post electroporation for HLA class I and TCRαβ expression (Figure 4-4 B). Although coupled CRISPR-mediated knockout of both HLA class I and TCRαβ was observed, there was an evident skew favouring B2M disruption (75% and 25% respectively grated on CAR19<sup>+</sup>). A TT-CAR19 vector control, however, confirmed that high levels of TRAC disruption were achievable (81.6% gated on CAR19<sup>+</sup>). The effect of skewed knockout efficiencies of the two targets led to, reduced double knockout in the population (5.8%), thereby limiting the final cell yield. This was rescued through magnetic bead mediated depletion of TCRαβ<sup>+</sup>/ HLA class I<sup>+</sup> cells rendering a highly homogenous population, showing 97% CAR19 expression with <1% carriage of TCRαβ<sup>+</sup>/ HLA class I<sup>+</sup> cells (Figure 4-4 C).

Both sgRNA expression cassettes contain a repeated 76bp CRISPR scaffold sequence (cr1), required for RNP complex formation. Such repeat sequences could lead to intramolecular recombination during reverse transcription of the lentiviral genome [241] resulting in loss of repeat sequences within the ΔU3 region of the proviral LTRs [221]. These recombination events may account for the observed skew between B2M and TRAC knockout within our multiplexed terminal-CRISPR vector. To this end the proviral 3’LTR was amplified by PCR and gel visualised, to determine the molecular integrity of the sgRNA expression cassettes (Figure 4-4 D).

Control reactions were performed with unmodified T cells, as well as TT-CAR19, and pCCL-CAR19 transduced cells. As expected no DNA band was seen in the unmodified negative control, whereas pCCL-CAR19 and TT-CAR19 transduced cells showed bands
of ~400bp and 750bp respectively, corresponding to the correct sized insertion within the LTR. Amplification of the proviral 3’LTR from cells transduced with the U6B2M<sub>cr1</sub>-H1TRAC<sub>cr1</sub>-CAR19 vector revealed three distinct bands, of ~1100bp, 1000bp, and 750bp. Analysis of band density using Image J software suggested that only 54% of the DNA constituted the predicted band of ~1100bp. Gel extraction and Sanger sequencing of the ~1100bp predicted band confirmed the presence of both sgRNA expression cassettes (Figure 4-4 E). However, similar analysis of the ~750bp band revealed sequence match to the B2M sgRNA CRISPR cassette alone, while the second cassette expressing the TRAC sgRNA was not present (Figure 4-4 F).
Figure 4.4 Multiplex terminal-CRISPR configuration containing tandem sgRNA expression cassettes:

A. Illustration of multiplex terminal-CRISPR plasmid DNA configuration expressing CAR19 from an internal hPGK RNA Pol II promoter, and two sgRNA sequences from tandem sgRNA expression cassettes. These sgRNA expression cassettes transcribe the B2M ex1-1 guide from a wild type U6 promoter, or the TRAC specific sgRNA from the wild type H1 promoter (U6B2M cr1 > H1TRAC cr1 > CAR19). Protospacer sequences contained a +1G or +1A nucleotide transcript start site for U6 and H1 RNA Pol III promoters respectively, with both cassettes containing the cr1 scaffold sequences.

B. Flow cytometry base analysis of primary human T cells transduced with
U6B2M_{cr1}>H1TRAC_{cr1}>CAR19 and electroporated with 100µg/ml SpCas9 mRNA. HLA class I staining is used as a surrogate for B2M knockout, with CD3 staining used to measure TRAC disruption. Both B2M and TRAC knockout is coupled to CAR19 expression with double knockout reaching 5.8% of the total cell population (red box). C. Simultaneous magnetic bead mediated depletion of TCRαβ, and HLA class I expressing cells allows stringent purification of this double knockout population (red box), <1% carriage. D. PCR amplification of the 3’ proviral LTR spanning the woodchuck hepatitis virus posttranscriptional regulatory element (WPRE), to the unique 5’ RNA (U5) sequence. Lane 1: unmodified cells (blank control), lane 2: TT-CAR19 transduced cells, lane 3: pCCL-CAR19 transduced cells, lane 4: U6B2M_{cr1}>H1TRAC_{cr1}>CAR19 transduced cells. E. Sanger sequencing data from the top band in lane 4 (~1137bp) aligned to reference sequence. F. Alignment of Sanger sequencing data from the bottom band in lane 4 (~750bp). Arrows represent the direction of the sequencing reads with red sections indicating gaps in the sequence alignment. All sequencing alignments were carried out in SnapGene software.

NGS analysis of the proviral LTRs, performed by Dr Gkazi was used to quantify the occurrence of the molecular skipping events. The results showed the presence of large deletions (>300bp) in the 3’ and 5’ proviral LTRs, which occurred in 31% and 18% of reads respectively [242] (Appendix D).

### 4.3 Incorporation of alternative sgRNA scaffold sequence reduced recombination events during reverse transcription

In order to reduce large stretches of sequence homology between sgRNA expression cassettes Adamson and colleagues (2016) engineered two modified sgRNA scaffold sequences (cr2 and cr3) [243]. These share no more than 20bp of continuous sequence with the original scaffold (cr1) described by Gilbert et al., (2014), while retaining ~80% sequence homology [244]. The cr2 scaffold sequence, contains a 5bp extension in the upper stem and 1bp extension in the second hairpin, as well as 7 additional nucleotide substitutions from the cr1 sequence (Figure 4-5 A i, and ii). In order to assay the efficiency of the alternative cr2 scaffold it was cloned into the TT-CAR19 vector (TT_{cr2}-CAR19) by SDM. This vector was used to transduce primary
human T cells from a healthy donor, before subsequent electroporation with SpCas9 mRNA (100µg/ml). Flow cytometry staining of TCRαβ showed high levels of TRAC knockout (94.6% gated on CAR19+) confirming the efficiency of the alternative cr2 scaffold.

Although multiple biological replicates would be required to elucidate any potential difference between TRAC knockout when using cr1 or cr2 scaffolds, the above data indicated that sgRNA containing the cr2 scaffold was able to form functional RNP complexes capable of high levels of TRAC disruption. This observation promoted the replacement the cr1 scaffold sequence upstream of the TRAC protospacer in the U6B2Mcr1>H1TRACcr1>CAR19 vector (U6B2Mcr1>H1TRACcr2>CAR19) (**Figure 4-5 B**). In order to determine if inclusion of alternative cr2 scaffold sequence reduces the occurrence of large deletions within the proviral sgRNA expression cassettes, U6B2Mcr1>H1TRACcr2>CAR19 vector was tested in healthy primary T cells. In this setting primary PBMCs were activated, 24hr prior to transduction at MOI 5. Transduction efficiency was measured four days post activation by flow cytometry showing 30% CAR19 expression, and cells were then electroporation with 100µg/ml SpCas9 mRNA. CRISPR mediated knockout of both B2M and TRAC, was measured by HLA class-I and CD3 staining respectively. At day 11 post activation SpCas9 treated cells were 17.8% HLA class I (59% gated on CAR19+) and 21.7% CD3- (69% gated on CAR19+), resulting in a double knockout population of 16.7% (56% gated on CAR19+) (**Figure 4-5 C**). Importantly, the knockout of B2M and TRAC loci no longer appeared skewed, and NGS analysis of the proviral LTRs, confirmed a greatly reduced frequency of both large and small deletions within the proviral LTRs (6% and 3% in both the 5’ and 3’ LTR respectively) [242](**Appendix D**).
Figure 4-5 Multiplex terminal-CRISPR vector configuration containing alternative scaffold sequences:

A. Diagram of sgRNA sequences, highlighting structural components of the 80bp cr1 scaffold (i) and noting the changes in the 92bp alternative cr2 scaffold (ii).

B. Plasmid DNA configuration of a multiplex terminal-CRISPR-CAR19 vector, expressing both B2Mex1-1 sgRNA with a cr1 scaffold and a TRAC sgRNA with a cr2 scaffold (U6B2M<sub>cr1</sub>H1TRAC<sub>cr2</sub> CAR19).

C. Protein knockout assessed by flow cytometry, from primary human T cells transduced with U6B2M<sub>cr1</sub>H1TRAC<sub>cr2</sub>CAR19 vector and electroporated with 100µg/ml SpCas9 mRNA. Staining for HLA class I and CD3 were used a surrogate markers for B2M and TRAC knockout respectively. Knockouts of
both targets no longer appears skewed, with a double negative population accounting for 16.7% of the total cell population (red box).

4.4 Application of minimal pol III promoters support efficient coupled multiplex genome editing

In order to minimise capacity constraints within the lentiviral LTRs [221], I reduced the tandem cassette size by substituting the full length Pol III promoters for the mini-U6, developed above, and the mini-H1 (mU6B2M>mH1TRACcr2) (Figure 4-6 A) [237, 242](Appendix D). Incorporation of minimal promoters reduced the size of the tandem sgRNA expression cassettes from 739bp to 426bp equating to >40% reduction. Functionality of the minimal multiplex terminal configuration was assessed in primary human T cells transduced with the vectors and subsequently electroporated with SpCas9 mRNA (100µg/ml). PCR amplification of the proviral LTRs confirmed the reduced size of the sgRNA expression cassettes with products corresponding to the 426bp cassette obtained. Similar to the data above, HLA class-I and CD3 antibody staining was used to observe both B2M and TRAC disruption respectively. At day 11 post activation (day 7 post electroporation), SpCas9 treated cells showed high levels of transduction (45% CAR19⁺) with knockout of both B2M and TRAC (66% and 73% gated on CAR19⁺) (Figure 4-6 B). Similar to the data described above using the cr2 scaffold, there did not appear to be a skew between B2M and TRAC knockout. This was further supported by NGS analysis of the proviral LTRs, where only low levels of large deletions (<5%) were detected [242](Appendix D).
A. Schematic of a multiplex terminal plasmid DNA configuration, expressing both B2Mex1-1 sgRNA with a cr1 scaffold and a TRAC sgRNA with a cr2 scaffold from minimal U6 and minimal H1 RNA Pol III promoters respectively ($mU6B2M_{cr1}>mH1TRAC_{cr2}>CAR19$). Use of minimal promoters reduce the size of the tandem sgRNA expression cassettes from 745bp, to 426bp. B. Flow cytometry-based assessment of protein knockout, from primary human T cells transduced with $mU6B2M_{cr1}>mH1TRAC_{cr2}>CAR19$ vector and electroporated with 100µg/ml SpCas9 mRNA. Staining for HLA class I and CD3 were used a surrogate markers for B2M and TRAC knockout respectively. Efficient knockout of both targets with a double negative population accounting for 27.5% of the total cell population (red box).

This chapter has thus far described the generation of three multiplex terminal-CRISPR configurations, all containing two tandem sgRNA expression cassettes within the ΔU3 regions of the 3’LTR (U6B2M$_{cr1}$H1TRAC$_{cr1}$, U6B2M$_{cr1}$H1TRAC$_{cr2}$, and mU6B2M$_{cr1}$mH1TRAC$_{cr2}$). These were then compared in multiple primary human T cells donors (biological replicates, n=4). In this instance frozen PBMCs from four
donors were thawed, and rested overnight before activation. Cells were subsequently transduced at MOI 5 with the different multiplex configurations, before electroporation with SpCas9 mRNA (100µg/ml). Flow cytometry based phenotyping indicated similar levels of B2M disruption in the CAR19 expressing population (U6B2Mcr1>H1TRACcr1 = 66% ± 9%, U6B2Mcr1>H1TRACcr2 = 63% ± 11%, and mU6B2Mcr1>mH1TRACcr2 = 74% ± 3%), showing no significant difference (p=0.3753). Similarly, TRAC knockout between U6B2Mcr1>H1TRACcr2 (81% ± 7%), and mU6B2Mcr1>mH1TRACcr2 (79% ± 3%) configurations showed no significant difference (p=0.8557). However, as observed previously in this chapter TRAC knockout was significantly reduced in cells transduced with the U6B2Mcr1>H1TRACcr1 (66% ± 4%) configuration compared to both U6B2Mcr1>H1TRACcr2 (p=0.0102) and mU6B2Mcr1>mH1TRACcr2 (p=0.0230) configurations (Figure 4-7 A). Additionally, this data demonstrated the ability to obtained comparable levels of double knockout using the minimal Pol III promoters compared to their full length counter parts (62% ± 3% and 56% ± 11% respectively, gated on CAR19+) (Figure 4-7 B). To confirm that the results obtained when delivering SpCas9 mRNA at 100µg/ml were not due to saturation of the nuclease, electroporations were repeated at a range of concentrations of SpCas9 mRNA (25, 50, 75µg/ml). These results showed no significant different between U6B2Mcr1>H1TRACcr2 and mU6B2Mcr1>mH1TRACcr2 configurations at the tested SpCas9 concentrations, supporting the observation that the minimal promoters function comparably to the wild type counterparts (Supplementary figure 3).

The multiplexed terminal-CRISPR configuration offers an important advantage over uncoupled sgRNA delivery, in allowing for highly stringent purification for two cell surface molecules using magnetic bead based depletion of residual non-disrupted populations while maintaining coupling to the expressed transgene. Thus in the primary human T cells transduced with the mU6B2Mcr1>mH1TRACcr2>CAR19 lentiviral vector, it was possible to carry out magnetic depletion of residual HLA class-I+ and TCRαβ+ cells using biotin conjugated anti-HLA and anti-TCRαβ antibodies, in combination with anti-biotin beads (Figure 4-7 C). Flow cytometry based phenotyping confirmed enrichment of the B2M/TRAC- population from 24.7% ± 4.6% to 99.4% ±
0.5% in the total CD45+ population (Figure 4-7 D). Sanger sequencing of the on-target TRAC and B2M loci corroborated the frequency of genomic aberrations at the molecular level, by TIDE analysis (Figure 4-7 E). As demonstrated in Chapter 3, the coupled nature of the vector ensured enrichment of the CAR19 transgene, from 48.3% ± 6% CAR19+ pre-depletion increasing to 95% ± 2.2% post depletion (Figure 4-7 F).

Figure 4-7 Generation of high homogenous HLA I-/TCRαβ⁻/CAR19+ populations with multiplex terminal configuration, containing minimal RNA Pol III promoters:
A. Comparison of the three multiplex terminal-CRISPR configurations, U6B2M<sub>cr1</sub>-H1TRAC<sub>cr1</sub>-CAR19, U6B2M<sub>cr1</sub>-H1TRAC<sub>cr2</sub>-CAR19, and mU6B2M<sub>cr1</sub>-mH1TRAC<sub>cr2</sub>-CAR19, in n=4 primary T cells donors electroporated with 50µg/ml SpCas9 mRNA. Flow cytometry based analysis at day 7 post electroporation with 50µg/ml SpCas9. HLA class I staining and TCRαβ staining, were used as surrogates for B2M and TRAC in turn. Data presented here was gated on CD45<sup>+</sup>/CAR19<sup>+</sup>. One-way Anova with Tukey’s multiple comparison test displayed no significant difference in B2M knockout (p=0.3753) across these constructs. TCRαβ knockout was significantly higher in both constructs containing the cr2 scaffold sequence compared to U6B2M<sub>cr1</sub>-H1TRAC<sub>cr1</sub>-CAR19 (U6B2M<sub>cr1</sub>-H1TRAC<sub>cr2</sub>-CAR19, *p = 0.0102, and mU6B2M<sub>cr1</sub>-mH1TRAC<sub>cr2</sub>-CAR19 *p = 0.0230). B. Histogram detailing the composition of double, single, and unedited cells in the CD45<sup>+</sup>/CAR19<sup>+</sup> population across n=4 donors treated with the three vector configurations. C. Representative FACS plots of pre-depletion and post HLA Class I and TCRαβ combined microbead depletion of mU6B2M<sub>cr1</sub>-mH1TRAC<sub>cr2</sub>-CAR19 treated cells. HLA Class I and TCRαβ knockout shown in total CD45<sup>+</sup> population. D. Histogram of detailing the makeup of the total CD45<sup>+</sup> population split into double, single, and unmodified cells, pre and post simultaneous HLA Class I and TCRαβ depletion across n=4 donors treated with mU6B2M<sub>cr1</sub>-mH1TRAC<sub>cr2</sub>-CAR19. E. Sanger sequencing data across the predicted cuts sites within the B2M and TRAC loci. Analysis by TIDE presenting the frequency of aberrant sequences at these loci, pre-depletion and post-depletion from n=4 donors treated with mU6B2M<sub>cr1</sub>-mH1TRAC<sub>cr2</sub>-CAR19 vector. Error bars represent SEM. F. Histogram of CAR19 expression in total CD45<sup>+</sup> population pre and post simultaneous HLA Class I and TCRαβ depletion in n=4 donors treated with mU6B2M<sub>cr1</sub>-mH1TRAC<sub>cr2</sub>-CAR19.

As detailed above, simultaneous disruption of CD52 and TRAC loci by TALEN genome editing has also proven advantageous in the field of universal CAR T cells [210]. Therefore to demonstrate the versatility of the multiplex terminal-CRISPR configuration, guides targeting CD52 (ex1-2) and TRAC were cloned under the transcriptional control of the mini-U6 and mini-H1 promoters respectively (mU6CD52<sub>cr1</sub>-mH1TRAC<sub>cr2</sub>-CAR19). Multiplex terminal-CRISPR lentiviral vector
targeting CD52 and TRAC was produced and used to transduce primary human T cells at MOI 5, prior to electroporation with SpCas9 mRNA at 100µg/ml (biological replicates, n=3). Disruption of both CD52 and TRAC loci, gave rise to a CD52⁺/TCRαβ⁺ population (28% - 45%). Magnetic bead mediated depletion of residual TCRαβ⁺ cells, increased the proportion of the double knockout population in the cell cultures to 71% - 83%, with <3% carriage of residual CD3 expression (Supplementary figure 4 A). The vectors ability to couple sgRNA and transgene expression results in transgene enrichment, after TCRαβ depletion (from 43% – 63% to 92% – 95% CAR19⁺) (Supplementary figure 4 B).

4.5 Chapter summary

In summary, this chapter uses the TT-CAR19 configuration to explore the architecture of the type III, Pol III U6 promoter leading to the development of a mini-U6 promoter devoid of nucleosomal spacer sequence. Furthermore, this mini-U6 promoter was shown to be as effective as its full length counterpart in both terminal-CRISPR vectors containing single and multiplexed sgRNA expression cassettes. Optimisation of the multiplex terminal-CRISPR vectors indicated the need for an alternative sgRNA scaffold (cr2), in order to reduce sequence homology between the two sgRNA CRISPR expression cassettes. These observations led to an optimised multiplex terminal-CRISPR vector containing tandem sgRNA expression cassette, with alternative scaffold sequences (cr1/2), as well as minimal Pol III promoters.

The optimised multiplex terminal-CRISPR vector, was able to demonstrate coupled expression of both B2M/ TRAC, and CD52/ TRAC guides to a CAR19 transgene, with further magnetic bead based depletions leading to highly homogenous cell products. Early therapeutic applications are envisaged in the context of lentiviral engineered T cells, but use of minimal Pol III promoter may also be helpful for alternative vector systems such as AAV where cargo size may be limiting.
Chapter 5  Optimisation of coupled and uncoupled cytidine base editing for primary human T cell engineering

5.1  Third generation cytidine deaminase base editor (BE3)

As described above (Section 1.6) base editing technology offers the ability to generate seamless base conversions to achieve site specific genomic correction, without the need to generate dsDNA breaks. These base editors can generally be split into two groups, cytidine deaminase base editors (CBE’s) or adenosine deaminase base editors (ABE’s). CBE’s can be employed to execute C>T conversions (G>A when the C is on the antisense DNA strand), while ABE’s facilitate A>G conversions (T>C when the A is on the DNA antisense strand). The combination of these tools enables the conversion of all four bases and therefore holds significant potential for the correction of pathogenic SNPs (>50% correction of pathogenic SNPs in the Clinvar database) [97]. Additionally, due to the reduced frequency of dsDNA breaks generated, this technology presents a safer option for multiplexed genome editing reducing the risk of large chromosomal translocation events between the edited loci.

This chapter will focus on the application and optimisation on a third generation, human codon optimised CBE (coBE3), for the simultaneous disruption of multiple loci in primary human T cells. A number of gene disruption strategies utilising base editing technologies have been described including, removal of start site [114], mRNA splice sites [115, 116, 123], as well as insertion of pre-mature stop codons [113, 245]. For modelling purposes the second generation 4-1BB CAR19 transgene will be used in conjunction with edits that allow the generation of universal CAR T cells capable of overcoming HLA barriers.

