

**Characterizing the granzyme-perforin pathway and its utility as a
cell-to-cell delivery system for cellular therapeutics**

by

Daniel Woodsworth

BSc (Honours), The University of British Columbia, 2008

A THESIS SUBMITTED IN PARTIAL FULFILLMENT
OF THE REQUIREMENTS FOR THE DEGREE OF

Doctor of Philosophy

in

THE FACULTY OF GRADUATE AND POSTDOCTORAL STUDIES
(Genome Science and Technology)

The University of British Columbia
(Vancouver)

June 2017

© Daniel Woodsworth, 2017

Abstract

Alongside small molecules and biologics, cell-based therapies are emerging as a third class of medical therapy. Additional sensors, actuators and control circuits would greatly expand the range of function and application of cellular therapeutics. To this end, a cell-to-cell delivery module has been developed by investigating and re-engineering the granzyme-perforin pathway of cytotoxic lymphocytes. A computational biophysical model of this process was developed and implemented using a spatial stochastic simulation algorithm, which indicated that hindered diffusion in the immune synapse is critical to ensure reliable granzyme internalization and that large amounts of granzyme escape the synapse, but should not have toxic effects due to rapid spatiotemporal dilution. Additionally, these results indicated that passive diffusion is sufficient for granzyme entry into the target cell, which motivated efforts to use granzyme as a molecular chaperone to transfer exogenous payloads from effector to target cells. Using a fluorescent protein payload, the subcellular localization of several granzyme B derived chaperones was characterized using fluorescence microscopy, and then their capacity to transfer the payload to target cells was evaluated in co-culture experiments. The results indicated that the motifs in granzyme B that are required for lytic granule loading are only functional and contiguous in the folded protein. Additionally, these experiments demonstrated that full length granzyme B is a suitable chaperone for delivering protein payloads to target cells via the granzyme-perforin pathway. Attempts were then made to use this system to deliver potent orthogonal toxins to apoptosis and lymphocyte resistant tumor cells. A range of granzyme B toxin fusion proteins were constructed, all of which retained toxic activity to varying degrees. To render target cells resistant to lymphocyte attack both small molecule and protein based inhibitors of apoptosis were tested in several cell lines, which delayed cell death, but did not stop it. Using effector target dose response curves, a moderate increase in target cell death was observed in cells targeted by lymphocytes expressing granzyme toxin fusion proteins, as compared to wild type lymphocytes, but the biological significance of this effect is uncertain. Approaches to improve this granzyme-perforin mediated delivery system and its therapeutic utility are discussed and explored.

Lay Abstract

Using biological cells as therapeutic devices has great potential. Cells are mobile in the human body, can be genetically programmed to take specific actions in response to environmental signals, and can be modified to have additional therapeutic functions that improve upon the cells' natural capabilities. Cytotoxic lymphocytes are components of the immune system that kill infected or malignant cells. These lymphocytes adhere to target cells and release two molecules, granzyme and perforin, into the region between the two cells, with perforin facilitating granzyme's entry into the target cell, whereupon granzyme kills the target cell. In this thesis I have taken preliminary steps towards adapting this pathway as a cell-to-cell molecular delivery system which could be incorporated into the cellular therapeutic devices described above. Using both computational biophysical models and experimental implementation, I have provided proof-of-principle that such an approach is feasible, although its therapeutic utility remains to be demonstrated.

Preface

The overall project was designed and conducted in collaboration with my supervisor Dr. Robert Holt. Considerations for specific chapters are as follows.

Sections of Chapter 1 and 5 were published as below. This paper was primarily written by me, with input and editing by Dr. Holt.

Woodsworth, D. J., & Holt, R. A. (2017). Cell-Based Therapeutics: Making a Faustian Pact with Biology. *Trends in Molecular Medicine*, 23(2), 104-115.

Chapter 2 was published as below. The conceptual approach was developed by myself, and my co-authors Valentin Dunsing and my committee member, Dr. Daniel Coombs. I developed the detailed model, wrote the code for implementation, and generated all figures. The data was analyzed by myself and Dr. Coombs. I primarily wrote the manuscript, with input and editing by Dr. Coombs.

Woodsworth, D. J., Dunsing, V., & Coombs, D. (2015). Design Parameters for Granzyme-Mediated Cytotoxic Lymphocyte Target-Cell Killing and Specificity. *Biophysical Journal*, 109(3), 477-488.

Chapter 3 has been submitted as a manuscript as below. The concept, approach and design was conducted by myself and Dr. Holt. I performed all experimental work, with the following exceptions. Lisa Dreolini conducted some of the molecular biology work: plasmid cloning, sequence verification and preparation, as well as western blotting. Libin Abraham aided me in the preparation of cells for microscopy, and image acquisition. I conducted all data analysis with input from Dr. Holt. I wrote the manuscript, with input and editing from Dr. Holt.

Woodsworth, D. J., Dreolini, L., Abraham, L., & Holt, R.A. Targeted cell-to-cell delivery of exogenous protein payloads via the granzyme-perforin pathway.

Chapter 4 is unpublished. I performed all experimental work in Chapter 4, except that Lisa Dreolini conducted some of the molecular biology work: plasmid cloning, sequence verification and preparation. The data was analyzed and interpreted by me, with input from Dr. Holt.

Table of Contents

Abstract	ii
Lay Abstract	iii
Preface	iv
Table of Contents	v
List of Tables	x
List of Figures	xi
List of Abbreviations	xiii
Acknowledgments	xvi
1 Introduction	1
1.1 Existing cell therapies	2
1.1.1 Regenerative medicine	2
1.1.2 Gene therapy	3
1.1.3 Mesenchymal stem cell therapy	5
1.1.4 Adoptive cell therapy in cancer	6
1.1.5 Engineered cellular therapeutics	9
1.2 Cytotoxic lymphocyte biology	12
1.2.1 Cytotoxic lymphocytes	12
1.2.2 The granzyme-perforin pathway	13
1.3 Overview of a lymphocyte-based delivery system	14
1.3.1 The granzyme-perforin pathway as a delivery module	14
1.3.2 Comparison with existing systems	15
1.4 Thesis overview	16

2	A computational biophysical model of the granzyme-perforin pathway	18
2.1	Introduction	18
2.2	Methods	19
2.2.1	Biophysical model: Geometry & molecular processes	19
2.2.2	Mathematical description: Spatial stochastic simulation algorithm	21
2.2.3	Hindered diffusion in the immunological synapse	23
2.2.4	Perforin oligomer diffusion and aggregation in the target cell membrane	24
2.2.5	Rate of granzyme translocation through perforin pores	25
2.2.6	Model parameterization and implementation	26
2.2.7	Validation of SSSA computational implementation	26
2.3	Results	32
2.3.1	Free diffusion in the synapse is incompatible with granzyme internalization	32
2.3.2	Pore formation is influenced by the amount of perforin released	34
2.3.3	The amount of granzyme internalized depends strongly on the rate of pore formation	35
2.3.4	Hindered diffusion critically influences pore formation and granzyme internalization	36
2.3.5	Dependence of pore formation and granzyme delivery on perforin insertion and diffusion in the target cell membrane	38
2.4	Discussion	38
3	Targeted cell-to-cell delivery of exogenous protein payloads via the granzyme-perforin pathway	43
3.1	Introduction	43
3.2	Results	44
3.2.1	Screening chaperone designs by assessing lytic granule colocalization using confocal microscopy	45
3.2.2	Transfer of fusion proteins from effector to target cells	46
3.3	Discussion	53
3.4	Methods	57
3.4.1	Computational identification of N-linked glycosylation motifs	57
3.4.2	Plasmids	57
3.4.3	Cell culture	57
3.4.4	Transfection	58
3.4.5	Flow cytometry	58
3.4.6	Microscopy	58
3.4.7	Image analysis	59
3.4.8	Cell labeling	59

3.4.9	Co-culture experiments	60
3.4.10	Statistical analysis	60
3.4.11	Western blotting	60
3.4.12	Crystal structure analysis	61
4	Efforts to use granzyme-perforin mediated delivery of orthogonal toxins to enhance cytotoxic lymphocyte killing of apoptosis resistant tumour cells	62
4.1	Introduction	62
4.2	Results	65
4.2.1	Development of orthogonal granzyme-toxin fusions for enhancing lymphocyte mediated cytotoxicity	65
4.2.2	Characterizing the lymphocyte resistance of the breast cancer cell line MCF-7	69
4.2.3	Efforts towards generating a lymphocyte resistant cell line	71
4.2.4	Effector dose response curves as a means of resolving small increases in YT target cell killing	75
4.2.5	Validating the GZB-HTK and GZB-NFSA toxin fusion proteins in MCF-7s	76
4.2.6	Predicted and measured enhancements of YT-killing of MCF-7s using granzyme-toxin fusion proteins	77
4.3	Discussion	82
4.4	Methods	86
4.4.1	Plasmids	86
4.4.2	Cell culture	87
4.4.3	Transfection	87
4.4.4	Flow cytometry	89
4.4.5	Metabolic activity assay	89
4.4.6	Direct toxin expression experiments	89
4.4.7	Cell labeling	90
4.4.8	Cytotoxic lymphocyte co-culture experiments	90
4.4.9	Isolation of YT Targeted cells using the FRET reporter	90
4.4.10	MCF-7 cell characterization	91
4.4.11	Lymphocyte resistance experiments	91
4.4.12	Enhancement of YT-Indy killing experiments	92
4.4.13	Drug reconstitution	92
4.4.14	Statistical analysis	92
5	Discussion, conclusions and future directions	94
5.1	The utility of the granzyme-perforin pathway as a delivery system	95
5.2	Further efforts to demonstrate the toxin-mediated enhancement of lymphocyte killing	98
5.2.1	Development of lymphocyte resistant target cells	98

5.2.2	Improving fusion protein granule loading and delivery	99
5.3	Future directions for granzyme-perforin delivery systems	100
5.4	Broader insights into cellular therapeutics	102
5.4.1	Context dependencies in cellular therapeutics	103
5.4.2	A framework for cell engineering	104
5.4.3	Grand challenges for cell engineering	105
5.4.4	The future of cellular therapeutics	109
Bibliography	110
A	Generating granzyme B mCherry fusion proteins expressed from genomic granzyme B locus using CRISPR/Cas9	131
A.1	Design of guide RNAs and donor template	134
A.2	pDN_GZB_E5-MCH plasmid	137
A.2.1	Plasmid map	137
A.2.2	Plasmid sequence	138
B	Mass spectrometry based investigation of granzyme B mCherry fusion protein regions of instability	141
C	Generation of a mutant EEF2 gene which protects cells from DTA when the two are co-expressed	146
D	Code	149
D.1	Core MATLAB scripts implementing computational biophysical model (Chapter 2) . . .	149
D.1.1	IMS_Ex.m	149
D.1.2	buildQ.m	164
D.1.3	calcDiffPropens.m	164
D.1.4	calcInterxnPropens.m	165
D.1.5	percDown.m	166
D.1.6	Qsort.m	167
D.1.7	swap.m	167
D.1.8	updateQ.m	167
D.2	Image filtration and colocalization analysis: MATLAB source code (Chapter 3)	168
E	Plasmids	172
E.1	Granzyme B derived chaperone-mCherry fusion protein plasmids (Chapter 3)	172
E.1.1	Base pdL plasmid map	172
E.1.2	Base pdL plasmid sequence	173
E.1.3	Coding sequence inserts	176

E.2	Granzyme B-toxin fusion protein plasmids and IAP-FRET plasmids (Chapter 4)	177
E.2.1	Base pdL plasmid map	177
E.2.2	Base pdL plasmid sequence	178
E.2.3	Base pMND plasmid map	182
E.2.4	Base pMND plasmid sequence	183
E.2.5	Source of component parts	187
E.2.6	Coding sequence inserts	187

List of Tables

Table 2.1	Parameters used in the immune synapse computational model.	27
Table 4.1	Fitted EC_{50} values for enhanced lymphocyte killing	81
Table E.1	Source of component parts of pdL vectors	187

List of Figures

Figure 2.1	Model geometry and molecular interactions	20
Figure 2.2	Mass action kinetics are accurately captured by the SSSA algorithm	28
Figure 2.3	Free diffusion described by the two-dimensional diffusion equation compares well with that simulated by the SSSA algorithm	30
Figure 2.4	Free diffusion with attenuation due to target cell membrane insertion is well described by the SSSA algorithm	31
Figure 2.5	Diffusion in the target cell membrane diffusion described by the two-dimensional diffusion equation compares well with that simulated by the SSSA algorithm	32
Figure 2.6	A two-dimensional reaction-diffusion system is well described by our SSSA	33
Figure 2.7	Time evolution of GZB and PFN in the immunological synapse without hindered diffusion	34
Figure 2.8	Effect of the amount of PFN released on pore formation probability and GZB internalization	35
Figure 2.9	Dependence of GZB internalization on rate of pore formation	36
Figure 2.10	Importance of hindered diffusion in pore formation and GZB internalization	37
Figure 2.11	Reduced diffusion of perforin in the target cell membrane reduces pore formation and GZB internalization	39
Figure 2.12	Impact of the rate of perforin insertion into the target cell membrane on pore formation and granzyme internalization	40
Figure 3.1	Design of payload delivery module chaperones	45
Figure 3.2	Subcellular distribution of candidate chaperone-mCherry fusion proteins	47
Figure 3.3	Automated image noise filtering pipeline	48
Figure 3.4	Quantitative assessment of candidate chaperone colocalization with lytic granules	49
Figure 3.5	Transfer of granzyme B mCherry fusion protein to target cells	50
Figure 3.6	Comparison of MCH payload transfer to target cells by the two granzyme B derived chaperones	51
Figure 3.7	Western blot confirmation of GZB-MCH fusion protein transfer to target cells	52
Figure 3.8	Spatial context of putative N-linked glycosylation sites in granzyme B	54

Figure 4.1	Granzyme toxin fusion protein design	65
Figure 4.2	GZB-DTA and GZB-PEA testing in HeLa cells	66
Figure 4.3	GZB-DTA testing in 293T cells	67
Figure 4.4	GZB-HTK testing in 293T cells	68
Figure 4.5	GZB-NFSA AND GZB-NFSB testing in 293T cells	69
Figure 4.6	Characterizing YT delivery of GZB to MCF-7s	71
Figure 4.7	Long term survival of YT-targeted MCF-7s	72
Figure 4.8	Small molecule inhibition of YT killing of 721 target cells	73
Figure 4.9	Long term viability of targets co-cultured with YTs in the presence of apoptosis inhibitors	74
Figure 4.10	Long term viability of YT-targeted target cells expressing IAPs	75
Figure 4.11	YT effector:target dose response of MCF-7s	76
Figure 4.12	Testing GZB-HTK and GZB-NFSA toxin/prodrug systems in MCF-7s	77
Figure 4.13	Estimates of toxin/prodrug enhancement of YT killing	80
Figure 4.14	Investigating the enhancement of YT killing of MCF-7s by GZB-NFSA/CB1954	83
Figure 4.15	Comparing estimates with experimental measurements of enhancement of YT killing	84
Figure A.1	Molecular characterization of genomically expressed granzyme B mCherry fusion proteins	133
Figure A.2	pDN_GZB_E5-MCH plasmid	137
Figure B.1	Western blot and gel of samples used for mass spectrometry experiments	142
Figure B.2	Mass spectrometry identified tryptic peptides from whole cell lysates	143
Figure B.3	Mass spectrometry identified non-tryptic peptides from whole cell lysates	145
Figure C.1	Evidence that mEEF2 co-expression with GZB-DTA restores protein synthesis function	148
Figure E.1	Base pdL vector	172
Figure E.2	Base pdL vector	177
Figure E.3	Base pMND vector	182

List of Abbreviations

AAV	Adeno-associated virus
ACT	Adoptive cell therapy
ALL	Acute lymphoblastic leukemia
BID	BH3 interacting-domain death agonist
BMT	Bone marrow transplant
CAR	Chimeric antigen receptor
CFP	Cerulean fluorescent protein
CFSE	Carboxyfluorescein succinimidyl ester
CL	Cytotoxic lymphocyte
CRISPR	Clustered regularly spaced palindromic repeats
CTL	Cytotoxic T-lymphocyte
DAPI	4
DCI	Dichloroisocoumarin
DMSO	Dimethyl sulfoxide
DTA	Diphtheria toxin A fragment
E:T	Effector:target
ER	Endoplasmic reticulum
ERSS	Endoplasmic reticulum signal sequence
ESC	Embryonic stem cell
FACS	Fluorescence activated cell sorting
FITC	Fluorescein isothiocyanate
FRET	Forster resonance energy transfer
FSC	Forward scatter
GCV	Ganciclovir
GFP	Green fluorescent protein
GILT	Glycosylation independent lysosomal targeting
GSL	Glycine-serine linker
GZB	Granzyme B

GZBSM	Granzyme B sort motif
GZBSS	Granzyme B signal sequence
HDR	Homology directed repair
HIV	Human immunodeficiency virus
HLA	Human leukocyte antigen
HSC	Hematopoietic stem cell
HSCT	Hematopoietic stem cell transplant
HTK	Herpes simplex thymidine kinase
IAP	Inhibitor of apoptosis protein
IFN	Interferon
iPSC	Induced pluripotent stem cell
IS	Immunological synapse
KAR	Killer activating receptor
KIR	Killer inhibiting receptor
LG	Lytic granule
MAD	Median absolute deviation
MCH	mCherry
MFI	Median fluorescent intensity
MHC	Major histocompatibility complex
MSC	Mesenchymal stem cell
MTOC	Microtubule-organizing center
NFAT	Nuclear factor of activated T-cells
NHEJ	Non-homologous end joining
NK	Natural killer (cell)
NTR	Nitroreductase
ODE	Ordinary differential equation
PB	PrestoBlue
PBS	Phosphate buffered saline
PCC	Pearson's correlation coefficient
PCR	Polymerase chain reaction
PDE	Partial differential equation
PEA	Pseudomonas exotoxin A fragment
PFN	Perforin
PI	Propidium iodide
RFP	Red fluorescent protein
scFV	Single chain variable fragment
SCID	Severe combined immunodeficiency
SELEX	Systematic evolution of ligands by exponential enrichment

SIN	Self inactivating
SMAC	Supramolecular activation complex
SSC	Side scatter
SSSA	Spatial stochastic simulation algorithm
STS	Staurosporine
SV	Subvolume
TALLEN	Transcription activator-like effector nuclease
TCR	T-cell receptor
TGF	Transforming growth factor
TIL	Tumour infiltrating lymphocyte
TNF	Tumour necrosis factor
TOX	Toxin
VC	Vehicle control
XIAP	X-linked inhibitor of apoptosis protein
YFP	Yellow fluorescent protein
ZFN	Zinc-finger nuclease

Acknowledgments

First and foremost, I would like to thank my supervisor Dr. Robert Holt. I walked into his office with a lot of ambitious ideas about engineering biology, and literally no idea what that looked like. More than anything, I thank him for looking past that and taking me on as his student. I am very grateful for the latitude he gave me to explore the many half formed ideas I had and have, and the patient and helpful answers to my near incessant questions and interruptions, especially near the beginning of my PhD. Rob gave me a unique combination of freedom and access to him that I feel extremely fortunate to have received. I also greatly respect the broad scope of his scientific ideas and interests, and his willingness to embark upon ambitious endeavors. The project that he and I eventually built would not have been possible without both of these traits. Some of the most enjoyable moments of my time in graduate school were sitting in his office, brainstorming or sketching out new ambitious ideas or applications—I wish we had the time and resources to pursue all of them. As I leave that office, I feel that he has been instrumental in giving me an approach with which to attempt to realize some of those ideas. My only regret is that we never got out for a bike ride.

I would also like to thank Dr. Daniel Coombs. In addition to sitting on my committee, we worked together to create the biophysical model of the immune synapse. I very much enjoyed this part of my time in graduate school. Again, some of my best memories are of standing in front of his chalk board trying to fit a few more symbols between the forest of equations already there. While I am obviously no mathematician, Dan was good enough to overlook that and I am grateful for our collaboration—whatever mathematical skill I retain is largely thanks to him. It has been a long time since Math 215, but I still remember the important lesson that a cooked turkey should be above room temperature.

I would like to thank the other members of my supervisory committee, Dr. Jonathan Choy, Dr. Phil Hieter, and Dr. Kelly McNagny for their guidance and suggestions, all of which were helpful, and many of which led to important additions to my project. Their patience with my often tight timelines is greatly appreciated. Dr. Eric Yung, a staff scientist in our lab has also helped me along much of my PhD. I appreciate his early efforts to construct some viral constructs, and teach me how to do so, as well as his willingness to hassle suppliers (although I think he may have enjoyed that).

Our lab has recently expanded, but there was a time when it was very small, and Scott Brown, Govinda Sharma and I have been together since then. Thanks to both of them for their advice, support and time spent kicking ideas around. Most of all, thanks to them for laughing. Graduate school is a

long road with many ups and downs, and they have both provided a combination of support and levity, as friends and colleagues. Kyla Cochrane and Chris May are new additions, but I have very much enjoyed our time together, and will be sad to leave them. Along with Jerry Tien, that student room has become a boisterous place, and while productivity may not be optimal, the crosswords, jokes, political conversations, movie quotes and occasional bits of science are totally worth it. Payal Sipahimalani does not sit in that student room, which is probably best for her sanity and productivity, which in turn benefits the rest of us. Without her patience and efficiency, many of us would be lost, broke, or have no access to our own data. Sadly this means I never got to see as much of her as I would have liked. Finally, I would like to especially thank Lisa Dreolini. She was immensely helpful in jumpstarting my project at a critical time. This would be a very different thesis without her efforts. Beyond that, I will miss our coffee conversations—at times interesting, at times hilarious. Her incredibly cheerful demeanor, ready laugh, friendship and support have meant a lot to me.

Five years is a long time, and I have many others to acknowledge: other and former members of the Holt lab, faculty in our building and others, and the many other students and staff that I have worked with or chatted with over the years. I especially thank the various members of the other labs at the GSC, many of whom I will miss. I would also like to thank the many good teachers I have had throughout my life, who have each contributed in some way to this. I think they are under appreciated in the role they play in getting us to a jumping off point.

Finally, I would like to thank my parents, Anne and Bob, and my sister Alexandra. Saying this is the first time I ever needed them would belie the incredible upbringing, nurturing, growth, education, guidance and love they have given me. Saying this is the first time I ever realized how much I needed them would be closer to the truth. I am very grateful and fortunate to have such a family, and I thank them.

Chapter 1

Introduction

The ability to deploy active therapeutic devices capable of engaging directly with the fundamental cellular and molecular causes of disease will be transformative for medicine. Using biological cells as the chassis for these devices has three fundamental advantages. First, cells maintain a homeostatic environment distinct from their surroundings, and integrate a wide variety of input signals to execute context-dependent actions. Second, there is incredible phenotypic and functional diversity across the various cell-types. Finally, this diversity is encoded in the genome of each cell, making it possible to encode logic in these cells in the form of additional, modified or deleted nucleic acid sequence. Thus, each of the distinct cell types in the human body is a potential basic chassis and platform from which to build a tailored cellular therapeutic.

However, this ecosystem is not necessarily optimal or functionally complete. Additional cellular sensors, effectors, and layers of control logic will be required to truly realize the potential for cell based therapies. Rather than *de novo* construction, these components can be best obtained by successfully leveraging the unique and varied functions of existing biological molecules or pathways. By altering, adding or removing elements of these processes, novel sensory and effector components can be generated that fill existing gaps in a cell's endogenous set of biological functions. These new components can then be inserted into a cellular chassis to create a cellular therapeutic with functions that are a composite of the original chassis and the new component derived from a repurposed pathway or molecule.

This thesis represents steps towards this vision. I have attempted to repurpose the granzyme-perforin pathway of cytotoxic lymphocytes as a cell-to-cell delivery module for therapeutic proteins. The overarching goal is that by using this pathway to deliver therapeutics, this module could be inserted in cytotoxic lymphocytes that are targeted to diseased tissue via lymphocyte surface receptors to yield a targeted and specific cellular delivery device, capable of trafficking throughout the body to find the desired site of disease, and deliver, in a cell-specific manner, a therapeutic to treat that disease.

1.1 Existing cell therapies

The history of cellular therapies is mainly rooted in the stem-cell research conducted over the last 60 years. This led to first bone marrow transplants, followed by hematopoietic stem cell transplants from other stem cell sources, for treatment of various hematological disorders and malignancies. The advent of immunosuppressive drugs allowed the therapeutic transplant of cells from other tissues, as well as bulk organ or tissue transplants. With the advent of recombinant DNA technology, the possibility of modifying cells *ex vivo* prior to administration emerged. These possibilities continue to expand with cheap and rapid DNA synthesis, improved methods for DNA delivery to a variety cell types, and, recently, rapid and facile methods for editing a cells genome with single base pair resolution. As a result, incorporating existing biological function into new cellular therapeutic devices is an approach that is gaining traction and application across a range of human disease.

1.1.1 Regenerative medicine

Current methods for treating diseases arising from dysfunctional or dying cells rely on: (i) systemic administration of therapeutics with often-poor specificity, (ii) surgical resection, or (iii) bulk tissue or organ transplantation. Rather than attempting to 'fix' these dysfunctional cells, the approach of regenerative medicine is to simply replace them with fresh, functionally identical cells, new 'parts' that retain appropriate function. The first challenge is obtaining these replacement cells. In a few instances they may be harvested from allogeneic donors, as in the case of Type I diabetes, where infusion of unmodified, donor-derived beta-islet cells is now a well-tested and often effective approach [1, 2], although achieving stable engraftment is still a substantial challenge [3]. More generally, supplementing tissue stem cell compartments with unmodified, donor-derived stem cells, including those of human embryonic origin, continues to be explored as a promising therapeutic approach for a broad array of pathologies, including neurological, hepatic, endocrine, and musculoskeletal disorders [4].

The recently developed methods for genetically reprogramming terminally differentiated autologous cells into pluripotent stem cells (iPSCs) [5], together with in vitro techniques for re-differentiating these iPSCs into a variety of mature, replacement cell types and tissues is bringing new possibilities to regenerative medicine and tissue engineering [6]. This technology is allowing generation of both stem cells and differentiated cell types from a potentially sick patient that may not have stem cells or differentiated cells available for harvesting, and for whom a suitable donor may not be available. Importantly these iPSC-derived cell products are autologous and therefore more likely to be immunologically compatible with the recipient. Furthermore, the *ex vivo* manipulations necessary for the derivation and re-differentiation of iPSCs provide new opportunities for additional genetic modification to enhance their therapeutic properties. For example, engineering replacement iPSC-derived beta-islet cells for immuno-resistance may reduce their sensitivity to the autoimmune mechanisms that eliminated their natural predecessors, making them more therapeutically relevant than unmodified replacement cells would be. Additional examples from the field of tissue engineering include generating a liver bud with

appropriate vascularization and three dimensional architecture [7], as well as recellularizing a decellularized heart scaffold with the aim of generating personalized whole organs for transplant [8]. Throughout regenerative medicine, major challenges remain, including sources of universal allogeneic cells, stable engraftment, avoiding rejection, and maintaining the viability of the graft over the long term [9].

1.1.2 Gene therapy

Gene therapy can be broadly defined as the delivery of genetic material to a patient's cells with the goal of modifying the genetic makeup of the cells for a therapeutic benefit. With exceptions, it has largely focused on inserting a working copy of a gene into a cell population in which the gene is damaged or absent. Functionally, it can be divided into two main approaches: gene delivery *in vivo* or *ex vivo*. The former approach typically relies upon the use of viral vectors as delivery platforms (although other approaches have been used including naked DNA). The former is challenging since it requires near perfect control of tissue specificity, either through viral tropism or tissues-specific promoters or other regulatory elements. Furthermore, post-insertion quality control is impossible. Finally, physically delivering sufficient viral vector to the diseased tissue is often difficult. Therefore, it is not surprising that the field of *in vivo* gene therapy has progressed most in diseases which are single gene disorders that occur in tissues where physical viral delivery is more anatomically accessible: notably the eye and liver [10].

While the original, and likely still ultimate goal of gene therapy was and is *in vivo* delivery, in many cases the technical considerations listed above have necessitated *ex vivo* gene delivery into autologous patient-derived cells, followed by transplantation of these cells back into the patient.

Due to the accessibility of the starting cell population from either peripheral blood or bone marrow, and the relative experience and familiarity that clinicians had developed from bone marrow transplant programs, gene therapy using hematopoietic stem cells (HSC) was the first, and is therefore the most mature, form of cell-based gene therapy [11]. Initial focus was on single gene disorders, mainly primary immune deficiencies (severe combined immunodeficiency disorder and Wiskott-Aldrich syndrome), as well as neurodegenerative storage disorders (adrenoleukodystrophy and metachromatic leukodystrophy) [12]. Early successes were reported in the treatment of pediatric patients with X-linked severe combined immunodeficiency (SCID-X1) [13]. These patients suffer from a total lack of lymphocytes, with concomitant susceptibility to infection. The underlying cause of the disease is a deficiency for the *IL2RG* gene, which codes for the gamma chain of several interleukin receptors that are necessary for lymphocyte development. Patients received CD34+ stem cells that had been transduced with a retro-viral vector carrying a functioning *IL2RG* gene. Initial reports were very promising, with the reappearance of lymphocytes in most children [13]. Unfortunately however, several patients developed leukemia [14]. This was determined to be due to insertional oncogenesis via viral vector inserted long-tandem repeat (LTR) activation of adjacent oncogenes. While these results significantly slowed the progress of gene therapy, several factors have led to a recent increase in enthusiasm. The success of self-inactivating (SIN) lentiviral vectors is likely to significantly increase the safety profile of these

therapies by reducing the likelihood of clonal expansion of a transduced cell. Lentiviral vectors have now been employed across a similar range of immunodeficiency diseases with good efficacy, and little evidence for clonal expansion [12]. They have been further used in the treatment of metachromatic leukodystrophy [15] and β -thalassemia [16], although oligoclonal expansion was reported in the latter.

Thus the current status of gene therapy is still in flux. Improvements in vector design have renewed enthusiasm for the field, but safety concerns remain. Gene-editing technologies, consisting of zinc-finger nucleases (ZFNs), TALENs and the CRISPR-Cas9 system have the potential to address some of these issues, by virtue of their ability to target DNA at a specific locus. This opens the possibility gene-correction of the endogenous diseased gene, rather than gene replacement by addition.

These gene editing systems all revolve around fusion proteins consisting of a DNA recognition domain joined to a nuclease domain. Unlike meganucleases or restriction endonucleases, the DNA recognition domains are programmable. ZFN technology was the first to be developed. The DNA binding domain consists of zinc fingers, which are ubiquitous protein domains that bind to short nucleic acid sequences [17]. By modifying small numbers of residues in the alpha helix of Cys₂His₂ zinc fingers, the triplet nucleotide sequence that the zinc finger binds to can be altered [18]. Chaining together multiple zinc fingers of varying triplet specificity results in a protein which binds to a defined, extended DNA sequence. This zinc finger domain is fused to the non-specific nuclease domain of the FokI enzyme.

Transcription activator-like effector nucleases (TALENs) are similar to ZFNs, with the main difference being that TALENs can bind arbitrary DNA sequence, and they are more modular. Each transcription activator-like effector (TALE) is a 36 amino acid motif that binds a single nucleotide, with the specificity of the TALE defined by a dipeptide at the protein-DNA interaction site [19]. Thus TALEs can be predictably strung together to target arbitrary nucleotide sequences. Furthermore, chaining TALEs is a fairly reliable process, unlike ZFNs whose DNA binding function typically require protein level optimization to resolve adjacent zinc finger interactions [19]. As with ZFNs, the TALE domain is fused to a non-specific FokI nuclease domain.

The clustered regularly interspaced short palindromic repeats (CRISPR) cas9 system is the newest gene editing technology. The cas9 protein complexes with a guide RNA (gRNA, which is a chimera of the two separate RNAs found in the prokaryotic endogenous system, consisting of a cas9 binding motif and a DNA base-pairing motif), which imparts the DNA binding specificity to the cas9 protein by forming a DNA:RNA heteroduplex, after which cas9 cuts the adjacent DNA region [20]. Targeting is a simple function of Watson-Crick base-pairing. The other great advantage of the CRISPR/cas9 system is that no protein engineering is required to target different loci: only new gRNAs are required. These features allow for robust predictable design, rapid feedback loops, and multiplex or library scale gene targeting [21].

Regardless of the gene editing system, after the fusion proteins are expressed in target cells via transfection or transduction, the DNA recognition domain binds in a sequence specific manner to its matching target DNA sequence, and then the fused nuclease introduces a double stranded break in

an adjacent, and predictable, region of the DNA. (Extensions of this approach that use nickases to introduce single stranded breaks are similar [22].) The double stranded break is then repaired by non-homologous end joining (NHEJ), and, if a repair template is present, homology directed repair (HDR) [23]. NHEJ is nonspecific and the end-processing of the two DNA ends results in nucleotide insertions or deletions (indels), making this approach suitable for knocking out genes [24]. If a DNA template is present that has homology to the two cut ends, the HDR pathway is also active [25], which results in the incorporation of the homology regions in the template, as well as any additional sequence between these two regions. This makes possible the introduction of additional sequence, either simple nucleotide modifications, or the insertions of whole genes [26].

These capabilities may be exploited in several ways. First it could simply be used to insert a gene at a known, 'safe harbor' location, for example the AAVS1 site, which would decrease the chances of gene activation or insertional oncogenesis [27]. Zinc fingers have also been used to correct the *IL2RG* gene in HSCs derived from a SCID-X1 patient [28]. Finally, gene therapy need not be limited to the addition or correction of genes, it can also be used in the context of removing genes for therapeutic purposes. Gene-editing of CD4 T-cells to knock out the CCR5 receptor used by HIV to gain entry to CD4 cells has even progressed to clinical trials, which demonstrated safety and improved viral control upon temporary antiretroviral withdrawal [29]. Genome editing still has outstanding questions that need to be answered prior to widespread therapeutic application, most importantly the frequency and impact of off-target activity, which is a known issue across all genome editing platforms [30–33]. In the case of gene correction, editing efficiency is also quite low, which poses a major challenge for *in vivo* gene-editing.

More broadly, in both traditional viral gene therapy, as well as newer non-viral gene therapy and genome editing strategies, additional challenges remain. These include: (i) accurate and efficient targeting and delivery of the genes to diseased cells; (ii) achieving modification or delivery to a sufficient fraction of cells for therapeutic effect; and (iii) avoiding iatrogenic effects, primarily insertional oncogenesis in the case of gene therapy and off-target editing in the case of gene-editing.

1.1.3 Mesenchymal stem cell therapy

Mesenchymal stem cells (MSCs) are emerging as promising therapeutic candidates for many diseases, due to MSCs' immunomodulatory effects, natural tropism for tumors and other sites of inflammation, multipotency, and relative ease of use. Both naturally, and following therapeutic implantation at sites of disease, MSCs have been observed to participate in tissue repair and regeneration in damaged or degenerative tissue, as well as induce a return to immuno-homeostasis in autoimmune diseases [34]. Hundreds of clinical trials are currently underway, with particular focus on two main applications: regeneration of bone-related injured tissue, as well as promoting an anti-inflammatory response [35].

In many cases their natural therapeutic properties have been augmented by genetic modifications that improve MSCs' homing to sites of disease [36], as well as their therapeutic activity once at these sites, including the heart [37], brain [38], and in tumors [39, 40] (which have now progressed to clinical

trials [41]).