5.2  Modelling coBE3 editing in a Jurkat E6.1 T cell leukaemia line

Briefly the structure of the BE3 fusion protein, used in this chapter is made up of a rAPOBEC1 joined via a 16AA linker to a D10A nCas9, which is in turn is fused to a UGI molecule via a 4AA linker, followed by an SV40 NLS separated via another 4AA linker (Figure 5-1 A). These components all play a vital role in the function of this fusion
protein, where the rAPOBEC1 elicits C>U changes (U’s in the DNA are read as T’s during DNA replication), the D10A nCas9 nicks the unedited DNA strand promoting correction from the uncut edited strand, the UGI prevents repair of U bases by the endogenous Uracil-DNA glycosylase (UDG) enzyme, and the SV40-NLS permits trafficking of the fusion protein to the nucleus. Once situated at its target site BE3 is able to perform these C>T changes most effectively within a 5bp editing window (4-8 nucleotides distal to the PAM).

In the first instance the DNA sequence for BE3 was synthesised by GeneART (ThermoFisher Scientific), using proprietary human codon optimisation, as this has been shown to increase mRNA expression and stability, leading to increased protein production within mammalian systems [99, 246]. The codon optimised BE3 (coBE3) sequence was then cloned into a second generation lentiviral vector plasmid by PCR based In-Fusion cloning. This allowed seamless insertion of the coBE3 transgene downstream of an internal spleen focus forming virus (SFFV) RNA polymerase II promoter and a translation initiation Kozak consensus sequence (GCCACC). Furthermore, an eGFP sequence separated by an internal ribosome entry site (IRSE) in-frame with the coBE3 transgene, provided an ergonomic way of tracking lentiviral vector expression (Figure 5-1 B). A second generation lentiviral vector expressing both coBE3 and eGFP was produced and titrated. This vector was subsequently used to transduce a CD3⁺ Jurkat E6.1 cell line at MOI 5, generating a CD3⁺eGFP⁺ population which was then bulk sorted by flow cytometry. Sorted CD3⁺eGFP⁺ cells stably expressing coBE3 could then be used as a model system for validating sgRNA sequences (Figure 5-1 C). In order to confirm that BE3 was able to generate C>T change, these cells were further transduced with our TT-CAR19 vector expressing a TRAC guide. This guide contains two C residues within the editing window that would promote a silent mutation and R>K AA transition. PCR amplification of the target site and subsequent Sanger sequencing confirmed that both these C could be edited.

Guides were then designed to disrupt the expression of the TCRαβ-CD3 complex, targeting either TRAC or conserved regions of TRBC 1 and 2. The online tool Benchling (https://www.benchling.com/) was used to select guides that contained C bases within this 5bp window, and provide in silico predicted on-target and off-target
scores. Benchling was able to predict guides that following base conversion would generate a premature stop codon, leading to formation of a truncated protein product. However, at the time of writing Benchling is unable to predict C>T changes that impact mRNA splicing. Therefore, the online tool Human Splicing Finder v3.1 (http://www.umd.be/HSF/) was used to detect the naturally occurring mRNA splice acceptor and donor sites within the target gene, and provide predicted scores for splice site disruption. Details of all BE3 compactible guides used in this project, can be found in Table 10. Two guides targeting TRAC and four targeting TRBC were cloned into a terminal-CRISPR-CAR19.

Terminal-CRISPR-CAR19 vectors expressing TRAC and TRBC sgRNA sequences were used to transduce the CD3+/eGFP+ Jurkat E6.1 cells at MOI 5. Flow cytometry based phenotyping was used to analyse eGFP, CD3, and CAR19 expression in these cells to assess TRAC or TRBC disruption. Although, all vectors were able to achieve high levels of CAR19 expression (96.9% ± 0.33 SEM), only the TRAC guide predicted to disrupt the exon 1 splice donor site (TRAC ex1 SD) demonstrated a reduction of CD3 expression (26%, Figure 5-1 D i). Additionally, of the four guides targeting conserved regions of TRBC 1 and 2, only guide TRBC ex1-2 exhibited any measurable reduction in CD3 expression (31.6%, Figure 5-1 D ii).
Figure 5-1 Modelling coBE3 activity in a CD3+ Jurkat cell line:

A. Schema of BE3 protein layout. B. Second generation lentiviral vector configuration expressing coBE3 from an spleen focus forming virus (SFFV) RNA Pol II promoter, and EGFP via an IRES sequence. C. Transduction of CD3+ Jurkat cell line for the constitutive expression of coBE3 (CD3+EGFP+coBE3+ Jurkat model). D. Transduction of coBE3 expressing Jurkat cell line with terminal-CRISPR-CAR19 vectors containing TRAC, and TRBC targeting sgRNA. i. Terminal-CRISPR-CAR19 vector expressing TRAC ex1 SD guide results in 26% TRAC knockout (n=1). ii. Terminal-CRISPR-CAR19 vector expressing TRBC ex1-2 guides promoted 31.6% TRBC knockout (n=1). Both TRAC and TRBC knockout were measured by CD3 staining. Flow cytometry for D i and D ii were performed at different times, and have been gated appropriately on unedited controls.
5.3 Efficient base editing by coBE3 mRNA in primary human T cells

The above modelling of coBE3 validated its ability to disrupt cell surface expression of the TCRαβ-CD3 complex in a Jurkat E6.1 cell line stably expressing coBE3 through disruption of either TRAC or TRBC loci. With this in mind the coBE3 transgene was cloned by In-Fusion into an mRNA synthesis plasmid from Trilink (pmRNA plasmid). This plasmid contains a T7 promoter, capable of recruiting the T7 RNA polymerase to catalyse RNA formation from downstream DNA, with a high fidelity. Additionally, this plasmid contains Trilink’s CleanCap technology that yields mRNA with a more natural Cap 1 post-transcriptional modification typical of multicellular eukaryotic systems. Furthermore, the coBE3 transgene in this vector encodes a poly-A tail increasing stability of mRNA transcripts. The pmRNA plasmid containing coBE3 was fully Sanger sequenced prior to mRNA synthesis.

To authenticate the effects of this coBE3 mRNA, a titration experiment was run in primary human T cells transduced with terminal-CRISPR-CAR19 vector, expressing either TRAC ex1 SD (Figure 5-2 A i) or TRBC ex1-2 (Figure 5-2 A ii) guides. Transduction efficiency for TRAC ex1 SD and TRBC ex1-2 expressing vectors were comparable, 61.5% and 63.7% respectively. Cells were next electroporated with 25, 50, 75, or 100µg/ml coBE3 mRNA and flow cytometry-based phenotyping was used to assess cell surface knockout. Both TRAC ex1 SD and TRBC ex1-2 expressed guides showed a positive correlation between increased coBE3 mRNA concentration and reduced TCRαβ-CD3 cell surface staining, plateauing at 50µg/ml (Figure 5-2 B). However, increased level of disruption was apparent with TRBC ex1-2 guide compared to the TRAC ex1 SD guide (61.7% and 17.4%, gated on CAR19*).

In order to assess any potential difference in genome editing efficacy between SpCas9 and coBE3, terminal-CRISPR-CAR19 vector expressing the TRBC ex1-2 guide (TTRBC-CAR19) was used to transduce primary human T cell donors (n=4). Transduction efficiencies ranged from 39.3% to 71.8% across the four donors (58.5% ± 6.5% SEM), with electroporation of 50µg/ml coBE3 mRNA resulting in 24.2% - 53.5% TRBC knockout (55.2 - 71% gated on CD45*CAR19*). Despite the relatively high knockout efficiency, electroporation of SpCas9 mRNA significantly increased levels of
TRBC knockout to 36.8% – 68.4% (85.4% - 90.9% gated on CD45^CAR19^), as shown in Figure 5-2 C (p= 0.0286).

Characterisation of the coBE3 editing profile at the genomic level was performed by Dr Gkazi, using NGS data ran on a MiSeq, followed by analysis using Freebayes to compile haplotypes within the TRBC ex1-2 protospacer. This data demonstrated the ability of coBE3 to execute the desired substitution and in turn, generate premature stop codons in the TRBC loci, however, editing outside the predicted editing window and non C>T changes were observed limiting the application of this tool for the correction of pathogenic SNPs (Figure 5-2 D).

![Diagram of TRAC and TRBC loci with knockout percentages and NGS reads](image)

Figure 5-2 Efficient TCRαβ knockout in primary T cells by targeted cytidine deamination:

A. Guides targeting the exon 1 splice donor site of TRAC (i.), or targeting a conserved region of TRBC 1/2 (ii.) were tested in primary human T cells. B. Primary human T cells transduced with terminal-CRISPR-CAR19 vectors, expressing either TRBC ex1-2,
or TRAC ex1 SD sgRNA sequences and electroporated with a titration of coBE3 mRNA (25, 50, 75, and 100µg/ml). Percentages reported here are based on flow cytometry analysis gated on CD45⁺/CAR19⁺ cells, with CD3 staining used as a surrogate marker for TRAC and TRBC disruption (n=1). C. Comparison of SpCas9 and coBE3 mRNA in n=4 primary T cells transduced with TTRBC-CAR19 and electroporated with 50µg/ml mRNA. SpCas9 mRNA presented with significantly increased knockout compared to coBE3 mRNA, analysed by Mann-Whitney U test (*p = 0.0286). Plotted values have been attained from flow cytometry data gated on CD45⁺/CAR19⁺.

D. Swimmers plot of NGS data at the on-target TRBC site quantifies C>T edits within the editing window, as well as indicating the occurrence of C>T edits outside the optimal window, and non-C>T base changes.

5.4 Application of coBE3, to CD52 and B2M genome editing

As described above in section 4.2 simultaneous disruption of multiple loci would be key to producing a truly universal T cell product. With this in mind, coBE3-compatible guides were designed for the disruption of both CD52, and B2M. A total of four CD52 guides were tested with the intention of creating premature stop codons, interfere with mRNA splicing, as well as changing the S-P-S AA sequence recognised by the chemotherapeutic monoclonal antibody Alemtuzumab (Figure 5-3 A). In the first instance these guides were tested in a terminal-CRISPR configuration. Both guides predicted to disrupt the exon 2 splice acceptor site (CD52 ex2 SA), and the S-P-S AA sequence (CD52 ex2 S-P-S) had no notable effect on CD52 protein expression in primary human T cells. Guide ex1 ISTOP, predicted to generate a premature stop codon in exon 1 was able to produce modest cell surface knockout (6.9% ± 0.4% gated on CAR19⁺) in primary T cell donors (n=3). However, guide CD52 ex1 SD aimed at removing the exon 1 splice donor site produced considerably higher knockout (35.6% gated on CAR19⁺) when tested in primary T cells (Figure 5-3 B).

With regards to B2M, two guides were tested either targeting the exon 1 splice donor site (B2M ex1 SD) or the splice acceptor site at the beginning of exon 2 (B2M ex2 SA) (Figure 5-3 C). As before, these guides were first cloned into a terminal-CRISPR configuration, before testing in primary human T cells. Of these two guides, only
guide B2M ex1 SD showed any appreciable disruption of HLA class I surface expression (26.1% of the CAR19+ population) (Figure 5-3 D).

**Figure 5.3 Optimisation of CD52 and B2M knockout with coBE3:**

A. Schematic of CD52 gene locus with tested guides marked up. B. Flow cytometry data from primary human T cells transduced with a terminal-CRISPR-CAR19 configuration expressing the CD52 ex1 SD guide and electroporated with 50µg/ml coBE3 mRNA. C. Schema of the B2M locus with both the ex1 SD and ex2 SA guide positions marked up. D. Visualisation by flow cytometry of B2M protein knockout in primary T cells transduced with terminal-CRISPR-CAR19 vector expressing the B2M ex1 SD guide sequence. HLA class I staining was used as a surrogate marker for B2M knockout.

5.5 Coupled multiplexed cytidine deamination using terminal-CRISPR vector

The multiplex terminal-CRISPR vector developed in chapter 4, has proven a robust method for the simultaneous delivery of two sgRNA sequences, and disruption of two loci by transient expression of SpCas9. With this in mind two version of the multiplex terminal-CRISPR-CAR19 configuration expressing either the CD52 ex1 SD guide or the
B2M ex1 SD guide in tandem with a second guide expressing the TRBC ex1-2 guide (mU6CD52_c1>mH1TRBC_c2>CAR19 and mU6B2M_c1>mH1TRBC_c2>CAR19). In order to assay the ability of these configurations to generate a double knockout population, primary T cells were isolated from a healthy donor and transduced with either mU6CD52_c1>mH1TRBC_c2>, or mU6B2M_c1>mH1TRBC_c2> vectors at MOI 5, resulting in 51.9% and 49.7% CAR19 expression respectively. These cells were then electroporated with 50µg/ml of either coBE3 mRNA, or SpCas9 mRNA as a positive control. Consistent with results from the previous chapter, this multiplex terminal-CRISPR configuration permitted high levels of simultaneous knockout, upon delivery of SpCas9 mRNA resulting in TCRαβ/CD52−, and TCRαβ/HLA class I− cell populations (64.1% and 66% respectively ,gated on CD45+CAR19+). Although, electroporation with coBE3 mRNA was able to promote knockout of both targeted loci, there was a marked impact on overall efficiency which was reduced by ~5 – 15 fold (TCRαβ/CD52− = 11%, TCRαβ/HLA class I− = 4%, gated on CD45+CAR19+) (Figure 5-4 A, B).

Figure 5-4 Multiplex terminal-CRISPR-CAR19 vectors for simultaneous disruption of two genomic loci:

A. Diagram of the tandem sgRNA expression cassettes placed within the ΔU3 region of the 3’LTR within the multiplex terminal-CRISPR-CAR19 vector. The CD52 ex1 SD guide is expressed for the mini-U6 promoter, while the TRBC ex1-2 guide is expressed via a mini-H1 promoter (mU6CD52_c1>mH1TRBC_c2>CAR19). Flow cytometry data from primary human T cells transduced with mU6CD52_c1>mH1TRBC_c2>CAR19 vector and electroporated with either SpCas9, or coBE3 mRNA at 50µg/ml. Electroporation...
with SpCas9 mRNA presented with efficient disruption of both \textit{CD52} and \textit{TRBC}, resulting in a double negative population of 64.1\%. Electroporation with coBE3 mRNA appears less effective, reaching only 11.2\% double knockout. B. \textit{layout of the tandem sgRNA expression cassettes in the mU6B2M}_{cr1}>mH1TRBC}_{cr2}>CAR19 configuration, expressing the B2M ex1 SD guide from a mini-U6 promoter, and the TRBC ex1-2 sgRNA from a mini-H1 promoter. Flow analysis from primary T cells transduced with the mU6B2M}_{cr1}>mH1TRBC}_{cr2}>CAR19, and subsequently electroporated with either SpCas9, or coBE3 mRNA at 50µg/ml. Superior protein knockout of both targets was observed with SpCas9 mRNA, compared to coBE3 mRNA resulting in double knockout populations of 66\%, and 3.6\% in turn. Flow cytometry data presented here is gated on the CD45\+/CAR19\+ population. Staining for HLA class I and TCR\(\alpha\beta\) was used as a surrogate for \textit{B2M} and \textit{TRBC1/2} respectively.

It was hypothesised that recovery of knockout efficiencies when using coBE3 with the multiplex terminal-CRISPR configurations may be achieved by titrating the amount of mRNA. To this end primary human T cells were transduced with either mU6CD52}_{cr1}>mH1TRBC}_{cr2}>CAR19 or mU6B2M}_{cr1}>mH1TRBC}_{cr2}>CAR19 vectors at MOI 5, resulting in 50.2\% and 49\% transduction respectively. These cells were subsequently electroporated with 25, 50, 75, or 100µg/ml coBE3 mRNA. Both these vectors demonstrated a positive correlation between coBE3 mRNA and increased knockout of both targets (Figure 5-5 A i, ii). Despite this, relatively modest frequency of double knockout events were detected reaching 18.3\% TCR\(\alpha\beta\)/CD52\- (Figure 5-5 B i), and 11.8\% TCR\(\alpha\beta\)/HLA class I\- (Figure 5-5 B ii) gated on CAR19\+ at 100µg/ml coBE3 mRNA.
Figure 5-5 Titration of coBE3 mRNA in primary human T cells transduced with multiplex terminal-CRISPR-CAR19 vector:

A. Primary human T cells transduced with either (i.) mU6CD52cr1>mH1TRBCcr2>CAR19, or (ii.) mU6B2Mcr1>mH1TRBCcr2>CAR19 vectors prior to electroporation with coBE3 at 25, 50, 75, and 100 µg/ml. Increased knockout of both targets was detected at higher coBE3 concentrations, reflected in the proportion of the double negative population (triangles) (n=1). B. Flow cytometry data presenting protein disruption at the highest coBE3 mRNA concentration (100µg/ml). All values presented here in this figure are gated on the CD45+/CAR19+ cells. Staining for HLA class I and TCRαβ were used as surrogate markers for B2M and TRBC respectively.

5.6 Reduced vector expressed sgRNA editing in the presence of chemically modified sgRNA

An alternative method of multiplexing CRISPR effects involves use of coupled and uncoupled sgRNA. In this setting TCRαβ-CD3 disruption was coupled to the vector through incorporation of the TRBC ex1-2 sgRNA expression cassette within the LTR
(TTRBC-CAR19), with additional edits carried out through delivery of uncoupled chemically modified sgRNA sequences (Figure 5-6 A). In order to optimise this strategy for the production of double negative cells, primary human T cells were transduced at MOI 5 with TTRBC-CAR19 vector, and subsequently electroporated 50µg/ml coBE3 in addition to CD52 ex1 SD sgRNA at 0, 20, 40, 60, and 80µg/ml. This titration demonstrated increasing CD52 cell surface knockout with increasing amounts of CD52 ex1 SD guide, ranging from 46.1% at 20µg to 70.8% at 80µg/ml. In direct contrast to this, TRBC knockout showed decreased disruption of TCRαβ with addition of CD52 ex1 SD guide (42.4% at 0µg to 3.51% at 80µg/ml in the CAR19+ population) (Figure 5-6 B i). Conflicting editing resulted in double knockout peaking at 20µg/ml of chemically synthesised CD52 ex1 SD sgRNA (5.5% gated on CAR19+) (Figure 5-6 B ii).

To confirm that this result was unrelated to the sgRNA sequence, the titration was repeated with CD52 ex1 SD guide being expressed from a terminal-CRISPR configuration (TCDS2SD-CAR19), and chemically synthesised TRBC ex1-2 ISTOP guide being delivered in an uncoupled fashion. Similar to the previous experiment, primary human T cells were transduced with TCDS2SD-CAR19 vector, before electroporation with 50µg/ml coBE3 and TRBC ex1-2 sgRNA (10, 20, 30, 40, and 50µg/ml). Consistent with previous results, knockout achieved from the loose guide (TRBC ex1-2) appeared at higher sgRNA concentrations, while vector expressed guide (CD52 ex1 SD) showed the opposite (Figure 5-6 C i). Similar to previous data double knockout efficiency peaked with the addition of 20µg/ml loose guide, reaching 29.1% of the CAR19+ population (Figure 5-7 C ii). These results imply a competition effect between the sgRNA constitutively expressed from RNA Pol III promoter within the terminal-CRISPR vector and chemically modified loose sgRNA.
Figure 5-6 Multiplex coBE3 editing through coupled and uncoupled deamination:

A. Illustration of multiplex coBE3 editing through combination of vector-expressed sgRNA, and uncoupled loose sgRNA. B. Titration of chemically modified CD52 ex1 SD sgRNA in primary human T cells transduced with TTRBC-CAR19 vector and electroporated with 50µg/ml coBE3 mRNA (n=1). i. competition effect on vector expressed TRBC knockout following addition of loose chemically modified CD52 sgRNA. ii. Flow cytometry plot showing highest proportion of double knockout population at 20µg/ml CD52 ex1 SD sgRNA, reaching 5.5%. C. Titration of chemically modified TRBC ex1-2 sgRNA in primary human T cells transduced with TCD52sd-CAR19 vector, and electroporated with 50µg/ml coBE3 mRNA (n=1). i. inverse correlation seen in vector expressed CD52 knockout when combined with increased concentration of loose TRBC ex1-2 guide. ii. Flow plot depicting highest proportion of...
double knockout observed when using 20µg/ml TRBC ex1-2 sgRNA, reaching 29.1%. All flow cytometry values in this figure are gated on CD45+/CAR19+ cells. TCRαβ staining is used as a surrogate for TRBC.

Results obtained in our lab, suggested that when using SpCas9 mRNA, knockout from the vector expressed guide was unaffected by the addition of loose guide. In this setting n=1 primary T cells were transduced with the TT-CAR19 vector, and electroporated with 100µg/ml SpCas9, as well as 0, 35, or 70µg/ml CD52 sgRNA. Flow cytometry confirmed that TRAC knockout remained unchanged (51.1% at 0µg to 49.7% at 70µg/ml). These results suggest that this competition between vector expressed guide and loose guide are restricted to coBE3 editing.

5.7 Multiplexed uncoupled editing of TRBC and CD52

It has been recently reported that effective multiplex base editing in primary human T cells can be achieved by delivery of BE3 mRNA with multiple loose sgRNAs [123]. In order to determine whether any such competition exists between loose guides, T cells were electroporated with both TRBC ex1-2 and CD52 ex1 SD guides at varying ratios. This was done at either a 1:3 (10:30µg/ml TRBC/CD52 sgRNA), 1:1 (20:20µg/ml TRBC/CD52 sgRNA), or a 3:1 (30:10µg/ml TRBC/CD52 sgRNA) ratios (Figure 5-7 A). There did not appear to be an appreciable difference between the ratios, with no apparent skew between TRBC/ CD52 knockout (1:3= 84.2%, 1:1= 85.1%, and 3:1= 81.4% double knockout) (Figure 5-7 B). These results confirmed the ability of coBE3 mRNA to generate high levels of multiplexed base editing, as well as suggesting that competition between guides does not appear to be in effect when guides are delivered together as chemically modified loose sgRNA.
Figure 5.7 Uncoupled multiplexed base editing:

A. Schema of uncoupled delivery of loose sgRNA co-delivered with coBE3 mRNA. B. Flow cytometry data from primary human T cells electroporated at 50µg/ml coBE3 mRNA, with both TRBC ex1-2 and CD52 ex1 SD sgRNAs at 1:3, 1:1, and 3:1 ratios. TCRαβ staining is used as a surrogate marker for TRBC.

5.8 Modification of guide scaffold sequence to improve sgRNA expression and RNP stability

To ascertain whether sgRNA expression levels from the vector are a limiting factor in the presence of chemically modified sgRNA, modifications can be made to the cr1 scaffold sequence reported to increase expression. Chen and colleagues (2013), demonstrated that increased guides expression from a U6 Pol III promoter could be achieved by making changes to the sgRNA scaffold sequence. They employ a U>A base pair change in the sgRNA scaffold stem-loop, to remove a potential premature Pol III termination signal (UUUU>UUUA). Additionally, they added a 5bp extension to the Cas9-binding hairpin, thereby improving RNP assembly [247]. Dang et al (2015), further characterised this guide scaffold sequence and indicated that changing the stem-loop, U to either a C or a G would improve CRISPR/Cas9 effects
To determine whether these modifications would improve RNP stability in our hands, site directed mutagenesis was used to convert the desired U>C, and extend the Cas9 binding hairpin by 5bp in our TTRBC-CAR19 configuration, now referred to as TTRBC_{C+5}-CAR19.

In order to validate this modified scaffold, primary T cells were transduced with a TTRBC_{C+5}-CAR19 vector and electroporated 3 days later with increasing amounts of coBE3 mRNA (25, 50, 75, and 100µg/ml coBE3 mRNA). Unlike the previous results shown in Figure 5-2 B using the cr1 scaffold where, flow cytometry-based phenotyping indicated that TRBC knockout had lower plateau at 25µg/ml, instead of 50µg/ml coBE3 mRNA (Figure 5-8 B). Additionally, the TTRBC_{C+5}-CAR19 configuration was able to support high levels of TRBC knockout reaching 54.3% TCRαβ—(80.3% of CAR19⁺) at 100µg/ml coBE3 mRNA (Figure 5-8 C). Moreover, this titration was performed in parallel with the addition of 10µg/ml CD52 ex1 SD and B2M ex1 SD loose sgRNA. Unlike previous data using the cr1 scaffold demonstrating heavily reduced editing from vector expressed guide in the present of uncoupled sgRNA, the TTRBC_{C+5}-CAR19 cells instead only showed minimal reduction in protein knockout from vector expressed TRBC guide with the addition of loose sgRNA (Figure 5-8 B). This was most apparent at 25µg/ml coBE3 mRNA where addition of loose guide reduced TRBC knockout from 73.1% to 63.1% (gating on CD45⁺CAR19⁺). Knockout of both CD52 ex1 SD and B2M ex1 SD sgRNA appeared to have reached maximal efficiency across the full-range of coBE3 mRNA titration (93%-94.5% CD52 knockout, and 92%-94.3% B2M knockout from 25-100µg/ml). Additionally, analysis of this data revealed an increasing proportion of triple knockout cells at higher concentrations of coBE3 mRNA (Figure 5-8 D), reaching 72.3% of the CAR19⁺ population at 100µg/ml (Figure 5.8 E).
Figure 5-8 Optimisation of sgRNA scaffold to increase expression and RNP stability:

A. Diagram of sgRNA secondary structure with (i.) cr1 scaffold, compared to (ii.) optimised scaffold with C+5 modification labelled. N = any base, green = nucleotide substitution, red = addition of nucleotide sequence.

B. Primary T cells transduced with TTRBC c+5-CAR19 vector, electroporated with coBE3 mRNA at 25, 50, 75, and 100 µg/ml alone (circles), or with 10 µg/ml CD52 ex1 SD and B2M ex1 SD sgRNA (squares). Values are based on CD45+CAR19+ population (n=1).

C. Representative
flow cytometry plot with 100µg/ml coBE3 mRNA, gated on CD45+ cells. D. Histogram detailing composition of CD45+CAR19+ cells with increasing concentrations coBE3 and the addition of both CD52 ex1 SD and B2M ex1 SD. E. Representative flow cytometry plot at 100µg/ml coBE3 mRNA with the addition of 10µg/ml CD52 ex1 SD and B2M ex1 SD sgRNA, gated on CD45+ cells. Staining for HLA class I and TCRαβ were used as surrogate markers for B2M and TRBC.

The application of novel genome editing tools has necessitated further investigation of anticipated molecular consequences. Genomic DNA was isolated from these cells and homolog sequences between TRBC 1 and 2 were amplified by PCR encompassing the predicted editing site and submitted for Sanger sequencing. Sequencing data was next analysed using the online tool EDITR (https://moriaritylab.shinyapps.io/editr_v10/), which assessed nucleotide conversions based on the expected levels of background in the sample and the height of the chromatogram peaks within the protospacer sequence [249]. Consistent with the protein knockout shown in Figure 5-8 B, base conversion at the TRBC loci appears to have reached its maximal efficiency at all tested coBE3 mRNA concentrations, resulting in only a slight positive correlation between base conversion at the DNA level and coBE3 mRNA concentration (Figure 5-9 A i). Similarly little to no effect on editing was observed at the genomic levels with the addition of chemically modified sgRNA. Additionally, this data indicated preferential editing of position C5 in the protospcacer compared to C6, reaching 33% and 24% respectively at 100µg/ml (Figure 5-9 A ii).

Additionally genomic DNA extracted from cells treated with CD52 ex1 SD and B2M ex1 SD were used to amplify the TRBC, CD52 and B2M targeted loci. PCR products were purified, and sent for Sanger sequencing. Consistent with the flow cytometry data shown in Figure 5-8 B, EDITR analysis revealed almost identical levels of C>T conversion at the TRBC loci in the presence of loose sgRNA (Figure 5-9 A i and iii). In agreement with flow cytometry based analysis, C>T deamination appeared saturated at the targeted cytidine within the CD52 exon 1 splice donor site (C7) across the coBE3 mRNA concentrations (87%-87% from 25-100µg/ml). However, at protospacer position C8 a modest increase in conversion, ranging from 72%-81% was observed
(Figure 5-9 B i and ii). Furthermore, position C2 in the CD52 protospacer laying outside the optimal BE3 editing window also showed high levels of C>T conversion ranging from 55%-63% between 25-100µg/ml coBE3 mRNA. The B2M protospacer contained three C nucleotides, within the optimal BE3 editing edit window (C4, 6, and 8). Relatively little editing was observed at C8 across all coBE3 mRNA concentrations (8% at 100µg/ml coBE3 mRNA). In contrast both C4 and C6 positions showed high levels of C>T editing across all concentrations, with an apparent editing preference for C6 over C4 (78% - 85%, and 69% - 74% respectively) (Figure 5-9 C i and ii).

Figure 5-9 Genomic analysis of C>T conversion at TRBC, CD52 and B2M loci:
A. Sanger sequencing-based EDITR analysis of the **TRBC** loci. i. Editing at protospacer positions C5 and C6 across the range of **coBE3** mRNA concentrations. ii. Representative EDITR output for **TRBC** with 100µg/ml **coBE3** mRNA alone, or (iii.) with CD52 ex1 SD and B2M ex1 SD sgRNA. B. EDITR analysis of C>T conversion at CD52 ex1 SD protospacer positions C7 and C8, within the optimal editing window. i. C>T conversion at C7 and C8 across the range of **coBE3** mRNA concentrations. ii. Representative EDITR output with 100µg/ml **coBE3** mRNA + CD52 ex1 SD. C. Sanger sequencing based EDITR analysis in the B2M ex1 SD protospacer sequence. i. editing at protospacer positions C4, C6, and C8. ii. Representative EDITR output from primary T cells electroporated with 100µg/ml **coBE3** mRNA + B2M ex1 SD. Optimal editing window (4-8bp distal to the PAM) is shown in a red box. Protospacer positions 1 and 20 are labelled. The boxes are coloured blue if the nucleotide in this position is determined to be the main peak in the chromatogram. If other peaks are presented at that same position, colours range from red to blue (red = low peak, blue = high peak). White boxes indicate that this nucleotide is not present at this position.

It has been reported that factors including the sequence context and relative position within the protospacer of the targeted C can impact the deamination efficiency of the rAPOBEC1 enzyme within the BE3 fusion protein [74, 97, 101]. The results above for both **TRBC** and **CD52** indicated a preference for C bases residing within a 5’ ACC 3’ motif compared to those in a 5’ CCT 3’ motif, where targeted C’s are within the optimal BE3 editing window. Furthermore, although the editing window for BE3 is generally considered as positions 4-8 distal to the PAM, substantial editing was observed at position C2 within the CD52 protospacer. Further investigation of the sequencing data revealed that C2 is positioned with a 5’ T nucleotide. Positioning of a C nucleotides with a 5’ T has been previously reported to increase editing even outside the optimal window [74]. In accordance with previously published data very little editing was detected within the B2M protospacer at position C8 which has a 5’ G nucleotide (5’ GCT 3’) considered detrimental to editing efficiency [74]. Additionally, this data also indicated preferential editing of C6 over C4 within the B2M protospacer sequence despite C4 having a 5’ T nucleotide preference by rAPOBEC1
This is likely due to their respective positions within the protospacer as generally C6 has been shown to possess a higher affinity for editing than C4 [101, 250].

To ascertain the optimal concentration of chemically synthesised sgRNAs for generation of triple knockout cells, a titration of loose guide was carried out in primary T cells transduced with TTRBC<sub>C<sub>6</sub></sub>-CAR19. Transduction resulted in 54.6% CAR19<sup>+</sup> cells, which were then electroporated with 50µg/ml coBE3 mRNA with the addition of 0, 2.5, 5, 7.5, 10, and 20 µg/ml of each CD52 ex1 SD and B2M ex1 SD sgRNA (Figure 5-10 A). Both CD52 ex1 SD and B2M ex1 SD guides showed a dose response in knockout when sgRNA concentration increased, plateauing at 10µg/ml with 87.8% and 88.9% respectively. As seen previously there was a reduction in vector expressed TRBC ex1-2 guide knockout at higher loose CD52 ex1 SD and B2M ex1 SD guide concentrations, however, this was not as pronounced as previously seen in Figure 5-6 A i using the cr1 scaffold. TRBC knockout ranged from 33.3% (59% of CD45<sup>+</sup>CAR19<sup>+</sup>, Figure 5-10 B i), to 20.6% (39.4% of CD45<sup>+</sup>CAR19<sup>+</sup>) at 0 and 20µg/ml sgRNA respectively. Optimal triple knockout was observed when 10µg/ml of each loose guide RNA was co-delivered, resulting in 44.8% TCRαβ<sup>+</sup>/CD52<sup>-</sup>/HLA class I<sup>+</sup> cells (gated on CD45<sup>+</sup>CAR19<sup>+</sup>) (Figure 5-10 B ii).
Figure 5-10 Uncoupled sgRNA titration with TTRBCC+5-CAR19 vector:

A. Primary T cells transduced with TTRBC<sub>C+5</sub>-CAR19 vector, electroporated with 50µg/ml coBE3 mRNA + CD52 ex1 SD and B2M ex1 SD guides at 2.5, 5, 7.5, 10, and 20µg/ml. Protein knockout values are based on CD45<sup>+</sup>CAR19<sup>+</sup> gating B. flow cytometry data depicting protein knockout for TRBC, CD52, and B2M at (i.) 50µg coBE3 mRNA alone, or (ii.) with 10µg/ml CD52 ex1 SD and B2M ex1 SD sgRNA. Cells have been gated on CD45<sup>+</sup> population. Staining for HLA class I and TCRαβ was used to assay B2M and TRBC respectively.

To elucidate any potential differences between BE3 editing using cr1 or C+5 scaffold sequences, primary T cells from four healthy donors were transduced at MOI 5 with TTRBC-CAR19 vector containing either the cr1 or C+5 scaffolds. Both cr1 and C+5 containing vectors demonstrated similar transduction efficiencies, 56.6% ± 8.1% and 54.2% ± 6.7%, respectively. Upon electroporation with 50µg/ml coBE3 mRNA, TRBC knockout efficiencies showed no significant difference between cr1 and C+5, when normalised on CAR19 expression (68.1% ± 4.1% and 72.6% ± 6.1%, p=0.5454) (Figure 5-11 A). Additionally, to check if the C+5 scaffold was affected by the presence of
loose sgRNA to the same degree as the cr1 scaffold, these cells were electroporated with 10µg/ml CD52 ex1 SD and B2M ex1 SD loose sgRNA, with 50µg/ml coBE3 mRNA. In the presence of loose sgRNA, vector expressed TRBC knockout was significantly reduced with both cr1 and C+5 scaffolds (p<0.0001, and p=0.0052 respectively). However, this reduction in vector expressed guide knockout appeared to be more distinct with the cr1 scaffold, compared to the C+5 scaffold, revealing a significant difference between the two (40.6% ± 2.6%, and 58.3% ± 2.8% of CAR19, p=0.001) (Figure 5-11 A). This improved TRBC knockout in the presence of loose sgRNA, translated to a significantly increased proportion of triple knockout cells when using the C+5 scaffold compared to cr1 in the CAR19+ population (54.1% ± 3.6% and 37.2% ± 2.5% respectively, p=0.0006) (Figure 5-11 B). As a control, cells were transduced with a pCCL-CAR19 vector and electroporated with 50µg/ml coBE3 mRNA along with 10µg/ml TRBC ex1-2, CD52 ex1 SD, and B2M ex1 SD guides. This approach resulted in significantly higher levels of triple knockout (80.1% ± 4.2%) compared to TTRBC-CAR19 with both cr1 and C+5 scaffolds (p<0.0001) (Figure 5-11 B). However, with this approach, there is no coupling between TRBC knockout and transgene expression. Consequently, TCRαβ depletion does not enrich for CAR19 expression as demonstrated with the terminal-CRISPR configuration resulting in a more heterogeneous population (Figure 5-11 C i, and ii).
A. Flow cytometry data from n=4 primary T cells donors transduced with TTRBC-CAR19, containing either cr1 or C+5 scaffolds and electroporated with 50µg/ml coBE3 mRNA alone or in conjunction with 10µg/ml CD52 ex1 SD and B2M ex1 SD sgRNA. No significant difference was seen in TRBC disruption between cr1 and C+5 scaffold when electroporated with BE3 alone (p = 0.5454). Both cr1 and C+5 scaffold sequences showed a significant reduction in TRBC knockout in the presence of loose sgRNA (cr1 **** p < 0.0001, C+5 ** p = 0.0052). Vectors containing C+5 scaffold, shows significantly increased TRBC knockout in the presence of loose sgRNA compare to the cr1 scaffold (***p = 0.001).

B. Histogram detailing the composition of these cell populations. Superior TRBC disruption in the present of uncouple CD52 ex1 SD and B2M ex1 SD sgRNA when using the C+5 scaffold leads to a significantly increased triple negative population compared to with the cr1 scaffold (***p = 0.0006). Triple knockout was significantly increase when delivering all three guide sequences as uncoupled chemically modified sgRNA (****p < 0.0001 for both cr1 and C+5). Flow cytometry values are based on CD45+/CAR19+ gating. Groups are compared by one-
way ANOVA with Tukey’s post hoc test. Error bars show ±1 SEM. **C.** Depletion of residual TCRαβ expression by magnetic bead separation. i. primary T cells transduced with pCCL-CAR19 vector, electroporated with coBE3 and uncoupled TRBCex1-2, CD52 ex1 SD, and B2M ex1 SD sgRNA. ii. Primary T cells transduced with TTRBC<sub>C+5</sub>-CAR19 electroporated with coBE3 and CD52 ex1 SD, and B2M ex1 SD sgRNA. TCRαβ-/CAR19+ population is shown in the red box. Plots are gated on CD45+ cells. Staining for HLA class I and TCRαβ are used as surrogate markers for B2M and TRBC respectively.

### 5.9 Minimal U6 promoter demonstrates compatibility with coBE3 base editing

The dynamics of guide expression differ between different RNA Pol III promoters, which in turn could impact the efficiency of base editing. Although, the mini-U6 promoter described in chapter 4 has demonstrated comparable knockout efficiencies to the wild type U6 in the context of SpCas9, this may not hold true for cytidine base editing technology. Therefore, a TTRBC<sub>C+5</sub>-CAR19 plasmid was produced with a mini-U6 promoter instead of the wild type U6 using SDM to assess differences between these configurations. Third generation lentiviral vectors were produced for TTRBC<sub>C+5</sub>-CAR19 configurations containing either wild type or mini U6 promoters (3.8 x10<sup>8</sup> and 2 x10<sup>8</sup> IU/ml respectively). These vectors were used to transduce primary T cell from three donors, achieving similar levels of CAR19 expression (74.9% ± 5.51% for wtU6 and 79.8% ± 5.9% for mini-U6 containing vectors). When electroporated with 50µg/ml coBE3 mRNA, there was no observable difference between the TRBC knockout achieved with either wild type U6 or mini U6 expressed sgRNA (77.6% ± 2.9% and 70.1% ± 10.39% of CD45+CAR19+) (**Figure 5-12 A**). Additionally, a similar level of reduction in TRBC knockout was seen in the presence of 20µg/ml CD52 ex1 SD loose guide between wild type U6 and mini U6 (53.1% ± 8.7% and 60.6% ± 7.8%) (**Figure 5-12 A**). Additionally, both wild type and mini U6 promoters demonstrated no observable difference between the proportion of the double knockout population after electroporation with coBE3 mRNA and CD52 ex1 SD sgRNA (46.3% ± 11.3% and 58.6% ± 8.3%, gated on CD45+CAR19+) (**Figure 5-12 B**). This data appears to support the idea that both wild type and mini U6 function equally in the context of BE3 editing, or produce similar levels of knockout with or without the addition of loose
chemically modified sgRNA (Figure 5-12 C). Carrying out this comparison between wild type and mini U6 promoters in a TTRBC-CAR19 vector containing the cr1 scaffold sequence revealed similar results (Supplementary figure 5 A, B)

![Figure 5.12](image_url)

**Figure 5.12 Comparison between wild type U6 and mini-U6 promoter:**

**A.** Primary T cell donors (n=3) transduced with TTRBC\textsubscript{C,5}-CAR19 either expressing the TRBC sgRNA from a wild type U6 or a mini-U6 promoter electroporated with 50µg/ml coBE3 mRNA alone, or with 20µg/ml CD52 ex1 SD sgRNA. *TRBC* knockout appears similar between wild type U6 and mini-U6 promoters with coBE3 alone, and coBE3 + CD52 ex1 SD sgRNA. **B.** Similarity can be observed in the proportion of double negative cells when using the wild type U6 compared to the mini-U6. **C.** Representative flow plot from one donor, showing TCRαβ/CD52 populations.