1.1.4 Adoptive cell therapy in cancer

Based on epidemiological evidence such as an increased incidence of some types of cancer in immunosuppressed transplant patients, the importance of the immune system in controlling and eradicating tumours has been acknowledged for some time [42, 43]. The observation that tumour infiltrating lymphocytes (TIL) could be expanded from surgically resected tumours, and that these TIL were reactive against tumour cells *in vitro* spurred initial interest in the therapeutic utility of these cells [44]. This observation was first exploited clinically by Rosenberg and colleagues: autologous lymphocytes from surgically resected tumour sections were expanded *ex vivo* by culture with high dose IL-2, and reinfused into melanoma patients [45]. As the approach has been refined, and in particular with the addition of prior lymphodepleting chemotherapy, the outcomes have improved—especially in melanoma patients, with a recent clinical trial reporting complete remission in 22% of patients with metastatic melanoma [46].

TIL therapy has several limitations. First, not all cancers have such a heavy mutational burden as melanoma, which limits the number of tumour neo-epitopes against which TIL have the potential to react, and thus limits the presence of TIL itself [47, 48]. Second, a suitable source of lymphocytes to expand is required, which is not always possible, either due to the nature of the cancer, or because the anatomical location of the tumour makes it unresectable [49]. Furthermore even in cases where a resection is available, only 30-40% of biopsies yield suitable T-cell populations [50]. Third, in many cases TIL are in a state of near total exhaustion and anergy [51]. This is due to a host of factors, including the immunosuppressive tumour microenvironment (further discussed below), as well as exhaustion due to repeated antigenic stimulus [52, 53]. This makes them challenging to expand *ex vivo* and of dubious utility and efficacy once administered back into the patient [50].

To address these issues, and to make adoptive transfer of T-cells a viable therapy for a wider range of cancers, it would be desirable to use cytotoxic lymphocytes isolated from peripheral blood as the starting source from which to expand T-cells. This approach would in theory allow for the generation of younger, less terminally differentiated and anergic T-cells, and would not be dependent on the availability of TIL, the tumour resection feasibility and quality, or on the successful expansion of a small starting population of tumour reactive TIL to clinically useable numbers [54]. However, this approach introduces a new challenge: in general, T-cells from peripheral blood are highly polyclonal, with only a small fraction having specificity for malignant cells [55].

A solution to this problem is to genetically modify the T-cells with an additional surface receptor specific for the tumour. This was first attempted by Rosenberg and colleagues using a MART-1 specific T-cell receptor (TCR) in melanoma patients [56]. TCR targeting allows for administered T cells to target intracellularly derived antigens presented on the tumour cell surface in the context of a peptide-MHC (major histocompatibility complex). The great advantage of this approach is that the potential target antigens encompass, in theory, the entire peptidome of the cell, thus increasing the theoretical

likelihood of finding a tumor specific antigen. There are however, several disadvantages. First, a TCR with appropriate reactivity is required, and current methods for TCR screening and discovery are relatively low-throughput [57]. Second, since TCRs recognize a peptide-MHC complex, a given TCR is only suitable for the subset of patients who have the appropriate HLA-type. Third, a key mechanism of immune escape by tumours is MHC-downregulation, and in this case tumours may be relatively invisible to TCR-targeted T-cells [58]. Finally, there is the potential for generating a chimeric TCR resulting from pairing between the original endogenous α -chain and the inserted β chain (or vice versa). This has the potential to create a TCR with unknown reactivity, possibly targeting self-antigens and resulting in graft-versus host effects. While this has been observed to cause lethal toxicity in mice [59], it has not been in humans [60]. Despite these challenges, TCR-targeted T-cells have been successfully used in a recent clinical trial using an NY-ESO-1 specific TCR in sarcoma and melanoma patients. While the results were positive, the improvement in survival was not dramatic [61].

The main alternative to TCR targeting is the use of chimeric antigen receptors (CARs). These are synthetic receptors consisting of an extra-cellular single chain variable fragment (scFV) fused to an intracellular domain consisting of various T-cell signaling components [62]. The scFV is a fusion of the variable region of the heavy and light chains of antibody, which retains the specificity of its parental molecule [63], which is selected to react against a surface expressed tumour antigen. The intracellular domains in the initial CAR designs consisted of a transmembrane and hinge region domain, a CD3 ζ domain, with the latter serving to provide stimulatory intracellular signaling, similar that of CD3 ζ in an endogenous TCR complex [64]. The coupling of epitope recognition with initiation of TCR signaling results in the modified T-cells targeting any malignant cell expressing the cognate antigen for the scFV. Since they were first reported [65], the design of these receptors has been improved upon significantly, most notably by the addition of either the CD28 or 4-1BB costimulatory domains, which has improved the CAR-mediated stimulus of T-cell proliferation and persistence [66].

Using CARs to target T-cells to tumours has several advantages. The targeting is MHC independent, meaning a given CAR is usable in all patients, it is unaffected by MHC downregulation, and targeting non-peptide epitopes, such as post-translationally modified proteins, lipids or carbohydrates, is feasible [67]. Furthermore, methods for scFV development to target a given antigen are relatively mature [63]. However, this presupposes the existence of a surface expressed tumour-specific target antigen for which a reactive antibody exists. Unfortunately, this excludes two major classes of tumour specific antigens: cancer/testis antigens and tumour neoantigens [68], thus greatly constricting the set of potential target antigens. Since these proteins are largely intracellular, the only way in which any part of them is found on the surface of a tumour cell is as small MHC-presented peptides. Efforts to develop antibodies recognizing peptide-MHC complexes were initially unsuccessful in that the antibodies were primarily reactive to the MHC itself (and thus potentially every patient cell), rather than the peptide-MHC complex [69]. Recently some success has been made in developing a peptide-MHC specific antibody, although porting it to a CAR eliminated its epitope specificity, which had to be repaired by lowering its affinity via rational mutation [70]. If these issues can be resolved, CAR targeting

of peptide-MHC would have the potential to greatly expand the range of cancers in which CAR T-cell therapy might be employed.

Clinically, CAR T-cell therapy has received the most attention for its use in hematological malignancies [71–73]. The CAR used in all cases targets the B-cell surface marker CD-19, resulting in B-cell aplasia and agammaglobulinemia which necessitates immunoglobulin transfusion [74]. The specifics of generating the CAR-T cellular product vary across the various academic centers that are currently investigating this therapy, but the general approach is as follows. A population of peripheral blood derived T-cells is selected for modification, which may be all T-cells, or skewed towards memory, naive, CD4 or CD8 compartments. The relative makeup of the starting population of cells has been shown to be of significant importance [75]. While the optimal distribution is by no means clear, recent work suggests that equal amounts of CD4 and CD8 T-cells from less differentiated subsets with greater proliferative capacity, such as naive or central memory cells, yield the best *in vivo* anti-tumour activity [75–78]. This starting population are then modified to express the CAR, typically using a retro or lentiviral vector, although other methods including transposons, and RNA transfection have been used. Modified cells are then activated and expanded using either agonistic antibodies against CD3 and CD28, or using irradiated feeder cell populations that express the cognate antigen for the CAR. A variety of proliferative interleukins are also included in the expansion, usually at least IL-2, although many others have been studied, most notably IL-7 and IL-15 [64]. Finally, the cell product is infused into the original (now lymphodepleted) patient from whom the cells were collected.

Optimizing the CAR design, *ex vivo* expansion protocols, and patient lymphodepletion preparative regimens has required substantial investment, but has paid off in the last five years, with CAR-T therapy in hematological malignancy achieving response rates that are truly spectacular, ranging from 70-90% complete remissions [76, 79–81]. With this level of efficacy, there is understandable enthusiasm for extending these results to other cancer types. Currently there are clinical trials underway using CARs that target a variety of antigens and corresponding tumour sites, including prostate specific membrane antigen (PSMA, prostate cancer), mesothelin (pancreatic cancer and mesothelioma, among others), GD2 (a ganglioside, used in neuroblastoma) and the oncogene HER2 (glioblastoma and sarcoma among others) [67].

In extending TCR- or CAR-targeted T-cell therapy to solid tumours, numerous additional obstacles will be encountered. Tumour specific antigens will be required, and despite significant investment there are still few validated targets about which the community is confident [82, 83]. Furthermore, on-target off-tumour toxicity, with outcomes as severe as patient fatality, remains an ongoing issue [84]. Perhaps most importantly, the ability of CAR-T cells to overcome the substantial capacity for immune evasion and suppression exhibited by tumours is an outstanding question. Chemokine mismatch is a common finding in solid tumours, which results in ineffectual CAR-T trafficking in, and extravasation from, the circulation to the tumour [85]. Many tumours are surrounded by an external stromal layer consisting of fibroblasts, myeloid cells and extracellular matrix, all of which can inhibit CAR-T penetration into the tumour proper [86, 87]. The tumour microenvironment is characterized by hypoxia [88], depletion

of key metabolites such as glucose and amino acids [89–91], and regulatory cytokines such as TGF- β , all of which inhibit T-cell proliferation and can induce T-cell anergy or conversion to a regulatory phenotype. These factors can be produced by malignant cells, but also by a variety of stromal cells such as myeloid derived suppressor cells (MDSCs), Tregs, and innate immune cells such as neutrophils [92]. In particular, Tregs suppress T-cell function and proliferation, as well induce T-cell apoptosis, through a host of effector functions [93], and their selective depletion can improve anti-tumour activity of CAR-T cells in mouse models [94, 95]. If a CAR-T cell survives this gauntlet and finds a tumor cell expressing the cognate antigen for the CAR—and assuming this antigen has not been deleted or downregulated, another mechanism of tumor immune evasion [96, 97]—the CAR-T cell encounters a target tumour cell that potentially expresses a variety of inhibitory, tolerogenic and pro-apoptotic ligands, most notably PDL-1 [98], although many others are under investigation [99]. While these challenges can appear insurmountable, there are active research programs that seek to address virtually all of these challenges [83], and with a variety of active clinical trials studying CAR-T therapy for solid tumours [54], the coming years should provide insight and clarification as to the broad applicability of CAR-T therapy.

More generally, and true of all cell-based therapies, significant hurdles are posed by the complexity associated with a therapy that is, currently, entirely personalized and manufactured separately for each patient. This makes regulatory approval, delivery and deployment, and public payment for these therapies all outstanding issues to be resolved.

Finally, several patients recently died in two waves in a CAR-T clinical trial run by one of the leaders in the field, Juno Therapeutics [54]. This illustrates the caution that will be required in moving forward with such complex, potent and only partially understood therapeutics.

1.1.5 Engineered cellular therapeutics

The therapies described in the previous sections are the first wave of cellular therapeutics. For the most part they focus on the addition or deletion of a single gene, potentially one that has been substantially engineered, as in the case of CARs. Moving forward, the full potential of cellular therapeutics will begin to be realized as layers of molecular function and control logic are built on top of the cellular chassis, and this process is well underway.

In the field of adoptive cell therapy, a variety of cell engineering is already being pursued. Conditionally active suicide switches are being included in adoptively transferred cells, to allow for eradication of the therapeutic product in case of adverse events [100–102]. Notably this is, in general, impossible for small molecules or biologics, unless a rapid inhibitor is available. In an inversion of this approach, a two-component CAR has been developed, that is only able to dimerize and activate downstream signaling pathways in the presence of an inert small molecule ligand [103]. This would allow for infusion of CAR-T cells, and then subsequent activation or deactivation of the effector cells as needed. This is a more nuanced approach than suicide switches as it might allow for dose titration and temporary interruptions in the therapeutic action of the CAR-T cells as needed, without fully destroying an expensive and personalized cellular product. To solve the problem of the immunosuppressive

tumour microenvironment, additional factors are being added to T-cells [104, 105]. Examples include secretion of the pro-inflammatory cytokine IL-12, as well as receptors that respond to IL-4, which is abundant in the microenvironment, by providing IL-2/IL-15 proliferative stimuli to the CAR-T cells [67].

Two recent combinations of gene editing and CAR-targeting have shown significant promise. The goal of the first was to generate CAR-targeted T-cells from an allogeneic donor. To do this, TALENs were used to knock out two genes in the donor cells: (i) TCR alpha, to eliminate potential graft-versus host reactivity; and (ii) the lymphocyte surface antigen CD52, which renders the cells resistant to the lymphodepleting anti-CD52 antibody alemtuzumab. This allowed for transient host lymphodepletion, thus avoiding graft rejection, while maintaining the viability of the incoming cellular product. After these modifications, the donor cells were virally transduced with a CD19-CAR coupled to the hybrid antigen RQR8 (a combination of CD20 and CD34), which renders the cells sensitive to rituximab, as a suicide switch for increased safety. This approach was tested in two pediatric patients with ALL who had progressive disease after multiple rounds of therapy, and thus was so lymphopenic that autologous CAR-targeted T-cells could not be generated. As of January 2017, the patients are in molecular remission [106]. A similar approach has been pursued using CRISPR-Cas9 to knock out three genes: (i) TCR-alpha to eliminate alloreactivity; (ii) beta-2-microglobulin, to eliminate donor T-cell MHC expression, and thus increase the persistence of allogeneic donor CAR-T cells in the recipient host; and (iii) PD-1, to decrease tumour microenvironment inhibition of CAR-T efficacy. In preclinical work these modifications resulted in increased persistence and efficacy [107], and a clinical trial is under way. Finally, preliminary work has shown the feasibility of using iPSCs to generate CAR-targeted T-cells, which in theory would provide a limitless supply of a highly controlled cellular product [108]. While in their early stages, these efforts may be the first steps towards an off-the shelf, universal CAR-T cell therapy, which would greatly alleviate the cost and administrative burdens currently faced by CAR-T therapy.

As discussed above, there are still relatively few cancer types for which a reliable, suitable CAR-T target antigen exists [82]. To solve this problem, combinations receptors are under development that target multiple antigens to implement simple Boolean logic gates [109]. While no single antigen may be unique to a tumour, it is more likely that the combination of two or three may be. In a particularly impressive work, Lim and colleagues have generated a novel synthetic receptor derived from the Notch signaling pathway. In its natural configuration, the extracellular domain of the Notch receptor binds to a delta ligand on another cell. This results in intracellular proteolytic cleavage of a transcription factor that activates downstream signaling. Somewhat surprisingly, this system is fairly modular: by simply retaining the core component of the Notch receptor that induces proteolysis, they were able to use a variety of novel external receptors (such as CARs, nanobodies and Myc tags), to trigger the release and activation of a variety of downstream transcription factors [110, 111]. Finally, the CAR concept has been inverted to create inhibitory CARs (iCARs), consisting of an scFV extracellular antigen recognition domain coupled to a CTLA-4 or PD-1 derived inhibitory intracellular domain. These receptors are

coupled with a conventional CAR, to confine the specificity of the CAR-T cell to tumour cells. This approach could be employed in the context of a tumour cell antigen target that is present on healthy tissue as well. In this case the iCAR would be specific for a second antigen found on the healthy tissue. When both antigens are present (healthy tissue), the T-cell would not be activated, while if only the tumour antigen were present, the CAR-T cell would be activated [112]. Together, these initiatives have the potential to enable combinatorial sensing of a variety of signals, and integration of these signals with minimal cross-talk.

Moving beyond the domain of adoptive cell therapy, several efforts are underway to construct relatively complex control logic, albeit in quite simple model cell lines. In a pioneering work, a cell classifier circuit was built to detect HeLa cells as a model cancer cell. A panel of 5 micro-RNAs was identified, whose abundance (high or low) differentiated between HeLa and healthy cells. Using a combination of the lac operator and miRNA binding sites, a classifier was built that was active only in the presence of the correct combination of miRNAs, and inactive otherwise. When the output of this classifier was set to be the pro-apoptotic protein Bax, and the whole circuit was transfected into mixed cell populations, only the HeLa cells exhibited substantial cell death [113]. Another substantial body of work has been undertaken by Fussennegger and colleagues, who have engineered cellular factories that have existing sensory and synthetic pathways joined to create novel cell-based therapeutics termed prosthetic gene circuits [114]. Preclinical work has seen investigators combine a uric acid sensor and urate oxidase (which degrades uric acid and is notably absent in humans) from two different bacterial species to create a prosthetic gene circuit that maintains uric acid homeostasis in experimental models of gout [115]. In another example, expression of an IL-22 receptor was placed under conditional control of a TNF-responsive promoter. IL-22 activity at the synthetic receptor resulted in production of anti-inflammatory cytokines IL-4 and IL-10. In this way, a cellular therapeutic implanted in a mouse model of psoriasis was able to suppress inflammation in a highly specific, targeted manner: only when both TNF and IL-22 were present was the circuit active [116]. Similar circuits have been designed and tested for diabetes [117], metabolic syndrome [118], thyroid disease [119]. These cellular therapeutics all have several common themes. First, they all rely upon mining the diversity of biological function to find useful parts (for example a light sensor, or an enzyme that degrades uric acid). Second, they include rewiring transcriptional logic circuits to combine the inputs and outputs of these novel parts to achieve the desired response. Finally, they are almost uniformly expressed in simple model cell lines that are encased in inert alginate gels and implanted in the body. These gels consist of a polymer matrix with pore sizes large enough to permit entry of crucial metabolites and exit of the cell-secreted therapeutic molecule, but small enough to block the host immune cells, thus preventing host rejection of the engineered cellular therapeutic [120]. This approach has several potential advantages over simple replacement of the diseased cell type or a small molecule therapy. First, immunologically compatible donor sources are often unavailable, and hESC or iPSC derived cells will not be available for all tissues for some time. Second, replacement may require implantation in a challenging physical location in order for function. Finally, replacement is not a viable option in the cases of increased metabolite levels,

or production of toxic metabolites due to dysfunctional enzymatic activity. Recombining biological pathways in a suitable chassis offers solutions to these problems. The chassis may be selected to operate in a more suitable physical niche, and be devoid of any pathogenic immunological markers. The pathways for sensing metabolite levels may be optimized, using multiple, engineered surface receptors. Similarly, the pathways for metabolite regulation (production or elimination) may also be optimized, for example to avoid toxic metabolite production. These applications are clearly years from being applied clinically. They have only shown efficacy in mouse models over short periods of time. The long term safety profile, and immunological reactivity of implanted cell-lines is very much in question, even if they are encapsulated. However, what these circuits represent are the early stages of a set of parts and approaches for combining sensors, effectors, and control logic into cellular therapeutics.

Incorporating existing biological function into new cellular therapeutic devices is an approach that is gaining traction and application across a range of human disease. I have attempted to add to this part set by developing a cell-to-cell therapeutic delivery module.

1.2 Cytotoxic lymphocyte biology

This thesis focuses on understanding and engineering the granzyme-perforin pathway, a key effector mechanism of cytotoxic lymphocytes. Before further discussing their application, I first provide some detail on their basic biology.

1.2.1 Cytotoxic lymphocytes

Cytotoxic T-lymphocytes are key elements of the adaptive immune response that are mainly responsible for the recognition and clearance of cells infected by intracellular pathogens, as well as tumour immunosurveillance [121]. T-cell identification of target cells is a complex process that hinges upon TCR engagement of a cognate peptide presented by cell surface MHC [122]. This interaction activates the key cytotoxic effector mechanisms of T-cells: [121] (i) the granzyme-perforin pathway; (ii) surface expressed death receptor ligands such as Fas (FasL); and (iii) cytokines, most importantly interferon gamma ($\text{IFN}\gamma$).

Unfortunately, primary T-cells are relatively difficult to manipulate genetically, with viral transduction often required, which is unsuitable for exploratory work involving iterations of design, testing, validation and optimization. Furthermore, maintaining primary lymphocytes in culture is more onerous than maintaining immortalized cell lines, and, due to their limited proliferation lifespan, any modifications made to primary cells will eventually be lost when the cells enter senescence. While immortalized T-T hybridoma cell lines do exist, none retain a functional granzyme-perforin pathway.

An attractive alternative to working with primary T-cells is to use natural killer (NK) cell lines as a model system. NK cells, which are the analogue of T-cells in the innate immune system, kill targeted cells using the same effector mechanisms as T-cells [123], and several NK cell lines exist with an intact granzyme-perforin secretion pathway, and intact target-specific cytotoxic function. NKs differ from

T-cells in that NK activation is a complex balance between activating and inhibitory receptor mediated signals, transduced by killer activating and killer inactivating receptors (KAR, KIR) respectively [124]. Importantly, this means that the cytotoxic machinery of NK cells can be mobilized against a target cell without the requirement of antigen receptor mediated activation, as would be the case in a T-cell. However, this antigen specificity can be imposed by expressing a CAR in NK cells [125], and indeed CAR-targeted NK cells have been used in clinical trials [126].

1.2.2 The granzyme-perforin pathway

In humans there are four granzymes: A,B,K and M, of which granzyme B (GZB) is the best characterized and most abundant [127]. GZB is a serine protease with a classical trypsin-like catalytic triad that initiates apoptosis in targeted cells [128]. Synthesized primarily in cytotoxic lymphocytes as a 247 amino acid precursor protein, GZB is directed to the endoplasmic reticulum by a signal peptide, which is subsequently cleaved, yielding the zymogen form of GZB, which is still inactive due to a N-terminal dipeptide. This proenzyme is sorted through the Golgi network in a pathway that involves the addition of mannose-6-phosphate, as well as the chaperone molecule serglycin, both of which promote localization of GZB to lytic granules (LGs), a type of specialized secretory lysosome. Once in the LG, the dipeptide is cleaved by cathepsin C, and the active form of GZB is safely sequestered in the acidic LG and stored there awaiting cytotoxic lymphocyte activation [127]. Despite these structural insights, the exact motifs responsible for GZB trafficking from synthesis through to the target cell cytosol are unknown. The other major component of this pathway is perforin, a long, thin protein that forms pores in targeted cells, and is stored in LGs along with granzyme-serglycin aggregates [129].

Granule synthesis occurs during cytotoxic lymphocyte development, and granules are prepositioned and ready for secretion upon target cell recognition [130]. Initial target cell interaction is mediated by integrins on surface of cytotoxic lymphocytes, a prominent example being LFA-1 [131]. T-cell activation results from antigen specific TCR recognition of a short peptide in the context of MHC [132]. More complicated is NK cell activation, which is a function of the relative balance between a host of inhibitory and activating receptors [124]. In both cell types, these surface receptor interactions result in the formation of the immunological synapse, a tight apposition between effector and target cell, with a peripheral ring of adhesion molecules (pSMAC) and a central region of target recognition molecules (cSMAC) [133]. Surface receptor ligand interaction (e.g. TCR-pMHC in T cells) results in activation of canonical cytotoxic lymphocyte intracellular signaling pathways, with a phosphorylation cascade converging on the assembly of the LAT signaling complex, which activates the MAPK/ERK pathway and initiates calcium influx into the cytotoxic lymphocyte via the PLC pathway [131, 134]. These pathways initiate cytoskeletal remodeling, with the microtubule organizing center (MTOC) polarizing to the immunological synapse, with lytic granules driven by dynein following the MTOC along microtubules [129]. Arriving at the synapse, surface molecules on the granules and cytoplasmic cell membrane facilitate docking, followed by fusion of the granule and membrane lipid bilayers, which results in the exocytosis of the lytic granule contents (including granzymes and perforin) into the synapse [129].

Perforin and granzyme diffuse across to the target cell membrane, into which perforin inserts, and then aggregates to form multimeric, transmembrane pores [135]. Historically it was thought that these perforin pores were directly responsible for target cell death, but it is now believed that physiological concentrations of perforin alone are not cytotoxic. Instead, the pores seem to be only briefly patent before membrane integrity is restored, with their main function being a conduit for passive diffusion of GZB into the target cell [136, 137]. Once in the cytosol, it is GZB that initiates apoptosis by cleavage of BH3 interacting-domain death agonist (BID) and caspases 3,7 and 8, which in turn activate the mitochondrial and caspase apoptosis pathways respectively [122, 135].

In summary, the synergistic activities of granzyme and perforin represent a unique pathway for transferring molecules from cytotoxic lymphocyte to target cell exclusively, as the immunological synapse confines granzyme and perforin between the two cells, and moreover, significant numbers of perforin molecules are required to form the pores required for granzyme's entry into the target cell.

1.3 Overview of a lymphocyte-based delivery system

The overall objective of this project is to develop a cell-to-cell therapeutic delivery system, that is built on a cytotoxic lymphocyte chassis, targeted by a CAR or TCR, and that uses the granzyme-perforin pathway to deliver a protein payload to a target cell. The specific focus of this thesis is to engineer the granzyme-perforin pathway as the delivery module of this system.

1.3.1 The granzyme-perforin pathway as a delivery module

In order to construct such a cell-to-cell protein transfer system, my approach is to use granzyme B as a molecular chaperone to mark the therapeutic payload for transfer to the target cell via the granzyme-perforin pathway. This will be achieved by fusing the payload to granzyme B (or derivatives thereof), such that the payload fusion protein will be expressed and packaged into LGs in preparation for release upon target cell encounter. Transfer of a granzyme-fluorescent fusion protein has been demonstrated previously, although the data is either in primary mouse cells [138] or of questionable validity [139]. In developing this system, several biophysical parameters must be considered for any potential therapeutic payload. First and foremost, the fusion protein must transit perforin pores that have been measured via electron microscopy to have an average luminal diameter of 13-20 nm [137]. Combining this data with the diameters of GZB and GFP (5 nm [137] and 3.5 nm [140] respectively), gives an approximate size restriction on potential payloads. Another significant constraint on the payload is that it must always be at the C-terminal end of the fusion protein, since GZB must be at the N-terminus, to ensure that the signal- and pro-peptides are appropriately processed. Therefore, any payload with critical motifs at its extreme N-terminus may have decreased or absent functionality at the C-terminus of a fusion protein. The stability of these fusion proteins in the harsh, acidic, proteolytic environment of lytic granules will also need to be assessed for each payload. Furthermore, for transit through the granzyme-perforin pathway, and functional activity once in the target cell, other factors such as fusion protein folding and

solubilization, external exposure of important signaling motifs, steric and electrostatic interaction, and charge distribution will affect the success of a particular fusion protein.

Thus on demand cell-to-cell protein transfer is enabled by a combination of the prepositioned lytic granules, the immunological synapse, and the granzyme-perforin pathway itself. The potential utility of this core function justify studying these systems in an attempt to repurpose them for various cell therapy applications.

1.3.2 Comparison with existing systems

This approach has a range of theoretical advantages over other related therapeutic modalities. Compared to biologics and small molecules, the sequestration of a therapeutic inside a delivery lymphocyte may well improve bioavailability, and enable the delivery of therapeutics that would be toxic if administered systemically. The combination of receptor targeting, and the confinement of the therapeutic in the immunological synapse may enable a level of specificity in the delivery of a therapeutic that is otherwise unattainable, except in the case of antibody based drugs. Antibody-conjugate therapeutics are a mature technology, and would likely have the same level of specificity as this approach. Interestingly, the best estimates for the amount of payload that would be delivered are fairly similar for cellular and antibody mediated delivery: on the order of hundreds of molecules per target cell [141, 142]. However, it is important to note that the estimate for granzyme delivery is taken from a computational study of the unmodified pathway, and as such the comparison should be interpreted with caution. A potential advantage of cellular delivery of a payload is that its bioavailability might be far greater. In the case of antibody-drug conjugates, the payload is exposed in transit to the target tissue, as opposed to intracellularly sequestered, which may greatly decrease the immunogenicity and clearance of the delivered payload. Furthermore, the potential for cellular control logic (for example a suicide switch, or molecular sensors) may allow for ongoing, post-administration control of the effects of the therapeutic.

Use of a cellular delivery system might be a useful bridge therapy to delay the decay of damaged tissues, but ultimately regenerative medicine and stem cell therapies will clearly be superior for actual replacement of damaged tissue. However, this assumes that an appropriate replacement cell population or tissue is available for all damaged tissues, which is not currently the case. Furthermore, depending on the location of the damaged tissue, and the properties of the incoming graft, implantation may be challenging or impossible. Conversely a cellular delivery system that is motile might be useful throughout the body for delivery of a therapeutic that partially regenerates damaged tissue, although this would require additional modification of the delivery lymphocyte (discussed below).

Many of these advantages would apply equally to gene therapy and mesenchymal stem cell therapy: reliable and specific activity at disease sites are challenges that both of these fields have struggled with. Again, the broad tissue distribution and target cell specificity of a lymphocyte delivery system might address both of these issues. In the case of lymphocyte delivery, presumably all delivered therapeutic payloads would have a half-life in the target cell. Depending on the therapeutic application, redosing might be required. In some cases this might be a disadvantage as compared to other, more permanent

types of therapy such as regenerative medicine, stem cells, iPSC, gene therapy, or MSC therapy. Alternatively this might be an advantage: the lack of permanent modification is a substantial safety benefit, and offers greater flexibility.

A significant challenge to using a cytotoxic lymphocyte chassis, and one which is not encountered in viral or mesenchymal technologies, is that unmodified cytotoxic lymphocytes will kill any target cell to which they deliver a payload. If used in an application in which the intent of payload delivery is to eliminate the target cell, as would be the case in a cancer or infectious disease context, this would not be a concern. However, for most other applications, such as delivery of a pro-survival or anti-inflammatory payload in the context of degenerative disorders, the delivery cell chassis would have to be rendered non-cytotoxic. This could be achieved through knockdown or knockout of the genes that code for lymphocyte cytotoxic effector proteins, or by reconstituting the granzyme perforin pathway in an inert cellular chassis. This is discussed further in Chapter 5 as a future direction, but is not a focus of this thesis, which is rather to provide proof-of-principle of granzyme-perforin mediated delivery.

1.4 Thesis overview

This thesis has three data chapters, followed by a final chapter of conclusions and discussion. As all of the chapters have or will be published as stand-alone manuscripts, discussion relevant to each chapter is presented at the end of that chapter. The final chapter is mainly concerned with ways to improve the delivery system, future directions for the project, and some broader insights and questions surrounding cellular therapeutics.

Chapter 2 is a computational biophysical study of the immunological synapse and the behavior of granzyme and perforin within the synapse. Based on my computational results, I question some of the core assumptions surrounding the mechanism of cytotoxic lymphocyte specificity and the immunological synapse, and suggest this specificity is the result of granzyme-perforin spatiotemporal dynamics, rather than immunological synapse geometry.

Chapter 3 demonstrates proof-of-principle that the granzyme-perforin pathway can be used to deliver a protein payload to a target cell population. I first designed a suite of granzyme B derived molecular chaperones, and fused them to mCherry as a model payload. I then screened these chaperones for their ability to load mCherry into lytic granules, using fluorescence microscopy. This generated two candidates, which I tested further to see if they were transferred to target cells. Using a model natural killer cell line, I demonstrate transfer of a granzyme B mCherry fusion protein to target cells.

Chapter 4 collects efforts to use this approach to deliver potent, orthogonal toxins to target cells that are resistant to lymphocyte cytotoxicity. I generate a variety of granzyme B toxin fusion proteins, and investigate their activity as fusion partners. I then attempt to generate lymphocyte resistant cell lines, efforts which are for the most part unsuccessful. Using an effector cell dose response curve, I attempt to demonstrate that effector natural killer cells armed with granzyme-toxin fusion proteins exhibit enhanced killing of target cells. I observe moderate effect sizes. I conclude this application

merits further investigation and optimization prior to any final judgment regarding its therapeutic utility.

Chapter 2

A computational biophysical model of the granzyme-perforin pathway

2.1 Introduction

Upon cytotoxic lymphocyte (CL) recognition of a target cell via surface receptor interactions, the so-called immunological synapse (IS) is formed – a region of tight proximity between the CL and target cell membranes in which two distinct killing pathways unfold. The first is the death receptor pathway, which is mainly thought to be important in the context of maintaining T-cell homeostasis and deleting autoreactive T-cells. Fas ligand expressed on the surface of the CL stimulates Fas receptors on the target cell, leading to receptor aggregation and activation of the extrinsic apoptosis pathway. The second main way in which CLs kill their targets is via exocytosis of lytic granules containing, among others, perforin and granzymes into the IS [129, 135, 143]. Here we confine our discussion to granzyme B (GZB), as it is the most important member of the granzyme family in inducing target cell death, possibly along with granzyme A, although this is controversial [144]. Perforin and GZB diffuse across the IS to the target cell membrane, where GZB achieves entry to the cytosol in a perforin-dependent manner. Once internalized, GZB, a serine protease, initiates apoptosis by cleavage of BH3 interacting-domain death agonist (BID) and caspase-3.

Exactly how perforin mediates GZB access to the target cell in the context of the IS has been the subject of debate for over two decades, with two principle models having been investigated [143]. The simpler model proposes that perforin creates pores in the target cell membrane, allowing GZB to diffuse into the cytosol of the target. The more complex theory suggested that perforin and GZB bind regions of the target cell membrane within the IS which are then rapidly endocytosed. Perforin pores form within the endosomes, allowing GZB to be released into the target cell. However, recent high resolution microscopic studies strongly support the simpler model whereby perforin monomers insert into the target cell membrane, and then combine to form multimeric pores, through which GZB can subsequently diffuse [136]. These pores have recently been observed and characterized using cryo-

electron microscopy [137]. It has also been demonstrated that perforin pores are rapidly repaired by the target cell, leaving only a short window of time for GZB to enter the cytosol [136].

Taken together, these studies raise interesting questions about the relative timescales for diffusion, pore formation and GZB delivery. Despite a large investment in experimental effort, we are aware of no existing theoretical consideration of this system that allows these questions to be resolved; previous theoretical work in which we developed analytic solutions for the concentration of a diffusing species in the synapse volume based on partial differential equations is restricted to a single diffusing chemical [145]. Here, we consider nonlinear kinetics of perforin aggregation and small numbers of multiple diffusing molecules. To accurately capture both these aspects of the problem, we apply a spatial stochastic simulation algorithm (SSSA). This method, although relatively time-consuming computationally, allows us to gain insight into this nonlinear system, and to obtain probability distributions of events in the model rather than just the mean behaviour, both of which are unobtainable with differential equation methods. Using this approach, we develop and analyze a mathematical model of GZB delivery via perforin pores. Our model allows us to show that perforin pore facilitated GZB entry into the target cell can support rapid, targeted killing. However, reliable pore formation requires previously unconsidered constraints on the rate of diffusive transport within the IS, which we hypothesize is due to molecular crowding in the synapse.

2.2 Methods

We seek to describe the dynamics of GZB and perforin (PFN) from their release from lytic granules, through their diffusion throughout the IS, to PFN pore formation and GZB internalization. We first provide a description of our biophysical model of this system, followed by its mathematical and computational implementation.

2.2.1 Biophysical model: Geometry & molecular processes

The IS is an irregular narrow region between the CL and target cell that has a very high aspect ratio: the radius of the enclosed region is on the order of microns, while the distance between the two cells is on the order of tens of nanometers [146]. Therefore we model the IS as a very flat, broad disc of radius $R = 3 \mu\text{m}$ and height $h = 20 \text{nm}$, as shown in Figure 2.1, with the CL membrane taken to be the upper surface of the disc, and the target membrane considered explicitly, immediately below the lower surface of the disc. Since $h \sim 20 - 40 \text{nm}$, and the diameter of GZB and PFN are $\sim 5 \text{nm}$ and $\sim 8 \text{nm}$ respectively [137, 147], we allow that molecules may escape through the synapse edge.