### 5.10 Detection of SpCas9 and BE3 protein expression by western blot

Protein expression and stability could be potential factors in the differences observed between SpCas9 and BE3 editing. Therefore, visualisation of the protein products
from SpCas9 and coBE3 mRNA overtime could provide valuable information. To this end, PBMCs were isolated from a healthy blood donor and activated via CD3/CD28. The cells were next transduced at MOI 5 with the TTRBC_C5-CAR19 vector, achieving 52.3% CAR19 expression. Subsequent electroporation’s were carried out with either SpCas9 or coBE3 mRNA, with sequential cell pellets collected at 0.5, 1, 2, 3, 7 days post mRNA delivery. Success of mRNA delivery was confirmed by on-target editing analysis by both flow cytometry (SpCas9 = 57.4%, and BE3 = 48.9%, TCRαβ knockout gated on CD45+) (Figure 5-13 A). This was also confirmed by Sanger sequencing analysis of the TRBC loci using the Synthego ICE tool for the SpCas9 treated sample (Figure 5-13 B), and EDITR for the BE3 treated sample (Figure 5-13 C).

Total protein was extracted from these cell pellets, and quantified via a BCA assay. Equal amounts of protein were loaded and subjected to SDS-PAGE and membranes were subsequently stained for the Cas9 protein in both SpCas9, and coBE3 mRNA treated cell lysates. In this instance staining for β-Actin was used as a loading control. Protein from unmodified, and transduced but non-edited cells were used as negative controls and displayed no Cas9 staining. Based on the titration presented in Figure 3-2 D, 15.6ng of purified SpCas9 protein was loaded as a positive control, presenting high levels of Cas9 staining. Similar to results shown in chapter 3 (Figure 3-2 E), the Cas9 protein could been detected up to 3 day post SpCas9 mRNA delivery (Figure 5-13 D). However, BE3 protein appeared considerably more transient, displaying a peak in expression at 0.5 days, which had mostly dissipated by 1 day post mRNA delivery (Figure 5-13 E).
Figure 5-13 Transient expression of BE3 protein:

A. Flow cytometry data from primary human T cells transduced with TTRBCCα5-CAR19 vector and electroporated with either SpCas9, or coBE3 mRNA. Coupled knockout of TRBC is observed with both mRNA. This data is gated on the CD45+ population. TCRαβ staining is used as a surrogate marker for TRBC. B. Sanger sequencing-based ICE analysis was used to quantify insertion-deletion mutations (indels) at the TRBC loci, when treated with SpCas9 mRNA. Genomic disruption appears consistent with protein knockout measured by flow cytometry. C. EDITR analysis of Sanger
sequencing data across the TRBC protospacer sequence measuring C>T changes. Editing at C5 and C6 appears similar to protein knockout measured by flow cytometry. D. Western blot from cells electroporated with SpCas9 mRNA at 0.5, 1, 2, 3, and 7 days post electroporation. Highest intensity of SpCas9 staining visible at 0.5 days, and had mostly dissipated by day 3. The wall of the well between the TD-no-EP sample and the empty well was damaged allowing spill over, resulting in actin staining in the empty well. Additionally, the bottom right hand half of the gel was split across the TD no EP sample resulting in the SpCas9 control being skewed above the appropriate ladder value (160kDa). E. Western blot from cells at 0.5, 1, 2, 3, and 7 days post electroporation with coBE3 mRNA. BE3 protein appeared highly expressed at 0.5 days, but had mostly dissipated by 1 day post electroporation. Both SpCas9 and BE3 were detected by staining the common Cas9 component. Unmodified (UT), and vector transduced alone cells (TD no EP), were used as negative controls. Purified SpCas9 protein (15.6ng), was used as a positive control. β-actin staining was used as a loading control.

5.11 Screening of large translocation events with SpCas9 and BE3

As detailed earlier in, simultaneous genome editing of multiple genomic loci has the potential to promote large chromosome translocations. Translocation events have been detected using ZFNs, TALENS, and CRISPR/Cas9 genome editing tools due to their propensity to generate dsDNA breaks [85, 123, 251]. To date there have been no transformative side effects in T cells harbouring these events [210]. However, their long term effects are widely unknown and present a risk factor for multiplexed genome editing. Use of emerging base editing technologies could be advantageous in this area, as genome editing is no longer subject to DSB-induced indel formation. Of note, Webber and colleagues (2019) demonstrate an absence of translocation events in two primary T cell donors simultaneously edited at three loci (TRAC, B2M, and PDCD1) when using BE3. In contrast editing with SpCas9, resulted in frequencies of up to 2% for some translocation events [123].

To assess potential translocation events between two of our previously edited genomic loci, TRBC and CD52, samples were generated from three primary T cell
donors, electroporated with 50µg/ml of either coBE3 or SpCas9 mRNA with 20µg/ml of both TRBC ex1-2 and CD52 ex1 SD loose sgRNA. These samples were analysed by flow cytometry to confirm protein knockout (Figure 5-14 A). Disruption of TCRαβ expression appeared comparable between SpCas9 and coBE3 (94.2% ± 3% and 92.9% ± 2.3% respectively). There appeared to be a reduction in knockout of CD52 when using SpCas9 compared to that of coBE3 (71.8% ± 4.9% and 92.9% ± 2.6% respectively), which was reflected in frequency of double knockout (70.8% ± 4.9% and 88.9% ± 2.9% respectively) (Figure 5-14 B). Reduced protein disruption from the CD52 ex1 SD sgRNA with SpCas9 is likely due to the predicted cleavage site residing near the beginning of the exon 1 intron sequence. Placement of the SpCas9 cleavage site just outside the exonic sequence (3’ of exon 1) may result in small indels not affecting CD52 protein expression.

EDITR analysis on PCR amplicons from samples treated with coBE3 revealed high levels C>T changes at both TRBC and CD52 loci. As shown previously C5 in the TRBC protospacer demonstrated preferential editing compared to C6 (68% ± 6.5% and 58.3% ± 3.9% respectively). Similarly, C7 in the CD52 protospacer also appeared to be preferentially edited compared to C8 (71.7% ± 6.9% and 59.3% ± 6.9% respectively). Consistent with previous observations in Figure 5-9 B ii, all donors presented with high levels of editing of C2 positioned outside of the optimal BE3 editing window (59.3% ± 8.5%).

Next, in order to visualise potential aberrations of multiplexed editing, primers were designed to amplify the four predicted translocations combinations (T1-T4) resulting from simultaneous breaks in TRBC and CD52 (Figure 5-14 C). PCR of genomic DNA extractions were visualised by gel electrophoresis demonstrating the presence of all predicted translocation events in SpCas9 treated samples. Conversely, bands were not present in the unmodified control samples, and importantly were mostly absent in coBE3 treated samples with some faint bands present in donor 3 for both T2 and T4 (Figure 5-14 D). DNA extracted from positive translocation bands seen in SpCas9 treated samples from donor 1 (Figure 5-14 D) were analysed by NGS confirming the presence of predicted translocation events (Supplementary figure 6 A).
However, in order to quantify the translocation events, genomic DNA of the samples was measured by ddPCR (performed by Dr. Gkazi). Similar to the gel images, unmodified samples show no evidence of translocations. Samples treated with SpCas9 revealed the presence of T1, T2, and T4 translocations (1.7% ± 1.4%, 0.4% ± 0.1%, and 0.4% ± 0.1% respectively). In comparison, two donor samples treated with coBE3 showed no evidence of translocations. However, donor 3 exhibited 1.5% and 2.2% translocation frequency for T2 and T4 in turn (Supplementary figure 6 B). Similar work is currently in development for assaying translocation events between TRBC, CD52, and B2M in both SpCas9 and BE3 edited cells.
Figure 5.14 Reduction of large translocation events with BE3 compared to SpCas9 editing in primary human T cells:

A. Representative flow cytometry data depicting protein knockout of TRBC and CD52 after editing with SpCas9 or coBE3. B. Histogram summarising TRBC and CD52 protein knockout in primary T cells donors (n=3). All values are gated on CD45+ cells. TCRαβ staining is used as a surrogate for TRBC. C. Schema of chromosomes 7 (grey) and 1 (blue), with a red line indicating the approximate location of TRBC ex1-2 and CD52 ex1 SD protospacer binding, as well as the four predicted translocation between these loci (T1-T4). D. PCR amplification from n=3 donors of the CD52 and TRBC loci across the protospacer binding site, as well as the four predicted translocation events.
(T1-T4). Positive bands seen in SpCas9 treated samples (T1-T4) confirm positive translocation events.

5.12 Chapter summary:

This chapter describes the application of cytidine base editing technology for multiplexed gene knockout in the pursuit of ‘universal’ CAR T cells capable of overcoming HLA barriers. Initial experiments performed in a Jurkat E6.1 cell line, stably expressing the coBE3 transgene were used to validate this approach for TRBC/ TRAC knockout. These earlier observations led to applications in primary human T cells, utilising coBE3 mRNA for the knockout of TRBC, CD52, and B2M. Despite encouraging results in a single knockout setting, several hurdles were encountered when attempting to multiplex this technology. These included relatively low knockout from multiplexed terminal-CRISPR configurations developed in chapter 4, and apparent competition effects between vector expressed and loose sgRNA, which were not encountered in the context of SpCas9 gene knockout. However, a number of titration experiments as well as inclusion of a modified guide scaffold sequence (C+5), appeared to mostly mitigate these issues. Additionally, this chapter presents compelling evidence that coBE3 generates cell products with lower frequency of large chromosomal translocation events compared to SpCas9 treated samples.
Chapter 6  Application of terminal-CRISPR for rTCR cellular immunotherapy

6.1  Recombinant TCR based immunotherapies for HBV driven HCC

As detailed in section 1.9.1, rTCR based immunotherapies provide an alternative to CAR based therapies and recognise peptides in the context of HLA molecules. This affords rTCR’s the capacity to survey intracellular peptides, opening up a diverse target range including neo-antigens, oncogenic transcription factors, and viral derived proteins.

rTCR based strategies have been applied clinically to a variety of malignancies including melanoma and leukaemia [162, 163]. The model chosen for rTCR based therapy here was hepatitis B virus (HBV) driven hepatocellular carcinoma (HCC). It has been long since established that HBV contributes significantly to the occurrence of HCC, due to integration of viral genomes in hepatocytes and chronic liver inflammation [252, 253]. HBV viral antigens are processed and presented by HLA molecules on the infected cell surface [254, 255]. As such naturally occurring HBV-specific T cells, have been detected that are capable of modulating the viral and tumour burdens in patients with HBV derived HCC [256, 257]. Despite this, these HBV specific T cell responses are often functionally exhausted in patients with chronic HBV infection as a result of persistent antigen stimulation [258-260]. Additionally, due to limited therapeutic options HCC has a relatively poor prognosis, which drives the need for novel treatment options for chronic HBV infection. Generation of T cells expressing rTCR’s specific for HBV derived antigens have been shown to be effective in a number of preclinical models [261-265]. Additionally, this approach has been translated to a number of patients, including after liver transplantation where metastatic lesions expressed distinct HLA molecules, thereby reducing the risk of hepatotoxicity [159]. Although, complete tumour clearance is yet to be observed with this treatment strategy there are a number of further clinical trials in development (clinical trial identifier: NCT03971747, NCT02719782, NCT02686372, NCT03634683).
Limits to rTCR based cellular immunotherapies, include cross-pairing between rTCR and endogenous TCR (eTCR) chains, resulting in novel dimeric complexes with unknown specificity [161, 164-167, 266]. Additionally, competition between eTCR and rTCR chains for shared cellular components of the multimeric CD3 complex, reduces or even negates rTCR cell surface expression. Approaches, to address TCR mispairing have focused on promoting exclusive rTCR pairing via additional disulphide bonds, use of high-affinity TCRs, murinisation of rTCR constant regions, hybrid TCR-CAR molecules, as well as development of single chain TCRs through covalently linking of the variable chain[170-177]. Additionally, nuclease mediated disruption of one or both eTCR chains by ZFNs, TALENS, and CRISPR/ Cas9 have been used to address both miss-pairing and competition for cellular components [46, 169, 178, 179, 266-268].

In this chapter, CRISPR based genome editing strategies were applied for the disruption of eTCR α and β chains. The terminal-CRISPR configuration developed in chapter 3 was used to couple disruption of eTCR chains to rTCR expression. This was facilitated through the use of emerging cytidine deaminase base editing technology described in the previous chapter for seamless base conversion, for the production of cellular immunotherapies.

### 6.2 Composition of HBV specific rTCR (S183-91)

Gehring and associates (2011), developed an rTCR targeting the HBV envelop surface antigen S183-91 (FLLTRILT), with variable (V) alpha (TRAV34) and V beta (TRBV28) regions restricted to HLA-A0201 [269]. The S183-91 rTCR was initially tested human constant (C) domains, in a terminal-CRISPR configuration expressing the TRAC guide validated in chapters 3. This configuration was able to achieve a transduction efficiency of 40.8% in primary human T cells. Electroporation with 100µg/ml SpCas9 mRNA appeared to increase the proportion of rTCR⁺ cells in the culture to 55.3%, as well as increasing MFI from 2.4x10³ to 3.2x10³.

Despite these initial encouraging results, rTCR chains containing human C domains have been shown to readily mispair with the endogenous TCR chains, resulting in
novel TCR molecules with unknown specificity [268]. Furthermore CRISPR/Cas9 editing technology relies on the recognition of a short 20bp protospacer sequence to direct nuclease mediated scission, therefore use of human C regions could result in off-target cleavage of the rTCR chains. Indeed, the first 15bp of the TRAC protospacer sequence used throughout this work shares complete complimentary to the recombinant TRAC sequence. PCR amplification of the recombinant TRAC sequence followed by TIDE analysis carried out on primary human T cells edited with SpCas9 revealed ~4% editing at this site.

The S183-91 specific rTCR utilised for clinical phase testing by Qasim and colleagues (2015), incorporated murine C regions within the rTCR to minimise mispairing events. An additional disulphide bond was introduced to stabilise rTCR chain pairing, and further reduce the ability of these recombinant chains to miss pair with the endogenous TCR chains [159]. Furthermore, use of these distinctive C regions, negates concerns around off-target editing of the introduced rTCR. Additionally, equimolar expression of both rTCR chains is achieved by expressing rTCR α and β chains as a single transcript, separated by the porcine teschovirus-1 2A (P2A) sequence. Inclusion of this sequence, promotes ribosomal skipping during translation, leading to self-cleavage and the formation of two mature protein products. Such 2A self-cleaving peptides have been isolated from many viral species since their initial discovery in the foot-and-mouth-disease virus in 1991 [270]. The P2A sequence used in this vector configuration has an N-terminus GSG linker, which has been reported to increase self-cleavage efficiency [271].

This S183-91 rTCR sequence was cloned, into a terminal-CRISPR configuration to couple expression of the murinized S183-91 rTCR with a TRAC targeting sgRNA (TTRAC-S183-91 rTCR) (Figure 6-1 A). PCR based Infusion cloning was used to transfer the S183-91 rTCR into a terminal-CRISPR configuration under the control of, an internal hPGK Pol II promoter with a translation initiation Kozak consensus sequence (GCCACC). The TRAC guide used in chapters 3 and 4 was cloned into this vector as described in Section 2.29.
6.3 TT-rTCR (S183-91) vector with transient SpCas9 delivery allowed enrichment of the eTCR⁻/rTCR⁺ population

Lentiviral vector production of the TT-rTCR (S183-91) construct, produced high titer virus of $2.34 \times 10^8$ IU/ml as measure by flow cytometry in a CD3⁺ Jurkat cell line. To test its efficiency, primary human T cells were activated and transduced with TT- rTCR (S183-91) vector at MOI 5, resulting in 56.3% rTCR expression. Unlike observations made when using humanised C domains, electroporation of these cell with 100µg/ml SpCas9 mRNA resulted in a distinct population of eTCR⁻/rTCR⁺ cells (53.4%). Because the eTCR was amenable to detection by anti-TCRαβ monoclonal antibody, magnetic bead mediated depletion of residual eTCRαβ expressing cells was possible. Notably, rTCR was not susceptible to these reagents and thus at the end of production, cells could be enriched as eTCR⁻/rTCR⁺ (86.2%), resulting in a highly homogenous product (Figure 6-1 B). Similar to results obtained with human C domains, an increased MFI of rTCR expression was observed post editing from $6.7 \times 10^3$ in the eTCR⁺/rTCR⁻ population, to $12.4 \times 10^3$ in the eTCR⁻/rTCR⁺ population.
Figure 6-1 Terminal-CRISPR configuration coupling TRAC sgRNA and rTCR (S183-91) expression:

A. Terminal-CRISPR plasmid DNA configuration, coupling the expression of a recombinant TCR (rTCR) against the Hepatitis B virus (HBV) envelope surface antigen 183-91 (S183-91), and TRAC specific sgRNA with a cr1 scaffold sequence expressed via a wild type U6 promoter. The rTCR is expressed as a single transcript with the rTCR α chain first, followed by the rTCR β chain separated by a P2A. These recombinant chains are composed of the T cell receptor α variable 34 (TRAV34), and the T cell receptor β variable 28 (TRBV28) domains, as well as either murine TRAC (muTRAC), or murine TRBC 1 (muTRBC1). The rTCR chains contain an additional cysteine-cysteine disulfide bonds between their murine constant regions through muTRAC mutation, T48C, and muTRBC1 mutation, S57C. B. Primary T cells transduced with the TT-rTCR (S183-91) vector. Flow cytometry confirms eTCR knockout isolated to the vector transduced rTCR+ population post electroporation with 100µg/ml SpCas9. Magnetic
bead-mediated depletion of residual eTCR expressing cells simultaneously enriches for rTCR expression resulting in a homogenous 86.2% eTCR- / rTCR+ population.

In order to assay in vitro functionality, Interferon-γ (IFNγ), Tumour Necrosis Factor alpha (TNFα), IL2, and C-C Motif Chemokine Ligand 4 (CCL4) responses of T cells were measured (execution by Dr. Stegmann) against HepG2 cells pulsed with a gradient of the target HBV peptide (HBV surface peptide S183-91, FLLTRILT). The S183-91 peptide was tested between 0.01-10⁶pmol, at 10 fold dilutions. Effector T cells were then incubated with target HepG2 cells overnight at a 1:1 ratio. Unmodified effectors (eTCR+/rTCR−) showed no cytokine responsiveness to HepG2 targets at any peptide concentration. However, both eTCR+/rTCR± and eTCR−/rTCR+ effector T cell groups showed increased cytokine responsiveness at higher peptide concentrations, with eTCR−/rTCR+ effector group consistently showing higher cytokine production than there unedited counterpart (eTCR+/rTCR+)(Appendix E). The use of a relevant control peptide (HBV core peptide C18-27, FLPSDFFPSV), as well as HepG2 cells alone demonstrated, no off-target activation of these effector groups confirming the specificity of the observed cytokine responsiveness.

6.4 Targeted cytidine deamination of endogenous TRBC, tolerates generation of rTCR cell products

Application of CRISPR/Cas9 in conjunction with our terminal-CRISPR lentiviral vector has proven efficient for the generation of CAR and rTCR cellular immunotherapies. However, there is evidence suggesting that dsDNA break formation has several pressing concerns, including formation of large indels spanning several kb, and increased p53 expression [84, 272-274]. Furthermore, multiplexed genome editing will be required to produce truly allogenic T cell products, by the removal of mismatched HLA molecules. Additionally, functional advantages maybe achieved via the KO of both eTCR chains, as well as immune checkpoints. However, as described previously simultaneous disruption of multiple loci incurs the risk of large chromosomal translocation events.
As described in chapter 5, use of base editing technology may help mitigate these concerns, and to this end this section will focus on the application of the coBE3 fusion protein for seamless genome editing. The TRBC ex1-2 guide validated in chapter 5 was cloned into a terminal-CRISPR configuration expressing the S183-91 HBV rTCR with murinized C regions (TTRBC-rTCR (S183-91)). The TRBC ex1-2 protospacer, contains duplex cytidine nucleotides within the BE3 editing window (13-17bp proximal to the PAM), capable of producing a premature stop codon in exon 1 of the TRBC 1 and 2 loci. Similar to the disruption of TRAC in the previous section, disruption of TRBC in theory removes eTCR surface expression, while also preventing miss paring between the rTCR α and the disrupted eTCR β chains. Additionally, the murinization of rTCR C regions further address the risk of rTCR cross pairing, and prevents CRISPR editing of the introduced rTCR chains (Figure 6-2 A).

Third generation lentiviral vector was generated for the TTRBC-HBV rTCR configuration, demonstrating the ability of this configuration to produce high viral titers (4.28x10^8/ml). This vector was then used to transduce primary human T cells at MOI 5, achieving 60.7% transduction efficiency. Electroporation of these eTCR+/rTCR⁻ cells with 50µg/ml coBE3 mRNA resulted in an eTCRαβ⁻ population of 53.9% (81.7% of CD45⁺rTCR⁻), measured by flow cytometry.

Consistent with observations made in the previous section, when disrupting the endogenous TRAC gene eTCR was susceptible to staining using the anti-TCRαβ monoclonal antibody, whereas the murine C regions in the rTCR prevented such staining. Therefore, magnetic bead-mediated depletion of the remaining eTCR⁺ population also enriched the rTCR expressing cells. This resulted in a highly homogenous population of eTCR⁻/rTCR cells (95%), with <1% residual eTCR expression as detected by flow cytometry. As shown previously, this eTCR⁺/rTCR⁻ population demonstrated increased rTCR expression compared to unedited but transduced cells (eTCR⁻/ rTCR⁻), as measured by MFI (6.4x10³ and 19.1x10³ respectively).

Following validation of the editing strategy in the context of rTCR expression, cell products were produced from three healthy T cell donors adhering to the time line
detailed in Figure 6-2 B. All three healthy donors were readily activated and transduced with the TTRBC-rTCR (S183-91) vector attaining 50.3%-59.8% rTCR expression at MOI 5 (Figure 6-2 C, and D i). In this instance, 1x10⁷ TTRBC-rTCR (S183-91) transduced cells from each donor were electroporated with 50µg/ml coBE3 mRNA before being placed into a G-Rex10 culture vessel.

All donors showed an appreciable T cell expansion between ~8-22 fold (7.9-22.1 x10⁷ cells). Base editing also resulted in emergence of rTCR⁺ populations, increasing in proportion to 59.8%-63.9% of the cultures (Figure 6-2 C, and D ii). Magnetic bead mediated eTCRαβ depletions where carried out, followed by culturing in a G-Rex10 for 72hrs. This resulted in yields ranging from 2.5- 5 x10⁷, and enrichment of eTCR⁻/rTCR⁺ populations. Characterisation of these cells by flow cytometry showed 93.9%-96.1% rTCR expression with ~0.3%-1% residual eTCR⁺ cells (Figure 6-2 C, and D iii). Furthermore, rTCR cell surface expression was measured by MFI using a one way ANOVA indicating a significant increase between eTCR⁻/rTCR⁺ and eTCR⁺/rTCR⁺ groups (p=0.0162) (Figure 6-2 D iv).
Figure 6-2 Vector design for generation of eTCR-/rTCR+ cells using coupled base editing:

A. Theoretical TCR chain pairing when introducing rTCR chains. B. Time-line of cell production. Human peripheral blood lymphocytes were isolated and activated with Transact (anti-CD3/CD28) before transduction and electroporation. After overnight hypothermic culture, cells were expanded in GRex flasks before cryopreservation on day 14. C. Flow cytometry phenotyping of unmodified and TTRBC-S183-91 rTCR
transduced cells. Delivery of 50µg/ml coBE3 mRNA by electroporation caused reduction of eTCR expression (38.1%) and emergence of eTCR\(^-\)/rTCR\(^+\) cells (red box). Magnetic bead-mediated depletion of residual eTCR\(^+\) T cells enriched eTCR\(^-\) populations, resulting in >99% eTCR\(^-\)/95.9% rTCR\(^+\). D. Expression of rTCR (S183-91) in three healthy donors. i. Transduction initially ranged from 50.3%-59.8%. ii. Exposure to coBE3 mRNA increased rTCR expression to 59.8%-63.9%. iii. eTCR\(\alpha\beta\) depletion enriched gene edited cells increasing rTCR levels to between 93.9%-96.1%. Three colours represent different donors. iv. Levels of cell surface expression of rTCR increased in eTCR\(^-\)/rTCR\(^+\) compared to eTCR\(^+\)/rTCR\(^-\) cells (one way ANOVA, p=0.0162).

*In vitro* functionality of the end of production cells were carried out. As previously described in Section 6.3, IFN\(\gamma\), TNF\(\alpha\), IL2, and CCL4 responsiveness of T cells to HepG2 cells pulsed with a gradient of the target HBV peptide (HBV surface peptide S183-91, FLLTRILT) from 0.01-10\(^6\)pmol was measured (Dr Stegmann). Both eTCR\(^+\)/rTCR\(^-\) and eTCR\(^-\)/rTCR\(^+\) effector T cell groups showed a positive correlation between increased peptide concentration and cytokine production. No cytokine responsiveness was observed from unmodified effectors (eTCR\(^+\)/rTCR\(^-\)) to HepG2 cells at any peptide concentration. In all three donors tested, cytokine production was higher in eTCR\(^-\)/rTCR\(^+\) T cells (*Appendix E*). Furthermore, no off-target cytotoxicity was detected against HepG2 cells alone, or when pulsed with a relevant control peptide (HBV core peptide C18-27, FLPSDFFPSV).

Further characterisation of these cells was performed by our collaborators at the Institute of Molecular and Cell Biology (IMBC), Agency for Science, Technology and Research (A*STAR) in Singapore. Employing an XCelligence impedance assay, a real time assessment of target cell killing with end of production effector groups at different target: effector ratios (1:1, 1:2, and 1:4) could be measured. This assay involved the seeding of HepG2 target cells in wells containing micro electrode arrays at the base. Proliferation of target cell on the electrode micro array impedes the current passing through them, resulting in increased index values, whereas apoptosis of target cells resulted in decreased index values. Control groups included target cells alone (HepG2 only), and non-transduced effectors, as expected displayed a progressive increase and plateau in index. In contrast, both effector groups exhibited
a transient rise and then decline in index, with more rapid reductions mediated by eTCR/rTCR⁺ cells compared to eTCR⁺/rTCR⁻ T cells at all E:T ratios. This data was summarised by calculation of the area under the curve (Appendix E).

Finally, collaborators at IMBC, A*STAR tested migration and target cell killing in a 3D microfluidics device. The system capture migration of effector T cells from a fluid channel into an adjacent collagen gel embedded with target HepG2-Env-GFP target cells. Phenotyping of effector groups post thaw, confirmed their eTCR and rTCR expression profiles and as excepted showed minimal cytokine expression before entering the 3D microfluidics device. Examination of effector T cells isolated from both outside and inside the collagen gel revealed that both eTCR⁺/rTCR⁻, and eTCR⁻/rTCR⁺ groups demonstrated an increased expression of TNFα inside the gel (20.9 and 38.8% respectively), compared to outside the gel (0.5% and 1.4% respectively) (Gated on rTCR⁺CD8⁺). Unmodified effector T cell (eTCR⁺/rTCR⁻) demonstrated minimal cytokine expression both inside and outside the gel region (Gated on rTCR⁻CD8⁺). Cytotoxic effects on HepG2-Env-GFP cells by eTCR⁻/rTCR⁺ effector was confirmed after 24hrs, whereas, eTCR⁺/rTCR⁻ showed no significant difference to eTCR⁺/rTCR⁻ or HepG2 alones controls. This was confirmed by direct visualisation of the microfluidics devices by confocal microscopy, showing greater clearance of HepG2-Env-GFP targets by eTCR⁻/rTCR⁺ effectors (Appendix E).

6.5 Molecular characterisation of base editing effects

As described previously in chapter 5, use of emerging base editing technology compelled investigation of the effects of coBE3 delivery. Analysis of the on-target editing at the endogenous TRBC1/2 loci was achieved by PCR amplification of shared sequences between both TRBC 1 and 2 surrounding the predicted editing sites. As described above, the TRBC ex1-2 protospacer contained duplexed cytidine bases within the BE3 editing window (4-8nt’s distil to the PAM) at positions C5 and C6. Sanger sequencing-based analysis of these positions was achieved using EDITR (Figure 6-3 A). High levels of C>T conversion (G>A sense strand) were captured at these positions (37% ± 6.8% and 24% ± 3.7% at C5 and C6 respectively), with remarkably little activity outside the editing window at nearby C residues (4.3% ±
3.1% C3, 2.3% ± 1.7% C2, and 5% ± 2.2% C1). Additionally, low levels of non C>T conversions were detected at both C positions (12.7% ± 0.5% C5, and 2.7% ± 0.5% C6) (Figure 6-3 B).

Figure 6-3 Molecular analysis of on-target DNA editing:

A. Sanger sequencing of on-target editing at TRBC 1/2 loci in eTCR’/rTCR+ cells. Representative EDITR analysis with wild type sequences and four possible bases conversions shown at each position and target G>A sites marked (C>A on opposite strand). B. Summary of EDITR data for n=3 donors at cytidine position 5 and 6 distal to the PAM, presented as C>T changes, non C>T changes, and no editing.

These findings were corroborated by NGS analysis across the predicted editing site (performed by Dr. Gkazi). Similar levels of C>T conversion were measured at both targeted C positions (40% ± 5% C5, and 32.3% ± 5.2% C6), compared to Sanger sequencing results. NGS analysis also confirmed low levels of non C>T conversion (11.3% ± 0.5% C5, and 2.3% ± 0.5% C6). Furthermore, although mostly seamless, a minority of reads exhibited small (<10bp, 8.4% ± 2.5%) and large (10-100bp, 8.2 ±
1.2%) indels, likely signatures of NHEJ following DNA cleavage encountered in a minority of cells as others have noted previously [74, 97, 123] (Appendix E).

### 6.6 Multiplex base editing of both endogenous TRAC and TRBC loci

Recent advances in the field of rTCR therapeutics have demonstrated that knockout of either the endogenous TRAC or TRBC loci, results in increased rTCR expression as well as reduced potential for mispairing. Additionally, simultaneous disruption of both endogenous TRAC and TRBC creates a non-completive environment, advantageous for rTCR’s with weak expression, while further reduces potential for mispairing [266, 268].

The results thus far in this chapter are consistent with these observing, showing increased rTCR expression with either SpCas9 mediated TRAC disruption, or TRBC1/2 knockout using coBE3. In this section, the potential advantages to the simultaneous disruption of both endogenous TRAC, and TRBC chains were investigated using CBE technology.

Several specific TRAC guides had been tested in chapter 5 within a CD3⁺/eGFP⁺ Jurkat E6.1 cell line that stably expressed coBE3. Of these guides, only TRAC ex1 SD showed any disruption of TCRαβ-CD3 protein expression on the cell surface. However, titration of coBE3 mRNA in primary human T cells constitutively expressing the TRAC ex1 SD guide from a terminal-CRISPR vector resulted in low levels of TCRαβ disruption (23% gated on CAR19⁺, at 100µg/ml). These results are consistent with published data from Webber and colleagues (2019), who observed low protein knockout using this sgRNA, despite relatively high genomic disruption [123]. However, an alternative guide sequence was also reported, predicted to disrupt the mRNA splice acceptor site at the beginning of exon 3 (TRAC ex3 SA), which demonstrated robust TRAC knockout in the context of primary human T cells (Figure 6-4 A).

Chemically modified sgRNA for the alternative TRAC ex3 SA guide was synthesised and tested in four healthy primary T cells against both the B2M ex1 SD and the CD52 ex1 SD guides. Electroporation of the cells with 50µg/ml coBE3 mRNA alongside 10µg/ml of each TRAC ex3 SA, B2M ex1 SD, and CD52 ex1 SD guides generated high
levels of protein knockout for all targets (95.4% ± 0.64% TRAC knockout, 91.6% ± 1% B2M knockout, and 91.7% ± 2.1% CD52 knockout), resulting in a triple negative populations of 85.1% ± 1.6% (Figure 6-4 B, and ii). Moreover, knockout with TRAC ex3 SA guide measured significantly higher than that achieved using B2M ex1 SD (p = 0.021), and CD52 ex1 SD (p = 0.024) guides (one way ANOVA, with Tukey multiple comparison test).

Having confirmed high levels of TCRαβ-CD3 cell surface disruption using the TRAC guide with coBE3, it was next tested in combination with the TTRBC-S183-91 rTCR vector in primary human T cells. In this instance fresh PBMCs from three healthy donors were activated and 24hrs later were transduced with TTRBC-rTCR (S183-91) vector at MOI 5 resulting in 44.3% ± 7.3% rTCR expression (eTCR+/rTCR±). Transduced or unmodified control cells were then electroporated, with 50µg/ml coBE3 mRNA, either alone or in the presence of 10µg/ml TRAC ex3 SA guide. As seen previously TRAC knockout alone resulted in high levels of TCRαβ-CD3 disruption (95.2% ± 1%). TTRBC-S183-91 rTCR transduced cells electroporated with coBE3 mRNA alone, demonstrated 26.3% ± 6.4% (62.7% ± 4.1% gated on CD45+rTCR+) TCRαβ-CD3 disruption, while the addition of chemically modified TRAC ex3 SA sgRNA increased TCRαβ knockout to 85.9% ± 3.6% (Figure 6-4 C).

Incorporation of murine C regions in the rTCR allowed discrimination in staining for eTCR and rTCR with the anti-TCRαβ monoclonal antibody, which was exploited for magnetic bead medicated depletion of eTCR+ cells. Depletions resulted in <1% residual eTCR expressing cells (Figure 6-4 C). Compared to unedited cells, expression of the rTCR measured by MFI, presented an observable increase when TRBC was removed. Interestingly, the addition of a TRAC disruption in combination with that of TRBC increased rTCR expression even further (Figure 6-4 D). Importantly, as the TRAC ex3 SA guide was uncoupled, eTCR depletion was unable to enrich the rTCR+ population, resulting in a more heterogeneous final cell product. As a result, following depletion, rTCR expression in the combined TRAC/TRBC knockout setting only reached a modest 32.2% - 44.7%, compared to 86.2% - 92.5% of the vector-coupled TRBC knockout cells (Figure 6-4 E).
Figure 6-4 Production of rTCR (S183-91) T cells by double knockout of both TRAC and TRBC:

A. Schematic representation of BE3 editing of the exon 3 splice acceptor site in the TRAC locus. Editing window (blue) of the BE3 ranges from 4-8bp distal of the PAM (red), covering the GA splice acceptor sequence. Removal of the mRNA splice acceptor will be resolved by exon skipping. B. Validation of TRAC ex3 SA guide in primary T cell biological replicates (n=4). Flow based analysis of protein knockout
after electroporation with 50µg/ml coBE3 mRNA, and 10µg/ml TRAC ex3 SD, B2M ex1 SD, and CD52 ex1 SD sgRNA. i. All guides resulted in high levels of disruption, with no significant difference observed between B2M, and CD52 knockout (p = 0.9976). TRAC knockout appeared significantly increased compared to B2M (*p = 0.0212), and CD52 (*p = 0.0235). Comparison between groups was carried out using a one-way ANOVA with Tukey’s post hoc test. Error bars represent ± SEM. Staining for HLA class I and TCRαβ, were used as surrogate markers for B2M and TRAC respectively. ii. Representative flow cytometry plots demonstrating the shift in eTCRαβ expression after electroporation with coBE3 + TRAC ex3 SA sgRNA. C. Primary human T cells (n=3) transduced with TTRBC-rTCR (S183-91) vector electroporated with either 50µg/ml coBE3 alone or in the presence of 10µg/ml TRAC ex3 SA sgRNA prior to eTCRαβ depletion (red box). Representative flow data from one such donor. D. Levels of cell surface expression of rTCR appears increased in double TRBC/TRAC knockout compared to TRBC knockout alone as measured by MFI. E. Expression of rTCR (S183-91) in n=3 healthy donors. i. Transduction initially ranged from 31.9-45.6%. ii. Exposure to coBE3 mRNA alone facilitated coupled TRBC knockout. Subsequent eTCRαβ depletion by magnetic beads facilitated enrichment of rTCR expression to 86.2-92.5%. iii. Electroporation of coBE3 mRNA with TRAC ex3 SA presented with high frequency of eTCRαβ disruption. Removal of residual eTCRαβ expressing cells by magnetic bead-mediated depletion did not enrich rTCR expression (32.2%-44.7%). Three colours represent different donors.

### 6.7 Chapter summary

The terminal-CRISPR platform allowed expression of both rTCR chains, while also permitting knockout of either eTCR C regions, using SpCas9 or BE3. Incorporation of murine C regions alleviated the issue surrounding unintended CRISPR genome editing of the rTCR chains, while the addition of an extra disulphide bridge helped mitigate concerns surrounding mispairing of eTCR and rTCR chains. Furthermore, incorporation of murine C domains within the rTCR chains, allowed discriminatory staining for eTCRαβ, facilitating their depletion by antibody coated magnetic beads, while leaving the rTCRαβ expressing cells untouched. Furthermore, disruption of the
endogenous TRAC or TRBC appeared to increased rTCRαβ expression intensity, as well as improving their performance on a number of HCC models of tumour elimination. The addition of a loose TRAC specific sgRNA facilitated the disruption of both TRAC and TRBC loci, further increasing the intensity of the rTCRαβ. Importantly, due to the uncoupled nature of the TRAC sgRNA, knockout of the eTCRαβ complex was no longer coupled to the Terminal-CRISPR vector, and therefore resulted in less homogenous products. These results would indicate that a delivery strategy that coupled both TRBC and TRAC sgRNA to expression of the rTCR would be preferable.
Chapter 7  Discussion

The overarching aim of this project was the optimisation of CRISPR/Cas9 based genome editing technology for the purposes of engineering universal T cell therapeutics, capable of overcoming HLA barriers. Particular focus has been given to the development of a third generation lentiviral vector expressing a transgene from an internal RNA Pol II promoter, and a CRISPR sgRNA expression cassette driven from an RNA Pol III promoter. The native lentiviral U3 region in the lentiviral LTRs contains strong promoter/enhancer elements required for the transcription of viral genomes, however, these sequences are removed from the 3’LTR in SIN gene therapy vectors to prevent vector mobilisation and transactivation effects [275-277]. Removal of these sequences, results in a ΔU3 region in the 3’LTR which is ~350bp shorter than the native. Tactical placement of a sgRNA expression cassette within the ΔU3 region of the 3’LTR facilitates its duplication to the 5’LTR during reverse transcription of the ssRNA lentiviral genome, while also circumventing possible interference effects with the internal RNA Pol II promoter [213, 219]. Use of this terminal-CRISPR vector configuration successfully links sgRNA and transgene expression, allowing robust CRISPR/Cas9 mediated editing of the transduced population upon SpCas9 mRNA delivery.

7.1  Coupled CRISPR/Cas9 editing and CAR19 transgene expression

Engineered T cells have proved a novel treatment option for patients with a variety of malignancies. However, these immunotherapeutic are generally provided in an autologous manner, or from a HLA-matched allogenic donor. However, the wider application of these patient tailored products is somewhat hindered by the need to recover enough starting material from patients, specialist infrastructure for their manufacture, batch to batch variability, and the economic burden they impose. Additionally, the clinical observation of a single leukaemic B cell transduced with a CAR19 expressing lentiviral vector during production of an autologous CAR19 T cell product, presents a potential risk factor for tumour relapse with CD19- escape [278].
Genome editing technologies have provided a route to engineer allogenic non-HLA matched T cells devoid of alloreactive markers, which provide an off-the-shelf universal immunotherapeutic. These gene edited universal T cell products are filtering into clinical phase testing, with their first use in the field of CAR T cells against B cells acute lymphoblastic leukaemia [210]. In that study TALEN genome editing was used to disrupt both TRAC and CD52 loci, with a CAR against the B cell marker CD19 expressed from a third generation lentiviral vector. In this setting gene disruption acts independently of CAR19 expression. Downstream removal of remaining TCRαβ expressing cells by magnetic bead mediated depletion is able to provide purities of >99% TCRαβ- T cells. However, a limitation of this genome editing strategy is the batch variability of the lentiviral transduction, which necessitates the administration of varying total cells doses. Due to the presence of residual TCRαβ expressing T cells in the cell inoculum, the total cell dose is of the utmost importance and can limit the delivery of high CAR T cells doses.

In this area the TT-CAR19 vector developed in chapter 3 has a distinctive advantage in its ability to produce highly purified TCRαβ+/CAR19+ cells, due to the coupling of the TRAC sgRNA expression in the terminal-CRISPR configuration. Additionally, presence of the SpCas9 endonuclease remains highly transient owing to its delivery as mRNA, which appears to be mostly dissipated by day 3 post electroporation. Furthermore, this strategy negates the need to produce bespoke clinical grade TALEN pairs or sgRNA reagents against individual targets, but instead requires only a single SpCas9 mRNA across multiple terminal-CRISPR vector applications. This TT-CAR19 platform has proven directly compatible with the clinically applicable automated manufacturing CliniMACS Prodigy device from Miltenyi Biotec [219](Appendix A).