Given that exocytosis of lytic granules is temporally synchronized [129], and that the time-scales of both exocytosis (on the order of milliseconds [148]) and diffusion across the synapse (calculated using the Stokes-Einstein relationship to be on the order of microseconds using the dimensions of the synapse given above) are much faster than pore formation (observed to be on the order of seconds [136, 149, 150]), we assume that GZB and PFN are instantaneously released from the CL membrane as an

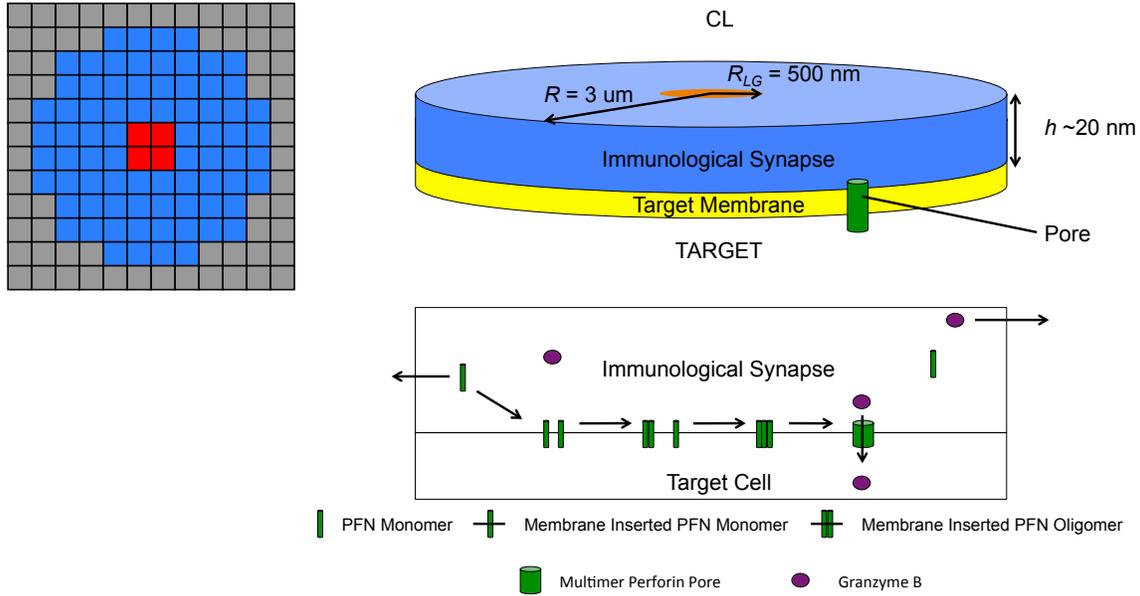


Figure 2.1: Model geometry and molecular interactions. We consider the synapse (blue) as a broad flat disc, with the upper surface the CL membrane, and the lower surface the target cell membrane (yellow). GZB (purple circles) and PFN (green cylinders) are released from a central lytic granule (red). We discretize this space into a two dimensional mesh of sub-volumes (upper left). The time evolution of the system is then governed by diffusive jumps between sub-volumes, and interactions between molecules within a sub-volume. These interactions encompass PFN membrane insertion and oligomerization leading to pore formation, followed by GZB internalization through pores (lower right).

initial bolus. The exact location of granule release, the so-called secretory domain, has been variously reported as both central [151] and in between the central and peripheral supramolecular activation complexes [152]. For simplicity, we assume GZB and PFN are released from a single lytic granule of radius $R_{LG} = 500 \text{ nm}$ [151] at the centre of the synapse. This assumption is also maximally conservative with respect to molecular escape from the synapse (see discussion below).

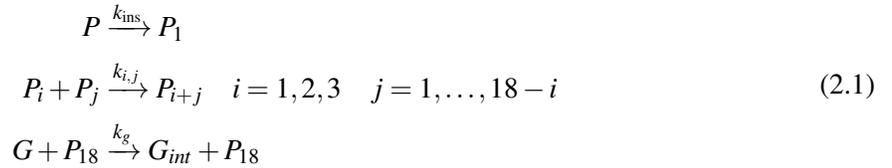
Within this geometry, we model the spatiotemporal dynamics of GZB and PFN by considering diffusive transport, as well as chemical interaction (schematically depicted in Figure 2.1). Both molecules diffuse throughout the synapse, eventually either escaping at its lateral edge or interacting with the target cell membrane, as described below. Due to the extreme aspect ratio of the synapse, the timescale for diffusive transport across the height of the synapse is short compared to all other relevant processes, and therefore we approximate diffusion in the synapse as two dimensional in the horizontal plane.

PFN monomers insert into the target cell membrane with rate k_{ins} , which we assume is slower than the diffusion limited rate, due to the energy requirements of lipid membrane displacement for perforin insertion. Membrane inserted monomers can then diffuse across the membrane and potentially combine to form pores. Based on an analysis of electron micrographs that indicates that pores consist of a ring

of 18-20 PFN monomers spanning the target cell membrane [137], we modelled pores as 18mers. We modelled the path to pore formation as a multistep, multi-pathway oligomerization process, which consists of monomer dimerization as well as monomer and dimer aggregation to form trimers. Monomers, dimers and trimers then combine with each other to form higher order oligomers. Since membrane diffusivity scales inversely with molecular size, higher order oligomers will be decreasingly mobile. They will also be sparsely distributed and therefore it is very unlikely that a higher order oligomer would encounter another higher order oligomer before a low order oligomer. We used this observation to simplify our model of pore formation by neglecting any interaction between two oligomers greater than a trimer: oligomers can only grow in size by combining with a monomer, dimer or trimer. We assume the rate of PFN oligomer aggregation is diffusion limited, and denote this rate $k_{i,j}$. We neglect reverse reactions for both membrane insertion and oligomerization.

Once an 18mer has formed, this becomes a pore through which GZB can diffuse, which occurs with rate k_g . We assume that this process is diffusion limited, and neglect the reverse reaction.

Symbolically we have the following reaction scheme:



where P and G represent synaptic PFN and GZB respectively, P_i denotes a membrane inserted PFN i -mer for $i = 1, \dots, 18$. G_{int} denotes internalized GZB, and k_{ins} , $k_{i,j}$ and k_g denote the rates of PFN membrane insertion, perforin oligomerization and GZB pore transit respectively. We assume that during the short timescale of the processes we are modelling, no GZB or PFN molecules are lost due to other processes such as irreversible non-specific binding of these proteins in the synapse, loss of P_j molecules in the target cell membrane due to endocytosis, or loss of functional activity of either molecule due to irreversible inactivation in the IS.

In summary, our model consists of a broad, flat, disc-shaped IS between CL and target cell. GZB and PFN are released as an instantaneous bolus from a single, centrally located lytic granule, whereupon they diffuse throughout the synapse, with the potential for any molecule to diffuse out of the IS through the lateral edges. Perforin inserts into the target cell membrane and oligomerizes to form pores, through which GZB can then diffuse.

2.2.2 Mathematical description: Spatial stochastic simulation algorithm

Since the numbers of certain molecular species (such as 18-mer pores) are very low, continuous models derived from mass action kinetics are inappropriate for describing our system, and stochastic methods are instead necessary. Furthermore, due to the localized release of molecules, combined with the relative sparsity of these molecules in the system, spatial effects are important, and homogenous stochastic

models are likely to be insufficient. In order to accurately model small numbers of molecules in space and time, we therefore applied a discrete spatial stochastic simulation algorithm developed by Elf and Ehrenberg [153, 154], which is an extension of the spatially homogenous next reaction method of Gibson and Delbruck [155], itself a computationally more efficient version of the original Gillespie algorithm [156].

The Gillespie method considers a system of reactions X_i that occur with rate constants r_i according to standard chemical kinetics, and asks two questions: (i) when does the next reaction occur? and (ii) which reaction occurs? By first calculating the probability that each reaction occurs, given by $a_i = r_i N_i$ (with N_i denoting the number of instances of reaction X_i), Gillespie showed that the answer to these questions may be obtained by randomly sampling two probability distributions:

$$t_{next} = \frac{1}{a_0} \ln\left(\frac{1}{n_1}\right) \quad (2.2)$$

$$X = X_i \quad \text{if} \quad \frac{1}{a_0} \sum_{j=1}^i a_j \leq n_2 < \frac{1}{a_0} \sum_{j=1}^{i+1} a_j \quad (2.3)$$

where n_1 and n_2 are random numbers between zero and one, and $a_0 = \sum_i a_i$. The time evolution of the system may then be obtained by first incrementing the time by t_{next} and updating the species numbers according to the stoichiometry of reaction X_i , then recalculating the probability distributions according to the updated species numbers, and finally resampling the new probability distributions.

The essence of the SSSA (the spatially inhomogeneous extension of the Gillespie algorithm) is to discretize the physical simulation space into sub-volumes (SVs) of length l , chosen to be small enough that the spatial distribution of species is approximately homogeneous within each sub-volume. This justifies using mass action chemical rate constants to describe the molecular interactions within a sub volume. Diffusion is modelled as another ‘reaction’ in which a molecule jumps from one sub-volume to an adjacent one with a rate constant of $d = nD/l^2$, where $n = 4$ are the spatial degrees of freedom for a diffusive jump, and D is the diffusivity. This mapping of diffusion to a reaction allows for the formalism of the Gillespie algorithm to be employed. The algorithm is then similar to that of the original Gillespie algorithm. An event occurs in the SV_i that has the lowest t_{next} , which is calculated by sampling a probability distribution analogous to those of Equation 2.2, with the probabilities a_i replaced by $s_i = a_i + d_i$. Having chosen when and in which SV an event occurs, the event is a chemical interaction if $n_3 < a_0/s_0$ ($s_0 = \sum s_i$, and n_3 is a random number distributed between zero and one) and diffusive otherwise. If the event is diffusive, which species diffuses is selected by weighting a random distribution by the number of each species within the subvolume, and the direction of diffusion is randomly selected. If the event is a chemical reaction, then the identity of the chemical reaction is determined as in the Gillespie algorithm. Following sampling of the probability distributions, the time and species number (in both the origin and destination SV if the event was diffusive) are updated, the probability distributions recalculated. The spatiotemporal evolution of the system is obtained by repeatedly iterating the above algorithm.

It should be noted that there are several nuances to the version of the SSSA presented by Elf and Ehrenberg that was implemented in this work that result in increased computational efficiency, and for details we refer the reader to their work.

We constructed a discretized IS with two-dimensional sub-volumes of side length l as shown in Figure 2.1. There are three regions: the central lytic granule which contains PFN and GZB initially, the rest of the synapse, and a region external to the synapse. This last region is present to allow for escape from the synapse and return from the exterior is prohibited. Finally, within each sub-volume, chemical reactions occur according to the chemical reaction scheme in Equation 2.1.

2.2.3 Hindered diffusion in the immunological synapse

We consider molecules in the synapse as roughly spherical particles diffusing freely in a bulk fluid, and so estimate their diffusivity from the Stokes-Einstein relation, $D_{\text{free}} = K_bT / (6\pi\eta_w r)$ where K_bT is the thermal energy, η_w is the solvent viscosity and r is the radius of the molecule. Additionally, the diffusivity is modulated by a hindered diffusion parameter, α , which is motivated by the observation of very high electron density in the IS, which we hypothesize is due to densely packed extracellular adhesion and signalling molecules. This molecular crowding in the IS has two effects, both resulting in a decrease in effective diffusivity. First, non-specific binding of GZB and PFN to intra-synaptic molecules decreases the total time during which GZB or PFN are free to diffuse. Second, the space occupied by intra-synaptic molecules is not available for GZB and PFN diffusion. The effects of the latter are accounted for by multiplying the free diffusivity by the volume fraction of the synapse still available for free diffusion: $(1 - f)$, where f is the volume fraction of the synapse occupied by the intra-synaptic molecules. To derive an expression to model the effects of non-specific binding we consider a molecule (GZB or PFN) diffusing in the synapse which is filled with other molecules (referred to hereafter as binders) filling the synapse at a number density ρ . We assume that the diffusing molecule binds a binder with a rate ρk_{on} and disassociates from it at a rate k_{off} . As these binders are attached to either the CL or the target cell, we assume that the diffusing molecule is immobile when bound to a binder. Thus the time the molecule spends free to diffuse or bound is proportional to the inverse of the associated binding and unbinding rates respectively, and the fraction of time (τ) that a diffusing molecule spends free and unbound is, after minor algebraic manipulation,

$$\tau = \frac{k_{\text{off}}}{\rho k_{\text{on}} + k_{\text{off}}} = \frac{k_D}{\rho + k_D} \quad k_D = \frac{k_{\text{off}}}{k_{\text{on}}} \quad (2.4)$$

where the second equality is obtained by dividing through by k_{on} , and k_D is the dissociation constant for nonspecific protein-protein interactions, which can be experimentally measured. To calculate ρ we note that it is equal to the number of molecules (N) in the synapse divided by the volume of the synapse (V_{syn}). We can approximate the number of molecules in the synapse by taking it to be the total volume of the synapse occupied by molecules (which itself is fV_{syn}) divided by the volume of an

average molecule in the synapse (V_{avg}):

$$\rho = \frac{N}{V_{\text{syn}}} = \frac{fV_{\text{syn}}/V_{\text{avg}}}{V_{\text{syn}}} = \frac{f}{V_{\text{avg}}}. \quad (2.5)$$

To obtain an expression for α , which relates the effective diffusivity of molecules in the synapse (D_{eff}) to the corresponding free (Stokes-Einstein) diffusivity (D_{free}), we multiply the volume fraction of the synapse still available for free diffusion ($1 - f$) by the fraction of time which molecules are able to diffuse freely (τ):

$$D_{\text{eff}} = \alpha D_{\text{free}} \quad \alpha = \frac{k_D}{\rho + k_D} (1 - f) \quad \rho = \frac{f}{V_{\text{avg}}}. \quad (2.6)$$

2.2.4 Perforin oligomer diffusion and aggregation in the target cell membrane

The diffusivity of proteins in cell membranes is a problem that has received considerable experimental and theoretical attention [157–159]. These efforts have demonstrated that the diffusivity varies with the total protein density in the membrane, as well as the radius of the diffusing protein, both of which are relevant for our consideration of a perforin oligomer of changing radius in the highly crowded IS. Recent experimental work has shown that at low densities, the diffusivity scales according to a Saffman-Delbruck [157] type relationship ($\ln[1/r]$, where r is the radius of the diffusing molecule), whereas in highly crowded membranes, diffusivity scales as $1/r$ [159]. This latter result was also arrived at independently by Gambin and colleagues [158], and is consistent with the Stokes-Einstein description. It is difficult to estimate whether the membranes at the immunological synapse are highly crowded or not in the terms of these experimental papers. We performed simulations using both relationships and found that pore formation and granzyme delivery were extremely improbable under the Stokes-Einstein description. Therefore, to be maximally conservative in terms of minimizing the requirement for hindered bulk diffusion in the synapse, we present simulations using the Saffman-Delbruck relationship to describe the diffusivity of a perforin oligo in the membrane,

$$D_j = \frac{K_b T}{4\pi\eta_l \xi} (\ln(\eta_l \xi / \eta_w r_{P_j}) - \gamma) \quad (2.7)$$

where $K_b T$ is the thermal energy, η_w and η_l are the solvent and membrane viscosity respectively, ξ is the thickness of the cell membrane, r_{P_j} is the radius of a perforin j -mer and γ is Euler's constant. The radius of a j -mer is taken to be the average of its long and short dimensions: $r_{P_j} = R_{PFN}(j + 1)/2$, where R_{PFN} is the radius of a perforin monomer. If two oligomers are in the same sub-volume, then they may combine to form a higher order oligomer, subject to the conditions in Equation 2.1. We take the rate constant $k_{i,j}$ for this process to be the diffusion limited rate for a particle (in our case the smaller

oligomer P_i) finding a trap (the larger oligomer P_j) on the surface of a membrane [160]:

$$k_{i,j} = \frac{2\pi(D_i + D_j)}{\ln(b/s)} \quad \text{where} \quad b = \sqrt{\frac{l^2}{\pi N_{P_i}}} \quad \text{and} \quad s = r_{P_i} + r_{P_j}. \quad (2.8)$$

Here l^2 is surface area of the sub-volume, which is the space the N_{P_i} i -mers explore in finding the j -mer trap, and s is the reaction radius, which we have taken to be the sum of the radii of the two oligomers. This is the maximum possible reaction rate, and so we are considering the fastest possible rate of pore formation in this model.

2.2.5 Rate of granzyme translocation through perforin pores

In order to derive a rate for GZB internalization through a pore, we consider two processes: (i) finding the pore; and (ii) translocating through the pore. The first step can be described by considering a particle in a cuboid with height h (the height of the synapse) and cross-section l^2 (the area of the sub-volume the particle is in). If we assume some diffusion driven flux Φ of particles hitting the bottom of the cuboid (namely the target cell membrane) then approximately $\Phi A_p/l^2$ particles per unit time will hit a pore, where A_p is the cross-sectional area of a pore. If we then divide this by the number of particles in the cuboid, we will have a rate for the particle finding the pore. To derive an expression for the flux Φ , we first note that lateral diffusion occurs when the particle jumps to the next sub-volume, while we wish to calculate the rate at which the particle hits the target cell membrane in those cases where the particle remains in the sub-volume. Therefore, the problem can be described by a one dimensional diffusion equation at equilibrium, with reflecting boundary conditions at the top of the cuboid (since we assume no particles are reabsorbed by the target cell) and absorbing boundary conditions at the bottom of the cuboid (since we are calculating the flux of all particles hitting the membrane). Solving the one dimensional diffusion equation allows us to calculate the flux Φ at the target cell membrane, which yields the rate $k_g = 3\pi R_p^2 D/l^2 h^2$ for GZB hitting a pore, where $R_p = 20\text{nm}$ is the pore radius [137].

To derive a rate for GZB translocation, we note that electron micrographs of PFN multimer pores show a relatively smooth and uniform pore, with no evidence of gates or obstruction to particle entry [137]. Therefore, we assume that pore transit can be modelled as a one-dimensional diffusive processes, with a rate that scales as $1/\xi^2$ where $\xi \sim 10\text{nm}$ is the thickness of the membrane [161]. Comparing this rate to that derived above for GZB finding a pore, and recalling that the the height of the synapse is $h \sim 20\text{nm}$ we see that GZB finding the pore will be the rate limiting step. Given that the GZB internalization process in our model is identical in form to an enzymatic reaction (Equation 2.1), it is conceivable that GZB might saturate pores, in which case Michaelis-Menten kinetics would be more appropriate than mass action kinetics. However, using the initial (maximal) concentration of GZB in our system, and a pore diameter of 20nm [137] we estimate the average number of GZB per pore to be on the order of 10^{-2} , and so we find that our assumption of mass action kinetics for pore translocation would be valid even for much higher GZB concentrations.

2.2.6 Model parameterization and implementation

To calculate the number of PFN and GZB molecules released into the synapse we first estimated the total number of these molecules in a cytotoxic lymphocyte. For perforin this was reported to be an average of 500 PFN molecules for CTLs and 3500 for NK cells [162], and so we selected an arbitrary intermediate value of 1500. We could not find any estimates of the number of GZB molecules per cell, so we instead used RNA expression data showing GZB mRNA copy number is ten times that of PFN [163] to set the number of GZB molecules at 15000. Given that mRNA copy number often correlates poorly with protein expression levels [164], that the experimental work quantifying perforin number per cell has several technical limitations, and the uncertainty surrounding the number of molecules released in a lymphocyte-target cell interaction, we use these values only as a starting point from which to subsequently explore the effects of PFN and GZB concentration (see Results).

To derive the hindered diffusion parameter α , we must estimate the volume of an average protein spanning the synapse (V_{avg}), the volume occupancy of these proteins in the synapse (f), and the dissociation constant for non-specific binding interactions of GZB or PFN with such an average protein (k_D). To strengthen our arguments about the importance of molecular crowding, we chose our parameters so as to reasonably maximize α and thus minimize the effects of hindered diffusion for a given parameter set. To estimate V_{avg} , we used an average molecular weight of 200 kDa as representative of abundant signalling and adhesion molecules found in the synapse (e.g. ICAM1, LFA1, CD45), and an average protein density of 1.35 g/cm^3 (valid for molecules larger than 20 kDa) [165], to calculate $V_{\text{avg}} = 250 \text{ nm}^3$. To estimate f we note that electron micrographs of the synapse show it as much more electron dense than the average density of the cytoplasm. Therefore we reasoned that the maximum estimates of the volume occupancy of the cytoplasm would be an appropriate lower bound on the volume occupancy of the synapse, choosing $f = 0.4$ [166]. Finally, again aiming to maximize α , we chose $k_D = 10^{-3} \text{ M}$, based on estimates of $10^{-6} - 10^{-3} \text{ M}$ for non-specific protein-protein interactions [167]. Finally, we used recent experimental data to estimate the rate of perforin insertion into the membrane as $k_{\text{ins}} = 2 \text{ s}^{-1}$ [168].

A summary of the parameters used in this study, with literature sources, is shown in Table 2.1. The model was implemented in MATLAB and then compiled as a standalone executable. The core source code is reproduced in the Appendix (Section D.1). For each parameter set, 100 independent simulations were conducted. Simulations terminated when no further diffusive or chemical events were possible. Data analysis, visualization and plotting was conducted using the R statistical analysis language.

2.2.7 Validation of SSSA computational implementation

The dynamics of perforin oligomerization and GZB translocation are nonlinear and no analytic solution exists that describes our model, which motivated our choice of the SSSA method to study this system. However, to ensure that our algorithm was correctly implemented, we sought to compare its predictions to those cases for which analytic solutions are tractable: chemical reactions according to mass action,

Symbol	Value	Units	Description	Reference
R	3	μm	IS radius	[146]
h	20	nm	IS height	[146]
ξ	10	nm	Thickness of cell membrane	[161]
R_{LG}	0.5	μm	Lytic granule radius	[151]
R_{GZB}	2.5	nm	GZB radius	[147]
R_{PFN}	4	nm	Perforin radius	[137]
R_p	10	nm	Perforin pore luminal radius	[137]
N_{PFN}	1500	-	Number of Perforin monomers released	[162, 169]
N_{GZB}	15000	-	Number of GZB molecules released	[162, 163, 169]
V_{avg}	250	nm^3	Avg. volume of synaptic molecules	[165]
f	0.4	-	Fractional synaptic occupancy	[166]
k_D	10^{-3}	M	Diss. constant for non-specific binding	[167]
k_{ins}	2	s^{-1}	Rate of perforin monomer insertion	[168]
T	310	K	Human body temperature	-
η_w	6.53×10^{-10}	$\text{kg } \mu\text{m}^{-1}\text{s}^{-1}$	Viscosity of water	-
η_l	$100\eta_w$	$\text{kg } \mu\text{m}^{-1}\text{s}^{-1}$	Viscosity of cell membrane	-
l	0.5	μm	Sub-volume side length	-

Table 2.1: Parameters used in the immune synapse computational model.

and molecular diffusion due to Brownian motion.

Chemical kinetics

We consider the simple case of a three species system given by $A + B \xrightarrow{k} C$, which can be described by the following set of ODEs:

$$\begin{aligned}
 \frac{dA}{dt} &= -kAB \\
 \frac{dB}{dt} &= -kAB \\
 \frac{dC}{dt} &= kAB
 \end{aligned}
 \tag{2.9}$$

To implement this limiting case in our SSSA model we set all diffusivities equal to zero, and initialized a single subvolume with 2000 perforin pores and 3000 granzyme molecules, and then simulated the time evolution of granzyme translocation through the perforin pores. Due to the nonlinearity of the above ODE system, there is no closed form analytic solution, so we solved the system numerically using `ode45` in Matlab, with the initial conditions as described above, and $k = k_g$. The results are shown in Figure 2.2, and show good agreement, indicating that our algorithm correctly describes chemical interactions.

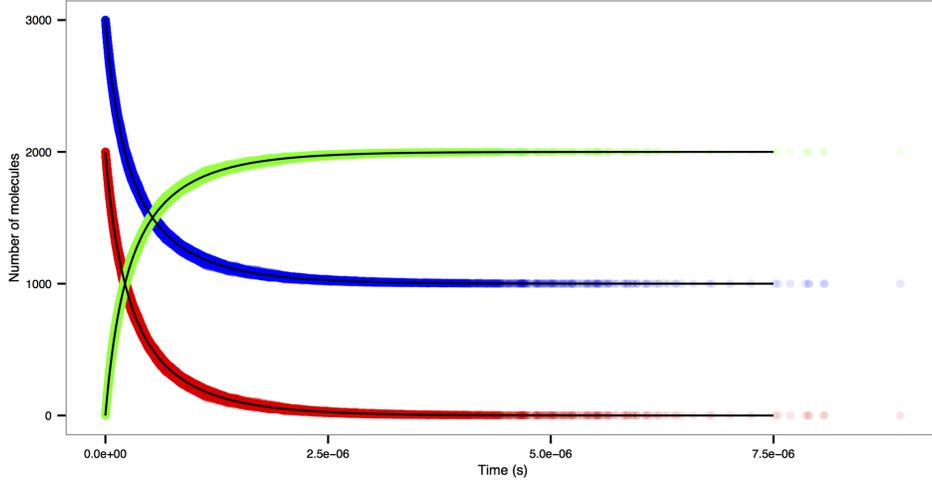


Figure 2.2: Mass action kinetics are accurately captured by the SSSA algorithm. A single sub-volume was initialized with 2000 perforin pores (red) and 3000 granzyme molecules (blue). This system was simulated using our SSSA, with all diffusivities set to zero, and all chemical rate constants set to zero, save the rate of granzyme translocation through perforin pores, which was set to k_g . Data from 100 independent simulations is plotted. Analytic solutions to this system were calculated from the corresponding ODE system in Matlab, and are plotted as black lines. The green data set is internalized granzyme.

Free diffusion in the synapse

We consider an initial bolus of N_o molecules released into the synapse from the centrally located lytic granule (of radius R_{LG}), and aim to describe the diffusive spread of these molecules. To model this process analytically, we sought solutions of the diffusion equation: $\partial u / \partial t = D \nabla^2 u$, where $u = u(x, y, t)$ is the molecular concentration and D the diffusivity. In order to naturally parameterize the square sub volumes, we chose to adopt a square geometry in which solve the diffusion equation, using cartesian coordinates and a square synapse of side length $2R$, chosen to be the same value as the diameter of the circular synapse in the SSSA algorithm. We assume that the aspect ratio is extreme so that the timescale of diffusion in the z -direction (across the synapse) is very short compared to that for xy plane, and spatial gradients in this direction may be neglected. Therefore we sought solutions to the diffusion equation in two dimensions over the domain $S = \{(x, y) | -R \leq x \leq R, -R \leq y \leq R\}$. Since we allow molecules to escape the synapse in our simulations, we adopt absorbing boundary conditions with $u(\pm R, \pm R, t) = 0$. Finally the initial state is taken to be $u(x, y, 0) = u_o = N_o / R_{LG}^2$ for $|x, y| \leq R_{LG}$, and zero elsewhere.

Using separation of variables, it is a fairly simple exercise [170] to find the solution for this problem in terms of sine functions:

$$u(x, y, t) = \sum_{m=1}^{\infty} \sum_{n=1}^{\infty} c_{m,n} \sin(\alpha_m(R+x)) \sin(\alpha_n(R+y)) e^{-D(\alpha_m^2 + \alpha_n^2)t}$$

where $\alpha_n = n\pi/2R$. Using the initial conditions and the orthogonality of sines, it is straightforward to obtain the coefficients:

$$c_{m,n} = u_o \frac{4 \sin(\alpha_m R) \sin(\alpha_n R) \sin(\alpha_m R_{LG}) \sin(\alpha_n R_{LG})}{R^2 \alpha_m \alpha_n}$$

To simulate simple diffusive spreading in our SSSA model, we initialized the sub volumes containing the central lytic granule with $N_o = 2000$ perforin molecules, spread evenly over the sub volumes. We set all rate constants to zero, and simulated the time evolution of the system 100 times, and calculated the average concentration and standard deviation for each sub volume, at each time point. We took a subset of this data ($y = 0.25$) that corresponds to a cross section of sub volumes that run as close as possible to the centre of the lytic granule and synapse. To confirm that this approach was representative, we repeated the subsetting for many different cross-sections, and in all cases the diffusive spread was Gaussian as expected. The concentration profiles from both our analytic and simulated data are plotted for various times in Figure 2.3. We found that our simulations reproduce the analytic solution well, and that the error introduced by the choice of sub volume size (initially $l = 0.5\mu\text{m}$) is reduced at smaller sub volume sizes. Since the agreement was already quite good for $l = 0.5\mu\text{m}$, we felt that the large increase in computational cost was not worth the minor correction in accuracy and all simulations in the main paper are presented with $l = 0.5\mu\text{m}$ as the sub volume dimension.

Free diffusion in the synapse with membrane insertion

We next considered a system as above, with the addition that perforin can insert into the target cell membrane. For our analytic solution the diffusion equation becomes $\partial u / \partial t = D\nabla^2 u - k_{ins}u$, where k_{ins} is the membrane insertion rate of perforin as in the main text. The analytic solution has a very similar form to that of the previous section, with the exponential term becoming $e^{-D(\alpha_n^2 + \alpha_m^2)t - k_{ins}t}$. To simulate this case we adopted the same approach as in the previous section, and set $k_{ins} = 2s^{-1}$, rather than zero. We then generated the analytic and simulated data in the same way, and the results are plotted in Figure 2.4. Again the analytic and simulated data compare well.

Diffusion and reaction in the target cell membrane

To test our SSSA further, we considered the problem of perforin monomers diffusing and reacting in the target cell membrane. Since our model does not allow membrane inserted perforin monomers to escape the synapse, we adopted reflecting (Neumann) rather than absorbing boundary conditions in deriving an analytic solution. In this case separation of variables yields a solution that is a series of cosines rather than sines:

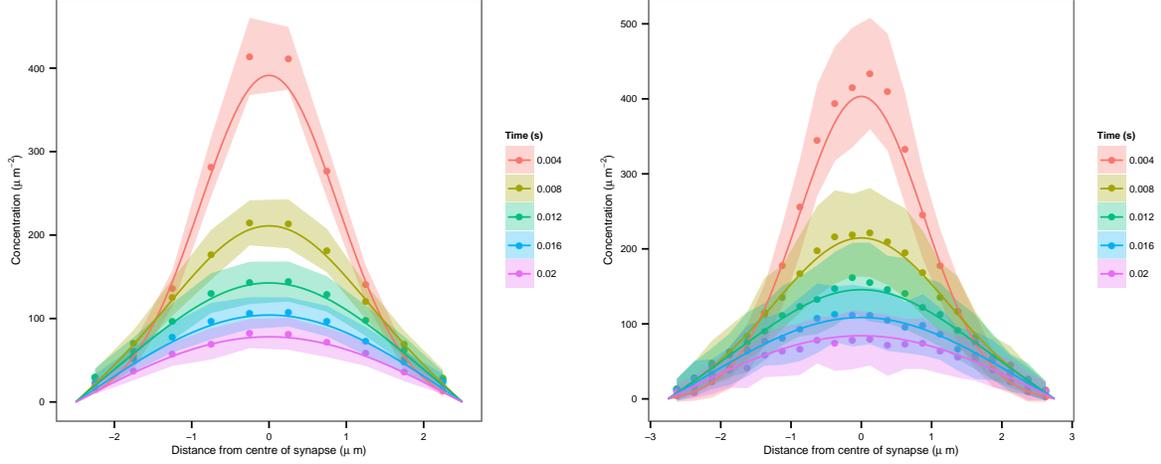


Figure 2.3: Free diffusion described by the two-dimensional diffusion equation compares well with that simulated by the SSSA algorithm. The subvolumes comprising the central lytic granule were initialized with 2000 perforin molecules spread across the subvolumes. All chemical rate constants were set to zero, resulting in free diffusion of perforin, and the system was simulated 100 times. We then plotted the concentration profiles at various times for the plane $y = 0.25$, with each dot representing the concentration in the subvolume with coordinates $(x, 0.25)$, where x is plotted on the x axis in the figure. The shaded region represents the standard deviation of the average concentration calculated from 100 independent simulations. The corresponding solutions to the two-dimensional diffusion equation were calculated in Matlab up to the twentieth term in the infinite sums, and are plotted as a solid line. Left panel: sub volumes of side $l = 0.5\mu\text{m}$. Right panel: sub volumes of side $l = 0.25\mu\text{m}$.

$$\begin{aligned}
 u(x, y, t) = c_{oo} &+ \sum_{n=1}^{\infty} c_n \cos(\alpha_n x) e^{-D\alpha_n^2 t} + \sum_{m=1}^{\infty} c_m \cos(\alpha_m y) e^{-D\alpha_m^2 t} \\
 &+ \sum_{n=1}^{\infty} \sum_{m=1}^{\infty} c_{n,m} \cos(\alpha_n x) \cos(\alpha_m y) e^{-D(\alpha_m^2 + \alpha_n^2)t}
 \end{aligned} \tag{2.10}$$

where $\alpha_n = n\pi/R$. Using the initial conditions and the orthogonality of cosines, we obtain the following expressions for the coefficients:

$$c_{oo} = \frac{u_o R_{LG}^2}{R^2}, \quad c_n = \frac{2R_{LG}u_o}{n\pi R} \sin(\alpha_n R_{LG}), \quad c_{n,m} = \frac{4u_o}{mn\pi^2} \sin(\alpha_n R_{LG}) \sin(\alpha_m R_{LG}).$$

The SSSA simulations were obtained in the same manner as in the initial case of free diffusion, except that the sub volumes were initialized with membrane inserted perforin monomers, rather than free perforin. Perforin oligomer aggregation was suppressed by setting all chemical interaction rate constants to zero. The resulting simulations and analytic solutions are plotted in Figure 2.5. Again the comparisons are reasonable and can be improved by decreasing the sub-volume size (not shown). Note that since the membrane inserted perforin monomers cannot escape, the system reaches an equilibrium

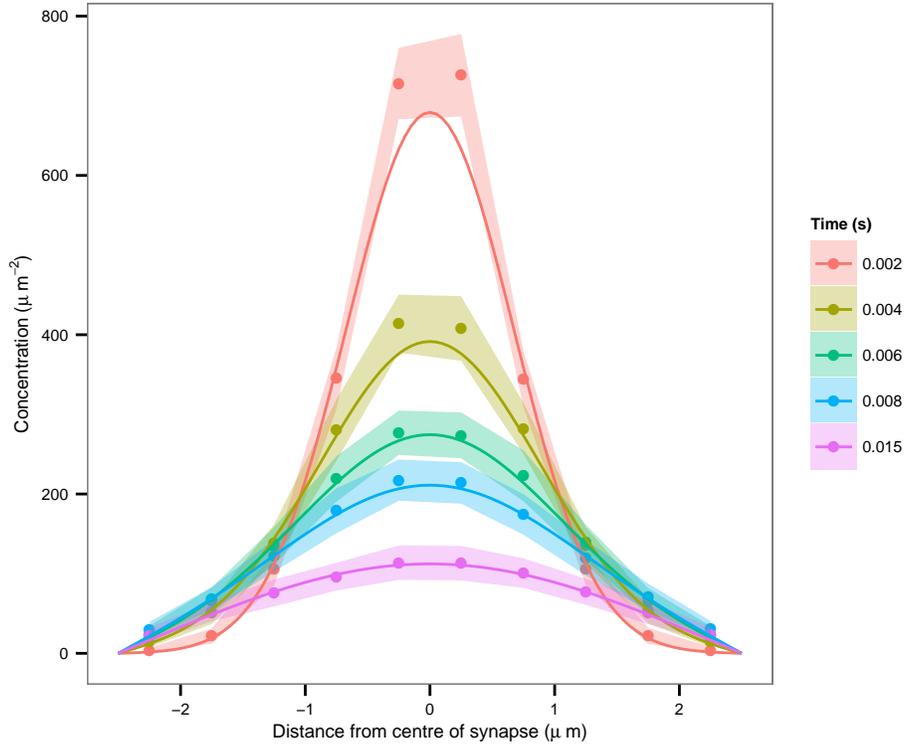


Figure 2.4: Free diffusion with attenuation due to target cell membrane insertion is well described by the SSSA algorithm. The data was generated and plotted in an identical manner to Figure 2.3 save the additional exponential multiplicative term $e^{-k_{ins}t}$ in the analytic solution, and the perforin membrane insertion rate k_{ins} was set at its value stated Table 2.1 (rather than zero as in the previous figures).

concentration given by the initial particle number (N_o) divided by the total area of the synapse. Since this area differs for the analytic ($2R^2$) and SSSA (πR^2) data, the equilibrium concentrations are slightly different.

Finally, we ensured that membrane reactions are correctly captured by our SSSA. We simulated the reaction $A + B \rightarrow C$ for membrane-bound diffusing molecules A and B which react to form an immobile complex C with bimolecular rate constant k . This system can be modelled using the following PDE system for the concentrations $A(x, y, t)$, $B(x, y, t)$ and $C(x, y, t)$:

$$\begin{aligned} \frac{\partial A}{\partial t} &= D_A \nabla^2 A - kAB \\ \frac{\partial B}{\partial t} &= D_B \nabla^2 B - kAB \\ \frac{\partial C}{\partial t} &= kAB. \end{aligned}$$

We simulated this system using our SSSA and compared the results to those obtained from numerical

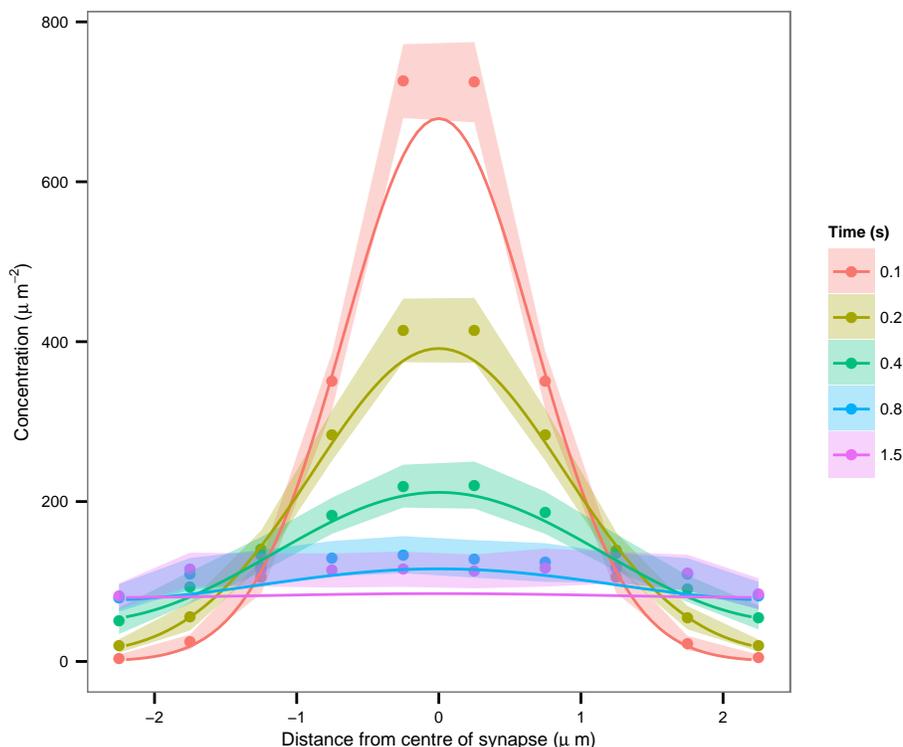


Figure 2.5: Diffusion in the target cell membrane diffusion described by the two-dimensional diffusion equation compares well with that simulated by the SSSA algorithm. The subvolumes comprising the central lytic granule were initialized with 2000 membrane inserted perforin monomers spread across the subvolumes. All chemical rate constants were set to zero, resulting in perforin monomers diffusing in the target cell membrane. The data was then generated and plotted in an identical manner to previous figures.

solution of the PDE system. We used a square domain of size $5\mu\text{m} \times 5\mu\text{m}$. In the SSSA, the domain was divided into square sub-regions of side $0.5\mu\text{m}$. We initialized 500 A molecules and 700 B molecules in the central four sub-regions of the domain and simulated up to a fixed end-time with varying values of k spanning two orders of magnitude. As shown in Figure Figure 2.6, the SSSA and PDE results agree well.

2.3 Results

2.3.1 Free diffusion in the synapse is incompatible with granzyme internalization

We initially built our model without including the hindered diffusion coefficient α . However, when we plotted the various molecular species against time, as in Figure 2.7, it became apparent that virtually all GZB, and a large majority of PFN, escaped the synapse before pore formation could occur. This

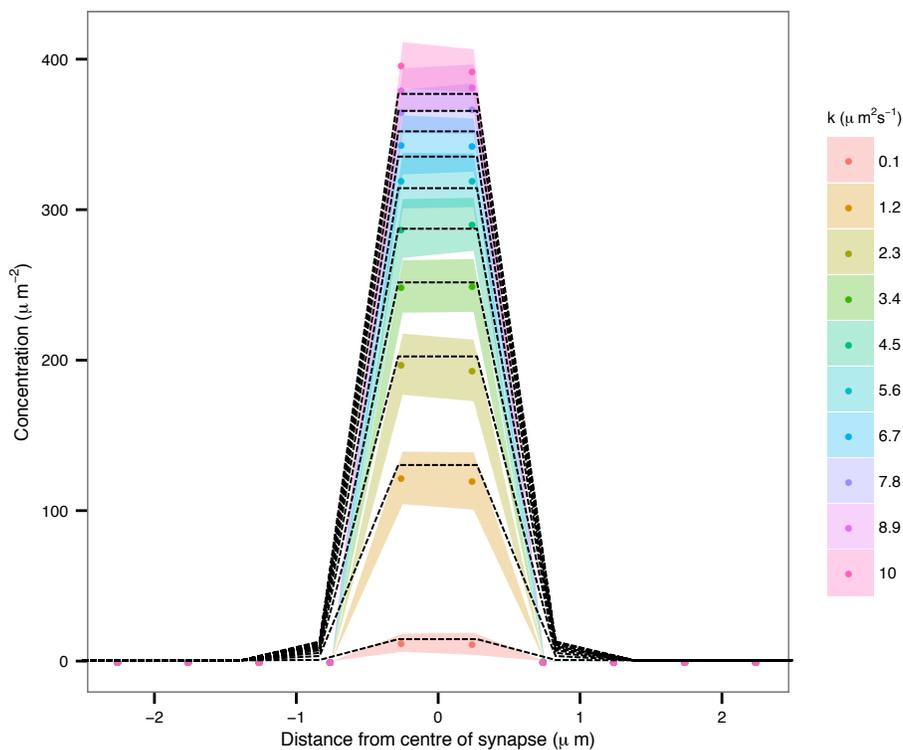


Figure 2.6: A two-dimensional reaction-diffusion system is well described by our SSSA. The SSSA was applied to the two-dimensional $A + B \rightarrow C$ system and concentrations of reaction product C were extracted for a line of sub-regions across the simulation domain, as in previous figures (coloured dots indicate averages and coloured regions \pm standard deviation over 100 simulations). The corresponding PDE system was solved using finite differences and the solution was then averaged over the regions corresponding to the same sub-regions of the SSSA to allow easy comparison (black dashed lines). SSSA and PDE results are shown for 10 different values of k spanning two orders of magnitude. We set $D_A = 2.9\mu\text{m}^2\text{s}^{-1}$ and $D_B = 2.7\mu\text{m}^2\text{s}^{-1}$

continued to be the case when we increased the numbers of both GZB and PFN by one to two orders of magnitude. Since lytic granule mediated cytotoxicity is known to be crucial for CL killing, this lack of internalization indicated that the aspect ratio of the synapse alone was insufficient to contain molecules in the IS. This motivated us to re-examine the physical environment of the IS. In contrast to a simple aqueous space, the IS contains a high density of signalling and adhesion molecules, giving it a characteristic electron dense appearance in electron micrographs. Incorporating a hindered diffusion parameter α (see Methods) to model the reduced diffusivity of GZB and PFN due to these conditions resulted in a marked decrease in the loss of both molecules, leading to increased pore formation probability and GZB internalization.

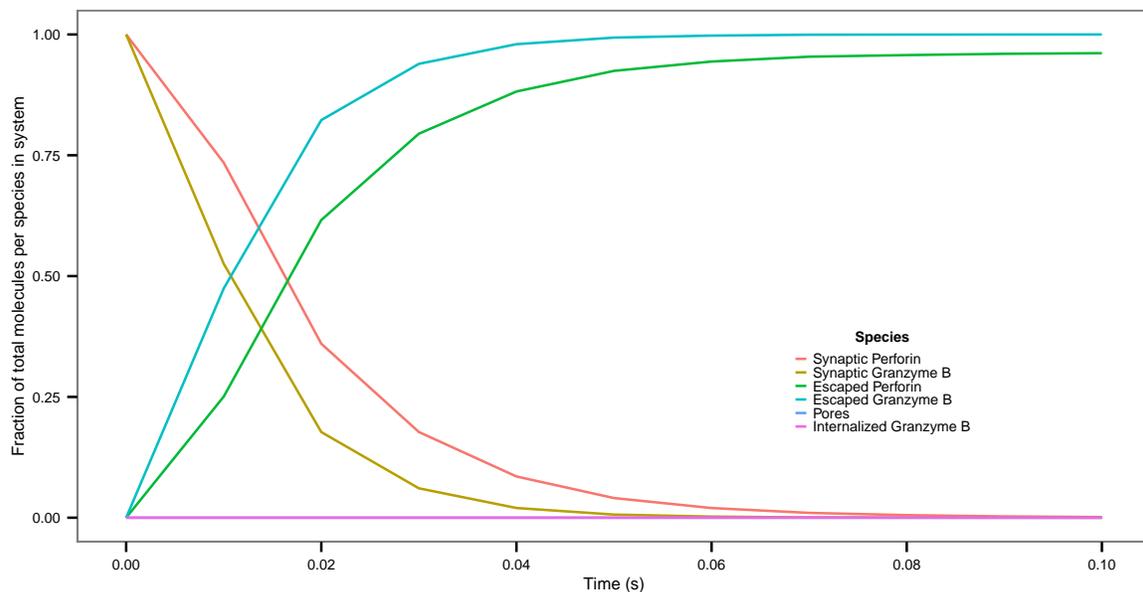


Figure 2.7: Time evolution of GZB and PFN in the immunological synapse without hindered diffusion. Absolute numbers for each species were normalized to the initial total amount of that molecule in the system. These results show that in the absence of hindered diffusion there is rapid and complete loss of GZB, and near complete loss of PFN, without pore formation or GZB internalization. Here, $N_{\text{PFN}} = 1500$.

2.3.2 Pore formation is influenced by the amount of perforin released

When we included a description of hindered diffusion in our model, the rate of GZB and PFN loss was dramatically attenuated, and occasionally pore formation did occur. However, the probability of pore formation was still only 0.04, which we consider incompatible with physiological expectations. Therefore we explored the importance of the amount of perforin released into the synapse. Interestingly, increasing PFN number did increase pore formation probability (Figure 2.8a), with this latter value increasing from close to zero at low PFN numbers, to unity at high PFN numbers. Most notably, in control simulations without hindered diffusion ($\alpha = 1$) at the maximum PFN value (6000N in Figure 2.8a), virtually no pore formation occurs, as compared to consistent pore formation for this maximum PFN value with hindered diffusion. This reinforces the argument that hindered diffusion is critical for the system to function. These results also introduce the recurrent finding that appreciable quantities of GZB are internalized if pore formation occurs, but even in these cases the majority of GZB still escapes the synapse. We return to these points in the discussion below.

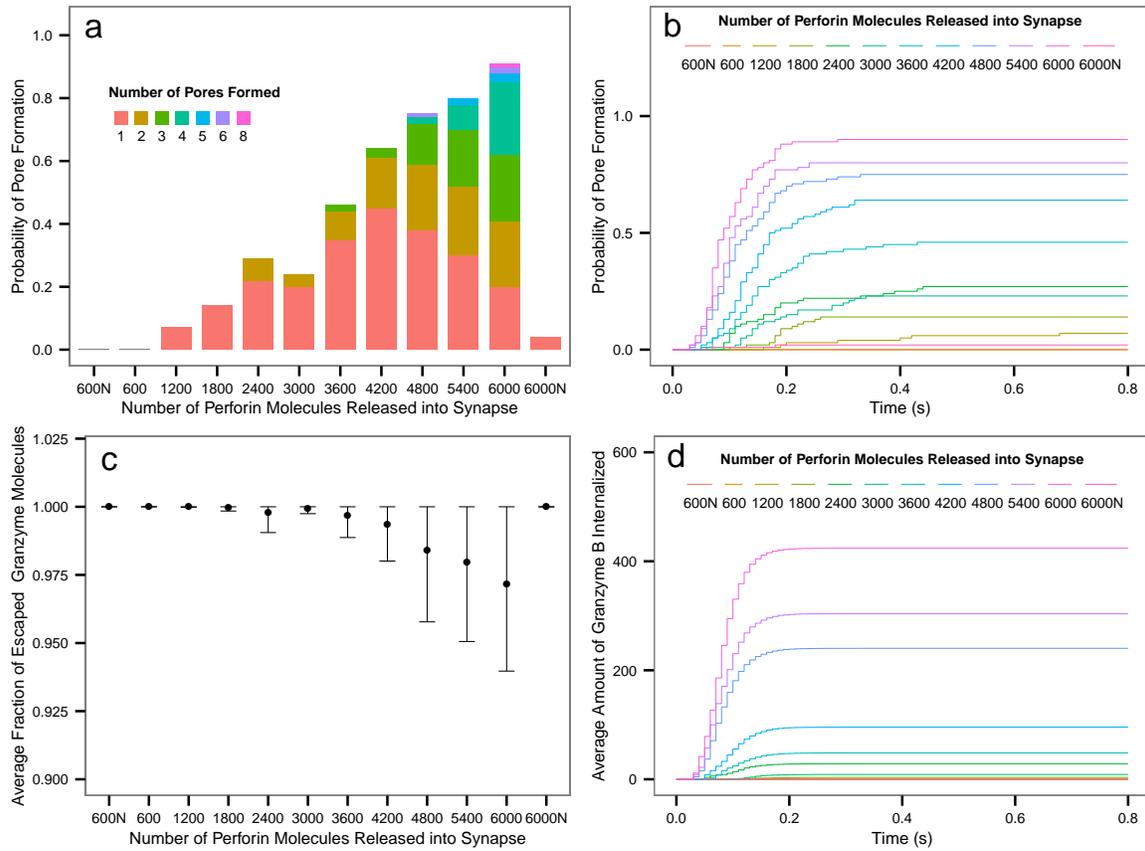


Figure 2.8: Effect of the amount of PFN released on pore formation probability and GZB internalization. (a,b) The probability of pore formation undergoes a transition from minimal pore formation to consistent pore formation over one order of magnitude of PFN number released into the synapse volume, but even at the maximum PFN value we consider, hindered diffusion is required for this effect. (c,d) Even when pore formation is certain, the majority of GZB still escapes the synapse. Each error bar represents standard deviation over 100 simulations. The baseline hindered diffusion is $\alpha = 0.306$, which is obtained from Equation 2.6 using values from Table 2.1 in the methods. The N suffix indicates no hindered diffusion ($\alpha = 1$).

2.3.3 The amount of granzyme internalized depends strongly on the rate of pore formation

Since the amount of PFN initially released had such a strong effect on the probability of pore formation and GZB internalization, we examined the potential correlation between the amount of GZB released and GZB internalized. Due to the uncertainty in the literature concerning the amount of GZB released, we present results over two orders of magnitude of N_{GZB} . As can be seen from Figure 2.9a while the amount of GZB internalized does increase with GZB released, it is a very modest effect. We hypothesized that the effect was so weak because GZB internalization is entirely dependent on pore formation: prior to pore formation, GZB is lost rapidly due to diffusive escape from the IS. Since

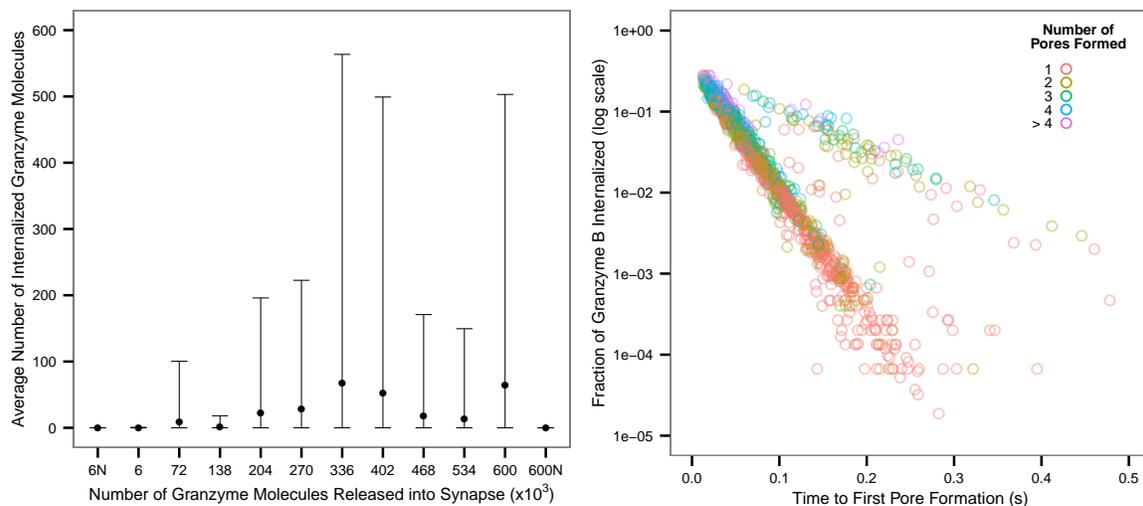


Figure 2.9: Dependence of GZB internalization on rate of pore formation. (a) Despite an increase of two orders of magnitude in the amount of GZB released, only a very modest increase in the average amount of GZB internalized is observed. The N suffix indicates no hindered diffusion ($\alpha = 1$) resulting in no GZB internalization. Each error bar represents standard deviation over 100 simulations. (b) Data from all simulations in this study is pooled, showing a tight dependency of GZB internalization on the rate of PFN pore formation. The three different clusters correspond to different parameter regimes.

the rate of loss is independent of total GZB number, increasing the total amount will not change the fraction of GZB preserved at the time of pore formation, but rather the absolute number. To confirm this, we pooled data from all simulations in our study, across heterogeneous parameter sets, plotting the fraction of total GZB that is internalized against the time to first pore formation (Figure 2.9b). A distinct negative correlation is observed, indicating that total GZB internalization is primarily a function of pore formation. Finally, we note the continued importance of hindered diffusion in maintaining appreciable levels of GZB internalization (Figure 2.9a).

2.3.4 Hindered diffusion critically influences pore formation and granzyme internalization

Even when varying other parameters across several orders of magnitude, hindered diffusion proved critical in all cases for pore formation and GZB internalization. Therefore, we sought to quantitatively investigate its importance by varying α from 0.1 to 1, corresponding to marked hindered diffusion and free diffusion, respectively. We chose to vary α rather than any of its constituent parameters from Equation 2.6 to explore its effect in an unbiased manner. As expected, at high levels of hindered diffusion, the loss of molecules is greatly attenuated, and pore formation occurs with a probability nearing unity, while when molecules diffuse freely, pore formation is rare (Figure 2.10a,b). Similarly, GZB internalization is significant at high levels of hindered diffusion and almost non-existent in the

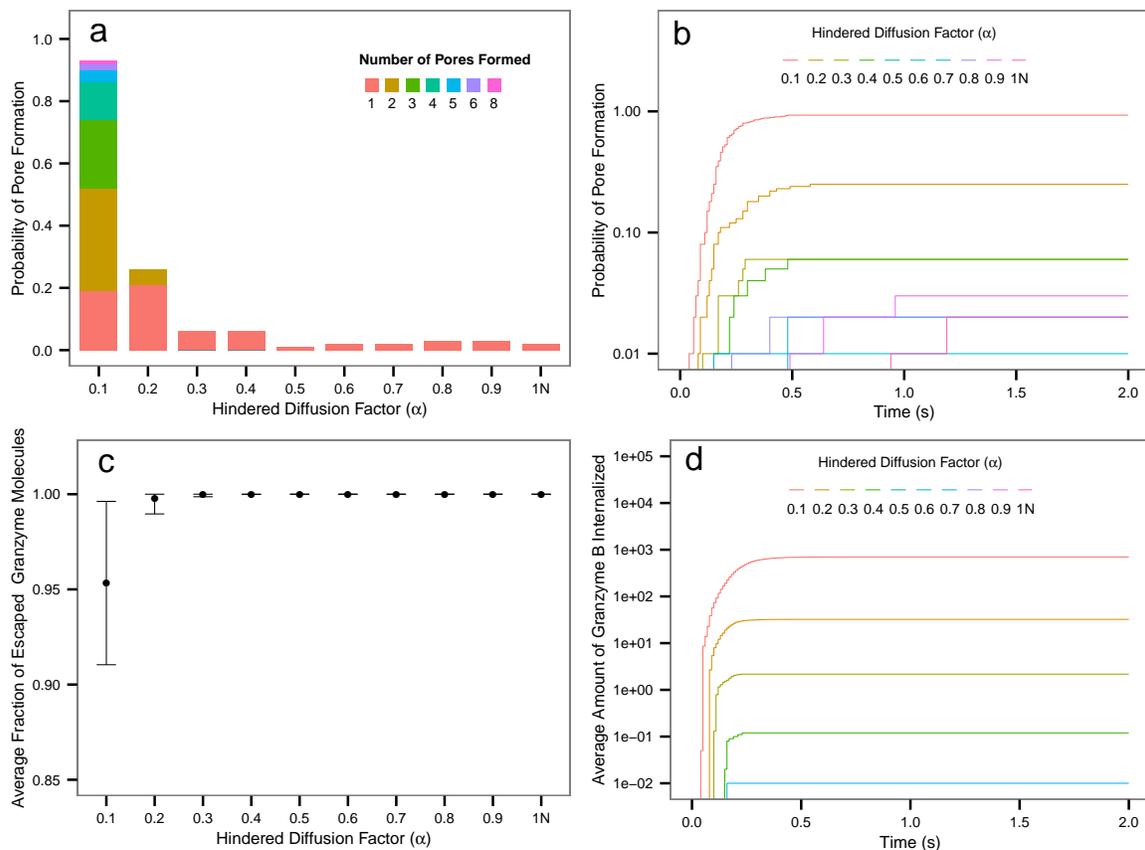


Figure 2.10: Importance of hindered diffusion in pore formation and GZB internalization. (a) Transition from consistent pore formation to minimal pore formation over decreasing levels of hindered diffusion. The importance of α is highlighted by the variability of pore formation probability (two orders of magnitude, (a,b)), and amount of GZB internalization (five orders of magnitude, (d)). While it is evident from (d) that appreciable GZB is internalized at high levels of hindered diffusion, (c) shows that the majority of GZB still escapes the synapse in these cases. Each error bar represents standard deviation over 100 simulations. The N suffix indicates free diffusion.

case of free diffusion (Figure 2.10c,d). Most notably, the variation in pore formation probability and GZB internalization extends over two and five orders of magnitude respectively, demonstrating the importance of hindered diffusion in creating an IS environment that is conducive to GZB internalization and effective target killing. We also note that even for $\alpha = 0.1$, corresponding to high levels of hindered diffusion, 95% of GZB escapes the synapse.

2.3.5 Dependence of pore formation and granzyme delivery on perforin insertion and diffusion in the target cell membrane

To investigate the effect of the diffusion coefficient for PFN oligomers in the target cell membrane, we repeated the analysis of the previous section, but with a tenfold-reduced membrane diffusivity. The results are shown in Figure 2.11. Comparing to Figure 2.10, we see approximately ten-fold lower pore formation and GZB internalization, but the same general pattern of results: pore formation and granzyme internalization depend heavily on a high level of hindered diffusion, and even in that case a large amount of GZB escapes. The main reason for this reduction in pore formation and granzyme internalization is that the lower perforin oligomer diffusivity strongly effects the ability of large j -mers of PFN to find each other and form pores, but free GZB is still escaping from the synapse at the same rate as in the simulations shown in Figure 2.10.

We also studied the effect of the PFN insertion rate into the membrane, k_{ins} . As one would expect, the likelihood of pore formation and GZB delivery increases with this parameter, but importantly, even at the highest value of k_{ins} we considered, successful delivery is strongly dependent on hindered diffusion (Figure 2.12).

2.4 Discussion

We have used a mathematical model to investigate important features of the granzyme-perforin pathway, a two-component system used by cytotoxic cells of the immune system to kill infected and malignant cells. Our key findings are as follows: (i) robust perforin pore formation and GZB delivery to target cells requires rapid pore formation *and* molecular crowding that hinders diffusive transport in the synaptic volume, thus slowing molecular escape; and (ii) even in regimes where we predict consistent formation of perforin pores and appreciable GZB internalization, the vast majority of GZB escapes from the synapse.

Historically, the potency and specificity of cytotoxic lymphocyte killing has been understood in the context of two observational constraints: effective target cell lysis with bystander sparing [171]. This has been explained by a model in which the IS volume is effectively sealed by extremely close proximity between the cell membranes at its edges, which would physically prevent the escape of secreted molecules and thus minimize bystander killing, while also ensuring adequate species concentrations in the synapse for target cell killing. Our first models of the IS, consisting of a simple aqueous environment, with no peripheral seal, clearly demonstrated that the high aspect ratio of the synapse in isolation was insufficient for molecular confinement and therefore CL function. Rather than adding such a seal, for which no mechanism has been even proposed, let alone observed, we considered alternative mechanisms for molecular entrapment. We noted the electron density of the IS on electron micrographs, and, assuming that this was due to a high density of signalling and adhesion molecules, we hypothesized that these molecules effectively hinder diffusion in the synapse, resulting in sufficient confinement of perforin and GZB in the IS for both bystander sparing and target cell killing.

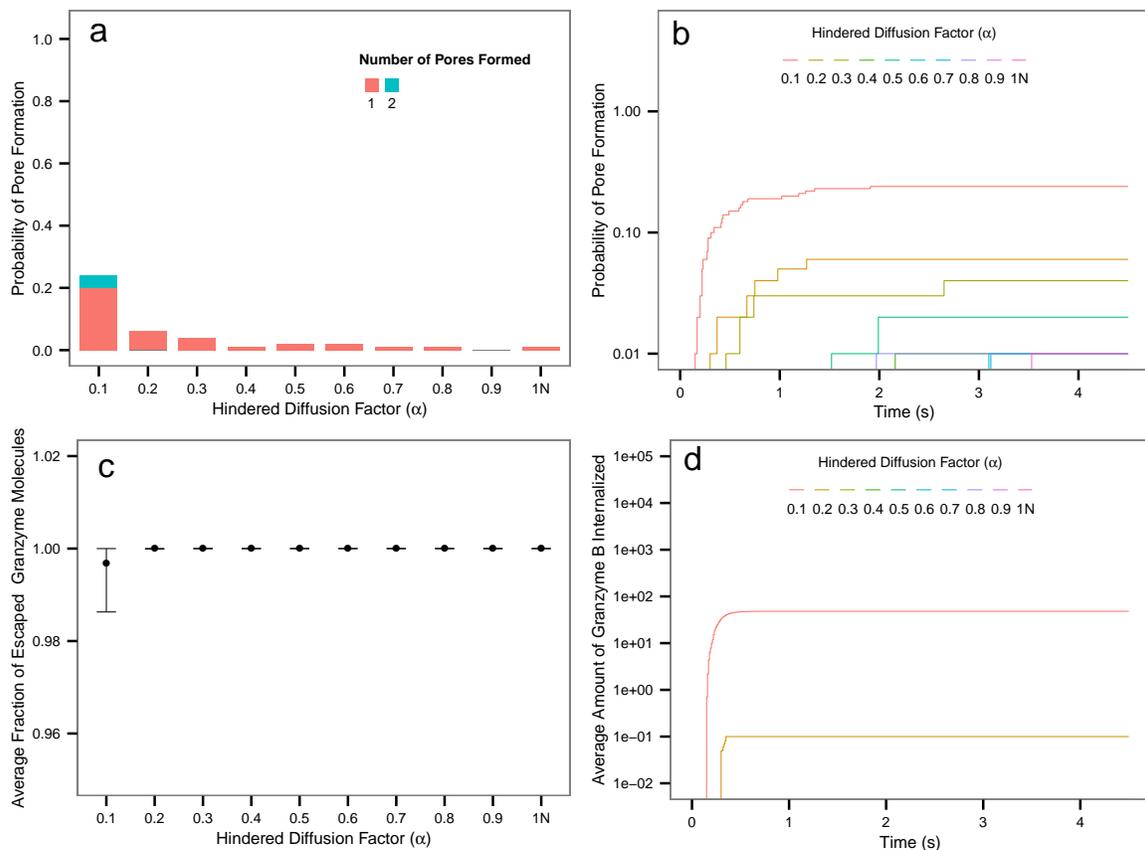


Figure 2.11: Reduced diffusion of perforin in the target cell membrane reduces pore formation and GZB internalization. In this figure, we have reduced the membrane diffusion constant for PFN by a factor of 10 compared to Figure 2.10. The main finding is that reduced membrane diffusion is detrimental to pore formation and GZB delivery. Each error bar represents standard deviation over 100 simulations. The N suffix indicates free diffusion.

When we updated our model to reflect this hypothesis, we found consistent pore formation and GZB internalization under various parameter regimes, indicating hindered diffusion is sufficient for target cell killing. However, even with rapid pore formation, the majority of GZB still escapes the synapse, along with a significant amount of perforin, raising the possibility of bystander killing due to these escaped molecules. We believe that this is highly unlikely to be a significant issue, because of the requirement for very high and localized concentrations of perforin to effectively form pores. Based on our model, such high concentrations are only fleetingly present at the point of release of a lytic granule, and so will likely not occur at any distance from the IS, let alone at even more dilute concentrations in the intracellular milieu. Even if a pore were to form, escaped GZB would be similarly diluted, making internalization of significant quantities of GZB highly unlikely. Additionally, cellular pore repair mechanisms act reasonably quickly [136], further reducing the potential for bystander killing in the absence of a simultaneous high local concentration of GZB. This requirement of spatiotemporal

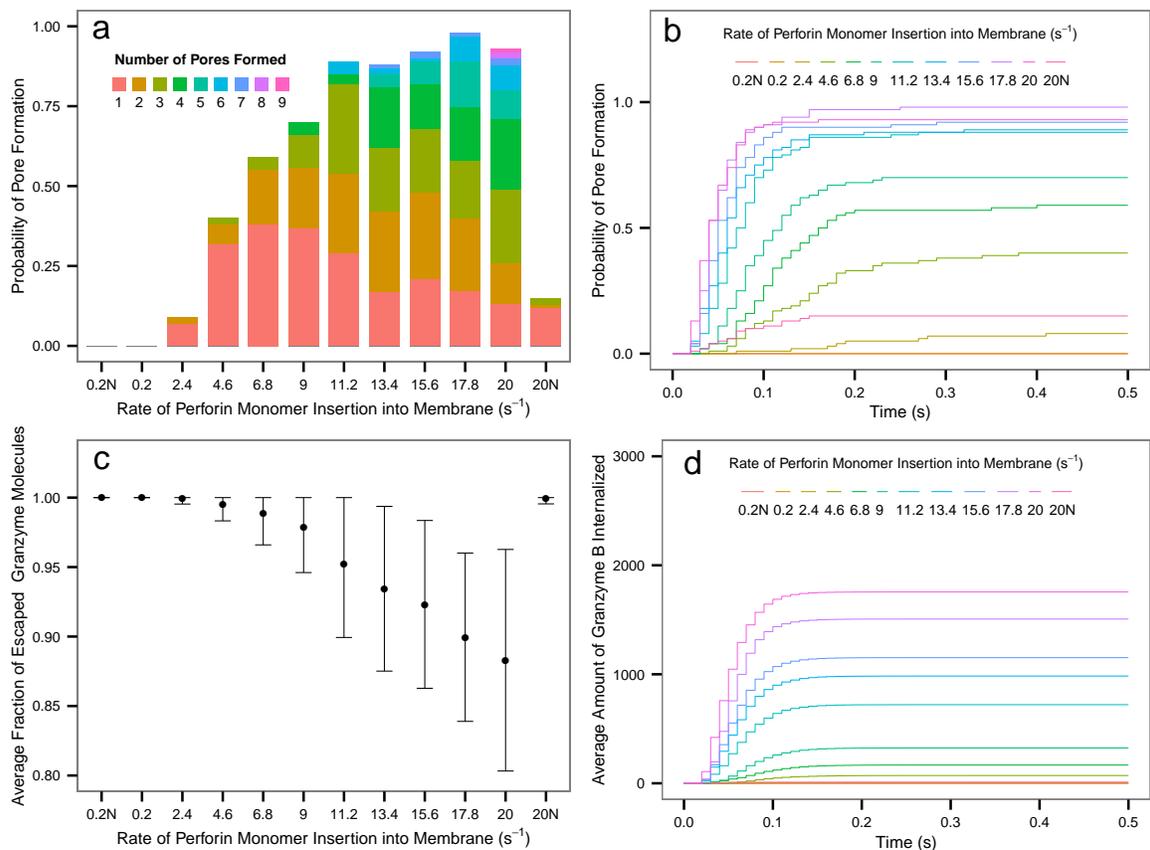


Figure 2.12: Impact of the rate of perforin insertion into the target cell membrane on pore formation and granzyme internalization. Pore formation and granzyme internalization increase dramatically with increasing membrane insertion rate. This is to be expected since a key determinant in these processes is the ratio of perforin escaped to perforin inserted. By increasing the rate of insertion, more perforin is preserved to make pores. Importantly, as with all other parameters, even at the highest rate studied, almost no pore formation or granzyme internalization occurs in the absence of hindered diffusion. Each error bar represents standard deviation over 100 simulations. The N suffix indicates free diffusion.

co-localization of high concentrations of perforin and GZB for pore formation and killing represents a very effective safety mechanism that avoids bystander killing, but nonetheless allows efficient targeted killing. If we consider the requirement of both a GZB and perforin signal as binary in space and time, but the strength of the signal as analog, and therefore multiplicative, we see that this two component system creates a filter which yields a signal that is very strong when co-localization occurs, but is rapidly attenuated when the two species are spatially or temporally disparate. As opposed to the ‘sealed’ IS model which conflates bystander sparing and target cell killing in a single mechanism (the seal), this hypothesis allows for the decoupling of pore formation, resulting from hindered diffusion, and bystander cell survival, resulting from the bimolecular filter.

Our hindered diffusion based model may also help explain the experimental observation that CTLs can kill multiple targets encountered simultaneously by polarizing lytic granules toward multiple targets [172]. Interestingly, a complete IS is not formed for every killed target. Under a sealed IS model, it is difficult to conceive how sufficiently high concentrations of GZB and perforin could be maintained in such a ‘multi-target-leaky’ IS to achieve target cell killing. However, in our model, where high concentrations are sustained by slowed diffusivity due to a crowded synapse, one could imagine that while the synapse is not a tight, organized apposition, there are still significant amounts of adhesion and signalling molecules present. In this case, despite an incomplete appearance, the synapse would still be crowded, with these molecules still slowing the escape of GZB and perforin, and thus enabling continued target cell lysis.