Additionally, CRISPR edited TCRαβ+/CAR19+ T cells showed superior anti-leukaemic effects when tested in an in vivo model, against CD19+/EGFP+/Luciferase+ Daudi tumour cells compared to their unedited TCRαβ+/CAR19+ effectors. Reduced tumour clearance in the TCRαβ+/CAR19+ effector group may in part be caused by xenoreactive GVHD, upregulating the early exhaustion marker PD1 compared to TCRαβ+/CAR19+ effectors. Similar xenoreactive GVHD has been observed in previous murine pre-clinical studies [279]. Interestingly phenotyping of the remaining Daudi
tumour cells isolated from the bone marrow of animals treated with TCRαβ+/CAR19+ cells showed a clear CD19−/CD20+ population, that was not detected in the CAR19− control group. These results are indicative of clinical reports from patients who experience CD19− relapsed [280]. As this CD19− population does not appear to occur in animals receiving CAR19− effectors, it could be speculated that CD19− tumour populations have a growth advantage in the presence of CAR19+ cell therapeutics. The more rapid tumour clearance in animals receiving the TCRαβ+/CAR19+ appears to have curbed the appearance of the CD19− tumour population. Although further investigation with increased number of animals would be warrant before any definitive conclusions can be drawn.

7.2 U6 RNA Pol III promoter exhibits redundancy of DNA spacer sequence

In addition to allowing efficient TRAC knockout coupled to expression of the CAR19 transgene, the TT-CAR19 vector configuration provided an ideal platform to explore the sequence and architectural characteristics of the U6 promoter used to drive sgRNA expression. To this end a number of U6 variants were designed which contain target deletions of the DNA spacer sequence, or essential promoter elements (DSE/PSE). By cloning of these variants within a TT-CAR19 lentiviral vector configuration, and using SpCas9 mediated TCRαβ disruption as a surrogate for TRAC sgRNA expression, it was determined that while both the DSE and PSE are required for effective promoter activity, the absence of the positional nucleosome DNA spacer sequences did not affect CRISPR/Cas9 knockout. It can therefore be speculated that this spacer sequence situated between the DSE and PSE is not fundamentally required for transcriptional activity from the U6 promoter. Titration of the SpCas9 endonuclease confirming equivalent CRISPR/Cas9 mediated knockout over a range of concentrations, further supports the speculation that this mini-U6 retains sgRNA transcriptional activity despite the removal of the spacer DNA sequence. Removal of this DNA spacer sequence reduced the overall size of the promoter by >50% (249bp compared to 111bp), and as such is referred to as a minimal or mini-U6.

In the DNA loop model of generic type III RNA Pol III promoters, the DNA spacer sequence is wrapped around a nucleosome which acts to bring DSE and PSE bound
transcription factors into juxtaposition, allowing cooperative binding between these proteins [239]. However, Mylinski et al (2001) showed that this model does not apply to all such promoters, and reported on the naturally compact DNA structure of the H1 promoter, with the DSE and PSE in direct proximity with no intervening sequence. Mylinski and colleagues suggest that the close proximity of the DSE and PSE in H1 promoter, allows direct interaction of DSE and PSE-bound transcription factors [237]. Therefore a similar model may hypothetically be applied to transcriptional activity from the mini-U6 promoter developed here. Previous reports have also demonstrated increased transcriptional activity from naked DNA templates containing either the h7SK or U6 promoters devoid of spacer sequence, although these sequences have previously been considered critical for transcriptional activity in the presence of chromatin [238, 281].

Importantly, results in a HEK 293T cell-line supports the functionality of the mini-U6 promoter, when delivered as either integrating lentiviral vector, or more transiently as plasmid DNA. These results confirm that the mini-U6 retains function in different mammalian cell types, while transcriptional activity is not dependent upon the context of the lentiviral LTR. These observations therefore expand the potential applications envisioned for this mini-U6 promoter, particularly in vector systems with limited cargo capacity, such as AAVs. Furthermore these results suggest that removal of homologous DNA spacer sequences from other commonly used type III RNA Pol III promoters, such as the h7SK promoter may be possible and could be applied to reducing the size of short RNA expression cassettes for multiple gene therapy applications.

7.3 Multiplex terminal-CRISPR configuration for simultaneous disruption of two loci

Targeted disruption of the TRAC, or TRBC loci allows the removal of the alloreactive TCRαβ complex from the T cell surface, thereby preventing non-HLA matched CAR T cells from initiating GVHD. However, these allogenic T cell products may still be recognised by the host immune system resulting in premature destruction of the immunotherapeutic cells. Previous reports have added additional disruption of CD52
providing resistance to the lymphodepleting chemotherapy alemtuzumab, and B2M for prevention of HLA class I antigen presentation [123, 210, 220]. In this report, development of multiplex terminal-CRISPR-CAR19 configurations capable of expressing B2M or CD52 sgRNA sequences, in addition to a TRAC specific sgRNA was undertaken. This was initially achieved through two tandem sgRNA expression cassettes within the ΔU3 region of the 3’LTR allowing simultaneous guide expression. Inclusion of two distinct RNA Pol III promoters (U6/H1) in these tandem cassettes reduced repetitive sequences that can lead to molecular inference due to reverse transcription of the lentiviral genome. Despite the use of both U6 and H1 promoters, the results presented in chapter 4 indicated that placing an identical cr1 scaffold sequence in both sgRNA expression cassettes was enough to lead to molecular interference and resulting in absence of the second sgRNA expression cassettes in proviral LTRs. Partial absence of the second sgRNA expression cassette used for transcription of TRAC guide resulted in prominent knockout of B2M compared to TRAC. Recombination events leading to deletion of repetitive sequences during reverse transcription have been well documented in lentiviral vectors, and have been shown to occur between RNA Pol III promoters in shRNA expression cassettes and between CRISPR/Cas9 sgRNA scaffold sequences [221, 240, 243]. The results in this report found that application of an alternative scaffold sequence (cr2) containing no more than 20bp stretches of identical sequence to that of cr1 scaffold appeared to mostly mitigate this issue [243].

Due to potential advantages there have been a number of attempts to insert genes of interest into the lentiviral LTRs. However, it has become apparent that the ΔU3 regions can only tolerate a limited cargo capacity of ~400bp containing non-repetitive sequences before compromising viral titers and transduction efficiency [221]. Applications of both mini-U6 and mini-H1 promoters in the multiplex terminal-CRISPR configuration reduce the overall size in the LTR from 745bp to 426bp and may help mitigate any potential issues with regards to cargo capacity within the LTR. Further multiplexing within this regions or application of larger transcripts, such as those required by other endonucleases, may necessitate use of these minimal RNA Pol III promoters.
Taken together, this system offers versatile, efficient and coupled delivery of a transgene and a combination of two distinct sgRNA sequences, in this case either B2M/ TRAC, or CD52/ TRAC guides. Stringent removal of residual TCRαβ+ cells, and where appropriate HLA class I, enables enrichment of CAR19 expression and generation of a homogenous, well defined cell therapy product. However, disruption of CD52 or B2M both have their limitations for avoidance of a host versus graft effect in the clinical setting. Immunotherapeutic devoid of CD52 expression are well suited to patients receiving the humanised monoclonal antibody alemtuzumab, however, use of this chemotherapy is not suitable for all patients and has been associated with a number of toxic effect [282]. Furthermore, removal of HLA class I prevents the presentation of self-antigens, thereby protecting a non-HLA matched allogenic graft from recognition by the host immune system. However, natural killer (NK) cells have been shown to target HLA class I negative cells, and could reduce the life span of these products in the patients [283]. To avoid NK-mediated rejection, expression of non-classical HLA class I molecules has been employed such as single-chain HLA-E fused to B2M due to its lack of antigen presentation [284]. However, HLA-E does not suppress all subsets of NK cell, and engineered cells would still be susceptible to killing by KIR2DL1-4+ NK cells. Alternatively, non-classical HLA-G can be employed to limit polymorphism and peptide presentation diversity [143, 285, 286]. However, this approach limits delivery methods to those that permit consecutive expression of the non-classical HLA molecule in gene edited HLA class I knockout cells, such as that possinle using the terminal-CRISPR configuration. Furthermore, disruption of HLA class I complex is normally achieved by targeting the non-polymorphic B2M subunit, which has been shown to play a role in iron metabolism [287]. This wider functionality of B2M, may prohibit its disruption in certain cell types. To avoid this issue, other groups have aimed to removal HLA class I by targeted removal of HLA-A, B, and C [286, 288]. To overcome NK cell rejection, while avioding B2M disruption Xu and associaited (2019), present the CRISPR/Cas9 mediated disruption of both HLA-A and HLA-B, but retain HLA-C expression [288]. Retention of HLA-C peptide expression has been shown to allow graft evasion from host CD8 T cell and NK cell rejection, while minimising the need to HLA match the patient with the donor.
As well as generating T cells that are devoid of alloreactive markers, a number of studies have strived to increase the potency/persistence of anti-tumour T cell responses. Targeted disruption of the early T cell exhaustion/activation marker, PD1, has been used to achieve this [123, 232]. Ren et al (2017), observed that CAR T cells with CRISPR-mediated disruption of PD1 expression were able to elicit superior \textit{in vivo} clearance in NSG mice engrafted with leukaemic and solid tumours [232]. Promising preclinical data has led to the first clinical application of CRISPR/Cas9 gene edited T cells incorporating a PD1 knockout [133]. However, genome-wide screening implicated PD1 in the suppression of oncogenic T cell signalling with mono- and bi-allelic deletion being observed in human T cell lymphomas, compelling the need for caution [289]. Furthermore, it has also been noted that lentiviral insertional mutagenesis of the \textit{Tet Methylcytosine Dioxygenase 2} (\textit{TET2}) enhanced the potency of a CAR19 T cell clone in a patient presenting with chronic lymphocytic leukaemia (CLL) [290]. Disruption of \textit{TET2}, appeared to lead to a less-differentiated central memory phenotype, which is believed to possess greater proliferative and anti-tumour potential. Similar observations have since been made in CAR T cells bearing lentiviral insertional mutagenesis of the \textit{Transforming Growth Factor Beta Receptor 2} (\textit{TGFBRII}) gene [291]. Despite these encouraging results, further investigation of the effects of clonal expansion in CAR T cells would need to be undertaken.

7.4 Coupled and uncoupled multiplex cytidine base editing

Genome editing tools including TALENs and CRISPR/Cas9, rely on targeted DNA cleavage and repair by NHEJ which results in the creation of indels leading to gene disruption. The approach has been associated, especially in multiplex applications, with large deletions, p53 upregulation and low frequency chromosomal translocation events, albeit with no attributable adverse effects reported [84, 85, 272-274]. CRISPR/Cas9-based, cytidine deaminase base editing technologies offer the possibility to disrupt gene expression, or correct SNPs with greatly reduced DSB formation. In this body of work a third generation human codon optimised cytidine base editor was used to disrupt gene expression of \textit{TRBC}, \textit{TRAC}, \textit{CD52}, and \textit{B2M} by the creation of premature stop codons or by altering critical mRNA splice sites.
Initial experiments indicated that the terminal-CRISPR configuration was compatible with BE3 and could produce high levels of knockout in primary human T cells. However, several hurdles were encountered when multiplexing this technology. Application of the multiplex terminal-CRISPR vector, appeared to show heavily decreased editing compared to cells treated with SpCas9. Additionally attempts to combine vector expressed sgRNA with uncoupled sgRNA resulted in an inverse correlation between disruption of the two targeted loci. These results implied a competition effect between coupled and loose sgRNA, plausibly due to fewer transcripts from the RNA Pol III-driven sgRNA compared to saturating amounts of loose sgRNA. Importantly, competition between coupled and uncoupled sgRNA was not observed when using SpCas9, implying that this observation was linked to the BE3. Modifications to the scaffold sequence that were reported to increase sgRNA expression and RNP stability appeared to somewhat mitigate these issues, resulting in recovery of high levels of multiplexed genome editing [248]. This is consistent with recently published data, demonstrating an up to 2 fold increase in BE3 activity in plant cells when using this optimised scaffold sequences [292]. Furthermore, the application of the mini-U6 promoter developed in chapter 4 functioned as well as its wild type counterpart, validating its use with base editing technology while indicating that reduced editing from the multiplex vector was not due to the used of minimal promoters. This competition effect has not previously been reported and could be important in serval delivery methods. Moreover, this has currently only been observed with BE3, and it is possible that changing the configuration or the deaminase of this editor could further resolve this issue.

As part of this work, the expression of both SpCas9 and BE3 proteins was assessed by western blot. These results indicated that the BE3 protein is far more transient than that of SpCas9, with BE3 dissipating by day 1 post electroporation whereas SpCas9 can still be detected at day 3. This difference in exposure time could in part be responsible for the differences observed between SpCas9 and BE3 editing in the terminal-CRISPR configuration. This difference in SpCas9 and BE3 protein expression may imply reduced stability of the BE3 fusion protein, although further work would be needed to test this. Alternatively, Grunewald et al (2019) reported that the
rAPOBEC1 used in BE3, was able to edit its own RNA sequence in a sgRNA independent manner [293]. Self-editing in principle could mutate the coBE3 mRNA, preventing prolonged expression of BE3 protein. Sequencing of the coBE3 mRNA in these samples would have to be undertaken to confirm that self-editing occurs in this setting. Furthermore delivery of BE3 mRNA, encoding rAPOBEC1 variants confirmed to have reduced RNA editing effects could help determine whether this transient expression is due to structural instability or self-editing of mRNA transcripts [294-296].

Previously published data from two primary T cells donors treated with BE3 and edited at three genomic loci (TRAC, B2M, PDCD1) was unable to detect large chromosomal translocation events between the three edited sites using ddPCR [123]. In this report, targeted PCR was used to amplify predicted translocations between TRBC, and CD52 loci. Editing with SpCas9 revealed the presence of matching bands for all predicted translocations across three primary T cells donors, with subsequent NGS analysis of these PCR products used to confirm their sequence. As expected these bands were mostly absent from BE3 treated samples. Analysis of these samples by ddPCR was also able to detect translocation events in the SpCas9 treated samples of T1, T2, and T4. No T3 events could be detected in these samples, indicating a discrepancy between the PCR and ddPCR. Of the samples treated with BE3, 2/3 showed no evidence of translocation events by ddPCR, with one donor showing the presence of both T2 and T3 events. These bands were faintly present by PCR on the agarose gel suggesting that these were likely to be genuine events.

The rapid development of tools enabling highly targeted base conversion through deamination effects promises tantalising opportunities, although in depth characterisation of desirable and unwanted effects in subsequent therapeutic applications have to be mapped. Existing CRISPR/Cas9 and Cas12a base-editors employing rAPOBEC1 (including coBE3) are known to mediate transcriptome-wide off-target DNA and RNA deamination in both protein-coding and non-coding regions [293-295, 297]. While these could be problematic, newer variants with more precise DNA restricted editing are already in development and should continue to evolve as ever more efficient, specific and non-toxic editing tools.
The field of base editing is currently undergoing a period of rapid expansion, with the application of different structural layouts, alternative deaminase, as well as evolved, or rationally designed variants of these deaminases. These editors are often developed with a specific niche in mind, including expanding/ narrowing of the editing window, decreased transcriptome-wide RNA editing, reduced off-target DNA editing, decreasing DSB indel formation, increasing expression levels, improved editing in a specific biological system, and changing the preferred editing sequence context [81, 97, 98, 100, 102, 104, 111, 119, 292, 294, 296]. Although expansion of available tools widens the potential applications of base editing technology and increases their safety profile, many of these tools have not been compared, or characterised in parallel which can make it difficult for investigators to choose the most applicable tool for their purposes. Furthermore, at the time of writing there are limited sgRNA design tools available for base editors, with the most recent, BEable-GPS (https://www.picb.ac.cn/rnomics/BEable-GPS/Tool) currently being the most in depth [298]. However, this tool is still limited to CBE’s and does not account for many of the existing variants.

7.5 CRISPR mediated base conversion allows discriminatory depletion of eTCRαβ, and enhanced rTCR (S183-91) potency

T cell immunotherapy against conventional tumour-associated targets such as NY-ESO-1 are being widely investigated, and recent reports indicate autologous T cells with additional CRISPR/Cas9 modifications may deliver improved longevity [133]. The first therapeutic use of autologous T cells modified to express HBV antigen (HBsAg) specific T cell receptors was seen in a subject with chemoresistant, extrahepatic, metastatic disease. In that case, HBV antigens were detectable in HCC metastases but not in donor-derived liver, following transplantation thereby reducing the risk of T cell mediated hepatitis. Gene-modified T cells survived, expanded and mediated a reduction in HBsAg levels and whilst efficacy was not established, there was no significant on- or off-target toxicity [159]. A small number of additional subjects have been treated subsequently, although the approach remains highly patient-tailored and extending to larger numbers of patients is logistically challenging and costly.
Emerging base editor technologies optimised in chapter 5 offer the prospect of highly specific C>T (G>A) base conversion that can be harnessed to create stop codons or modify splice sites to disrupt gene expression, as well as offering the prospect of advanced T cell engineering while minimising the potentially deleterious effect of DSB formation. This report shows the application of rAPOBEC1 BE3 deaminase technology for the generation of engineered T cells, which subsequently undergo precise removal of endogenous TCRαβ and uniformly express rTCRαβ specific for HBsAg (S183-91). The resulting product was homogenous and exhibited enhanced rTCRαβ intensity, greater levels of cytokine production and antigen specific functional integrity in three models of HCC tumour clearance. Discriminatory staining allowed depletion of residual eTCRαβ T cells while rTCRαβ populations remained untouched, providing critical advantages, especially for allogeneic treatment strategies where the eTCRαβ is alloreactive.

Furthermore, the incorporation of murine C domains in these rTCR chains could be potentially immunogenic, although it has been previously reported that it is unlikely that these regions are determinants in the generation of human anti-mouse antibodies [299]. Likewise, the BE3 configuration employs bacterial and rodent derived elements, but given how transiently this fusion protein is expressed during ex-vivo culture is therefore unlikely to be problematic in vivo.