From a biophysical standpoint, this model consists of four interacting processes, each with different timescales: (i) diffusion, whether hindered or free, in the synapse which influences the rate of molecular escape at the lateral edges of the synapse; (ii) the rate of perforin insertion into the target cell membrane; (iii) the rate of diffusion and aggregation of perforin oligomers in the target cell membrane which influences the rate of pore formation; and (iv) the rate that granzyme finds perforin pores. The relative timescales of these processes determine whether sufficient granzyme and perforin are retained in the synapse, for a long enough time, to allow for pore formation and granzyme internalization. In reality, the first three are the rate limiting steps: once pore formation occurs, granzyme internalization is very rapid.

There are myriad effects that could influence these processes such as the volume occupancy of the synapse, specific or non-specific interactions of the two species with each other or other molecules, spatial variations in the width or height of the synapse, active membrane processes at the target cell involving surface molecules or the cytoskeleton, and the possible presence of a heretofore unobserved physical seal at the edges of the synapse, to name but a few. Given the sparse quantitative data regarding these effects, rather than attempt to incorporate them into first-principles descriptions for the three processes listed above, we have systematically investigated the influence these processes have on lymphocyte function. Our quantitative model allowed us to delineate the sensitivity of the granzyme-perforin pathway to these parameters by calculating the effect that varying a parameter has on the probability of pore formation, the key determinant of cytotoxic lymphocyte killing in our model. One way to quantitate this is to use a metric of maximum difference in pore formation probability normalized to the fold change in the parameter value that was varied (ΔP_{max}), with the normalization allowing for comparison between parameters with different units. Using this approach we investigated the degree of hindered diffusion ($\Delta P_{max} = 0.335$, Figure 2.10), the rate of perforin insertion ($\Delta P_{max} = 0.162$, Figure 2.12), and the diffusivity of perforin in the target cell membrane ($\Delta P_{max} = 0.069$, Figure 2.11), and found that the most critical parameter of these three for pore formation and granzyme internalization is the degree of hindered diffusion in the IS. This can be most clearly seen by noting that when hindered diffusion is replaced by free diffusion in our model, pore formation is dramatically attenuated or eliminated, even at the extreme values of the other parameters we investigated.

While our model is appealing in its mechanistic simplicity, there are certainly others that are possible, such as a peripheral ‘seal’, or transient, localized confinement of perforin in the target cell membrane [173–175] enhancing pore formation. We investigated this latter hypothesis (Figure 2.11) and found that our model predicts significantly reduced target cell killing due to slowed PFN pore formation. Importantly none of these models are mutually exclusive, and further computational and experimental work to delineate the relative importance of these mechanisms, as well as to further characterize this important system, would be welcome. To investigate the ‘sealed’ IS model using our computational implementation, the most reasonable approach would be to use a spatially-dependent hindered diffusivity that is highest at the IS boundary as a model for the ‘seal’. Experimentally, there are two testable characteristics which would help distinguish between the two models. The first is the maximum size of molecule that can enter or exit the synapse, which could be tested by adding fluorescently tagged inert polymers of increasing size to the extracellular milieu of a CL-target cell conjugate, and using single-molecule microscopy to determine the maximum size of molecule that enters the synapse. Comparing this maximum size with the geometry of the synapse could provide insight into the nature of a peripheral seal. Second, our model rests on the notion of hindered diffusion, which implies that the diffusivity of a molecule in the synapse should be well below free diffusion. To test this, fluorescently tagged molecules might be observed within the synaptic region, and their diffusivity measured using fluorescence recovery or single particle tracking. A diffusivity similar to that of free diffusion in an aqueous environment would argue against our model.

In closing, we note that there is rapidly increasing excitement surrounding cancer immunotherapies in general [176], and recognition of the central role that cytotoxic lymphocytes play in these modalities. As the granzyme-perforin pathway is crucial to the lytic capability of these cells, we believe that a better understanding of the pathway may enable rational design of improved effector mechanisms for cell based therapies that circumvent apoptosis resistant malignancies. In particular, our model’s prediction that free diffusion of GZB is sufficient for its internalization was a crucial finding that enabled our approach to using GZB as a molecular chaperone for cell-to-cell delivery via the granzyme-perforin pathway.

Chapter 3

Targeted cell-to-cell delivery of exogenous protein payloads via the granzyme-perforin pathway

3.1 Introduction

With their ability to sense and integrate a wide range of signals, actively move to specific tissue compartments, and actuate context-dependent responses, engineered cell-based therapeutics are emerging as the next major class of medical intervention [177]. Chimeric antigen receptor T-cells (CAR-Ts) are highly effective in treating hematological malignancies [79–81], and many mesenchymal stem cell therapies [178] are at various stages of development for use in cardiac [36, 37], neurological [38], and malignant [39–41] disease. These advances are a result of recombining the diverse functionality of biological systems [179, 180] to generate new functional biological molecules and pathways [110, 111, 115, 116, 119]. Current cell-based therapeutics are limited by the small repertoire of available modules and there is an unmet need for additional sensory and effector components for engineered cell therapies.

Cytotoxic lymphocytes have exceptional utility as cellular therapeutics because they are targetable, expandable, and amenable to genetic manipulation [181]. Cytotoxic lymphocytes possess a unique effector mechanism, the granzyme-perforin pathway, one of the main ways in which they kill target cells [121]. The main components of this pathway are the serine protease granzymes, and the pore forming protein perforin, both of which are stored in membrane bound secretory lysosomes, or lytic granules, in the cytosol of cytotoxic lymphocytes [182]. Upon target cell recognition, the cytotoxic lymphocyte forms a tight apposition with the target cell, forming the immunological synapse. Surface receptor signaling results in the endocytic release of granzymes and perforin from the lytic granules into the synapse between the two cells [129]. Perforin then inserts in the target cell membrane and oligomerizes to form transient pores, through which granzyme diffuses into the target cell [136, 137, 139].

Finally, granzyme cleaves caspases and BID to initiate target cell apoptosis. Importantly, surrounding bystander cells typically do not receive appreciable quantities of granzymes [142, 171, 183]. In summary, surface receptor mediated target cell recognition results in specific, cytoplasm-to-cytoplasm intercellular transfer of granzymes to that same target cell.

Granzyme B (GZB) is a well studied effector molecule that transits the granzyme-perforin pathway. Here I engineer GZB-derived chaperones and trace chaperone mediated trafficking of a functional fluorescent protein payload through this pathway from effector to target cells. This constitutes a cell-to-cell transfer module that can be used in cellular therapeutics to deliver ectopic protein payloads to targeted tissues or cells.

3.2 Results

Design of granzyme B derived molecular chaperones

Granzyme B is synthesized as a pre-pro-protein, with an 18 amino acid N-terminal ER signal peptide, followed by an inhibitory dipeptide, followed by the rest of the protein [182]. Upon initiation of translation, the ER signal peptide directs the nascent protein to the ER, where it is co-translationally inserted into the ER. As the protein is synthesized in the ER, an N-glycan is added, which targets the protein to the Golgi network once synthesis is complete. In the Golgi, the N-glycan is further phosphorylated. This phosphosugar moiety on granzyme B then binds to the mannose-6-phosphate receptor, which targets the protein to lytic granules, where it is sequestered until target cell recognition, resulting in granzyme B release into the immune synapse [129]. Importantly, recent work has shown that following release into the immune synapse, the trafficking of GZB to the target cell membrane and entry into the target cell via perforin pores is likely a result of passive diffusion only [136, 137, 142].

I used this information to guide my design of chaperones for granzyme-perforin mediated delivery. Since the steps in this process that are downstream of lytic granule exocytosis appear to be passive, I hypothesized that a chaperone that successfully directed a payload to be loaded into lytic granules would also be sufficient for payload delivery to the target cell. In designing such a chaperone, I adopted two strategies: rational and empirical.

For the rational design, I sought to develop a set of minimal granzyme B domains that would shuttle a protein payload to lytic granules. I took this set to be an ER localization domain, and a lytic granule localization domain. For the former, I used the GZB ER signal peptide (GZBSS). For the latter, I used a 53 amino acid motif surrounding two computationally predicted N-linked glycosylation sites (GZBSM). The final rational design consisted of GZBSS at the N-terminus, followed by the model payload, followed by GZBSM (Figure 3.1). For a model payload, I selected crmCherry (hereafter mCherry or MCH), a derivative of the mCherry red fluorescent protein, that is known to be stable in the acidic environment of lysosomes [184].

The behaviour of chimeric proteins consisting of domains derived from multiple proteins that have

been rearranged is unpredictable. Therefore, I also selected full length granzyme B as an empirical chaperone. My rationale for this choice was that if there were unknown domains within granzyme B other than the region surrounding the N-linked glycosylation sites that were necessary for lytic granule loading, or if the necessary domains are adjacent to the N-linked glycosylation sites only in the tertiary structure of granzyme B, then they would be captured in the full length protein. To keep granzyme B in as native a form as possible, I fused the MCH payload to the C-terminus of GZB, with the two proteins connected by a flexible glycine serine linker (Figure 3.1).

As controls, I also generated two additional constructs: MCH alone, and GZBSS followed by MCH (Figure 3.1).

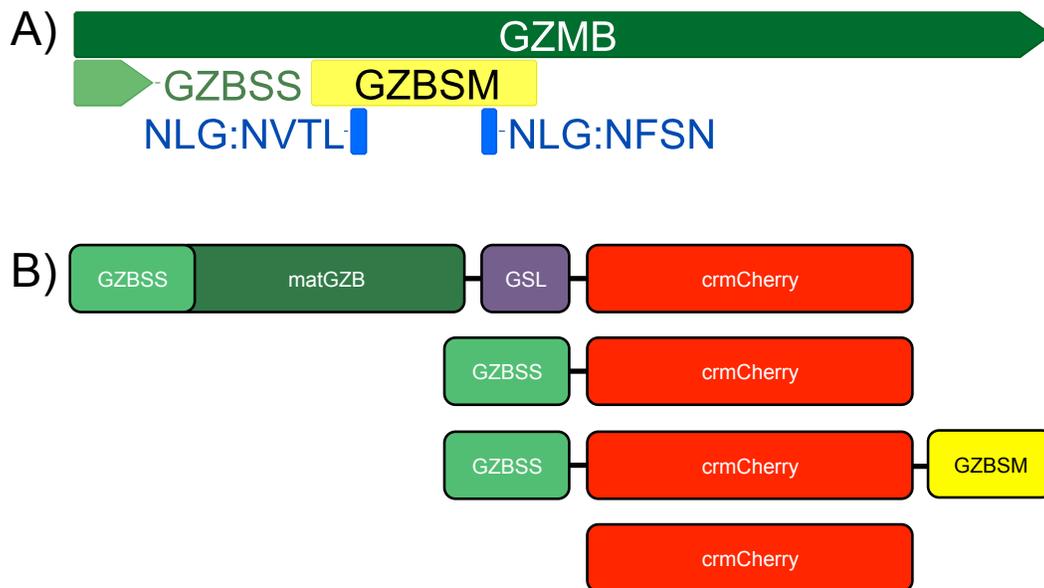


Figure 3.1: Design of payload delivery module chaperones. (a) Granzyme B as a model protein that transits the granzyme-perforin pathway. The full length coding sequence is shown in green, with the ER signal peptide in light green (GZBSS). The two putative N-linked glycosylation motifs are shown in blue, with the encompassing putative sort motif (GZBSM) in yellow. (b) Schematic of the constructs used in this study. mCherry (red), an RFP protein was used as a model payload, and a flexible glycine serine linker (GSL, purple) was used to join granzyme B to mCherry.

3.2.1 Screening chaperone designs by assessing lytic granule colocalization using confocal microscopy

Since my hypothesis was that lytic granule loading of a payload would be sufficient for payload delivery, I first investigated the subcellular localization of the chaperones using confocal microscopy. I expressed the candidates in the natural killer cell line YT-Indy (hereafter YT), which retains a functional granzyme-perforin pathway and has well characterized target cell lines [185]. Following enrichment for mCherry+ cells via cell sorting, I stained the cells for the lysosomal and lytic granule marker Lamp1

and then acquired images of the cells using confocal microscopy. As expected, the Lamp1 distribution was punctate in nature, but the MCH distribution was highly variable (Figure 3.2).

Due to the range of phenotypes observed in the images, I sought to evaluate the degree of payload (MCH) colocalization with Lamp1 in an unbiased manner. To do this I developed a semi-automated image filtration and analysis pipeline. The algorithm is illustrated in Figure 3.3 and seeks to eliminate both local background and bleed, as well as pixel noise. This is achieved using both local and global image information to filter each pixel. This filtering is critical to enable quantitation of colocalization, as it eliminates the background noise from the regions of the image in which there are no cells, as well as regions that are adjacent to granules that have moderate signal intensity, both of which could give a spurious contribution to any quantitative metric of colocalization. The efficacy of this method can be observed by examining the progression of the three columns from top to bottom of Figure 3.3: note that pixel intensities of the punctate structures remain relatively intact, whereas the binarized images (showing the extent of the background signal) change from containing large homogeneous structures to puncta that closely resemble those in the pixel intensity images.

Using this approach, I quantified the colocalization between MCH and Lamp1 in these filtered images using the Manders M1 coefficient and Pearsons Correlation Coefficient (PCC) (Figure 3.4). Both metrics indicated that MCH alone had a low degree of colocalization with lytic granules, which would be expected as the lytic granules are small dense granules, and unfused MCH is distributed throughout the cytosol. GZBSS-MCH had high colocalization with Lamp1, with a perinuclear and membrane distribution, consistent with it entering the secretory pathway. GZB-MCH also had high Lamp1 colocalization, but with punctate cytosolic distribution consistent with granule loading. GZBSS-MCH-GZBSM exhibits a partially punctate granular distribution (similar to GZB-MCH), but also a moderate intensity, diffuse cytosolic distribution (similar to unfused MCH). The mixed phenotype of GZBSS-MCH-GZBSM suggested this chaperone may not be effective at loading payloads into lytic granules. These imaging results suggested that of the two chaperone candidates, GZB had the most potential.

3.2.2 Transfer of fusion proteins from effector to target cells

I next characterized the capacity of the two candidate chaperones (GZBSS-MCH and GZBSS-MCH-GZBSM) to facilitate transfer of the payload through the granzyme perforin pathway to target cells. To do this I conducted a series of co-culture experiments, again using mCherry as a fluorescent reporter payload that was easily traceable. Effectors expressing a variety of mCherry fusion proteins were co-cultured with fluorescently labeled target cells, and then analyzed for evidence of mCherry in the target cell populations. I used the B-cell lymphoblastoid cell line 721.221 (hereafter 721) [186] as target cells, as they are a well known YT target.

I started by testing GZB-MCH, GZBSS-MCH and MCH alone. GZB-MCH was the chaperone the microscopy images had suggested was most likely to load payloads into lytic granules, while MCH alone was clearly not loaded into lytic granules, and thus a good control. GZBSS-MCH was included to confirm that it was not being loaded into lytic granules and hence would not be transferred to target

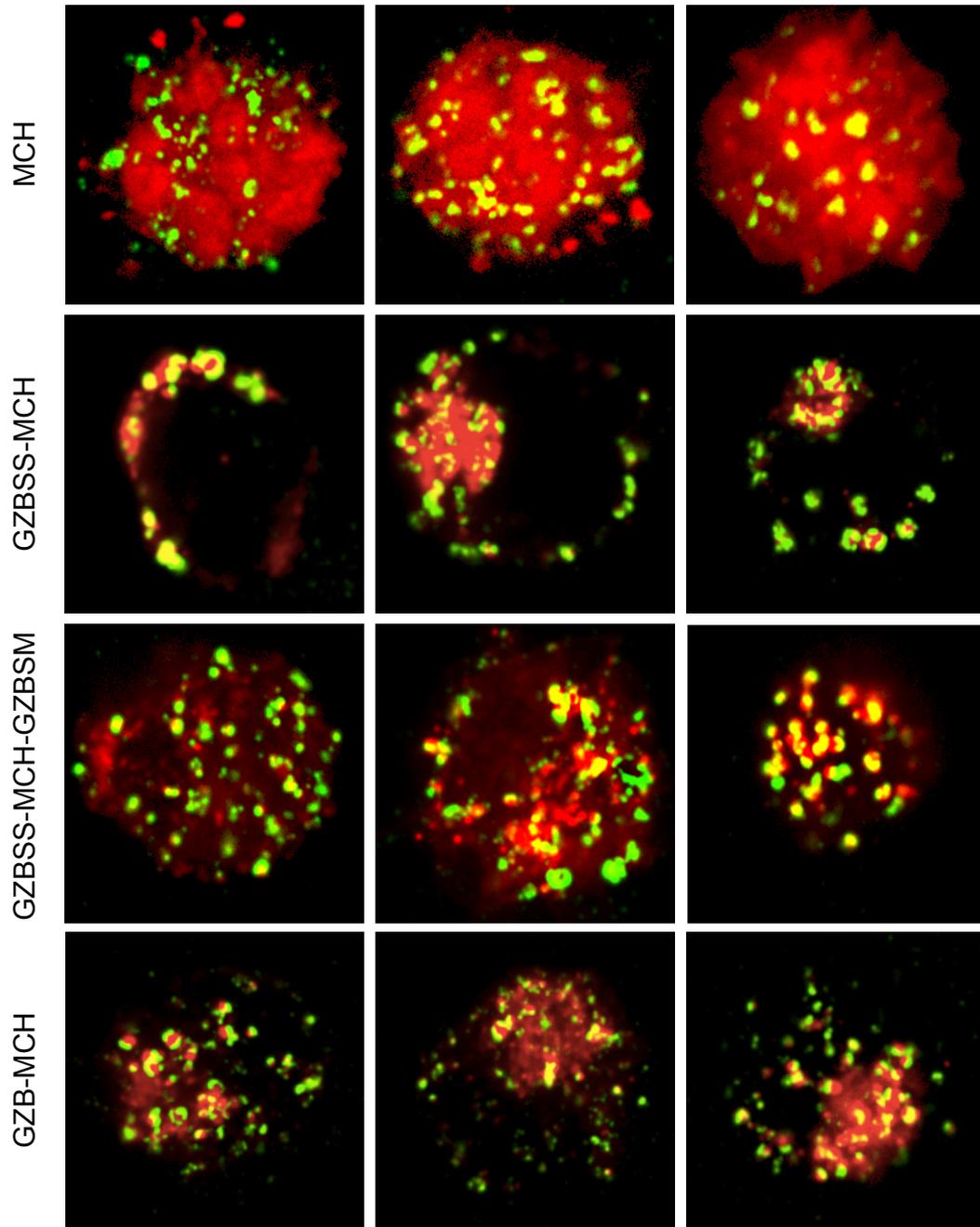


Figure 3.2: Subcellular distribution of candidate chaperone-mCherry fusion proteins. YTs expressing the candidate fusion proteins (labeled at left) were stained for the lytic granule marker Lamp1 and imaged using confocal microscopy. Shown are merged red (mCherry) and green (Lamp1) channels for three representative cells for each sample.

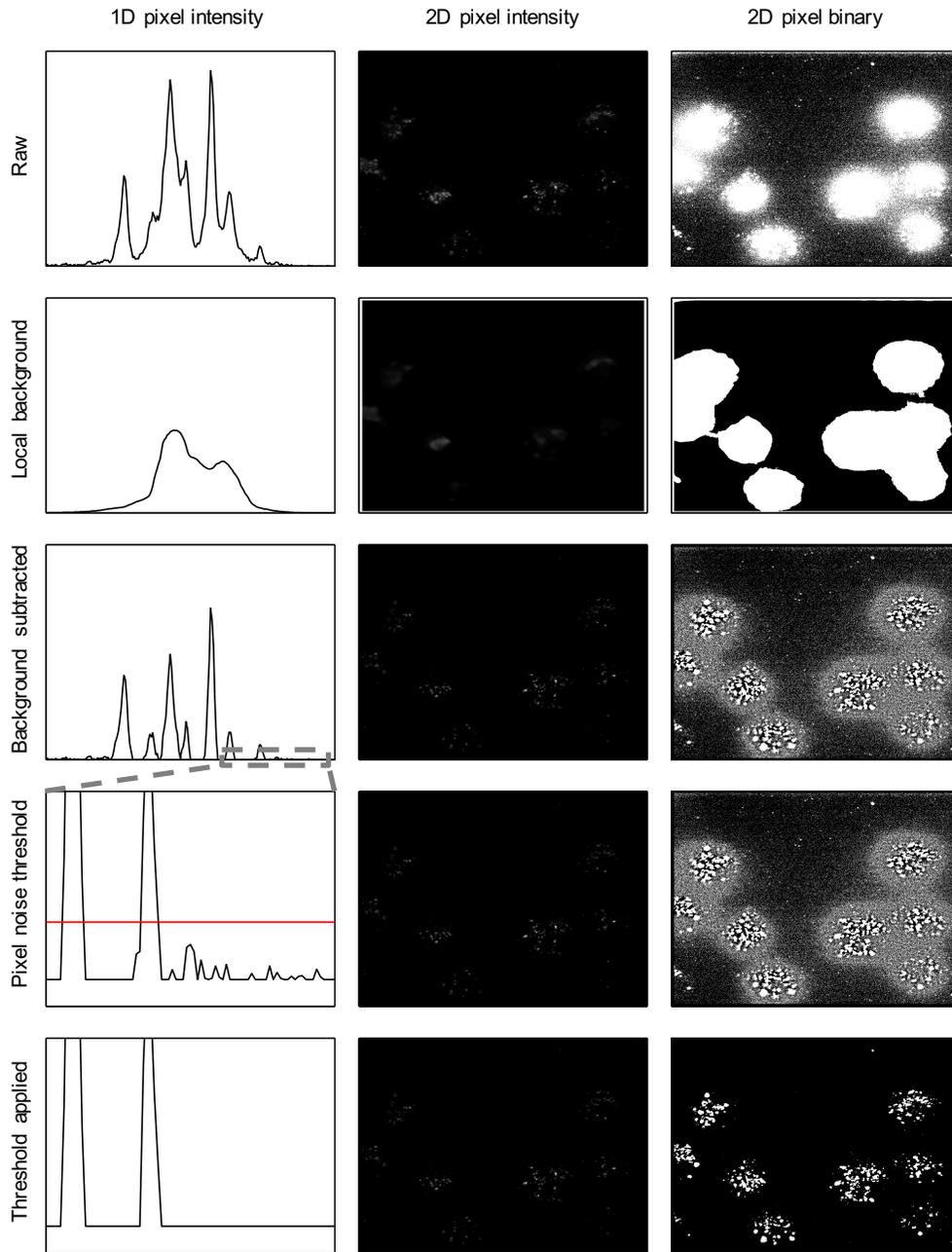


Figure 3.3: Automated image noise filtering pipeline. All panels are derived from a single channel of a single image. The first column consists of pixel intensity traces of a single horizontal line of pixels through the corresponding whole two dimensional image, shown in the second column. The third column shows binarized versions of the middle column: all pixels with intensity greater than 0 are set to 1. The first row is the raw image data. The second row is the local background of the image. The third row is the background subtracted image, literally the second row subtracted from the first. The fourth row is identical to the third, except the first column plot has been enlarged to a small region (gray dashed lines) to better show the pixel noise (small fluctuations near zero). The horizontal red line is the threshold that will be applied to filter pixel noise. The bottom row is the final processed image.

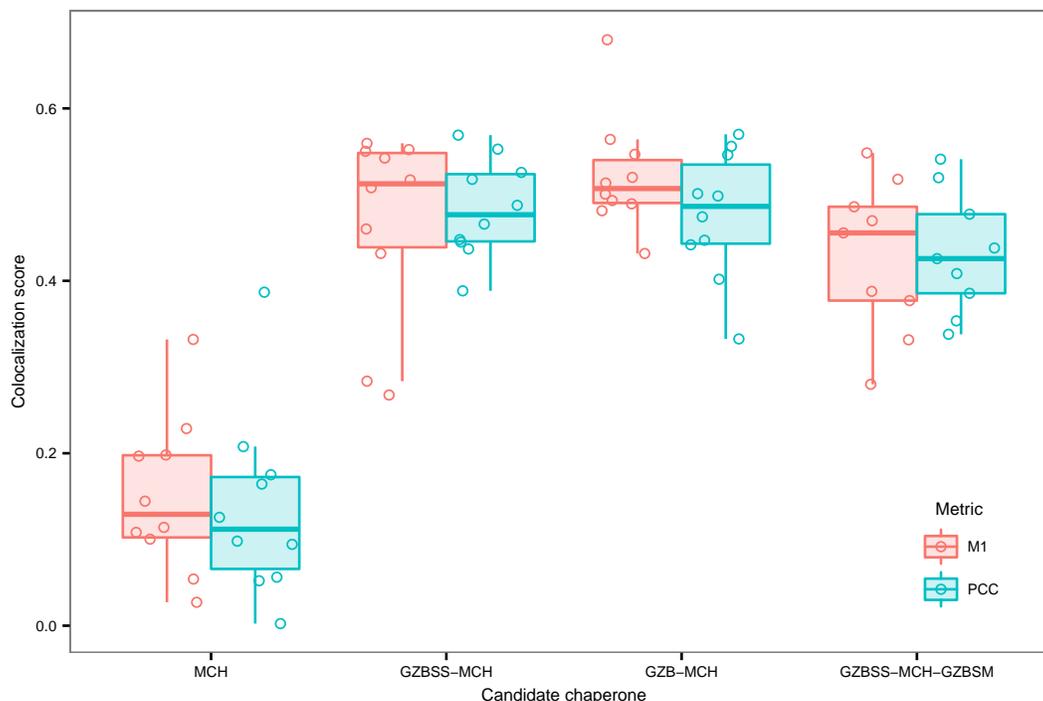


Figure 3.4: Quantitative assessment of candidate chaperone colocalization with lytic granules. Two colocalization metrics were calculated: Manders M1 (quantifying the fraction of pixel intensity of mCherry positive pixels that also contains Lamp1 signal), and Pearson correlation coefficient (PCC, quantifying the degree to which red and green pixel intensities are correlated). These were calculated using both channels from each image. Each circle is the score of a single image, and each image contained between 2 and 5 cells. Overlaid are box and whisker plots.

cells. YT cells were transfected with plasmids coding for these chaperones and then FACS enriched for RFP+ cells. The various effector cell types were separately co-cultured with 721 target cells that had been labeled with a fluorescent dye, to distinguish between the effector and target cells. This mixed cell population was then analyzed via flow cytometry (Figure 3.5). Target 721 cells that were co-cultured with YTs expressing GZB-MCH show an increase in mCherry signal, as compared to 721s alone, 721s co-cultured with unmodified YTs, and 721s co-cultured with YTs expressing either MCH or GZBSS-MCH. Notably, this increase is most prominent in the dead cell fraction (DAPI+) of the target cells. Since the majority of these cells are dying due to YT attack, DAPI positivity can be viewed as a proxy for YT targeting. Therefore, the increase in MCH signal in dead (DAPI+) targets co-cultured with YTs expressing GZB-MCH, but not MCH or GZBSS-MCH, suggests that GZB-MCH is transferred to target cells specifically via chaperone mediated trafficking through the granzyme-perforin pathway.

I then investigated if the rationally designed chaperone (GZBSS-MCH-GZBSM) would perform similarly to GZB-MCH. I conducted the same type of experiment as above, comparing YTs expressing GZBSS-MCH, GZBSS-MCH-GZBSM and GZB-MCH. I selected GZBSS as the comparator so that all constructs would have the same N-terminal ER signal peptide and potential for secretion. This would

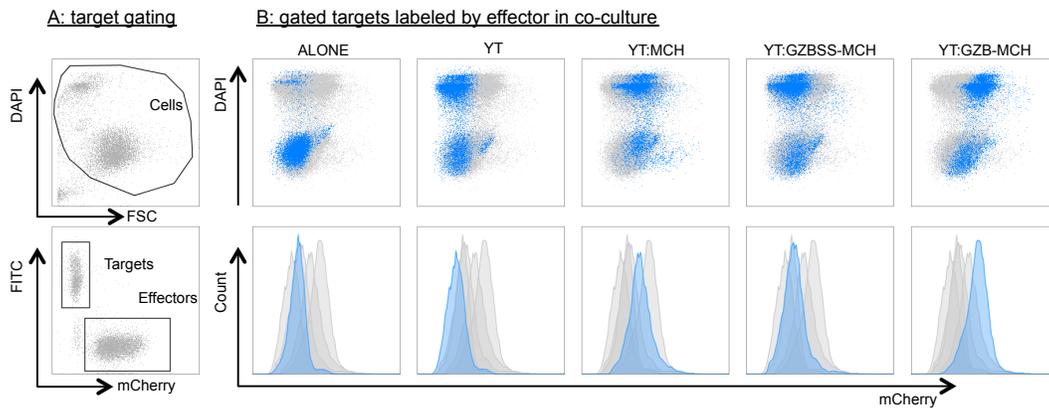


Figure 3.5: Transfer of granzyme B mCherry fusion protein to target cells. YTs expressing various mCherry fusion proteins were co-cultured with CFSE labeled target 721 cells, and the mixed population was analyzed via flow cytometry. **(a)** Gating strategy for isolating target cells. Debris was eliminated (top panel) and then FITC+ targets selected (bottom panel). Not shown is an intermediate hierarchical gating step in which doublets are excluded using forward and side scatter width vs height gates. **(b)** Target 721 cell mCherry fluorescence. Each column is labeled by the effector that was present in the co-culture, but only target cells are plotted, using the gating from (a). Each column contains the same data showing target cell populations from all co-cultures, but only a single target population is highlighted in blue, which corresponds to the effector partner that was present in the co-culture partner.

allow us to differentiate between non-specific mCherry signal in target cells, and mCherry signal in target cells resulting from chaperone mediated transfer. The data (Figure 3.6) is consistent with the initial co-culture experiments, and indicates that GZB transfers MCH to target cells, but GZBSS or GZBSS in combination with GZBSM does not. The fact that GZBSS-MCH-GZBSM did not transfer MCH to target cells is interesting. Given the mixed phenotype observed in YTs expressing this construct in the microscopy images, I thought there was a possibility that it might also traffic to target cells. The fact that it does not provides useful information concerning the nature and location of the granzyme B motifs responsible for its trafficking through the granzyme-perforin pathway, which I discuss below.

Finally, to confirm at the protein level that GZB transfers the MCH payload to target cells, I repeated the above experiments, with the additional, post co-culture, step of collecting live and dead (DAPI- and DAPI+ respectively) 721 target cells from each co-culture via FACS. Data from this sort is shown in Figure 3.7a, and it is consistent with the flow cytometry data from previous experiments. Whole cell lysates from the sorted target cell populations were then size-separated by gel electrophoresis, and probed for mCherry via western blot (Figure 3.7b). A prominent 60 kDa band consistent with GZB-MCH is observed in lysates of 721s co-cultured with YTs expressing GZB-MCH. There is also a background band of approximately 30 kDa consistent with unfused mCherry in the lysate of 721s co-cultured with YTs expressing GZBSS-MCH. This is not unexpected as the ER signal peptide in GZBSS-MCH directs mCherry to the secretory pathway [187], resulting in extracellular mCherry,

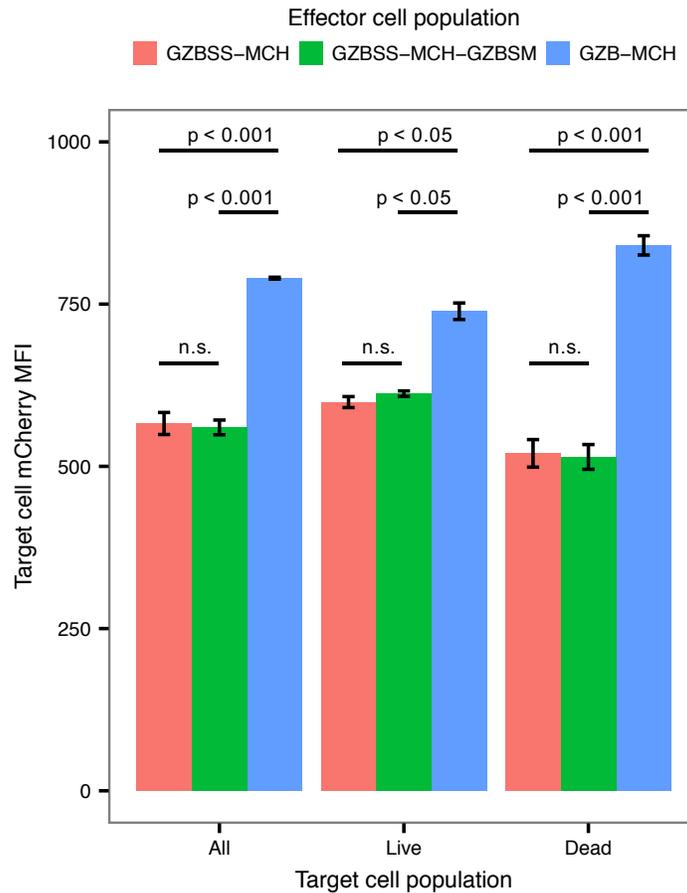


Figure 3.6: Comparison of MCH payload transfer to target cells by the two granzyme B derived chaperones. YTs expressing GZBSS-MCH (red), GZBSS-MCH-GZBSM (green) or GZB-MCH (blue) fusion proteins were co-cultured with CFSE labeled target 721 cells, and the mixed population was analyzed via flow cytometry. The same gating strategy from Figure 3.5 was used and only target cells are plotted. Live and dead cells were selected using a DAPI vs MCH dot plot. Bar plots show mean fluorescent intensity of the RFP channel (MCH MFI) of 721 target cells, with error bars denoting the standard deviation of duplicate samples. p-values were calculated using Tukey's HSD test applied to the results of a single factor ANOVA that was conducted for each target cell population separately.

some of which is likely taken up by the target 721s. However, if non-specific uptake were the main mechanism of MCH transfer from effector to target cell for all samples, then I would not expect to see any difference between 721s co-cultured with YTs expressing GZB-MCH compared to 721s co-cultured with YTs expressing GZBSS-MCH. Instead, of the DAPI+ 721 samples, only the sample from targets that were co-cultured with YTs expressing GZB-MCH has detectable amounts of mCherry protein, and this band is detected at approximately 60 kDa, the expected size of GZB-MCH. Conversely, there is no detectable analogous 30 kDa band corresponding to MCH in the lysates from DAPI+ targets

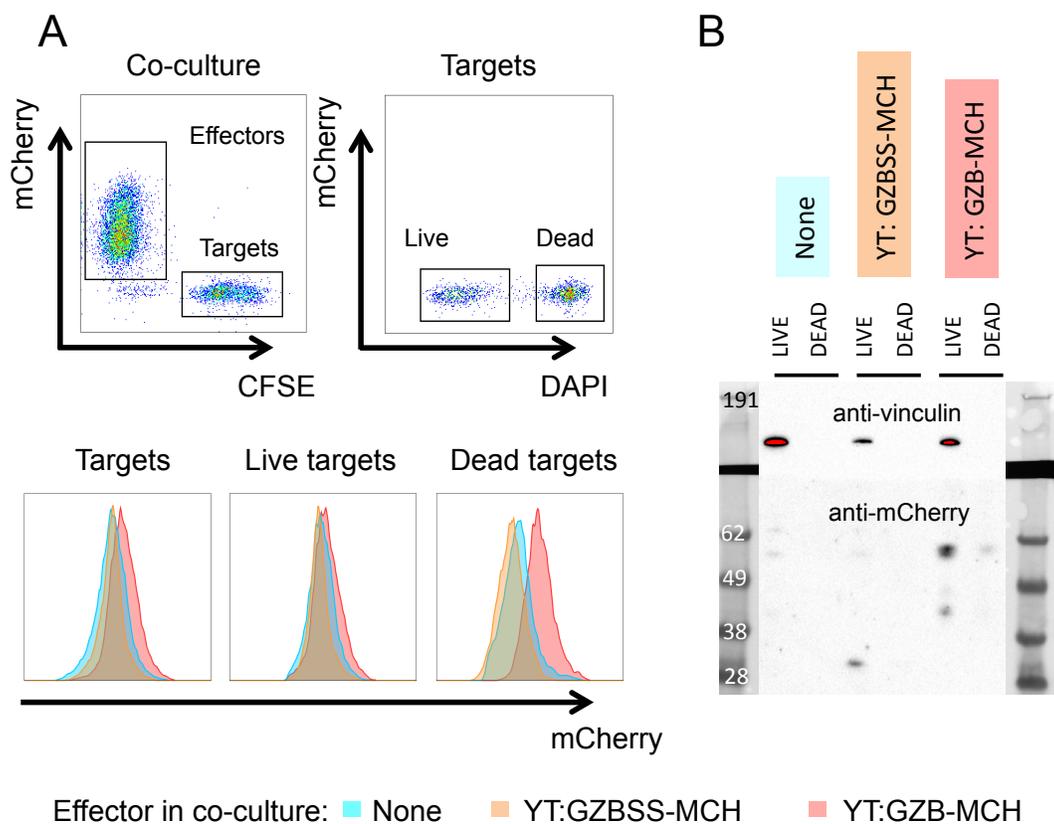


Figure 3.7: Western blot confirmation of GZB-MCH fusion protein transfer to target cells. (a) FACS sort data. YTs expressing either GZB-MCH or GZBSS-MCH were co-cultured with CFSE labeled target 721 cells, stained with DAPI and FACS sorted. Target cells were first selected (upper left panel), and then divided into live and dead (upper right), which were sorted separately and analyzed in (b). The bottom panel shows the mCherry fluorescence of targets (bottom left), live targets (bottom middle) and dead targets (bottom right), for 721s co-cultured alone (blue), with YTs expressing GZBSS-MCH (orange), or co-cultured with YTs expressing GZB-MCH (red). (b) Western blot of sorted target cell populations from (a). Equal cell-equivalent amounts of whole cell lysates of sorted target 721 populations were separated by gel electrophoresis, transferred to blots and probed for mCherry and vinculin (as a loading control). Expected protein sizes: MCH = 30 kDa; GZB-MCH = 60 kDa; vinculin = 130 kDa. Numbers displayed are sizes in kDa of the protein ladder.

co-cultured with YTs expressing GZBSS-MCH. That the putative GZB-MCH band in the DAPI+ 721 sample co-cultured with YTs expressing GZB-MCH is even detectable is noteworthy given the actual amount of protein loaded is quite small, as demonstrated by the lack of a vinculin loading control band. This is despite equal cell numbers for all lanes being sorted and lysed, and is because the DAPI+ dead cells are apoptotic and rapidly degrading which results in a loss of protein.

Taken together these results and analysis suggest that while there is some background, non-specific 721 uptake of MCH from the co-culture media, YTs expressing GZB-MCH specifically transfer the

fusion protein to targeted 721s, while YTs expressing GZBSS-MCH and GZBSS-MCH-GZBSM do not specifically transfer MCH to targeted 721s. Thus GZB appears to be a suitable chaperone protein for delivery of protein payloads via the granzyme-perforin pathway.

3.3 Discussion

Cellular therapeutics that repurpose and recombine biological function in a cellular chassis are transforming medicine [177]. These efforts will rely heavily on the development of modules and systems that perform specific sensory, computational and effector functions [179, 180]. I report here efforts to develop a cell-to-cell delivery module for cellular therapeutics, by repurposing the granzyme-perforin pathway of cytotoxic lymphocytes. The results support the use of granzyme B as a molecular chaperone for inserting protein payloads into this pathway and facilitating payload delivery to target cells.

I hypothesized that lytic granule loading of a payload would be sufficient for transfer to a target cell, and that loading could be achieved by fusing a chaperone to the payload. I designed two candidate chaperones derived from granzyme B, fused them to mCherry and investigated their subcellular localization in the natural killer cell line YT-Indy.

All constructs containing an N-terminal ER signal peptide (GZBSS-MCH, GZB-MCH and GZBSS-MCH-GZBSM) exhibited a high degree of colocalization with Lamp1. This result is best understood by considering the biological distribution of Lamp1 and the cellular compartments in which the colocalization occurs. The primary route of newly synthesized Lamp1 follows the secretory pathway to exosomes at the cell membrane, and is then recaptured in early endosomes and eventually fuses with nascent lysosomes [187, 188]. In the case of GZBSS-MCH, the ER signal peptide would direct the protein to the secretory pathway, so GZBSS-MCH is expected to be found co-localized with Lamp1 in a perinuclear distribution in the ER and Golgi and in punctate granules at the cell membrane, but not in cytoplasmic lytic granules. This is what I observe in cells expressing GZBSS-MCH, in contrast to those expressing GZB-MCH, in which the observed colocalization is primarily in cytoplasmic puncta, consistent with lytic granules. These observations are supported by the co-culture experiments that indicated that GZB-MCH was transferred from effector to target cell, but not GZBSS-MCH. This interpretation is predicated on the assumption that entry into the target cell via perforin pores is passive, which while historically controversial, is supported by most recent experimental and theoretical data [137, 142, 150, 182, 189].

I postulated that combining an ER signal peptide with a putative N-linked glycosylation motif would be sufficient for payload delivery, but the results from the co-culture experiments clearly refuted this. That GZBSS-MCH-GZBSM did not transfer to target cells has several interesting implications surrounding the intracellular trafficking of granzyme B. The first is that the putative N-linked glycosylation sites I computationally identified, and their flanking amino acids, are insufficient for granule loading. While I cannot rigorously exclude the possibility, I do not believe that these results are simply due to faulty glycosylation of GZBSM, since this process occurs cotranslationally and only depends on

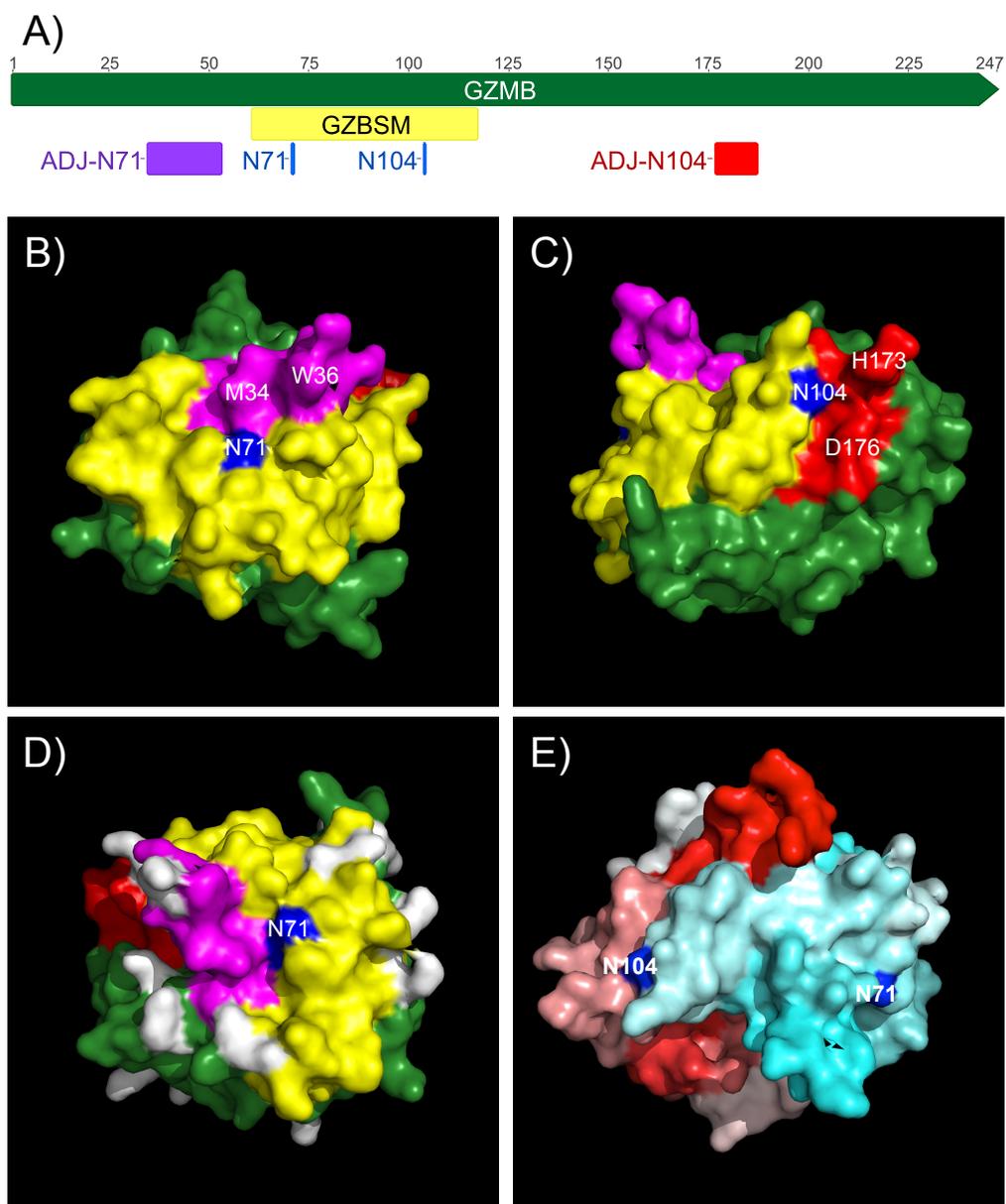


Figure 3.8: Spatial context of putative N-linked glycosylation sites in granzyme B. (a) Schematic of primary amino acid structure of granzyme B. The coloring corresponds to the crystal structures below. Numbers are amino acid residues. (b-c) Three dimensional crystal structure of granzyme B, colored as in (a), highlighting the potential importance of surface exposed residues that are immediately adjacent (ADJ-N71, purple, (b); ADJ-N104, red, (c)) to the N-linked glycosylation sites (blue), but are not contained within the GZBSM (yellow). Note that these regions are quite far from the GZBSM in primary amino acid space, as shown in (a), and in (b,c) by the labeling of representative amino acids in these regions. (d) Lysines have been colored in white, to show their inverted triangular pattern surrounding the N71 putative glycosylation site. (e) Location of putative N-linked glycosylation sites (blue) throughout the protein. The residues have been colored from red through white, to cyan, according to their position in the primary amino acid sequence.

local sequence [190]. This suggests then that GZBSM is not being phosphorylated, likely because the binding domain for the GlcNAc-1-phosphotransferase, which adds a phosphate group to the mannose of the N-linked glycan in the Golgi [187], is not faithfully recapitulated in GZBSS-MCH-GZBSM. This could either be due to a lack of actual amino acids that are present elsewhere in the full length GZB protein, or that the phosphotransferase binding domain is conformation dependent, as has been suggested elsewhere in the literature [191], or both. Much work has been invested into characterizing this domain, but its exact nature remains elusive. These results suggest that whatever the exact composition, its constituent residues are likely distributed throughout the primary amino acid sequence, and hence were not captured in GZBSM, which is why it failed to facilitate transfer of the MCH payload to target cells. This conclusion is consistent with the location of the asparagine residues within the context of GZBSM and the full length granzyme B protein. As shown in Figure 3.8, both of the N-linked glycosylation sites I computationally identified are located immediately adjacent to residues that lie external to GZBSM. In particular, N104 is located at a junction in which residues on one side of N104 are located within GZBSM, while those on the other side are located in the other half of the GZB protein (Figure 3.8a,c). Also of note is that the other N-linked glycosylation site (N71) is surrounded by a triangular pattern of lysines (Figure 3.8d), a pattern which some experimental data suggests is the phosphotransferase binding site [192, 193].

In summary, these results argue that granzyme B trafficking to lytic granules requires residues or domains beyond those immediately flanking the putative N-linked glycosylation motifs. In particular, this data implies that the GlcNAc-1-phosphotransferase binding domain is not a contiguous amino acid sequence, but rather a conformation dependent motif composed of residues located throughout the length granzyme B.

This analysis implies that full length granzyme B is necessary for delivery of a payload to a target cell. Both the flow cytometry and western blot data from the co-culture experiments demonstrated that it is also sufficient: YTs expressing GZB-MCH transfer it to 721 target cells. Importantly, these same data also indicated a background level of accumulation of mCherry signal in target cells co-cultured with YT cells expressing the comparator constructs (GZBSS-MCH and MCH alone). While this might initially appear to undermine the utility of this system, I in fact believe the opposite: it demonstrates the need for specific, cell-to-cell delivery, the activation of which is controlled by surface receptor interactions. In the case of comparator effector populations expressing GZBSS-MCH or MCH, the mCherry signal is the same in both live and dead target cells, indicating a non-specific effect. If these effector cells were used to deliver the payload, the specificity of delivery would be at best localized. However, in the case of the GZB-MCH expressing effector population, there is a significant increase in the RFP signal in dead target cells, indicating that YT-targeted 721s specifically received the most GZB-MCH. Furthermore, my observations consistently have been that MCH is much brighter than GZBSS-MCH, which is in turn brighter than GZB-MCH. If the transfer were non-specific and occurred at roughly equal rates for all mCherry fusion proteins, then I would expect to see 721s co-cultured with YTs expressing MCH alone to display the greatest increase in RFP signal, followed by those

co-cultured with GZBSS-MCH and finally those with GZB-MCH. Instead, I see the opposite: with the greatest increase in RFP signal in cells co-cultured with YTs expressing GZB-MCH, despite this fusion protein having the dimmest fluorescent intensity. Together this data suggests there is a basal level of background accumulation of the mCherry payload in all cases, but substantial, target specific transfer of the payload in the case of YTs expressing GZB-MCH.

The successful transfer of GZB-MCH highlights two unique and highly desirable features of the granzyme-perforin pathway: modularity and prepositioning. The first is important in that all that is required to deliver a payload is to fuse it to the chaperone. In principle, no further modifications are required, regardless of the payload. This modularity suggests that the system might be widely applicable as a means of cellular delivery, either in cytotoxic lymphocytes, or in the long term, in other, orthogonal, highly engineered cellular chassis. The second advantage is that, as opposed to producing a payload in response to target cell recognition using transcriptional control, a presynthesized payload loaded into a lytic granule can be released on the timescale during which the immune synapse remains intact, and hence cell-to-cell specificity is maintained.

A critical consideration in using this system is the stability of future payloads in the harsh environment of the lytic granule, which is acidic and contains many proteases. Some desirable payloads may not be as stable in this environment as mCherry. This might limit the range of applications such a system could be used for, although it is possible that the payload could be engineered to increase its ability to survive the lytic granule, for example by removing a protease cleavage site. The size of the payload is also important, since the internal lumen of the perforin pore has been observed to be 10-20 nm [137], which sets an upper limit on the payload size. However, the diameter of granzyme B is only 5 nm [147], leaving an appreciable window for a variety of payloads.

Any eventual application would also have to consider the native cytotoxic effector mechanisms of the lymphocyte chassis. Unmodified cytotoxic lymphocytes are appropriate vehicles to deliver payloads to target cells with the intent of killing them, as would be the case with tumour cells. However, for other applications—for example delivery of pro-survival factors in degenerative diseases, or deficient enzymes in metabolic diseases—the granzyme-perforin delivery functionality would have to be decoupled from the delivery cell cytotoxicity. This may be possible either through attenuation or knockout of the native effector mechanisms in a cytotoxic lymphocyte, or by reconstituting the pathway in an independent, non-cytolytic cell chassis. The granzyme chaperone itself should be readily catalytically inactivated, as with other serine proteases [194].

I have repurposed the granzyme-perforin pathway as a cell-to-cell delivery module for cellular therapeutics. By facilitating targeted transfer of arbitrary payloads with single-cell precision, this system is an important addition to the part set of synthetic immunology.

3.4 Methods

3.4.1 Computational identification of N-linked glycosylation motifs

The granzyme B coding sequence was downloaded from NCBI RefSeq gene (accession NG_028340.1). I then used NetNGlyc 1.0 [195] to predict putative N-linked glycosylation sites, of which there were two, 33 residues apart. Since the NX(S/T) consensus sequence is necessary but insufficient for glycosylation, and the glycosylation occurs co-translationally [196], it follows that local sequence context surrounding the consensus site is critical. Therefore, I extracted a 53 amino acid domain from GZB, extending from 10 amino acids N-terminal of the first putative glycosylation site, to 10 amino acids C-terminal of the second site. Intriguingly, this domain was also present in human granzyme H.

3.4.2 Plasmids

A custom mammalian expression vector was used in this work. This pdL vector was constructed in house, based on a pcDNA3.1(+) (Thermo Fisher Scientific) backbone. Specifically, the mammalian and bacterial selectable markers and all origins of replication are derived from pcDNA3.1(+), corresponding to bases 1670 (CGATTTTCGGCCTATTGGTTA...) to 5396 (...TAAACAAATAGGGGTTCCGC). A custom expression cassette was cloned into this backbone. This cassette consisted of eukaryotic and prokaryotic promoters and ribosomal binding sequences, followed by the open reading frame, followed by eukaryotic and prokaryotic transcriptional termination sites. For the mammalian promoter I used the CAG promoter for its ability to drive high levels of expression in a variety of tissues. The sequence was amplified from pEMS1157 [197]. This was followed by a hybrid T7 prokaryotic promoter, taken from pCMVTnT (Promega). This was followed by consensus Shine-Dalgarno and Kozak sequences. Following this is the open reading frame, which varies by plasmid. Following the end of the coding sequence, there is a BGH polyA sequence, and then a T7 terminator (with both sequences taken from pcDNA3.1(+)). Restriction enzyme cleavage sites flank all components to facilitate subcloning.

The vector map and full plasmid sequence for the base pdL vector is in the Appendix, along with the full coding sequence for all plasmids used. All plasmids were constructed through a combination of PCR, synthesis and restriction/ligation cloning. All PCR amplicons and coding sequences were sequence verified.

3.4.3 Cell culture

YT-Indy and 721.221 cells were a gift from Judy Lieberman (Harvard University). YT cells were cultured in RPMI 1640 media, supplemented with 20% heat inactivated fetal bovine serum, 1X GlutaMAX, 1mM sodium pyruvate, 10 mM HEPES, 0.1 mM beta-mercaptoethanol. 721 cells were cultured in DMEM, supplemented with 10% heat inactivated fetal bovine serum and 1X GlutaMAX. All cell culture reagents were purchased from Thermo Fisher Scientific.

3.4.4 Transfection

YT cells were electroporated using the Neon system (Thermo Fisher Scientific), using the 100 μ l tip. 6×10^6 cells were washed once in PBS, and resuspended in Buffer R along with 20 μ g plasmid DNA, in a final volume of 110 μ l. The extra volume ensures no bubbles are generated in aspirating the cell mixture into the electroporation tip. Critically the plasmid DNA must be of a concentration of at least 1 μ g/ μ l, and it must be prepared using an endotoxin free method. The quality of the plasmid prep greatly influences the electroporation efficiency as well as the post-electroporation viability. The apparatus was prepared as in the manufacturer's manual, using the E2 electrolytic buffer. The electroporation conditions were 3×10 ms pulses at 1250 V. The electroporated cells were then immediately added to 5 ml media spread across two wells of a 6 well plate.

3.4.5 Flow cytometry

Cells were harvested and resuspended in PBS supplemented with 10% complete media and 1 μ g/ml of DAPI (Sigma) as a viability stain. If cells were to be sorted, they were passed through a 35 μ m nylon filter (BD Falcon). Cells were kept on ice and then analyzed on a BD Fortessa II, or sorted on either a BD Aria III or Fusion. For sorting, cells were sorted into complete media. In all flow cytometry experiments two initial gating steps were used. Debris was excluded by excluding cells at the bottom left corner of a PI vs FSC-A (forward scatter area) gate. Doublets were excluded using a hierarchical gating scheme: all cells with a wider pulse width signal were excluded first in FSC-W vs FSC-H (forward scatter width vs height) and then SSC-W vs SSC-H (side scatter width vs height). All flow cytometry data was analyzed in FlowJo.

3.4.6 Microscopy

Transfected cells were first FACS sorted for moderate intensity RFP+ cells. 2.5×10^5 cells resuspended in 50 μ l complete media (RPMI-1640 supplemented with 10% fetal calf serum, 2 mM glutamine, 1 mM pyruvate, 2 μ M 2-mercaptoethanol, 50 U/ml penicillin and 50 μ g/ml streptomycin) were adhered to 0.01% poly-L lysine (Sigma cat # P4707) coated, pre-cleaned 12 mm coverslips (#1.0, Fisherbrand cat # 12-0545-80) for 15 minutes at 37 $^{\circ}$ C. Cells were fixed with 2% paraformaldehyde (Electron Microscopy cat # 15710) for 15 minutes, washed with PBS, and then permeabilized with 0.1 % Triton X-100 (Sigma cat # T8787) for 1 minute. Samples were washed with PBS, and then blocked in 10% goat serum in PBS (blocking buffer) (Jackson Immunoresearch Labs cat # 005-000-121) for 1 hour. Subsequently, cells were stained with polyclonal mouse anti-Lamp1 primary antibody (Abcam cat # 24170) at 1:250 dilution in blocking buffer for 1 hour. Samples were washed with PBS, and then stained with AlexaFluor 488 conjugated goat anti-mouse secondary antibody (Thermo Fisher Scientific cat # A-11008) at 1:1000 dilution in PBS for 45 minutes. After washing with PBS, coverslips were then mounted on glass microscope slides using Prolong Diamond (Thermo Fisher Scientific cat # P36961) overnight. All steps were completed at room temperature unless otherwise noted.

The following day samples were imaged using a spinning disk confocal system (3i Intelligent Imaging Innovations) based on an inverted Zeiss Axiovert 200M microscope equipped with 100 NA 1.45 Oil Plan Fluor objective and a QuantEM 512SC Photometrics camera. 10 images were acquired for each sample, with each image containing 2-5 cells in the field of view. All exposure parameters were kept constant across all samples.

3.4.7 Image analysis

Image filtering was done using a custom script written in MATLAB. The green channel (Lamp1) was filtered as follows. The localized background of the image was calculated for each pixel as the median intensity of a 25x25 pixel square centered on that pixel. This background pixel intensity was subtracted from the original pixel intensity. Any pixels with negative intensity after this step were set to zero. This step aids in distinguishing small punctate structures from one another. Next, a pixel noise threshold was calculated as follows. First the median absolute deviation (MAD) of all nonzero pixels from the raw image was calculated. From this the standard deviation of the pixel intensity was approximated as 1.4 times the MAD, which is a reasonable estimate of the pixel noise. Finally the noise threshold was taken as 6 times this value (that is $6 \times 1.4 \times MAD$). Any pixels in the background subtracted image whose intensity were below this value were set to zero. The red channel (mCherry) was filtered in the same way, except localized background was not subtracted. The MATLAB script implementing this algorithm is in the Appendix.

Colocalization analysis was also conducted in MATLAB. For paired red and green channel images, with pixel intensities R_{ij} and G_{ij} respectively, Pearson's correlation coefficient was calculated as

$$PCC = \frac{\sum_i \sum_j (R_{ij} - \bar{R})(G_{ij} - \bar{G})}{\sqrt{\sum_i \sum_j (R_{ij} - \bar{R})^2 \sum_i \sum_j (G_{ij} - \bar{G})^2}} \quad (3.1)$$

where \bar{G} and \bar{R} are the mean pixel intensities. The Manders M1 coefficient was calculated as

$$M_1 = \frac{\sum_{i,j} c_{ij} R_{ij}}{\sum_{i,j} R_{ij}}, \quad c_{ij} = \begin{cases} 1, & G_{ij} > 0 \\ 0, & G_{ij} = 0 \end{cases} \quad (3.2)$$

These colocalization scores were calculated separately for each sample of each image, and then plotted using RStudio.

3.4.8 Cell labeling

Cells were fluorescently labeled with CFSE (eBioscience) following the manufacturer's protocol, except that only 1 PBS wash prior to labeling was done and only 1 media wash after labeling was done.

3.4.9 Co-culture experiments

YTs were transfected with mCherry fusion proteins and 48 hours later FACS sorted for viable RFP+ cells. The following day 4×10^5 YT effector cells were combined with 1×10^5 CFSE labeled target 721 cells at a 4:1 effector:target (E:T) ratio in a final volume of 500 μ l YT media in 5 ml polystyrene round-bottom tubes (BD Falcon). The cell suspension was gently pelleted by spinning it at $200 \times g$ for 15 seconds. The tubes were then incubated at 37 °C for 90 minutes, and then prepared for flow cytometry or FACS sorting as above.

3.4.10 Statistical analysis

mCherry median fluorescent intensity was tabulated for each target cell population using flow cytometry data from above. For each target cell population, a single factor analysis of variance was conducted to determine if the MCH MFI means were the same for all effector cell populations using the model $MFI \sim \text{EffectorPopulation}$. I then used these results as input for a Tukey's HSD test of the difference between sample means within each target cell population. Statistical tests were conducted in R, using the `aov` and `TukeyHSD` commands respectively.

3.4.11 Western blotting

3×10^4 cells were sorted into PBS in microcentrifuge tubes. Cells were kept on ice thereafter. Cells were then pelleted, resuspended in 10 μ L PBS and lysed directly by adding 10 μ L 2X Laemmli sample buffer. Samples were incubated at 95 °C for 10 minutes and then stored at -20 °C.

For blotting, samples were boiled again at 95 °C for 10 minutes and then loaded onto pre-cast 4-12% Bis-Tris polyacrylamide gels (Thermo Fisher Scientific). Proteins were size separated by gel electrophoresis by running the gel at 150 V for 75 minutes. Proteins were transferred to a nitrocellulose membrane using a standard wet transfer, at 300 mA for 2 hours.

The blot was cut horizontally at 100 kDa, and then was blocked in TBS-T with 5 % skim milk powder at room temperature for 1 hour, and then incubated with primary antibody in sealed pouches at 4 °C overnight. The primary antibodies used were rabbit anti-mCherry (Biovision cat # 5993-100, lot 1A085993) and rabbit anti-vinculin (Abcam cat # EPR8185, lot GR82271-16), as a loading control. The dilutions were 1:500 (mCherry) with 5% skim milk powder, 1:10000 (vinculin) with 2% skim milk powder, both in TBS-T. Blots were then washed with TBS-T, and incubated with horseradish-peroxidase conjugated goat anti-rabbit secondary antibody (Santa Cruz Biotechnology cat # sc-2004, lot H1015) for 1 hour. The dilution was 1:5000 in TBS-T, with 5% skim milk powder (anti-mCherry), and 2% skim milk powder (anti-Vinculin). Finally, the blots were washed with TBS-T, and then developed using Bio-Rad Clarity Western ECL (enhanced chemiluminescence) Substrate reagent (Bio-Rad cat # 170-5061), following the manufacturers protocol. Blots were imaged using a Bio-Rad Chemidoc MP Imaging System, with exposure times ranging from 1 to 100 seconds.

3.4.12 Crystal structure analysis

To visualize the location of the various motifs of GZB in the three dimensional protein, I downloaded the granzyme B crystal structure from the Protein Data Bank (accession 1FQ3) and rendered the base crystal structure and custom annotations using PyMOL (Schroedinger, LLC). Surface exposed residues adjacent to the N-linked glycosylation sites were determined by first selecting all residues that were within 15 Å of the glycosylated residue. I then selected the subset of these residues that were surface exposed, using a custom PyMOL script written by Jason Vertrees.

Chapter 4

Efforts to use granzyme-perforin mediated delivery of orthogonal toxins to enhance cytotoxic lymphocyte killing of apoptosis resistant tumour cells

4.1 Introduction

It is well known that tumours frequently elude the immune response [198]. This is partially due to tumor cell evasion of cytotoxic lymphocyte recognition, via mechanisms that include downregulation of antigen processing and surface MHC expression (for CTL targeting) [199] and upregulation of KIR receptors (for NK attack) [200]. The advent of chimeric antigen receptor therapy has begun to address this aspect of the problem [201–203]. However, independent of evading recognition by lymphocytes, tumour cells have often been shown to be resistant to lymphocyte cytotoxic effector mechanisms [64, 198, 204–206]. Disruption of apoptosis pathways mediated by death receptors (such as the Fas system) has been found across a range of cancers, and has been implicated in carcinogenesis as well as apoptosis resistance [207, 208]. Downregulation of the executioner caspases 3 and 7 is widely observed and correlates with poor survival [209–218]. Inhibitor of apoptosis proteins (IAPs) are consistently overexpressed in tumors, have been shown to initiate hematological malignancies *in vivo*, are responsible for metastatic potential, have been found to cause resistance to adoptively transferred lymphocytes, and are being actively pursued as small molecule targets, with these efforts having progressed to clinical trials [219–227]. Direct inhibition of granzyme by overexpressed serpins is well characterized [228, 229]. Most importantly, overexpression of XIAP [230], survivin [231], and serpinb9 [232] confer apoptosis resistance to tumor cells, disrupt key nodes in apoptotic pathways and are directly and specifically responsible for the resistance of these cells to lymphocyte mediated cytotoxicity, despite

effective targeting, both *in vitro* and *in vivo*. Thus, apoptosis resistance and specifically resistance to lymphocyte-induced apoptosis is a real, unsolved challenge in the field of cancer therapy.

I have developed the granzyme-perforin pathway as a delivery module for cellular therapeutics (Chapter 3). Briefly, by fusing protein payloads to granzyme B and expressing these fusion proteins in cytotoxic lymphocytes, granzyme B acts as a molecular chaperone that inserts the payload into the granzyme-perforin pathway, resulting in the payload fusion proteins being loaded into lytic granules in the cytosol. Upon target cell recognition, the cytotoxic lymphocytes release these fusion proteins, along with other granzymes and perforin, into the immunological synapse between the lymphocyte and target cell. Perforin forms a transient pore in the target cell membrane, through which the granzyme-payload fusion proteins diffuse. Here, I sought use this approach to deliver potent toxins to lymphocyte resistant tumour cells. My hypothesis was that by expressing granzyme B-toxin fusion proteins in cytotoxic lymphocytes, these toxin fusion proteins would be transferred to targeted cells, resulting in enhanced killing of lymphocyte resistant target cells. This approach would have several advantages: (i) TCR- or CAR-mediated specific delivery of potent toxins to tumour cells, minimizing off-target toxicity; (ii) sequestration of toxins inside the delivery lymphocyte would enable selection of toxins that could not be administered systemically, either due to toxicity or poor bioavailability.

There is one significant caveat to selecting the problem of killing apoptosis and lymphocyte resistant tumour cells as an application in which to establish the utility of the granzyme-perforin mediated delivery system. Unlike the delivery of pro-survival payloads with the intent of rescuing target cells, in the case of tumour cells, the intent of any payload delivery to a tumour cell is target cell death. This simplifies the experimental implementation of the latter application, since unmodified cytotoxic lymphocytes can be used as a cellular chassis, which would be inappropriate for pro-survival payload delivery as the lymphocytes would naturally deliver their own endogenous cytotoxic payload, killing the target cell. In the case of target tumour cells, the purpose of the payload is to augment native lymphocyte cytotoxicity. Unfortunately however, this also implies that for any enhancement of target cell killing to be measurable, it is critical to generate a model system in which the target cells induce lymphocyte reactivity and degranulation, but are completely resistant to the effects of that degranulation, which is a significant challenge in an *in vitro* system.

With this potential concern in mind, I searched for toxins that might be capable of both killing apoptosis and lymphocyte resistant tumour cells, as well as amenable to cell-based granzyme-perforin delivery. These two requirements implied several criteria which guided my selection of toxins. First and foremost, the toxin ought to have an orthogonal mechanism of action, that is one that is at least partially independent of cytotoxic lymphocyte killing mechanisms, to maximize the toxin's additive effect, and minimize the chance that a lymphocyte resistant tumour cell might also be resistant to the toxin. Additional criteria for compatibility with granzyme-perforin mediated delivery dictated that the toxin be: (i) genetically encodable (so that it may be fused to granzyme); (ii) relatively small (so that it may translocate through perforin pores as a granzyme fusion); (iii) act in the cytosol of targeted cells. Based on these criteria I selected a suite of candidate toxins for study: the diphtheria toxin A fragment

(DTA), pseudomonas exotoxin A (PEA), herpes simplex virus thymidine kinase (HTK), and the *E. coli* nitroreductases *nfsA* and *nfsB* (NFSA, NFSB).

Diphtheria and pseudomonas toxins are bacterial exotoxins secreted by *Corynebacterium diphtheriae* and *Pseudomonas aeruginosa* respectively. These toxins are composed of three major domains, each responsible for either membrane binding, membrane translocation, or ribosomal inhibition. The latter is the main mechanism of toxicity of these two proteins, and is a result of inhibition of elongation factor 2 by ADP-ribosylation, thus preventing polypeptide elongation during protein synthesis [233]. Both DTA and PEA have been used in recombinant immunotoxin fusion proteins in clinical trials, wherein the membrane binding domain is replaced by a single chain antibody or cytokine to target the toxin to a specific cell population [234].

Herpes simplex thymidine kinase and the *E. coli* nitroreductases both exert their toxic effect by activating otherwise inert prodrugs [235]. HTK phosphorylates the prodrug ganciclovir, which once activated acts as a nucleoside analogue and thereby terminates DNA synthesis [236]. *nfsA* and *nfsB* metabolize the prodrug CB1954 to potent DNA alkylating agents, resulting in DNA crosslinking and interruption of DNA synthesis [237, 238]. These toxin/prodrug systems have two important features. First, the toxins themselves are not toxic to the delivery cell, obviating the need for additional chassis modifications to protect the host delivery lymphocyte from the toxin it carries. And second, prodrugs activated in one target cell may kill adjacent target cells, a process known as the bystander effect. This is advantageous in that the toxicity of any single toxin armed lymphocyte can be greatly amplified, and by extension, near total killing of a target cell population can be achieved even in absence of near total tumour cell targeting by delivery lymphocytes. This might be particularly relevant in the context of an immunosuppressive tumour microenvironment, in which the lifetime of active cytotoxic lymphocytes (prior the microenvironment inducing lymphocyte anergy, conversion to a regulatory phenotype, or apoptosis) may be quite short. In this case successful lymphocyte targeting of only a fraction of tumour cells prior to succumbing to the tumour microenvironment might still result in a substantial anti-tumour effect.

Here I report my efforts to enhance lymphocyte cytotoxicity by delivery of these toxins via the granzyme-perforin pathway. I generated granzyme-toxin fusion proteins, attempted to create lymphocyte resistant target cells, and attempted to demonstrate enhanced target cell killing mediated by granzyme-toxin delivery. My main findings are that granzyme-toxin fusion proteins have a variable, toxin dependent, activity, and that fully lymphocyte resistant target cells are likely required for compelling demonstration of the efficacy of this approach. However, using a dose response curve experimental design, I was able to demonstrate significant enhancement of target killing by granzyme delivered toxins, pending repetition of the experiment.