Removal of one or both eTCR chains enhances expression of the introduced rTCRαβ, reduces the risk of aberrant cross-pairing, and allows discriminatory enrichment of engineered T cells. The strategy also opens the door to generating ‘universal’ allogeneic T cells from healthy HLA-mismatched donors by reducing the risk of graft versus host disease. In the case of HBV S183-91 rTCR, cells from healthy HLA-A201 donors could be further edited to disrupt mismatched HLA-B and C molecules while additional multiplexed editing may promote persistence and anti-tumour effects (PDCD1), or confer resistance to immunosuppression (CD52). Ultimately, pre-manufactured banks of eTCR/rTCR⁺ T cells specific for groups of dominant HLA/peptide combinations could provide treatment options for a large number of subjects.
7.6 Concluding remarks

Genome editing technologies have seen significant advancements with the development of a number of new platforms including the RNA guided CRISPR/Cas9 platform that has augmented targeting capabilities and maximised accessibility. The development of the RNA-guided CRISPR/Cas9 system has allowed breakthroughs across broad scientific disciplines through the inherent ability to create precise genomic modifications, generation of model systems, as well as prevention and correction of disease. Most noteworthy, however, has been their successful implementation in the oncology field accelerating clinical treatment development. Rapid advancements in CRISPR/Cas9 genome editing technologies have witnessed the emergence of deaminase-based tools capable of producing specific edits in a seamless fashion. Innovative refinements aim to increase their safety profile and endow them with clinically relevant features. Although this report focuses on the development and optimisation of the terminal-CRISPR configuration with the type II SpCas9 endonuclease, this platform is expected to be readily compatible with other type II or type V CRISPR/Cas species. Clinical applications for these vectors are currently envisioned with the hope to propel the transition of these allogeneic T cells from bench to bedside.
Chapter 8  References


200
119. Liu, Z., et al., Efficient and precise base editing in rabbits using human

120. Zuo, E., et al., Cytosine base editor generates substantial off-target single-

121. Ryu, S.M., et al., Adenine base editing in mouse embryos and an adult mouse

122. Lee, H.K., et al., Cytosine base editor 4 but not adenine base editor generates

123. Webber, B.R., et al., Highly efficient multiplex human T cell engineering
5222.

124. Sanjana, N.E., O. Shalem, and F. Zhang, Improved vectors and genome-wide

125. Wang, W., et al., CCR5 gene disruption via lentiviral vectors expressing Cas9

126. Yu, S., et al., Simultaneous Knockout of CXCR4 and CCR5 Genes in CD4+ T Cells
via CRISPR/Cas9 Confers Resistance to Both X4- and R5-Tropic Human


128. Kim, S., et al., Highly efficient RNA-guided genome editing in human cells via

129. Charlesworth, C.T., et al., Identification of preexisting adaptive immunity to


Supplementary figures

Key:
Red = SPH
Blue = OCT
Green = PSE
Pink = TATA box
Purple = +1 transcription start (G)

A. Wild type U6: 249bp
GAGGGCCTATTCCCATGATTCCCTCATATTGCATATACGATAACAGGGCTGTTAGAGGATAATTAGAATTATTTGAGCTAACAAACAGATATTAGTACAAAAATACGTGACGTAGAAAGTAATTATTTCTTGGGTAATTGCGAGGTTCTTAAAATTATGGGACTATCATATGTTACCCGTAACCTTTGAAAGTATTTCGATTTCGTCTTATGATGTTTTAAAATGGACTATCATATGTTACCGTACCCGTAACCTTTGAAAGTATTTCGATTTCGTCTTATGAT

B. Minimal U6: 111bp
GAGGGCCTATTCCCATGATTCCCTCATATTGCATATACGATAACAGGGCTGTTAGAGGATAATTAGAATTATTTGAGCTAACAAACAGATATTAGTACAAAAATACGTGACGTAGAAAGTAATTATTTCTTGGGTAATTGCGAGGTTCTTAAAATTATGGGACTATCATATGTTACCCGTAACCTTTGAAAGTATTTCGATTTCGTCTTATGATGTTTTAAAATGGACTATCATATGTTACCGTACCCGTAACCTTTGAAAGTATTTCGATTTCGTCTTATGAT

C. TATA alone: 46bp
TTTCGATTTCGTTGAAAGGAGCAGAAGACCG

D. PSE+TATA: 75bp
ATACGATAACATGTTACCTTTGAAAGTATTTCGATTTCGTCTTATGATGCTTCTTTGGGTAATTGCGAGGTTCTTAAAATTATGGGACTATCATATGTTACCCGTAACCTTTGAAAGTATTTCGATTTCGTCTTATGATGTTTTAAAATGGACTATCATATGTTACCGTACCCGTAACCTTTGAAAGTATTTCGATTTCGTCTTATGAT

E. SPH-OCT+TATA: 82bp
GAGGGCCTATTCCCATGATTCCCTCATATTGCATTTTCGATTTCGTTGAAAGGAGCAGAAGACCGAAACCG

Supplementary figure 1 DNA sequences for all tested U6 variants:
A. Annotated sequences of wild type U6 promoter. B. Minimal U6 promoter, devoid of spacer sequences. C. Removal of both DSE, and PSE. D. Removal of DSE. E. Removal of PSE. All sequences are shown in the 5’-3’ orientation.
Supplementary figure 2 Comparison of wild type and minimal U6 promoters in HEK 293T cells:

293T cells were transduced at MOI10 with either the U6B2M<sub>cr1</sub>&gt;H1TRAC<sub>cr2</sub>&gt;CAR19, or miniU6B2M<sub>cr1</sub>&gt;miniH1TRAC<sub>cr2</sub>&gt;CAR19 vectors. Cell were then electroporated with 50µg/ml SpCas9 mRNA using the Lonza 4D nucleofector with program DG130. Mann-Whitney U statistical test revealed no significant difference between B2M knockout, using either wild type of minimal U6 (p = 0.686).
Supplementary figure 3 Comparison of wild type and minimal promoters across a gradient of SpCas9 mRNA concentrations:

n=4 primary T cell donors were transduced with either U6B2M_{cr1}⇒H1TRAC_{cr2}⇒CAR19, or miniU6B2M_{cr1}⇒miniH1TRAC_{cr2}⇒CAR19 multiplex terminal-CRISPR vectors. Subsequent electroporation’s were carried out at 25, 50, or 75µg/ml, followed by quantification of knockout (A) B2M, (B) TRAC, and (C) B2M/TRAC knockout by flow cytometry. HLA class I staining and TCRαβ staining, were used as surrogate markers for B2M and TRAC respectively.
Supplementary figure 4 Multiplex terminal-CRISPR for combined CD52 and TRAC disruption:

A. Representative FACS plots of pre-depletion and post TCRαβ microbead-mediated depletion of miniU6CD52<sub>cr1</sub>\(\rightarrow\)miniH1TRAC<sub>cr2</sub>\(\rightarrow\)CAR19 treated cells. CD52 and TCRαβ knockout is shown in the total CD45<sup>+</sup> population. B. Histogram detailing the makeup of the total CD45<sup>+</sup> population split into double, single, and unmodified cells, pre and post TCRαβ depletion across n=3 donors treated with miniU6CD52<sub>cr1</sub>\(\rightarrow\)miniH1TRAC<sub>cr2</sub>\(\rightarrow\)CAR19 vector. TCRαβ staining, was used as surrogate marker for TRAC disruption.
Supplementary figure 5 Comparison between wild type U6 and mini-U6 promoter:

A. n=3 primary T cell donors transduced with TTRBC_{CR1}-CAR19 either expressing the TRBC sgRNA from a wild type U6 or a mini-U6 promoter, electroporated with 50µg/ml coBE3 mRNA alone, or with 20µg/ml CD52 ex1 SD sgRNA. TRBC knockout appears similar between wild type U6 and mini-U6 promoters with coBE3 alone, and coBE3 + CD52 ex1 SD sgRNA. B. Similarity can be observed in the proportion of double negative cells when using the wild type U6 compared to the mini-U6. TCR\(\alpha\beta\) staining, was used as surrogate marker for TRBC disruption.
Supplementary figure 6 Translocation events between TRBC and CD52 loci:

A. Next generation sequencing analysis of PCR produces seen in Figure 5-14 D, in the donor 1 SpCas9 treated samples. Alignment of sequencing data confirms that bands observed on the agarose gel are the predicted translocation events. Measured at % NGS reads that align to the predicted sequence. B. Quantification of translocation events by digital droplet PCR in n=3 unmodified, SpCas9, or coBE3 treated primary T cell donors. Red boxes indicate outlier values seen in the donor 3 coBE3 treated sample.
Appendices


CRISPR mediated base conversion allows discriminatory depletion of endogenous T cell receptors for enhanced synthetic immunity.

Authors: Roland Preece, Andrea Pavesi, Soragia Athina Gkazi, Kerstin A. Stegmann, Christos Georgiadis, Zhi Ming Tan, Jia Ying Joey Aw, Mala K. Maini, Antonio Bertoletti, Waseem Qasim

1 Molecular and Cellular Immunology Unit, UCL Great Ormond Street Institute of Child Health, London, WC1N 1EH
2 NIHR Great Ormond Street Hospital Biomedical Research Centre, 30 Guilford Street, London WC1N 1EH
3 Institute of Molecular and Cell Biology (IMCB), Agency for Science, Technology and Research (A*STAR) 61 Biopolis Drive, Singapore 138673
4 UCL Division of Infection and Immunity, The Rayne Building, 5 University Street, London, WC1E 6EJ
5 Program Emerging Infectious Diseases, Duke-NUS Medical School, Singapore.
6 Singapore Immunology Network (SigN), Agency of Science technology and Research (ASTAR), Singapore
Abstract

Background: Emerging base editing technology exploits CRISPR RNA guided DNA modification effects for highly specific C>T conversion and this can be used to efficiently disrupt gene expression in human cells. These tools may enhance synthetic T cell immunity by tightly restricting specificity, addressing HLA barriers and promoting persistence. We report coupled lentiviral delivery of a Hepatitis B virus specific recombinant TCR (rTCR) and an RNA guide for simultaneous cytidine deamination mediated disruption of endogenous TCRs (eTCR).

Methods: Activated T cells were stably modified using a third generation self-inactivating vector encoding a HBV specific rTCR and a CRISPR RNA guide expression cassette under the control of a U6 PolIII promoter embedded in the vector 3'LTR. Codon optimised base editor-3 (coBE3) was supplied as mRNA by electroporation. Selective removal of residual eTCR expressing cells by magnetic bead depletion provided a product enriched for rTCR+ T cells.

Results: Coupled expression of rTCR and discriminatory depletion of eTCR resulted in enrichment of HBV specific populations from 55.6% (SD ±4.9%) to 94.8% (SD ±1.2%). Intensity of rTCR expression increased between 1.8-2.9 fold (compared to cells retaining their competing eTCR) and cells exhibited increased cytokine production, migration and killing of a HBV antigen-expressing hemoma cell line in a 3-D microfluidic model. Molecular signatures confirmed seamless conversion of C>T (G>A) had created a critical stop codon in the TCR beta chain constant gene (TRBC), with no notable activity at predicted off-target sites.

Conclusions: Targeted disruption of eTCR by cytidine deamination and discriminatory enrichment of antigen specific T cells offers the prospect of enhanced, more targeted T cell therapy, modelled here against Hepatitis B associated hepatocellular carcinoma. The base conversion platform is ideally suited for multiplexed editing to address wider barriers and the generation of a universal cell therapy.
Lay Summary

White blood cells called T cells mediate powerful antiviral effects that can be used to target liver cancers linked to Hepatitis B virus infection. We report new techniques that change the DNA code in T cells and reprogram them to only recognise cells that show a particular Hepatitis-B flag on their surface. Ultimately such approaches could allow banks of healthy donor T cells to be created and used in multiple patients.
Introduction

T cells redirected with recombinant T cell receptors (rTCR) are being widely investigated (1-5). Limitations include unpredictable ‘off-target effects’ due to TCR cross reactivity (for example cardiac toxicity following therapy with MAGE-A3 rTCR) and concerns that endogenous TCR α and β chains may miss-pair with rTCR chains and give rise to novel dimeric complexes with unpredictable specificities (3, 6-13). These limitations may have been somewhat mitigated using models to predict rTCR cross-reactivity and TCR mispairing has been addressed by promoting exclusive rTCR pairing via additional disulphide bonds, use of high-affinity TCRs, murinisation of rTCR constant regions, hybrid TCR-CAR molecules, as well as development of single chain TCRs through covalently linking the variable domains (14-23).

Also of note is the importance of rTCR assembly on the cell surface as a multimeric complex with CD3 chains, as competition from the endogenous TCR (eTCR) for the shared components can limit cell surface expression (24). Competition for such cellular components can be addressed either by overexpression of CD3, disruption of eTCR by RNA interference, or nuclease mediated genetic disruption of eTCR chains. Previously, zinc finger nucleases (ZFNs), transcription activator-like effector nucleases (TALENs), and clustered regularly interspaced short palindromic repeats (CRISPR)/ CRISPR-associated protein 9 (Cas9) have all been used to disrupt one or both TCRαβ/CD3 (13, 25-29). These genome editing approaches also reduce the likelihood of mispairing, but existing nuclease based approaches all result in double stranded DNA breaks and may create large indels, trigger translocation events, and increased activation of p53 pathways (30-34). Recently, the first report of autologous anti-tumour therapy with T cells edited using Cas9 to disrupt TCR expression and PD1 noted readily detectable chromosomal translocations in the infused products (35), and similar aberrations were found after TALEN editing of T cells modified to express anti-CD19 chimeric antigen receptors (36).

We now report the application of emerging APOBEC mediated cytidine deamination (base editing) technology for efficient and seamless base conversion to introduce a premature stop codon in homologous regions of T cell receptor beta constant 1 and 2 (TRBC 1/2) chain (37, 38). Here a CRISPR guided nickase Cas9 (D10A), fused to a rat APOBEC1 deaminase, operates within a 13-17 bp window proximal from the protospacer adjacent motif (PAM) sequence. The inclusion of a C-terminus fusion comprising a uracil DNA glycosylase inhibitor (UGI)
(derived from Bacillus subtilis bacteriophage PBS1) inhibits uracil DNA glycosylase and blocks uridine excision and promotes conversion to thymidine as cells replicate. High levels of C>T conversion and low levels of Indels have been reported for this third generation base editor (BE3) (39-41). Here we investigated BE3 in the context of engineering T cells against Hepatitis B virus surface antigen, an important target in the treatment of hepatocellular carcinoma (HCC) (42, 43). HBV viral antigens are processed and presented by major histocompatibility complex (MHC) molecules on the surface of infected cells (44, 45), and naturally occurring HBV-specific T cells, can engage with peptides presented in the context of HLA, to moderate viral and tumour burdens (46, 47). Nevertheless, such HBV specific T cell responses can become exhausted during chronic HBV infection (48-50) and synthetic HBV-specific T cells can be generated through the expression of recombinant T cell receptors (rTCR) (51-55). The approach has already been tested clinically in HBV associated HCC, (1, 56) with further studies planned.

Lentiviral vector delivery of an rTCR specific for Hepatitis B peptide S183-91 presented by HLA-A2, incorporating murine constant regions, and coupled to a CRISPR single guide RNA (sgRNA) targeting TRBC1/2 resulted in high levels of targeted cytidine deamination after transient delivery of mRNA encoding coBE3. Thereafter, discriminatory removal of residual eTCR+ cells was achieved using magnetic bead mediated depletion using the anti-human TCRαβ monoclonal antibody. Consequently rTCR expression is enriched, as the murine constant regions lack the specific epitope recognised by this antibody. Phenotypical and functional assessments, including migration and killing effects in a 3-D microfluidic model verified immunotherapeutic effects following genome editing and molecular analysis of both DNA and RNA effects was performed to examine editor effects.
Materials and Methods

CRISPR guide RNA:
Guide sequences compatible with coBE3 targeting homologous sequences in TRBC1 and 2 were designed using the CRISPR design tool, Benchling (https://benchling.com) and provided an on-target editing score for predicted activity for each cytidine around the editing window (39). TRBC1/2: C_{0.8} C_{1.7} A_{21.9} C_{21.4} AGCUCAGCUCCAG (anti-sense, numbers indicate predicted editing scores for the specific cytidine base). Predicted exonic off-target binding requires at least 3 mismatches within the protospacer, and a full list of potential off-target sites are provided in Supplementary Tables 1.

Lentiviral construct for rTCR and sgRNA delivery:
Lentiviral design for coupled transgene and guide RNA expression has been previously described (57). Briefly, rTCR HLA-A0201/HBs183-91 was cloned under the control of an internal PKG promoter and a CRISPR guide expression cassette was embedded in the lentiviral 3’LTR. This comprised a 5’ Pol III promoter (U6) a sgRNA specific for TCRB1/2. Vector stocks were produced in 293T cells by transient transfection with third generation packaging plasmids and concentrated by ultracentrifugation prior to storage at -80°C.

Primary human lymphocyte culture and modification:
Peripheral blood mononuclear cells (PBMCs) were isolated by ficoll density gradient and subsequently activated with TransACT reagent (130-111-160, Miltenyi Biotech) at 10ul/ml. TexMACS medium (130-097-196, Miltenyi Biotech) with 3% human AB serum (GEM-100-512-HI, Seralabs) and 100U/ml Proleukin IL-2 (Novartis), was used for all lymphocyte cell culture. Transduction with lentiviral vector was performed 24hr post activation at a multiplicity of infection (MOI) of 5. Electroporation of coBE3 mRNA was performed at day 4 post activation, after which cells were cultured in a G-Rex®10 (P/N 80040S, WILSONWOLF). Lymphocytes were cultured for 11 days post activation and magnetically depleted using anti-TCR a/b-biotin (130-098-219, Miltenyi Biotech) followed by incubation with anti-biotin microbeads ultrapure (130-105-637, Miltenyi Biotech) and separation through LD columns (130-042-901, Miltenyi Biotech). Cells were rested overnight before phenotyping and cryopreservation.

Phenotyping Flow cytometry
Flow cytometry was performed on a 4-laser BD LSRII (BD Biosciences), with subsequent analysis executed using FlowJo v10 (TreeStar). Cells were stained according to manufacturer instructions with Mouse TCR β constant-APC (Clone H57-597, Biolegend, Cat 109211), Human TCRα/β-PerCP vio 700 (Clone REA652, Miltenyi Biotec, Cat 130-113-540), PD1-PE (Clone PD1.3.1.3, Miltenyi Biotec, Cat 130-117-384), CD4-VioBlue (Clone REA623, Miltenyi Biotec, Cat 130-114-534), and CD45-VioGreen (Clone REA747, Miltenyi Biotec, Cat 130-110-638).

**Antigen specific responses**

Target (T) HepG2 cells were pulsed with HBV env\textsubscript{183-191} (FLLTRILTI, JPT) and core\textsubscript{18-27} (FLPSDFFPSV, JPT) peptide at gradient concentrations for 1h at 37°C. Cryopreserved effector (E) T cells (eTCR+/rTCR-, eTCR+/rTCR+, and eTCR-/rTCR-) were thawed and cultured at E:T ratio of 1:1 and 0.1µg/ml Brefeldin A (Sigma) was added before overnight coculture. A Fortessa X20 flow cytometer (BD) and using FlowJo v10 (TreeStar) was used to analyse phenotype and function of effector T cells groups, and included intracellular staining with TNFα FITC (clone MAb11, BD biosciences, Cat 502906), MIP-1β PE (clone D21-1351, BD biosciences, Cat 550078), IL-2 PerCP-eFlour710 (clone MQ1-17H12, eBioscience, Cat 46-7029-42), GranzymeB AF700 (clone GB11, BD biosciences, Cat 560213), IFNg V450 (clone B27, BD biosciences, Cat 560371), and surface staining with CD3 BUV395 (clone UCHT1, BD biosciences, Cat 563546) and mouse TCR β constant-APC (Clone H57-597, Biolegend, Cat 109211)

**Electroporation of base editor mRNA**

CleanCap™ coBE3 mRNA (TriLink biotechnologies) incorporated one nuclear localisation signal (C terminus), had a Cap 1 structure, and was purified by high-performance liquid chromatography (HPLC). Electroporation used a Neon transfection system (ThermoFisher Scientific) and following electroporation T cells were subject to overnight incubation at 30°C before restoration to 37°C.

**Molecular characterisation of ‘on-target’ DNA editing**

Genomic DNA extraction was performed using DNeasy Blood and Tissue Kit (69504, QIAGEN) and PCR sequencing undertaken using primers for TRBC1/2. TRBC forward: AGGTCGCTGTGTTTGAGC, TRBC reverse: CTATCCTGGGTCCACTCGTC. Sanger sequencing data
(Eurofins Genomics) was analysed using EditR ([https://moriaritylab.shinyapps.io/editr_v10/](https://moriaritylab.shinyapps.io/editr_v10/)) (58). In addition amplified products were library prepped for NGS using a Nextera XT kit (Illumina, Cambridge, UK). After the library preparation, individually barcoded samples were pooled in a MiSeq using a 500-V2 nano-cartridge. Demultiplexed fastq files were uploaded to Galaxy (59) for trimming and alignment. NHEJ signatures were analysed using Pindel (60), haplotypes were analysed using Freebayes (61). Figures were created in R.

### Screening for rTCR RNA editing effects

Total RNA was extracted using a QIAamp RNA Blood Mini kit (Qiagen, 52304) for TCR library preparation and sequencing as previously described (62, 63). rTCR RNA was reverse transcribed using a murine TRBC specific primer (TGGACTTCTTTGCGTTGAC). Following ligation of an oligonucleotide containing the Illumina SP2 primer and unique molecular identifiers (UMI), products were amplified using primers specific to the murine constant alpha and beta chains (CGTTGATCTGGCTGTCGAA and TTGACCCACCAAGACAGCTC, respectively). Finally, libraries were built in two further steps of amplification during which the SP1 sequencing primer, indices and Illumina adaptors are added. Part of the primers used in these were also specific for the constant regions (ACACTCTTTCCCTACACAGCTCTTCCGATCTNNNNNNGCCAATGCACGTTGATCTGGCTGTCGAA & ACACTCTTTCCCTACACAGCTCTTCCGATCTNNNNNNGCCAATCCGTTGACCCACCAAGACAGCTC). The final purified libraries were verified using Tapestation (Agilent) and Qubit (Thermo Fisher Scientific), multiplexed and sequenced on a MiSeq system (Illumina) using 500-V2 cartridges (Illumina). Fastq files were demultiplexed using Demultiplexor ([https://github.com/innate2adaptive/Decombinator](https://github.com/innate2adaptive/Decombinator)) (64). Using Galaxy tools (59), the demultiplexed fastq files were trimmed (Trim Galore and Trimmomatic) and aligned (Bowtie2) to the relative TCR HBV gene map. Aligned files were interrogated for the frequency of the reference sequence per base around the CDR3 (100bp total window).

### Xcelligence impedance assay:

Target HepG2 cells were seeded (100k per well) in the dedicated device (E-Plate VIEW 16, ACEA Biosciences Inc.) and let them grow for 24h. Impedance measurement was acquired with an interval of 15 minutes by an array of electrodes located at the bottom of the plate.
Different T cell preparations and Effector-to-Target ratios (E:T) were added in the well after 24h, and the impedance signal was recorded for the subsequent 72h. Three different donors were tested in triplicate conditions.