4.2 Results

4.2.1 Development of orthogonal granzyme-toxin fusions for enhancing lymphocyte mediated cytotoxicity

For developing my granzyme-toxin fusions, I used the same fusion protein design (shown in Figure 4.1) that was validated and successful previously, as discussed in detail in Chapter 3. I fused the toxins at the C-terminus of full length granzyme B, with the two components separated by a flexible glycine-serine linker. I used full length granzyme B (including the pre and pro peptides at the N-terminus) so that the fusion protein would be appropriately processed and loaded into lytic granules. This has two important implications: (i) by virtue of its sequestration in lytic granules, the toxin domain of the fusion protein ought to have less, and perhaps no, detrimental effect on the delivery cell; and (ii) since the inhibitory dipeptide of GZB is only removed by cathepsin C once in lytic granules, if the fusion proteins were directly expressed in a non-lymphocyte cell type, the N-terminal inhibitory dipeptide ought to remain intact, keeping the granzyme B domain catalytically inactive.

Using this design I generated a variety of granzyme B-toxin (GZB-TOX) fusion proteins. To facilitate delivery cell tracking and cell sorting, I included a C-terminal GFP protein, separated from the granzyme-toxin fusions by a 2A ribosomal skipping sequence. This configuration results in a single mRNA transcript coding for GZB-TOX-2A-GFP in a single open reading frame. However, the 2A sequence results in ribosomal skipping during translation, yielding two separate protein products, GZB-TOX and GFP [239].



Figure 4.1: Granzyme toxin fusion protein design. The overarching design is, from N- to C-terminus: full length granzyme B (shades of blue) fused to a toxin (red) via a glycine serine linker (GSL, purple), followed finally by GFP (green). The polyprotein shown is synthesized as a single polypeptide. The 2A peptide (grey) induces ribosomal skipping, resulting in two mature proteins: GZB-TOX and GFP. Full length granzyme B was used, consisting of, from left to right (N- to C- terminal), the ER signal sequence (ERSS), the inhibitory dipeptide (IN) and finally mature granzyme B (matGZB)

Prior to testing the potential for cell mediated delivery of these toxin fusion proteins, it was first important to determine if the toxins retained their activity as C-terminal fusion proteins. I did this by directly expressing these toxin fusions in a variety cells, and then assaying their viability. Importantly, as noted above the propeptide included in full length granzyme B ought remain uncleaved in these non-lymphocyte cell lines, keeping granzyme B inactive and non-toxic. As such any toxicity observed in these transfected cells would be expected to be a result of the toxin domain.

Bacterial inhibitors of protein synthesis: diphtheria and pseudomonas toxins

I first studied two bacterial toxins: the A fragment of diphtheria toxin (DTA) and pseudomonas exotoxin A (PEA). I tested the toxicity of GZB-PEA and GZB-DTA fusion proteins in Hela cells, using propidium iodide (PI) staining as a measure of cell death (Figure 4.2). Hela cells were transfected with plasmids coding for GFP, GZB-DTA and GZB-PEA. Untransfected Hela cells were treated with the apoptosis inducer staurosporine (STS) as a positive control. 48 hours after transfection, cells were stained with propidium iodide and analyzed via flow cytometry. I found that GZB-DTA was substantially more potent than GZB-PEA, and therefore selected GZB-DTA for further testing.

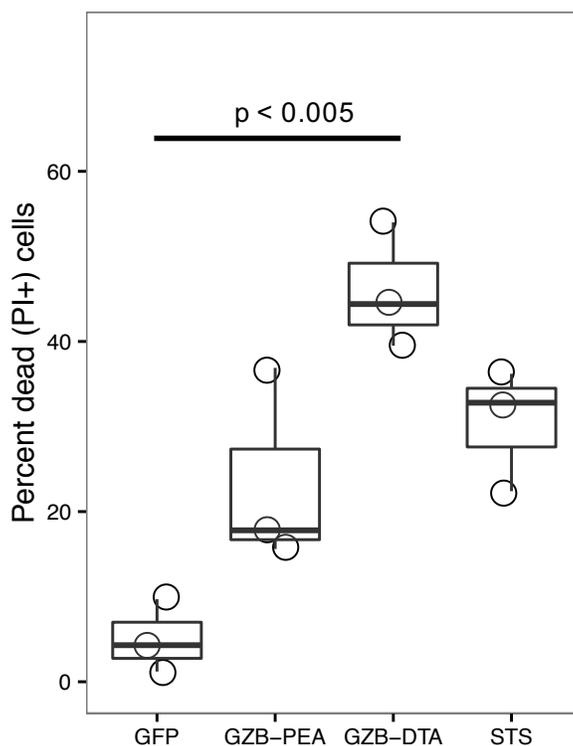


Figure 4.2: GZB-DTA and GZB-PEA testing in Hela cells. Data shown is the dead cell percentage 48 hours after transfection, as assessed by PI staining. Each circle is an independent replicate.

To confirm that the observed toxicity was not due to granzyme, but rather DTA, I expressed GFP, GZB and GZB-DTA in 293T cells and assessed their viability (Figure 4.3). I also included untreated cells as a negative control, and cells treated with the microtubule inhibitor colcemid as a positive control. 48 hours after transfection, cell viability was assayed using a fluorescent metabolic activity assay (PrestoBlue). This reagent contains a cell permeable resazurin dye, which is not fluorescent. The reducing environment of live cells reduce this dye to produce resorufin, which emits red fluorescence when excited by green light. Dead cells lose their reducing environment, and thus do not activate resazurin. Therefore, the fluorescent intensity of a sample of cells that is stained with PrestoBlue is proportional

to the number of live cells in the sample. Using this assay allowed indirect measurement of cell death throughout the experiment, rather than only at the time of measurement, since samples with appreciable cell death will have fewer cells at the time of assay, resulting in a lower fluorescent signal. This is important as dead cells eventually disintegrate to the point that they are not measurable when using, for example, viability dye stains such as DAPI or PI.

Raw fluorescence values were obtained for all samples. I then subtracted the colcemid readings from all other values, and corrected the values for transfection efficiency (quantified by flow cytometry to be 70 %). Finally I normalized the viability of each treatment condition to the viability of the untreated cells. The results show that the viability of 293Ts transfected with GZB alone was comparable to GFP, while 293Ts transfected with GZB-DTA exhibited substantial cell death.

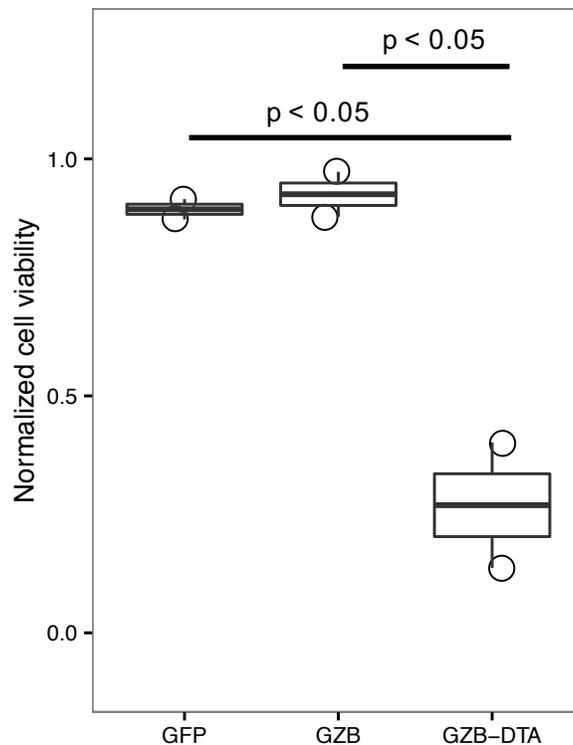


Figure 4.3: GZB-DTA testing in 293T cells. Shown is the viability of each transfected cell population, normalized to untransfected controls. Each circle is a biological replicate, which as an average of technical triplicates.

The herpes simplex virus thymidine kinase ganciclovir system

I next investigated the herpes simplex derived thymidine kinase ganciclovir (HTK/GCV) system, using a similar approach to what was used in testing GZB-DTA. 293T cells were transfected with plasmids coding for GZB or GZB-HTK and then treated with either ganciclovir (GCV) or vehicle control. I

also included cells treated with colcemid (COL) as a positive control. 5 days after transfection, the cells were imaged via brightfield microscopy, and then trypsinized, reseeded in plates, and assayed using a fluorescent metabolic activity assay. I normalized each sample's viability to the viability of cells transfected with GZB and treated with GCV. The results demonstrate both that the GZB-HTK fusion protein can activate GCV, and illustrate the synergy of the HTK/GCV prodrug system: GCV or HTK in isolation is not toxic, while the combination is highly toxic Figure 4.4.

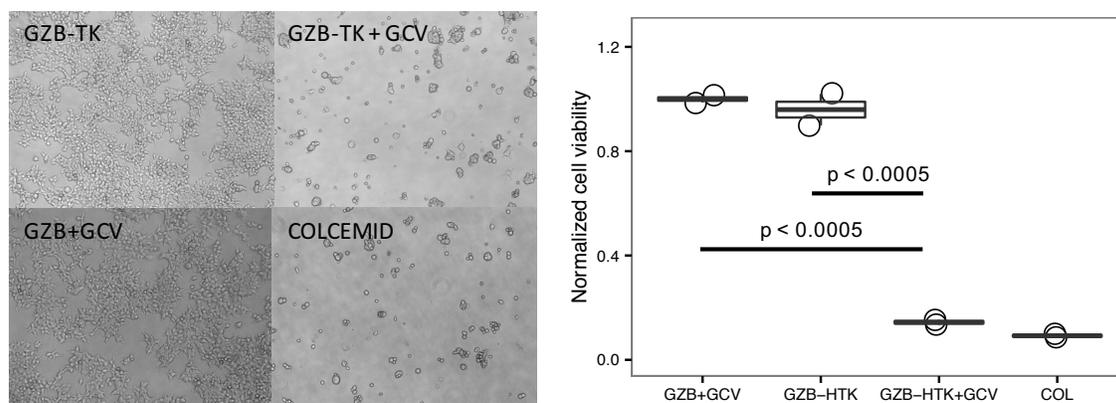


Figure 4.4: GZB-HTK testing in 293T cells. At left, optical microscopy images are shown of samples that were then used to produce the data plotted at right. Shown is the viability of each transfected cell population, normalized to the viability of cells transfected with GZB and treated with GCV. Each circle is a biological replicate, which as an average of technical triplicates. GCV = ganciclovir, COL = colcemid.

The *E. coli* nitroreductase CB1954 system

Unlike the HTK/GCV system, the nitroreductases (NTR) *nfsA* *nfsB* are not well characterized. Therefore, rather than a simple binary comparison of drug compared to vehicle control, I chose to assess the activity of the GZB-NTR fusion proteins using a dose response approach. 293Ts were transfected with plasmids coding for HTK (as a negative control, since HTK should not activate the NTR prodrug CB1954), NFSA, NFSB, GZB-NFSA, GZB-NFSB, and then treated with increasing levels of CB1954. 5 days after transfection, the cells were reseeded in plates and assayed using a fluorescent metabolic activity assay. The raw values were corrected for media background and then normalized to the values measured for each transfected cell population (e.g. HTK, NFSA, etc) that were treated with 0 μ M CB1954. The results, shown in Figure 4.5, have three key findings: (i) both nitroreductases exhibit a loss of activity as C-terminal granzyme fusions as compared to their unfused counterparts; (ii) GZB-NFSA retains sufficient activity to activate CB1954 to produce a toxic effect, while GZB-NFSB does not, and is indistinguishable from the negative control HTK population; (iii) 50 μ M is tentatively an appropriate dose at which CB1954 alone is non-toxic, but becomes so in the presence of GZB-NFSA. Since I did not have biological replicates for this experiment, it is impossible to draw firm conclusions,

but the data was sufficient to select GZB-NFSA as a candidate worth pursuing further.

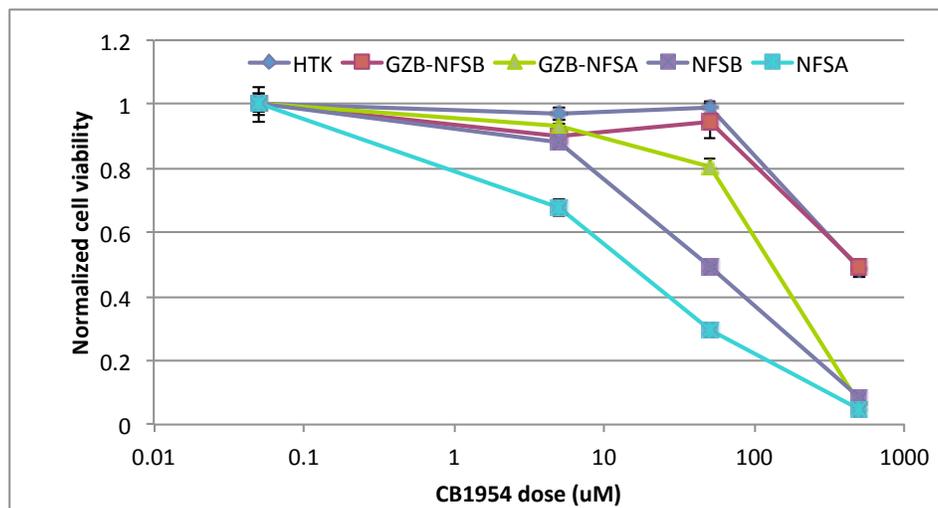


Figure 4.5: GZB-NFSA and GZB-NFSB testing in 293T cells. Shown is the viability of each transfected cell population as a function of the CB1954 dose. Each data point has been normalized to the 0 µM viability for each transfected plasmid (e.g. HTK, NFSA, etc.). Data plotted is the average of technical triplicates. Error bars shown represent the standard deviation of the triplicates. Note that the CB1954 dose is plotted on a log scale, and to accommodate this the 0 µM data points are falsely plotted at 0.05 µM.

Taken together the results in this section demonstrate that a variety of toxins retain their activity as C-terminal granzyme fusions, albeit to varying degrees, and that these GZB-TOX fusions kill target cells when expressed directly in them. This encouraged us to develop a lymphocyte resistant model cell-line in which to further test the granzyme-toxin fusions.

4.2.2 Characterizing the lymphocyte resistance of the breast cancer cell line MCF-7

As a model system for apoptosis resistance, I selected the breast cancer cell line MCF-7 as a target cell. MCF-7s lack caspase 3, one of two key executioner caspases in the apoptotic cascade [240]. Caspase 3/7 downregulation in tumour cells is widely observed, correlates with poor survival, and thus MCF-7s have been frequently used to study the chemotherapeutic and apoptosis resistant phenotypes that can result from caspase downregulation [209, 210]. I used the natural killer cell line YT-Indy (hereafter YT) as a model for cytotoxic lymphocytes, as it has an intact granzyme-perforin pathway [185].

I am aware of only one report in the literature in which MCF-7s were used as a target cell for YT-Indy [241], which provides data only that obliquely suggests that YTs target MCF-7s. Therefore, I first set out to characterize MCF-7's susceptibility to YT attack. I co-cultured CFSE labeled MCF-7s with YTs, and then stained the mixed cell population with propidium iodide (PI) and analyzed the cells via flow cytometry. After excluding debris and doublets and gating on CFSE+ target cells, I found that there was minimal change in the viability of MCF-7s cultured with YT's as opposed to those cultured

alone (89.5% compared to 92.9%). This is in stark contrast to the 721.221 (hereafter 721) cell line which is known to be killed by YTs. Using an identical experimental design as was used for MCF-7s, YTs killing of 721s was measured to be roughly 50% (721 alone viability of 92.2% compared to 42.8% viability of 721s co-cultured with YTs).

To exclude the possibility that this apparent lack of cell death was due to mechanisms of resistance involving the granzyme perforin pathway, I used a granzyme B FRET reporter (Sharma et al., in preparation) to confirm that granzyme B does indeed enter MCF-7s. This reporter consists of CFP and YFP fluorescent proteins separated by the granzyme B consensus cleavage substrate. In the absence of granzyme B there is a CFP to YFP FRET transfer signal. The presence of granzyme B results in the cleavage of the substrate, separating the FRET partners, which results in an increase in the CFP signal, and a decrease in the FRET transfer signal. I cultured target MCF-7 cells expressing this FRET reporter either alone, or in the presence of YT effector cells, and then analyzed the target MCF-7 cells for granzyme B status via flow cytometry (Figure 4.6). MCF-7s cultured with YTs exhibited a clear FRET-shifted population as compared to MCF-7s cultured in isolation, indicating that granzyme B is transferred by YTs to MCF-7s. The fraction of FRET-shifted MCF-7s (9%) in Figure 4.6 is at the lower end of the range I have measured, which can be as high as 30%. This percentages are substantially lower than those observed for other target cell lines such as 721s, which commonly exhibit approximately 60-70% death in similar YT co-culture conditions. The potential reasons for this difference are many. An obvious possibility is that the complement of surface receptors on 721s produce a much more robust YT response than do those on MCF-7s. MCF-7s are also very large and adherent cells, so it is possible that the YTs are less able to fully explore the physical environment of the co-culture. An important final note is that the relatively low percentage of YT-targeted MCF-7s (10-30%) is an important consideration when interpreting YT-MCF-7 co-culture experiments, as only these cells will receive the granzyme-payload fusion proteins, which, depending on the application, puts an upper limit on the potential observable effect size.

The viability of the FRET+ cells (98.2%) was essentially the same as both the overall viability of the target MCF-7s co-cultured with YTs (97.1%) and the viability of MCF-7s cultured in isolation (98.1%). This suggested that MCF-7s might be resistant to YT attack. To confirm this, I investigated the long-term viability of YT-targeted MCF-7s. To do this I co-cultured YTs with MCF-7s expressing the granzyme B FRET reporter, and then isolated viable, FRET shifted, granzyme B positive, MCF-7s, via cell sorting. As controls, I also sorted viable, non-shifted (FRET-) MCF-7s, and dead MCF-7s. I cultured these sorted cells for one week, and then assessed their viability using a commercial fluorescent metabolic activity assay (PrestoBlue). After measuring raw fluorescent intensity, I subtracted the background signal measured from the dead cell wells from the other two populations, and then normalized the fluorescent values to the FRET- population. The resulting data shows that all MCF-7s that were FRET shifted ultimately died (Figure 4.7). Combining these results, my working model for YT killing of MCF-7s is that the absence of caspase-3 significantly slows MCF-7 cell death, but that YT effector mechanism are ultimately sufficient to initiate target cell death, and therefore MCF-7s are not

lymphocyte resistant.

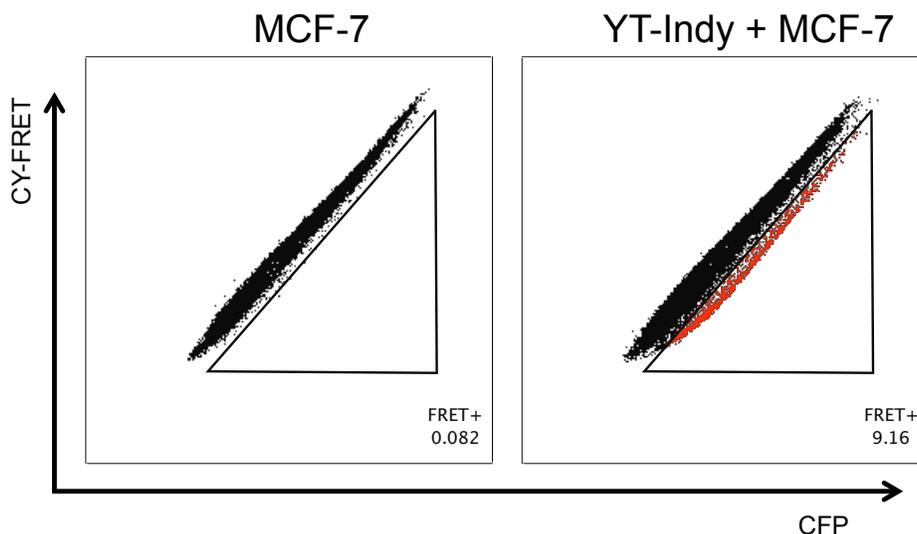


Figure 4.6: Characterizing YT delivery of GZB to MCF-7s using a FRET reporter. After excluding debris and doublets, YFP+ target cells were selected, which are plotted here. The CFP to YFP FRET transfer signal (CY-FRET) signal is plotted on the y-axis, and CFP signal on the x-axis. The triangle gate is set on the MCF-7 alone population (left), and copied to the co-culture population (right). FRET-shifted cells are plotted in red.

In summary I found that: (i) YTs target MCF-7s; (ii) granzyme B is transferred from YTs to MCF-7s; and (iii) MCF-7 cell death due to YT attack is slow, but occurs nonetheless, meaning that unmodified MCF-7s are not suitable as a model lymphocyte resistant target cell line.

4.2.3 Efforts towards generating a lymphocyte resistant cell line

Based on the results of the previous section it was evident that in order to observe enhanced lymphocyte killing due to orthogonal toxin delivery to MCF-7s, the cell line would have to be rendered fully resistant to lymphocyte attack by further modifying the apoptosis pathway. I attempted to do this both by treating target cells with small molecule inhibitors of apoptosis, as well as by overexpressing various genes that inhibit apoptosis in target cells. I tested the effects of these modifications in two cell lines, MCF-7 as well as the B-cell lymphoblastoid cell line 721 [186], as they are a well known YT target.

Small molecule inhibition of apoptosis

I first conducted experiments to see if the combination of the granzyme B inhibitor dichloroisocoumarin (DCI) [242] and the pancaspase inhibitor Q-VD-OPh would attenuate YT killing of 721 target cells. I co-cultured CFSE labeled target cells with YTs, and treated these co-cultures with either DMSO as a vehicle control, or the two inhibitors. I then stained the cells with propidium iodide and quantified target cell viability via flow cytometry. The results clearly show that the inhibitors essentially abrogate

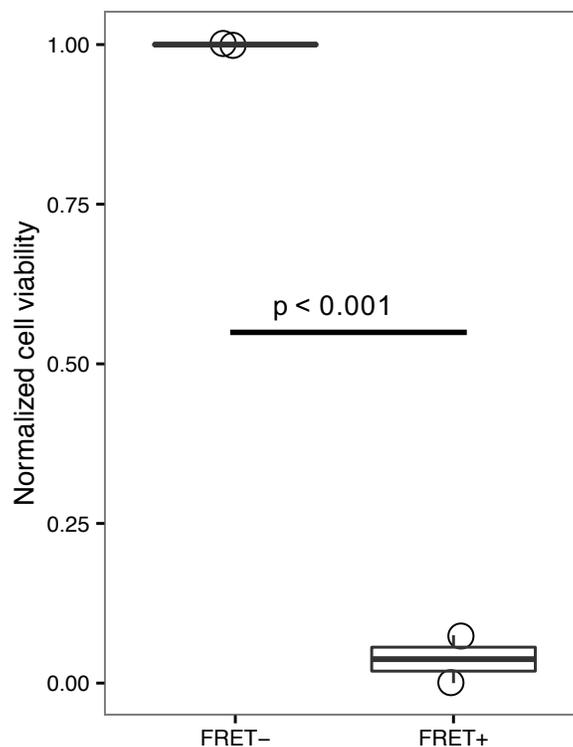


Figure 4.7: Long term survival of YT-targeted MCF-7s. Normalized MCF-7 viability post YT co-culture and week long culture in isolation is shown. Each circle is a biological replicate, which as an average of technical duplicates.

YT killing of 721s (Figure 4.8).

Since my experiments in Section 4.2.2 demonstrated that post co-culture viability is not necessarily indicative of long term survival, it was necessary to determine if this was the case for DCI and Q-VD-OPh inhibition of YT killing. As target cells I used 721s or MCF-7s, both expressing the granzyme FRET reporter (Section 4.2.2). I co-cultured target cells with YTs in the presence of Q-VD-OPh. I was unable to use the granzyme B inhibitor DCI since it has a fluorescence spectra that overlaps with CFP, a component of the FRET reporter which is required for isolating YT-targeted cells. After the co-culture, I stained the mixed population with propidium iodide (PI) and then isolated three target cell populations via cell sorting: PI-FRET-, PI-FRET+, and PI+. After one week in culture I assessed the viability of these cells using a commercial metabolic activity assay (PrestoBlue). I subtracted the readings from the PI+ cells from the other data, and then normalized the remaining values to the PI-FRET- values for each cell type, with the results shown in (Figure 4.7). Similar to my results for MCF-7s without inhibitors, virtually all YT-targeted (FRET+) 721s and MCF-7s ultimately die.

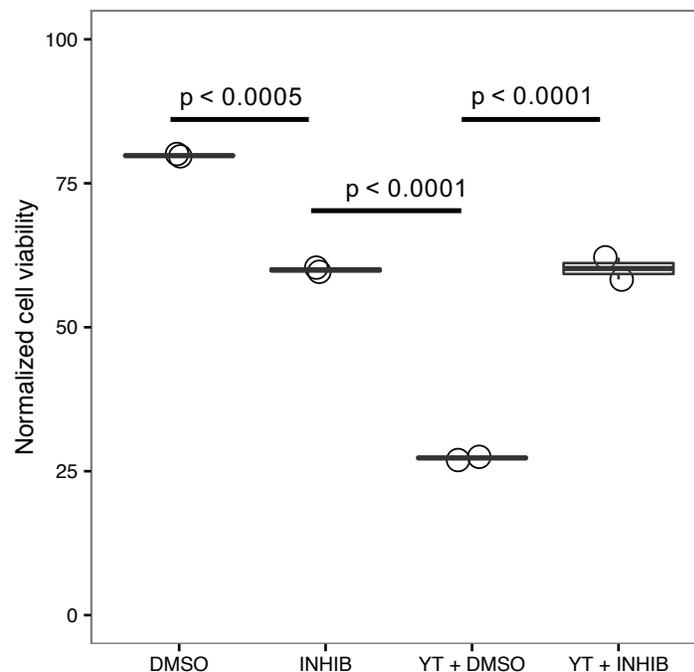


Figure 4.8: Small molecule inhibition of YT killing of 721 target cells. After excluding debris and doublets, CFSE+ 721 target cells were selected. Plotted is the percent of the 721 population that is dead (PI+). Each circle is a biological replicate.

Overexpression of inhibitor of apoptosis proteins

I also investigated the potential of rendering target cells resistant to lymphocyte attack by overexpressing inhibitor of apoptosis proteins (IAPs). Overexpression of IAP family members blocks the activation of executioner caspases 3 and 7 by granzyme or other intrinsic activators, and IAP upregulation is a natural apoptosis resistance mechanism of tumors [209, 222, 243]. I chose to focus on two IAP proteins survivin (SURV) and X-linked Inhibitor of Apoptosis Protein (XIAP), as these have been experimentally shown to render cells resistant to NK-mediated cytotoxicity, despite effective NK targeting, engagement and degranulation [230, 231]. My hope was that the more biologically relevant mechanism, as well as the potential for ongoing production of the IAPs in the target cells might be sufficient to render them resistant to lymphocyte attack.

I constructed plasmids that polycistronically co-expressed an IAP protein (either XIAP or SURV) along with the FRET reporter. I used the same 2A peptide design as for the granzyme toxin fusion proteins. In this case, each polyprotein consisted of (from N- to C-terminus) IAP-2A-FRET. As in the toxin case, the 2A peptide causes ribosomal skipping, resulting in separate IAP and FRET proteins. I transfected MCF-7s with plasmids coding for XIAP-FRET, SURV-FRET, and FRET (as a control), and co-cultured these target cells with YTs. Following the co-culture I stained the mixed cell population with PI, and isolated PI- FRET+ target cells. As controls, I also collected live and dead (PI- and PI+)

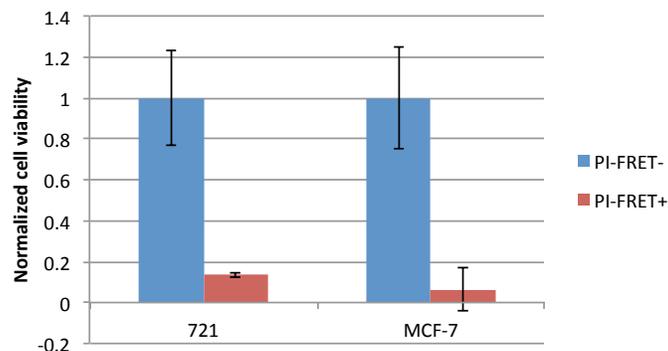


Figure 4.9: Long term viability of targets co-cultured with YTs in the presence of apoptosis inhibitors. Normalized target cell viability post YT co-culture and week long culture in isolation is shown. Data shown is the mean of technical duplicates, and the error bars are the standard deviation.

unmodified MCF-7s. After one week in culture I measured their viability using a fluorescent metabolic activity assay, and normalized all viabilities to the values measured for the PI- MCF-7s. The results in Figure 4.10 indicate that the IAP proteins do not provide any increase in MCF-7 resistance to YT-attack as compared to MCF-7s expressing FRET alone.

I was somewhat surprised by this result, as there are reports in the literature that overexpression of IAP proteins in target cells can protect them *in vitro* from lymphocyte mediated cytotoxicity [220, 230, 231]. However, there are several possible reasons for this discrepancy. My model system uses human effector and target cells, while much of the data in the literature uses murine cells. More generally, comparing different immortalized cell lines, even within species, is challenging, and none of the experiments in the literature used either MCF-7s or YT-Indys. Furthermore most of the experiments in the literature only follow the viability for a matter of hours, and thus it is very possible that, as with our experiments, the IAPs are actually only slowing death rather not stopping it. In the experiments which did measure the viability of the cells over the long term, the protective effect sizes diminished dramatically, and were only preserved at extreme effector to target ratios.

In summary, my efforts at rendering target 721 and MCF-7 cells resistant to YT killing were ultimately unsuccessful. Despite in many cases slowing it, neither small molecule inhibition of caspases, nor IAP mediated inhibition, ultimately prevented YT-induced target cell death. While I did not have the biological replicates in these experiments necessary to rigorously exclude the possibility that these treatments might render target cells fully resistant to YT killing, the effect sizes are sufficiently large and reproducible across multiple experiments that I did not think it worthwhile to continue pursuing this avenue of experiments.

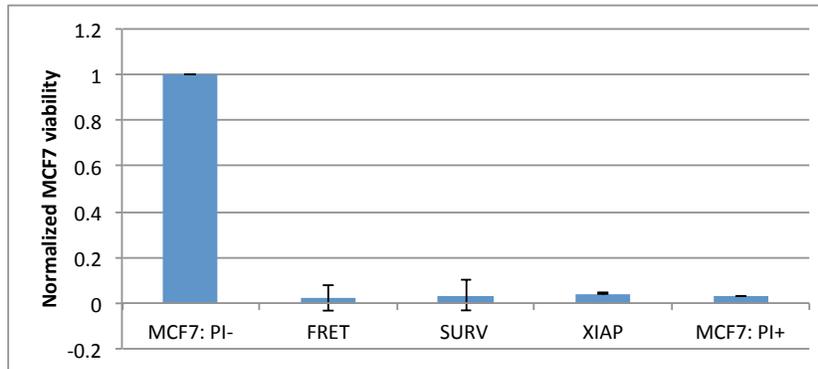


Figure 4.10: Long term viability of YT-targeted target cells expressing IAPs. Normalized target cell viability post YT co-culture and week long culture in isolation is shown. Data shown is the mean of technical duplicates, and the error bars are the standard deviation.

4.2.4 Effector dose response curves as a means of resolving small increases in YT target cell killing

Based on the results of the preceding three sections it was evident that the only way to enhance YT killing of MCF-7s was to deliver a payload which would have toxic effects that extend beyond the cell to which the payload is delivered. This is because all YT-targeted MCF-7s ultimately die, so any additional payload toxicity in the targeted cell would not be observable. In light of this I chose to proceed using the toxin/prodrug systems and the bystander effect they provide. Observing this effect would only be possible at effector target ratios at which there are significant populations of both YT-targeted and non-targeted MCF-7s, the former as a pool of toxin delivered and prodrug activating cells, and the latter as a pool of cells that are susceptible to the activated prodrug. To meet these requirements I chose to proceed with my experiments by attempting to deliver granzyme-toxin fusion proteins in the context of an effector cell dose response experiment using MCF-7 target cells. Using a dose response curve has the added benefit of resolving smaller effect sizes than simple binary comparison. To confirm such an approach was feasible, I characterized the effector dose-response behavior of fluorescently labeled MCF-7s subjected to YT attack by co-culturing the two cell types at varying effector target (E:T) ratios. I then isolated the MCF-7s via cell sorting. The total MCF-7 population was isolated, rather than only YT-targeted MCF-7s, since the goal of this experiment was to determine appropriate conditions for future toxin-delivery experiments. Since YT-targeted MCF-7s die, these experiments would require a pool of non-targeted MCF-7s to be present as targets for a toxin activated prodrug bystander effect. As controls I included MCF-7s that were not co-cultured with YTs, and dead MCF-7s. After one week in co-culture, I assessed their viability using a metabolic activity assay, with the results show in (Figure 4.11). At higher E:T ratios there was near total MCF-7 death, while at lower E:T ratios it was quite minimal. Based on my results above, this minimal cell death is likely due to the low absolute number of YTs in the co-culture.

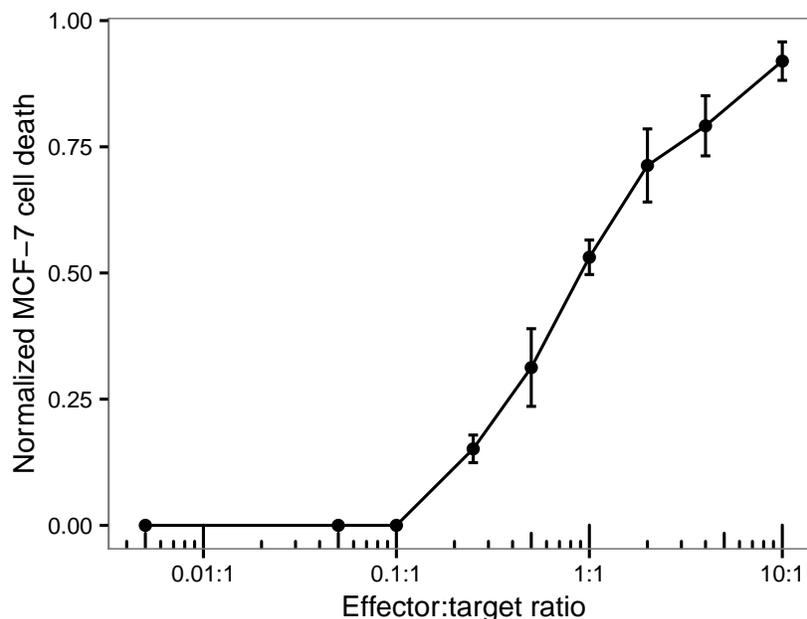


Figure 4.11: YT effector:target dose response of MCF-7s. Target cell death post YT co-culture and week long culture in isolation is shown. Fluorescent values were normalized to those from MCF-7s not co-cultured with YTs, and then these values were subtracted from 1 to transform the data from viability to cell death. Data points are the mean of biological duplicates, which are the mean of technical triplicates. Error bars are the standard deviation of the biological duplicates.