**3D microfluidics device:**

Briefly, dissociated HepG2-Env-GFP target cells were mixed with collagen type I gel (rat tail, Corning) and injected into the dedicated region of the 3D cell culture chip (DAX-1, AIM Biotech), before gel polymerization, following a previously developed protocol (49, 54, 65). R10 media with 3µM of DRAQ7 (Biolegend) cell-impermeable nuclear dye was then added to the media channels to hydrate the gel, and chips were incubated at 37°C. T cells were stained with 3µM Cell-Tracker Violet BMQC (Thermo Fisher Scientific) and were injected into one of two media channels flanking the gel region before overnight incubation. 3D confocal images were acquired daily with a high content imaging system (Phenix, Perkin Elmer). T cells from the liquid channel were collected by manual pipetting; after, collagenase solution was injected into the device to retrieve the immune cells migrating in the hydrogel region for flow cytometer analysis on a 4-laser BD LSRII (BD Biosciences).

**Statistics**

One-way ANOVA, comparing the mean from all groups against each other, followed by a Tukey’s multiple comparison post-hoc test was used to determine significance between groups in Figure 1F ii, and Figure 3D. Two-tailed impendent t-test was used to assay difference in data distribution in Figure 4B. Statistical analysis was performed using GraphPad Prism software, version 8.0.0.
Results.

Base conversion disrupts endogenous TCR expression and allows enrichment of T cells expressing rTCR.

A third generation self-inactivating (SIN) lentiviral vector encoding both a HBV envelope specific rTCR (HLA-A0201 restricted, S183-91 (FLLTRILT) (66)) and a sgRNA specific for TRBC 1/2 was generated. The guide cassette was embedded in the 3’LTR of the vector with a U6 type III RNA polymerase promoter (Figure 1A) as previously described (57). Upon electroporation of codon optimised (co) base editor (BE3) mRNA, the sgRNA mediated highly targeted base conversion of duplex cytidine nucleotides within exon 1 of TRBC 1/2 loci. Single or double stop codons were thus generated within a 13-17bp window proximal to the nCas9 (D10A) PAM sequence (Figure 1B). Disruption of endogenous TCR β chain expression eliminated expression of eTCRαβ, and the inclusion of murine constant regions within the rTCR addressed the possibility of aberrant cross pairing of rTCRβ with endogenous TCRα (Figure 1C). Healthy T cells were readily activated and transduced with 50-60% rTCR expression (Figure 1E and F i eTCR+/rTCR±) and exposure to coBE3 resulted in disruption of eTCR expression and simultaneous emergence of rTCR+ populations, increasing in proportion to approximately 60-65% of the cultures (Figure 1 E and F i eTCR±/rTCR±). Furthermore, because eTCR was amenable to detection by anti-TCRαβ monoclonal antibody, magnetic bead mediated depletion of residual eTCRαβ expressing cells was possible. Notably, rTCR (constructed with murine C domains) was not susceptible to these reagents and thus at the end of production, cells could be enriched as eTCR-/rTCR+, resulting in a highly homogenous product (>99% eTCR-/ 95.9% rTCR+) (Figure 1E and F i). There was also a significant increase in the mean florescence intensity (MFI) of rTCR in eTCR-/rTCR+ cells compared to eTCR+/rTCR±, suggesting enhanced cell surface expression of rTCR in the absence of eTCR which may otherwise have competed for CD3 chains during assembly (one way ANOVA, P=0.0162) (Figure 1Fii).

Hepatitis B antigen specific responses of eTCR-/rTCR+ T cells

Three different in vitro assessments of antigen specific functionality were undertaken. Firstly, production of cytokines including interferon-γ, TNFα, IL-2 and CCL4 was determined by flow cytometry in T cells responding to HepG2 cells pulsed with the irrelevant control peptide (HBV core C18-27, FLPDFFPSV) or a gradient of Hepatitis B target surface envelope peptide (S183-
91, FLLTRILT) concentrations. In all three donors tested, cytokine production was higher in eTCR-/rTCR+ T cells in response to target S183-91 peptide (Figure 2A i and ii), with no notable response to the control C18-27 peptide (data not shown). Next, we investigated effector function at different target:effector ratios in a previously described XCelligence impedance assay and calculated the relevant cell indices over 96 hours. An increased index indicated HepG2 target cell proliferation, whereas cell death or apoptosis resulted in a reduced index, signifying higher levels of effector T cell activity (Figure 2B). Control groups included target cells alone (HepG2 Only), and non-transduced effectors, where as expected there was a progressive increase and plateau in index. In contrast, both effector groups exhibited a transient rise and then decline in index, with more rapid reductions mediated by eTCR-/rTCR+ cells compared to eTCR+/rTCR± T cells at all E:T ratios (Figure 2C i). Overall effector function was calculated by area under the curve as shown in Figure 2C ii, reflecting the increased cytotoxicity of the eTCR-/rTCR+ effector cell group compared to their unedited counterparts (eTCR+/rTCR±).

Finally, migration and target cell killing by engineered T cells was determined in a 3D microfluidics device. The system captured migration of effector T cells from a fluidics channel to a collagen gel embedded with target PreS1-GFP-HepG2 cells (Figure 3A). Phenotyping of effector T cells confirmed rTCR expression (Figure 3B) and minimal cytokine expression in the absence of stimulation after thawing. Cytokine expression profiles were compared between cells recovered from inside or outside the gel area (Figure 3C). Both eTCR+/rTCR± and eTCR-/rTCR+ effector groups presented higher levels of IL2, INFγ, and TNFα expression within the gel consistent with migration of rTCR+ T cells to their targets. Killing of HepG2 cells by eTCR-/rTCR+ cells was confirmed within 24hours whereas eTCR+/rTCR± cells at this time point were comparable to control eTCR+/rTCR- indices and the HepG2-alone control groups (Figure 3D). Direct visualisation revealed greater clearance of HepG2 cells after co-culture with eTCR-/rTCR+ T cells (Figure 3E).

**Molecular characterisation of base editor effects**

The application of novel genome editing tools has necessitated further investigation of anticipated and unexpected molecular consequences after cell engineering. There is an established experience of lentiviral mediated effects, including their propensity to integrate into transcriptionally active genes (67-69) and we did not re-examine these aspects. However,
the effects of coBE3 were characterised in depth. Firstly, direct sequencing of ‘on-target’
editing at TRBC 1/2 loci in TCRαβ depleted eTCR-/rTCR+ T cells was undertaken and analysed
using EDITR, with cytidine at positions 15 and 16 proximal to the PAM of particular interest
(Figure 4A i and ii). High levels of C>T conversion (G>A sense strand) were captured at these
positions (24 ± 3.7% and 37 ± 6.8% at C15 and C16 respectively), with remarkably little activity
at other nearby C residues (4.3 ± 3.1% C18, 2.3 ± 1.7% C19, and 5 ± 2.2% C20). Next generation
sequencing revealed similar levels of C>T conversion at both positions C15 (32.3 ± 5.2%) and
C16 (40 ± 5% at C16). Although mostly seamless, a minority of reads exhibited small (<10bp,
8.4 ± 2.5%) or large (10-100bp, 8.2 ± 1.2%) indels, and these were the likely signatures of NHEJ
following DNA breakage and repair in a minority of cells (Figure 4Aiii) as others have noted
previously (39-41).

In-silico analysis of sgRNA binding and possible off target activity was undertaken. Six genomic
loci with the highest scores for off-target activity were interrogated directly by NGS in three
different donors (Figure 4B). We found very low levels (<1%) of conversion activity at these
sites, and only one intronic site exhibited C>T changes higher than in respective control, non-
edited, sample (Figure 4B). Recent reports in cell lines have also suggested that promiscuous
rAPOBEC1 RNA deamination (including by BE3) can arise following plasmid mediated
expression of base editors (70-72). In the T cell context, and with coBE3 transiently expressed
by mRNA electroporation we investigated if regions directing antigen receptor specificity
might be affected. Analysis of RNA from T cells exposed to coBE3 focussed on high throughput
interrogation of TCR hypervariable regions (TCRVα and TCRVβ CDR3 regions). Analysis of
samples collected at serial time points, from 24hrs to 8days, found no obvious evidence of
aberrant deamination compared to controls (99-100% cysteines unmodified) and intact
sequence integrity of HBs183-91 rTCR was verified (Figure 4C). Thus, while on-target
deamination and creation of TCR-stop codons was highly efficient, there was no notable
activity at sites of potential interest at either the DNA or RNA level for coBE3.
Discussion
T cell immunotherapy against conventional tumour-associated targets such as NY-ESO-1 are being widely investigated, and recent reports indicate autologous T cells with additional CRISPR/Cas9 modifications may deliver improved persistence and longevity (35). Emerging base editor technologies offer the prospect of highly specific C>T (G>A) base conversion that can be harnessed to create stop codons or modify splice sites to disrupt gene expression, and offer the prospect of advanced T cell engineering.

We previously reported the first therapeutic use of autologous T cell modified to express HBsAg specific T cell receptors in a subject with chemoresistant, extrahepatic, metastatic disease. In that case, HBV antigens were detectable in HCC metastases but not in donor derived liver (following cadaveric liver transplantation) thereby reducing the risk of T cell mediated hepatitis. Gene-modified T cells survived, expanded and mediated a reduction in HBsAg levels and whilst efficacy was not established, there was no significant on- or off-target toxicity (1). A small number of additional subjects have been treated subsequently, although the approach remains highly patient-tailored and extending to larger numbers of patients is logistically challenging and costly. Similar hurdles are being addressed in the arena of haematological malignancies through the generation of ‘universal’ T cells expressing chimeric antigen receptors (CARs) from non-HLA matched healthy donors. As such, depletion of endogenous TCR and other antigens by genome editing has allowed HLA barriers to be circumvented, and ongoing trials suggest that such universal CAR T cells can expand and persist sufficiently to induce molecular remission (36). The editing tools applied to date have included TALENS and CRISPR/Cas9, and rely on targeted DNA cleavage and repair by non-homologous end joining (NHEJ) which results in the creation of indels leading to gene disruption. The approach has been associated, especially in multiplex applications, with large deletions, p53 upregulation and low frequency chromosomal translocation events, albeit with no attributable adverse effects reported (30-34). Application of CRISPR guided base conversion to create stop codons or alter critical splice site to disrupt gene expression offers the possibility of seamless gene disruption with greatly reduced likelihood of translocation events. We report the application of APOBEC deaminase technology for the generation of engineered T cells, which are then rendered devoid of endogenous TCRs and uniformly express rTCR specific for an epitope in HBsAg. The resulting product was homogenous and
exhibited enhanced rTCR intensity, greater levels of cytokine production and antigen specific functional integrity in models of HCC elimination. An ability to discriminate and selectively process and deplete eTCR T cells while rTCR populations are untouched provides critical advantages, especially for strategies when allogeneic donor cells bearing potentially alloreactive eTCRs can be eliminated. Non-human protein sequences within constructs have the potential to be immunogenic, although murine TCR constant regions are considered unlikely determinants in the generation of human anti-mouse antibodies (73). Likewise, the BE configurations employ bacterial and rodent derived elements, but expression is transient during ex-vivo culture and unlikely to be problematic in vivo.

The rapid development of tools enabling highly targeted base conversion through deamination effects promises tantalising opportunities, although in depth characterisation of desirable and unwanted effects in subsequent therapeutic applications have to be mapped. Existing CRISPR/Cas base-editors employing rAPOBEC1 (including coBE3) are known to mediate off-target DNA edits and transcriptome-wide RNA deamination in both protein-coding and non-coding regions (70-72, 74). While these could be problematic, newer variants with more precise DNA restricted editing are already in development and should continue to evolve as ever more efficient, specific and non-toxic editing tools. Our analysis of possible off-target sgRNA activity in three donors found minimal base conversion effects at the top six predicted DNA sites, of which five were intronic. Importantly, examination of RNA found only very minor perturbations and C>U conversions of the CDR3 variable regions no greater than in control cells. Such changes could otherwise redirect the specificity of the introduced TCR, and would risk causing autoimmunity or off-target T cell effects.

Conclusion:
Removal of eTCR enhances expression of introduced rTCR, reduces the risk of aberrant cross-pairing, and allows discriminatory enrichment of engineered T cells. The strategy also opens the door to generating ‘universal’ allogeneic T cells from healthy HLA-mismatched donors by reducing the risk of graft versus host disease. In the case of the rTCR specific for HBs183-91, blood from healthy HLA-A201 donors could readily be further edited to disrupt mismatched HLA molecules creating immunologically stealthy cells. Additional multiplexed editing of T cell exhaustion markers may promote enhanced persistence and anti-tumour effects. Ultimately,
pre-manufactured banks of eTCR-/rTCR+ T cells specific for groups of dominant HLA/peptide combinations could provide treatment options for large numbers subjects.

Disclosures

WQ holds interests unrelated to this project in Autolus Ltd and Orchard Therapeutics.
WQ received unrelated research funding from Cellectis, Servier, Miltenyi, Bellicum.

Funding:

Supported by NIHR (RP-2014-05-007), NIHR Blood and Transplant Research Units (BTRU) and Great Ormond Street Biomedical Research Centre (IS-BRC-1215-20012). National Research Foundation (NRF-CRP17-2017-06) Singapore.
The views expressed are those of the author(s) and not necessarily those of the NHS, the NIHR, or the Department of Health.
Figure 1: Vector design for generation of eTCR-/rTCR+ cells using coupled base editing. A) SIN lentiviral genome, expressing rTCR against the HBV S183-91 envelope antigen, and a sgRNA targeting homology regions of TRBC 1 and 2 genes in the 3’LTR (TTRBC-S183-91 rTCR). rTCR comprised human TRAV34 and TRVB28 variable regions fused with murine constant regions.

B) Reverse Transcription and Genomic Integration: LTR duplication (3’U3/CRISPR cassette is duplicated in the 5’LTR).

C) Unmodified T cells: eTCR+/ rTCR-
Lentiviral transduced: eTCR+/ rTCR-
Lentiviral transduced TRBCKD T cells: eTCR+/ rTCR+

D) Electroporation with coBE3 mRNA
Endogenous TCR:
Possible recombination:

E) Unmodified eTCR+/ rTCR-
TTRBC-S183-91 rTCR+ eTCR+/ rTCR±
coBE3 mRNA eTCR±/ rTCR±
eTCR depletion eTCR-/ rTCR+

F) i. TTRBC-S183-91 rTCR+ eTCR±
   coBE3 mRNA eTCR±/ rTCR±
   eTCR depletion eTCR-/ rTCR+

   ii. 
   
   iii. 

   Figure 1: Vector design for generation of eTCR-/rTCR+ cells using coupled base editing. A) SIN lentiviral genome, expressing rTCR against the HBV S183-91 envelope antigen, and a sgRNA targeting homology regions of TRBC 1 and 2 genes in the 3’LTR (TTRBC-S183-91 rTCR). rTCR comprised human TRAV34 and TRVB28 variable regions fused with murine constant regions.
Both chains were expressed as a single transcript using a self-cleaving P2A sequence. B) Schematic representation of BE3 editing of exon 1 of TRBC 1/2 loci. Editing window (Blue) of the BE3 ranged from 13-17bp proximal of the PAM (Red) with conversion of Tyr codons to create premature stop codons. C) Theoretical TCR chain pairing when introducing a rTCR. D) Schema of cell production. Human peripheral blood lymphocytes were isolated and activated with Transact (anti-CD3/CD28) before transduction and electroporation. After overnight hypothermic culture, cells were expanded in GReX flasks before cryopreservation on day 14. E) Flow cytometry phenotyping of unmodified and TTRBC-S183-91 rTCR transduced cells. Delivery of coBE3 mRNA by electroporation caused reduction of eTCR expression (38.1%) and emergence of eTCR-/ rTCR+ cells (Red box). Magnetic bead mediated depletion of residual eTCR+ T cells enriched eTCR- populations, resulting in >99% eTCR-/ 95.9% rTCR+. F) Expression of S183-91 rTCR in three healthy donors. i) Transduction initially ranged from 59.8-63.9% in cells exposed to both vector and BE3, and after enrichment of gene edited cells, rTCR levels increased to between 93.9-96.1%. Three colours represent different donors. ii) Levels of cell surface expression of rTCR increased in eTCR-/rTCR+ compared to eTCR+/rTCR± cells (one way ANOVA, p=0.0162).
Figure 2: Cytokine anti-HepB responses of rTCR T cells after base editing and removal of eTCR. A) Cytokine responses of effector T cells to HepG2 cell line pulsed with target HBV surface peptide (S183-91, FLLTRILT), and appropriate control HBV core peptide (c18-27
FLPSDFPSV) (n=3). i) Histograms of TNFα responses to HepG2 target cells pulsed with 1µM target peptide (Gated on rTCR+CD8+). ii) Cytokine responsiveness across three donors at difference concentrations of target peptide (S183-91). Control peptide (C18-27) was used to ensure specificity of response (Gated on rTCR+CD8+, unmodified cells gated on rTCR-CD8+).

B) Schematic depiction of XCelligence impedance assay showing cancer cells (green) seeded in wells with micro electrode array (yellow), in the present of effector T cells (blue). Where T cells recognise cancer cells, this leads to cell death (brown) and reduced impedance resulting in lower cell index values, and AUC. C) XCelligence data from three health donors across different effector: target ratios (1:1, 1:2, and 1:4). (i) Presents cell index over time, all donors showed increased cell index with decreased effector: target ratio. Both HepG2 alone (Red), and eTCR+/rTCR- (Orange) show steadily increasing cell index over time. Whereas eTCR+/rTCR± (Purple) and eTCR-/rTCR+ (Green) groups show an initially increased cell index, followed by a marked decline. (ii) Summary data of area under the curve (AUC). Increased AUC values were observed at the lower effector: target ratios, with the eTCR-/rTCR+ consistently presenting with the lowest AUC values.
Figure 3: Effector T cell migration and target cell killing in a 3D microfluidics device. A) Schematic of 3D microfluidics device, detailing migration of effector T cells (Grey) from the fluidics channel to the adjacent collagen gel containing target PreS1-GFP-HepG2 cells (Green), and inducing cell death (Red). B) Flow cytometry based phenotyping of effector T cells used for tumor killing. C) Cytokine expression (%). D) Normalised % of dead cells. E) Fluorescence images showing live and dead HepG2 cells.
in this assay post cryopreservation (Gated on CD45+). C) Cytokine responsiveness of effector T cells isolated from either outside the collagen gel (Grey), or inside the collagen gel (Blue) after 24hrs (Gated on rTCR+CD8+, unmodified cells gated on rTCR-CD8+). D) Normalised killing of target HepG2 cells in response to effector T cells groups. HepG2 alone (Orange) and 20% DMSO (Red) were used at negative and positive controls respectively. Increased cytotoxicity was observed between eTCR-/rTCR+ (Purple), compared to HepG2 alone (p<0.0001), eTCR+/ rTCR- (Green, p=0.0017), and eTCR+/ rTCR± (Blue, p=0.0001). E) Visualisation of a region within the collagen gel. Addition of 20% DMSO resulted in cell death (Red), while HepG2 target cells alone to high viability (Green). Addition of effector T cells (Blue), resulted in different degrees of target cell killing between different effector groups (scale bar= 100µm).
Figure 4: Molecular analysis of on/off-target DNA editing, and fidelity of CDR3α and β regions of rTCR mRNA transcripts. A) Sanger sequencing of on-target editing at TRBC 1/2 loci in eTCR-/rTCR+ cells. i) Representative EDITR analysis with wild type sequences and four
possible bases conversions shown at each position and target G>A sites marked (C>A on opposite strand). ii) Summary of EDITR data for three donors at cytidine position 15 and 16 proximal to the PAM, presented as C>T changes (Blue), non C>T changes (Green) and no editing (Grey). iii) NGS sequencing analysis of on-target TRBC loci, quantification and characterisation of indels after BE3 editing found only low levels of small (<10bp) or larger indels (10-100bp). B) Box plots showing NGS analysis at the top 6 in silico predicted off-target editing for the TRBC 1/2 sites, with comparison of unedited eTCR+/rTCR- and edited eTCR-/rTCR+ groups (n=3). Larger dots represent outliers, in all cases ≤1.3% conversion. Two-tailed independent t-test between unmodified (eTCR+/rTCR-), and edited (eTCR-/rTCR+) samples showed for donors 1 p = 0.518, donor 2 p = 0.001, and donor 3 p = 0.103). D) Serial examination of RNA from rTCR HBs183-91 for 8 days post coBE3 mRNA delivery found no evidence of promiscuous deamination, with fidelity CDR3α and CDR3β regions maintained. Amplicon positions are marked above for C residues and schematic highlights hypervariable CDR3α and CDR3β regions that confer HLA-peptide specificity. CDR3 regions were mapped as a Heatmap in R using the gplots library for C>T conversion rates at the marked sites (TCR Clone software: 75).
References:


61. Garrison E, Marth, G. Haplotype-based variant detection from short-read sequencing 2012 [ ]