4.2.5 Validating the GZB-HTK and GZB-NFSA toxin fusion proteins in MCF-7s

The requirement for a bystander effect—namely delivered payloads having a toxic effect that extends beyond the cell to which the payload was delivered, necessary to show enhanced lymphocyte killing since YT-targeted MCF-7s uniformly die—constrained the rest of my work to the two toxin/prodrug systems I had evaluated: GZB-HTK/ganciclovir and GZB-NFSA/CB1954, as both ganciclovir and CB1954, once activated in a cell, can produce toxicity in adjacent cells. Prior to testing the potential for cell mediated delivery of these toxin fusion proteins, it was first important to investigate if the granzyme-toxin fusion proteins were capable of killing caspase-3 deficient MCF-7s. I did this by directly expressing these toxin fusion proteins in MCF-7s and assessing their viability. MCF-7s expressing either GFP or GZB-HTK fusion proteins were treated with either ganciclovir or vehicle control. After one week I assayed the cell viability using a metabolic activity assay (Figure 4.12a). I did not have biological replicates for this pilot experiment, but the results suggest that neither ganciclovir nor GZB-HTK treatment alone produces MCF-7 toxicity, while the GZB-HTK fusion protein in combination with ganciclovir is toxic to MCF-7s.

I validated the GZB-NFSA fusion protein in a very similar manner, with the additional step of testing multiple CB1954 doses, since this system is less commonly used, and so suitable doses were not clearly available from the literature. I transfected MCF-7s with plasmids coding for either GFP

or GZB-NFSA fusion proteins and treated these cells with several concentrations of CB1954. After one week I assayed the cell viability using a metabolic activity assay (Figure 4.12b). Again, without biological replicates, I cannot make any definitive conclusions, but these experiments suggest two important findings to guide my next set of experiments. Unlike ganciclovir, CB1954 seems to have an inherent toxicity to MCF-7 cells, regardless of NFSA activation, at doses above 1 μM (19% and 59% at 10 μM and 50 μM respectively). However, GZB-NFSA appears to increase this toxicity (by 25% and 57% at 10 μM and 50 μM respectively), with the greatest increase measured at 50 μM . Despite this I chose to proceed with 10 μM as the working concentration of CB1954 for future experiments, since it offered the best balance of low baseline toxicity but significant GZB-NFSA activated toxicity.

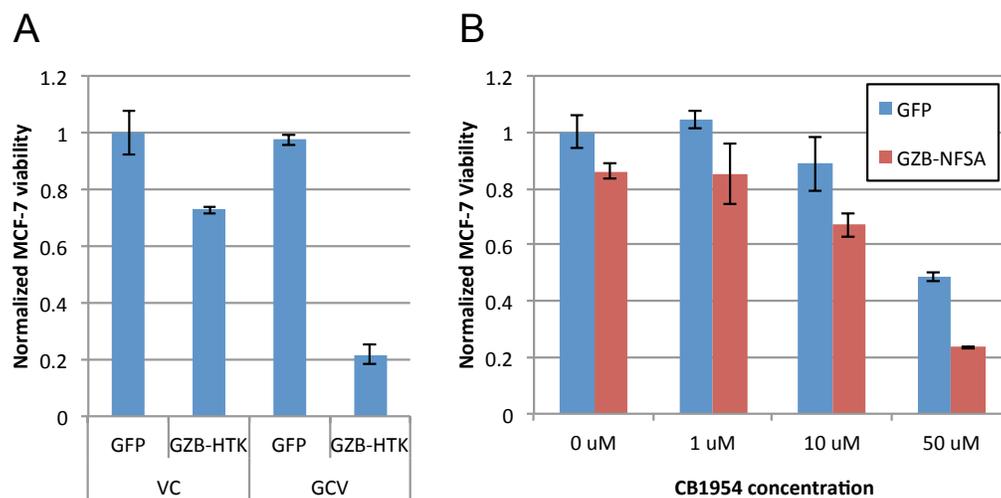


Figure 4.12: Testing GZB-HTK and GZB-NFSA toxin/prodrug systems in MCF-7s. **(A)** GZB-HTK. MCF-7s were transfected with plasmids coding for either GFP or the GZB-HTK fusion protein. Five thousand transfected (GFP+) cells were then sorted into each well of a 96 well plate. The cells were then treated with either 5 μM ganciclovir (GCV), or vehicle control (VC). Cell viability was assessed using a metabolic activity assay. Raw fluorescent values were normalized to the GFP + VC value. Error bars represent the standard deviation calculated from triplicate wells. **(B)** GZB-NFSA. MCF-7s were transfected with plasmids coding for either GFP (blue) or the GZB-NFSA (red) fusion protein. Five thousand transfected (GFP+) cells were then sorted into each well of a 96 well plate. The cells were then treated with a range of CB1954 doses. Cell viability was assessed using a metabolic activity assay. Raw fluorescent values were normalized to the GFP 0 μM value. Error bars represent the standard deviation calculated from triplicate wells.

4.2.6 Predicted and measured enhancements of YT-killing of MCF-7s using granzyme-toxin fusion proteins

Having validated the granzyme B toxin fusion proteins directly in MCF-7s, I attempted to use YTs to deliver these fusions to target MCF-7 cells. This would serve as proof of principle that cell based delivery via the granzyme perforin pathway could be used for practical applications such as toxin

delivery to apoptosis resistant tumour cells in the context of adoptive cell therapy.

To do this, I designed an effector dose response experiment, in which I co-cultured various types of granzyme-toxin fusion expressing YT effector cells with target MCF-7 cells at a range of effector target (E:T) ratios. I selected this approach for several reasons. First, at high effector target ratios, most MCF-7 target cells are killed, and so the potential window for increased YT killing due to toxin delivery is small. Similarly, the GZB-NFSA fusion only provides around a two-fold increase in CB1954 toxicity. Small effect sizes are much easier to resolve as a shifted curve, rather than a binary comparison. Furthermore, all YT-targeted MCF-7s do die eventually (Section 4.2.2). Therefore, any additional toxin-mediated killing would necessarily be via the bystander effect, in which toxins delivered to a YT-targeted MCF-7 activate prodrugs that then diffuse to and kill adjacent, non-YT-targeted MCF-7s. Thus the largest toxin-mediated increase in YT killing should be expected at intermediate E:T ratios, where there is a large pool of both targeted MCF-7s with toxins capable of activating prodrugs, as well as a large pool of non-targeted MCF-7s that have the potential to be killed by activated prodrugs. Finally, given that most solid tumours are large masses, and that the achieved E:T ratios in clinical adoptive cell therapy are reported to be well below unity [244], I feel this is a biologically realistic experimental design in which to test the utility of using granzyme-toxin fusions to enhance effector cell killing of tumour cells.

I first used the results from previous sections to predict what enhancement in YT killing could be reasonably expected, and then conducted the dose response experiments to determine if any enhancement actually occurs.

Estimated toxin/prodrug mediated enhancement of YT killing

It is challenging to provide an accurate prediction of the magnitude of the effect size that would be expected, but an upper bound can be estimated by considering the results of Section 4.2.5 and Section 4.2.4. To derive a useful expression that can be parameterized from the data available from these experiments, I assume that there are three potential causes of MCF-7 cell death: YT killing, baseline prodrug toxicity, and toxin activated prodrug toxicity. I further assume that these effects are independent (i.e. there is no synergy between them), and that they act in an order of precedence determined by their application to the target MCF-7s, that is first the YT killing, then the baseline drug toxicity, and then the activated prodrug toxicity. Finally I assume that a given effect is only exerted on whatever fraction of cells were not already killed by effects earlier in the order of precedence. Under these assumptions, the fractional death D of MCF-7s in the co-culture can be taken to be:

$$D = \gamma + (1 - \gamma)\beta + (1 - \gamma)(1 - \beta)\tau \quad (4.1)$$

where γ is the fraction of cells killed by YT-Indys, β is the fraction of cells that die due to baseline prodrug toxicity, and τ is the fraction of cells that die due to toxin activated prodrug toxicity. These values are taken to represent the cell death that occurs when the effect is applied in isolation. The multiplicative prefactors that appear in front of β and τ account for the attrition due to effects that are

earlier in the order of precedence: these later effects can only kill the fraction of cells not already killed by a previous effect. This assumption allows for estimating γ , β and τ from previous experiments in which the respective effects were applied in isolation: γ can be estimated from Figure 4.11, while β and τ can be estimated from Figure 4.12. Specifically:

$$\beta = 1 - \frac{V_{G,D}}{V_{G,VC}} \quad \text{and} \quad \tau = \alpha * \left(1 - \frac{V_{T,D}}{V_{T,VC}}\right) \quad (4.2)$$

where $V_{G,D}$ is the viability of MCF-7s transfected with GFP and treated with the prodrug (either GCV or CB1954), $V_{G,VC}$ is the viability of MCF-7s transfected with GFP and treated with vehicle control, $V_{T,D}$ is the viability of MCF-7s transfected with the granzyme B toxin fusion protein (either GZB-HTK or GZB-NFSA) and treated with the corresponding prodrug (GCV or CB1954 respectively), and $V_{T,VC}$ is the viability of MCF-7s transfected with the granzyme B toxin fusion protein and treated with the vehicle control. I have included α to model the efficiency with which MCF-7s that have received the granzyme-toxin fusion activate the prodrug. This is almost certainly a function of the effector: to target ratio, likely initially increasing as increasing amounts of the prodrug activating toxin are delivered, and then decreasing at high effector target ratios as the MCF-7s are so heavily attacked by the YTs that they die more rapidly and so produce less active prodrug for a shorter period of time. However, since I have no way of estimating the form of this dependence, in order to proceed I set $\alpha = 1$. I have also ignored any toxicity associated with granzyme B toxin fusion protein expression directly in MCF-7s (a small amount of which is evident in Figure 4.12), which is justified since here I am estimating cell death in a co-culture experiment, in which the only MCF-7s that contain the toxin fusion protein will also receive a YT-hit from which they will already die, and thus are already accounted for by γ .

Using the data from Figure 4.12 to estimate the viabilities V in Equation 4.2 gives $\beta = 0.03$ and $\tau = 0.78$ for the GZB-HTK/GCV system, and $\beta = 0.11$ and $\tau = 0.25$ for the GZB-NFSA/CB1954 system (assuming $10\ \mu\text{M}$ CB1954). This allows estimation of the death D from Equation 4.1 for any given γ , which is a function of the effector to target ratio, and can be estimated from Figure 4.11. In Figure 4.13 I have plotted these estimates for wild type effectors, as well as effectors delivering both toxin/prodrug systems. Notably, the death due to the GZB-HTK/GCV system is very high even for low effector to target ratios. This is not due to baseline ganciclovir toxicity, which is actually very low ($\beta = 0.03$), and rather because of a very high toxin activated prodrug toxicity ($\tau = 0.78$), and, critically, the assumption of perfect prodrug activation efficiency in the presence of any GZB-HTK (that is setting $\alpha = 1$). Again it is likely that at low E:T ratios, this conversion would not be perfect (i.e. $\alpha < 1$), and so the total cell death would be lower than is plotted. Nevertheless these plots demonstrate the potency of the GZB-HTK/GCV system, as can also be seen in Figure 4.12.

Finally to provide a summary metric for the estimates of maximum possible enhancement of YT killing due to toxin delivery, I selected an intermediate E:T ratio of 0.5:1, for which $\gamma = 0.31$ in Figure 4.11. I then calculated the absolute increase in cell killing $\Delta = D_A - D_Y$, and the fractional fold increase in cell killing $f = D_A/D_Y$ using Equation 4.1, where D_A is the total cell death due to all effects,

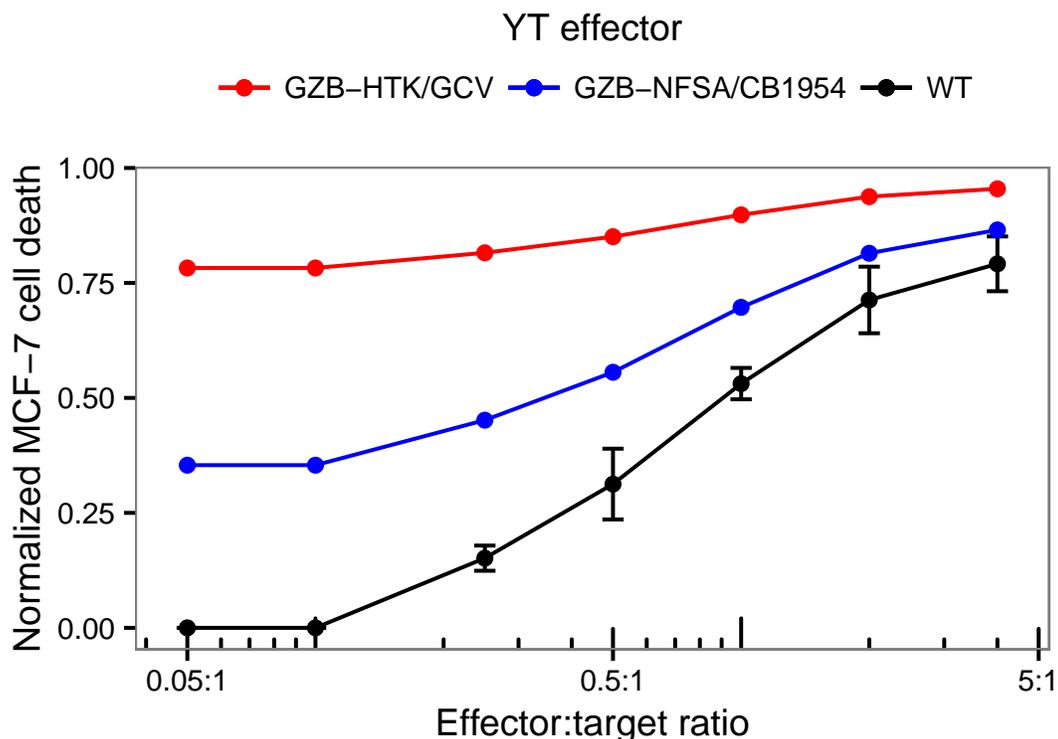


Figure 4.13: Estimates of toxin/prodrug enhancement of YT killing. Effector dose response curves for MCF-7s co-cultured with YTs. These were experimentally determined for wild type YTs (black, from Figure 4.11), and calculated using Equation 4.1 for YTs expressing GZB-HTK (red), and GZB-NFSA (blue), assuming treatment of the MCF-7s with the corresponding prodrug.

and $D_Y = \gamma$ is the cell death due to YT killing only, i.e. Figure 4.11. These were $\Delta = 0.54$ and $f = 2.72$ for the GZB-HTK/GCV system and $\Delta = 0.24$ and $f = 1.78$ for the GZB-NFSA/CB1954 system. These calculations provide a very rough estimate of the upper bound on the increase in cell killing that I would expect to see experimentally, and suggest that the GZB-HTK/CB1954 system might be very effective.

Experimental measurement of enhancement of toxin/prodrug enhancement of YT killing of MCF-7s

Based on my granzyme-toxin validation work, I set out to characterize the potential for enhancement of YT killing provided by both GZB-HTK and GZB-NFSA. However, when I transfected YTs with GZB-HTK, I observed massive genotoxicity in the YTs due to GZB-HTK expression. Since I have expressed a variety of granzyme B fusion proteins in YT cells and this was the first observation of genotoxicity, I am fairly confident this is due to the HTK domain. Given that HTK has been expressed in lymphocytes previously as a suicide gene in cell therapies [245], I was surprised by this result and repeated the experiment several times, but the finding was consistent. It is possible that the difference between my

Sample	EC_{50}	$\sigma_{EC_{50}}$	p-value
YTN-CB1954	0.30	0.037	0.0110
YTN	0.52	0.03	0.0035
YT-CB1954	1.12	0.015	0.0006
YT	3.32	0.65	-

Table 4.1: Fitted EC_{50} values for enhanced lymphocyte killing. Samples are listed in increasing EC_{50} (decreasing potency). $\sigma_{EC_{50}}$ is the standard error of the estimate for EC_{50} resulting from the fit. The p-values are the Tukey HSD adjusted p-value comparing the sample to its next-nearest neighbor.

observations and those in the literature are due to expression levels: those in the literature primarily use viral transduction, while I used electroporation, likely resulting in much higher expression levels. Since HTK catalyzes the synthesis of ADP from ATP, it is mechanistically conceivable that extremely high levels of this foreign kinase (not subject to autoregulatory mechanisms) could sufficiently deplete ATP levels to cause cell death, although there are no reports of this in the literature. Another possibility is that the HTK domain could have dislodged the inhibitory dipeptide at the catalytic site of GZB, resulting in GZB-mediated toxicity.

Fortunately this effect was not observed with GZB-NFSA expression, allowing me to proceed with effector dose response experiments. YTs expressing GZB-NFSA, as well as unmodified YTs as a control, were co-cultured with fluorescently labeled MCF-7 target cells at a range of effector to target ratios. Target MCF-7 cells—consisting of YT-targeted and YT-naive cells—were then isolated from the mixed cell population via FACS, sorted into 96 well plates, and treated with CB1954, or DMSO as a vehicle control (giving all wells equivalent DMSO concentration). Finally, the viability of these cells was measured using a metabolic activity assay. The results, shown in Figure 4.14, have three main findings. First, there is a clear E:T ratio dose response: higher YT-cell numbers result in higher levels of MCF-7 cell death, for all conditions, as expected from my pilot experiments. Second, there is increased MCF-7 cell death in samples that were either co-cultured with YTs expressing the GZB-NFSA toxin and then treated with vehicle control, or in samples that were co-cultured with unmodified YTs and then treated with CB1954. Again, I expected this based on experiments in which I directly treated MCF-7s with this toxin/prodrug system, and found intermediate toxicity due to either component of the system. Third and most importantly, the highest level of MCF-7 cell death is in samples that were co-cultured with YTs expressing the GZB-NFSA toxin, and then treated with the CB1954 toxin. Conversely, the lowest amount of MCF-7 cell death is in target cells co-cultured with unmodified YTs and not treated with CB1954.

To confirm that these results were statistically significant, I fit the data from each co-culture condition with a sigmoid logistic function of the form $y = (1 + e^{m(EC_{50}-x)})^{-1}$, where EC_{50} is the E:T ratio at which target cell death reaches half of its maximum value. The fit was done using a non-linear least-squares algorithm, and the results are shown in Table 4.1 and Figure 4.14. Using the EC_{50} estimates

and associated standard errors resulting from the fit, I conducted a single factor ANOVA test with the null hypothesis that all EC_{50} s resulted from the same distribution, which I rejected with $p = 3.6 \times 10^{-5}$. To determine which EC_{50} s were different, I conducted a post-hoc Tukey's HSD test, and found that all EC_{50} s were different with $p < 0.05$ minimum (see Table 2.1).

I also compared these results with the estimates of maximum enhancement of cell killing that I derived above. I was surprised to find that the experimentally measured $\Delta = D_A - D_Y$ was even greater than the theoretical predictions (Figure 4.15). It is possible that this is partly due to the higher level of wild type YT killing in the data used for the theoretical predictions (from Section 4.2.4)—which is likely due to minor variations in experimental conditions, as well as cell health, functionality and passage number. However, this is unlikely to be the only cause of this effect, since the two different YT populations in this co-culture experiment were treated identically and concurrently. The other likely factor is that the assumptions made to calculate the estimate are wrong. In particular, this comparison might suggest that the three sources of cell death (YT-killing, baseline CB1954 toxicity and toxin-activated CB1954 toxicity) are not independent, and are rather this is an indication that they are acting synergistically to enhance YT-killing. However, the simplistic nature of Equation 4.1 and its lack of validation, along with the fact that this discussion is based on comparing data from two separate experiments, makes any firm conclusions inappropriate.

Together, these results demonstrate a moderate enhancement in YT-killing of MCF-7s due to some combination of GZB-NFSA expression, CB1954 treatment, and possibly synergistic activity between the two, although the latter is uncertain. Furthermore, repetition of this experiment has been confounded by technical issues—flow cytometry sorting errors and YT effectors with an exhausted, non-cytotoxic phenotype—and therefore I cannot make any final definitive conclusions at this time regarding the enhancement of YT-killing by the GZB-NFSA/CB1954 system. Efforts are ongoing in the lab to remedy this.

4.3 Discussion

Cell based therapies are in the process of transforming medicine [177, 180]. In cancer therapy, adoptive cell therapy using tumour infiltrating lymphocytes and CAR targeted T-cells can be a curative treatment in large fractions of patients with malignant melanoma [181] and B-cell malignancies [246], respectively. However, many patients still fail to respond to these treatments, for reasons which have not been fully elucidated [247], but are sure to include the immunosuppressive tumour microenvironment, antigen escape, and resistance to cytotoxic effector mechanisms.

Here I have attempted to begin to address one aspect of the issue of apoptosis and lymphocyte resistance by using cytotoxic lymphocytes to deliver potent orthogonal toxins via the granzyme-perforin pathway. The approach I used to do this was using granzyme B as a molecular chaperone to insert the toxins into the granzyme-perforin pathway, as I had previously shown this was possible using a fluorescent protein (Chapter 3).

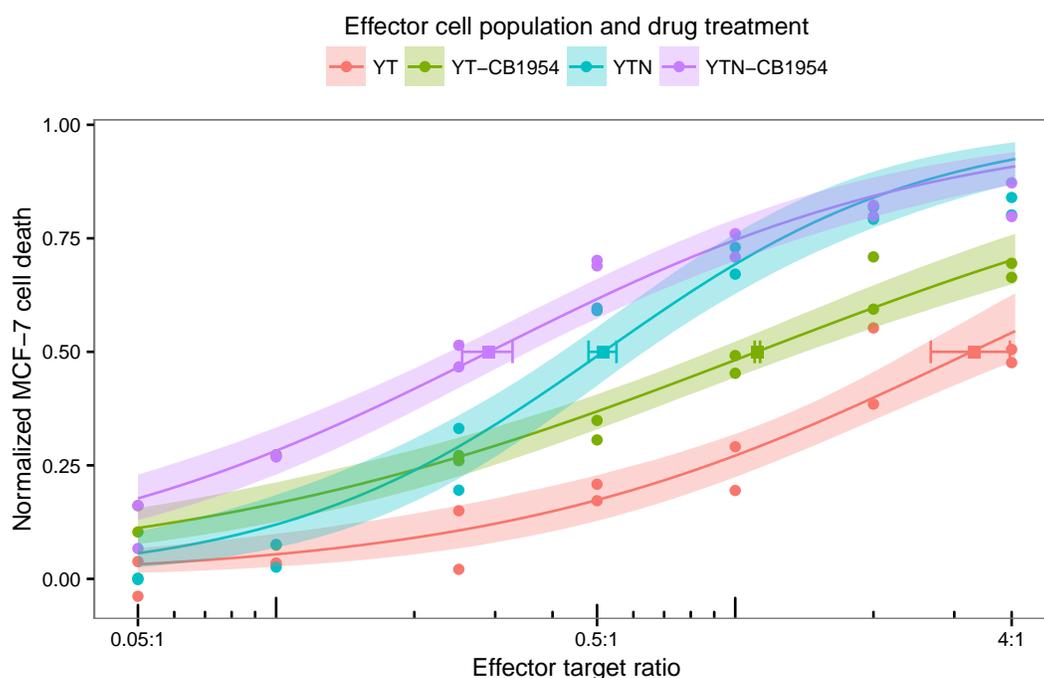


Figure 4.14: Investigating the enhancement of YT killing of MCF-7s by GZB-NFSA/CB1954. Fluorescently labeled target MCF-7s were co-cultured with YTs (YT) or YTs expressing GZB-NFSA (YTN). The target cell number was fixed, and the effector cell number adjusted for the range of E:T ratios. After the co-culture, the mixed cell population was sorted via FACS. After debris and doublet gating, target MCF-7 cells were gated upon using two colour discrimination (one colour for each cell type). One thousand target MCF-7 target cells were sorted into each well of a 96-well plate. These cells were then treated with either 10 μ M CB1954 or vehicle control (0.1% DMSO). Cell viability was assessed one week later using a fluorescent metabolic activity assay. Raw fluorescent values were normalized to those derived from wells containing untreated MCF-7s, and then these values were subtracted from 1 to transform the data from viability to cell death. Each data point is a reading from a single well of a plate. Solid lines are the fitted logistic functions, with the shaded region denoting the 95% confidence bands for the fitted function. The solid squares with horizontal error bars are the fitted EC_{50} and associated error the estimate. YT = wild type YT; YTN = YT expressing GZB-NFSA; DMSO = vehicle control treatment after co-culture; CB1954 = 10 μ M CB1954 treatment after co-culture.

I have generated a variety of granzyme-toxin fusion proteins, and shown that the toxins retain their activity as C-terminal fusions to granzyme B. However, my results also demonstrated that substantial attenuation of toxin potency can occur, as was the case with the nitroreductase fusion proteins GZB-NFSA and GZB-NFSB.

MCF-7s are widely used as a model cell line both for breast cancer, as well as apoptosis resistance, due to their deficiency in caspase 3. I characterized their susceptibility to killing by the natural killer cell line YT-Indy. I found that while MCF-7s are resistant to YT-killing over the duration of typical

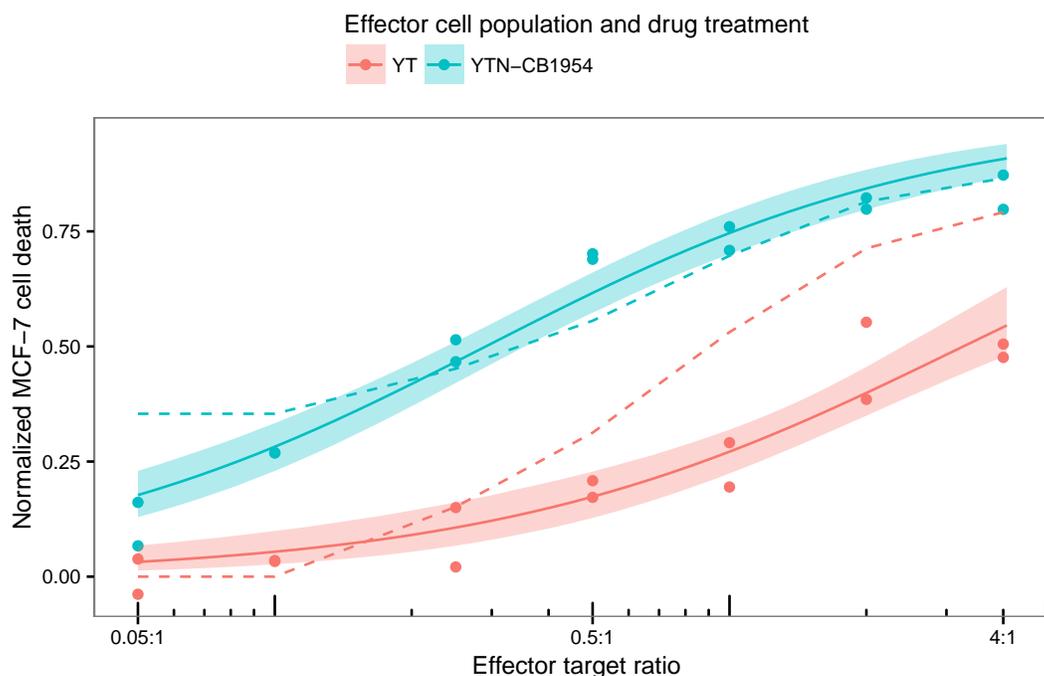


Figure 4.15: Comparing estimates with experimental measurements of enhancement of YT killing. Comparison of the experimental data (solid lines, circles and shaded bands, all as in Figure 4.14) and estimates from Equation 4.1 (dashed lines).

co-culture and assay timescales (several hours to a day), ultimately YT-targeted MCF-7s do die (within a week). I encountered similar results in my attempt to render target cells resistant to YT-attack by inhibition of apoptosis. When using both small molecule inhibition of caspases, as well as overexpression of IAP proteins, I found these modifications delayed target cell death due to YT attack, but did not actually prevent it. This recurring theme of delayed target cell death is relevant to the wider community studying both apoptosis resistance and methods to overcome it, in that long term evaluation of cell viability is critical.

These results, and the failure of multiple approaches to render target cells fully lymphocyte resistant, raises two important points. First, it is natural to wonder why, despite reports in the literature (Section 4.1), inhibition of apoptosis was insufficient to render MCF-7s lymphocyte resistant. In this regard it is important to note that I did not directly functionally verify that the apoptotic cascade was inhibited, for example by monitoring caspase activity. However, both IAP overexpression and small molecule inhibition did delay cell death on the timescale of hours, so it is fairly clear that these treatments were exerting some pro-survival activity. Further, as discussed above, most of the reports of *in vitro* IAP-mediated lymphocyte resistance only measured target cell viability after short timescales (at which point our data and that in the literature is congruent), and the few data points in the literature from experiments that monitored the target cell viability over several days have small effect sizes

(Section 4.2.3). Beyond these specific reports, much of the evidence for IAP involvement in cancer comes from either mouse models or patient samples [219–227]. In these *in vivo* contexts, it is possible that IAP upregulation provides a proliferative effect rather than, or in addition to, resistance to immune mediated toxicity. Also, as it is the dominant therapeutic modality, much of the evidence for IAP mediated tumour resistance to apoptosis is in the context of chemotherapy [248–250], which is not directly comparable to lymphocyte resistance.

More broadly, the challenges I encountered in generating a fully lymphocyte resistant cell-line speak to the complexity of the cytotoxic lymphocyte-tumour cell interaction, and the difficulty in modeling it *in vitro*. My results indicate that simple modification of one core pathway, no matter how central, is perhaps insufficient to prevent target cell death in a highly artificial co-culture environment, where the effector cells are free to attack over an extended period of time, with none of the metabolic, soluble or cell-mediated inhibitory factors encountered in a tumour. A more physiologically relevant model system might include multiple modifications to a variety of components of the apoptosis pathway, perhaps with redundancies. These results may also suggest that in the context of cytotoxic lymphocyte therapy of tumour cells, apoptosis and lymphocyte resistance are secondary issues. Certainly inhibition of cytotoxic lymphocyte trafficking to the site of the tumour, inhibition of lymphocyte recognition of tumour-cells, and inhibition of lymphocyte binding and engagement to tumour cells are all well characterized immunoresistance mechanisms [96]. Furthermore the immunosuppressive tumour microenvironment, containing regulatory T-cells and myeloid derived suppressor cells, as well as tumour upregulation of PD-1 and IDO expression are key considerations [105, 251]. Perhaps these mechanisms are the core issues that need to be addressed. More likely it is some combination of all three factors: the tumour microenvironment, lymphocyte tumour cell recognition and engagement, and target cell susceptibility to lymphocyte toxicity. My results seem to suggest that simply focusing on the latter in isolation, especially in an isolated co-culture system is perhaps insufficient.

The limitations of my model system are likely part of the reason why the effect sizes I observed in my final series of experiments are modest. When I attempted to use the GZB-NFSA/CB1954 system to enhance YT-killing of MCF-7s, I was aiming at a small therapeutic window. Any YT-targeted MCF-7s I knew would ultimately die, and the therapeutic index of the GZB-NFSA/CB1954 system in MCF-7s is also very small. I conducted a series of dose response experiments and found that MCF-7s co-cultured with YTs expressing the toxin fusions and treated with CB1954 exhibited greater cell death than those that were co-cultured with unmodified YTs and not treated with CB1954. However, the relative contribution of CB1954 treatment and GZB-NFSA delivery to target cells, and the potential synergy thereof, is unclear. This, along with the moderate magnitude of the enhancement of cell killing, I attribute at the very least to my lack of a YT-resistant cell line, as well as the small therapeutic index of NFSA as a granzyme fusion. It is also possible that the presence of a host immune system might increase the efficacy of GZB-NFSA/CB1954, in that even moderate levels of bystander killing could result in exposure of new antigens, providing a degree of antigen spread, and thus restarting the host immune response.

Unfortunately, CB1954 has a dose dependent toxicity in MCF-7s even in the absence of NFSA, which results in a small therapeutic window. Several of the granzyme-toxin fusion proteins I developed have significantly greater therapeutic windows, in particular GZB-HTK/ganciclovir and GZB-DTA. Viral transduction of GZB-HTK might reduce or eliminate the genotoxicity of this fusion protein upon expression in YTs, which would make this system very attractive. Since GZB-DTA does not produce a bystander effect, using it would require a target cell line that was fully lymphocyte resistant. Furthermore, the delivery chassis would have to be protected from DTA toxicity. This is feasible, as a mutant form of EEF2—recall DTA inhibits protein synthesis by ribosylating EEF2 (Section 4.1)—has been reported, the overexpression of which renders cells resistant to DTA toxicity [252]. I have shown that co-expressing this mutant EEF2 along with GZB-DTA abrogates the toxicity of GZB-DTA (as shown in the Appendix). If these two issues were resolved, DTA would be another attractive payload to move forward with.

In conclusion, I have demonstrated that granzyme-toxin fusion proteins retain their activity to varying degrees. Total, long-term resistance to lymphocyte attack in an artificial co-culture system is challenging to achieve, and perhaps not as relevant as other elements of tumour immunoresistance. Finally, I have shown measureable enhancement in MCF-7 target cell killing by YTs expressing GZB-NFSA. The effect sizes are moderate, the synergy of the toxin/prodrug interaction unclear, and any firm conclusion is pending repetition of the experiment. I am hopeful that with a more relevant target cell line, and potentially a toxin with improved activity as a granzyme fusion, this system may be shown to be worth pursuing for further development in the context of adoptive cell therapy of cancer.

4.4 Methods

4.4.1 Plasmids

Two plasmid systems were used for this work. The MND plasmids are based on a lentiviral transfer vector, and have been described elsewhere [253, 254]. They use the MND promoter to drive high levels of expression in hematopoietic lineages. The pdL vector is a custom mammalian expression plasmid I constructed in house, based on a pcDNA3.1(+) (Thermo Fisher Scientific) backbone. Specifically, the mammalian and bacterial selectable markers and all origins of replication are derived from pcDNA3.1(+), corresponding to bases 1670 (CGATTCGGCCTATTGGTTA...) to 5396 (...TAAACAAATAGGGGTTCGC). A custom expression cassette was cloned into this backbone. This cassette consisted of eukaryotic and prokaryotic promoters and ribosomal binding sequences, followed by the open reading frame, followed by eukaryotic and prokaryotic transcriptional termination sites. For the mammalian promoter I used the CAG promoter for its ability to drive high levels of expression in a variety of tissues. The sequence was amplified from pEMS1157 [197]. This was followed by a hybrid T7 prokaryotic promoter, taken from pCMVTnT (Promega). This was followed by consensus Shine-Dalgarno and Kozak sequences. Following this is the open reading frame, which varies by

plasmid. Following the end of the coding sequence, there is a BGH polyA sequence, and then a T7 terminator (with both sequences taken from pcDNA3.1(+)). Restriction enzyme cleavage sites flank all components to facilitate subcloning.

A variety of coding sequences were inserted into these two base vectors. The coding sequences fall into two general categories: toxin related and inhibition of apoptosis related. For the granzyme B toxin (GZB-TOX) fusions, the basic structure was always: full length granzyme B, followed by a glycine-serine linker, followed by the toxin, followed by a P2A ribosomal skipping peptide, followed by GFP. For plasmids expressing inhibitor of apoptosis genes the inserted coding sequence had the general form of: inhibitor of apoptosis protein (either XIAP or Survivin), followed by the same P2A peptide, followed by a granzyme B FRET reporter. This last component consists of CFP, followed by the consensus granzyme B cleavage substrate, followed by YFP. It was constructed in house, and the details have been published ([255], Sharma et al., in preparation). In all cases restriction enzyme cut sites flank all components.

For clarity and completeness, actual plasmid information is organized as follows. Plasmid maps and full plasmid sequences for the base MND and pdL vectors are in the Appendix (Section E.2). The full coding sequence that was inserted into these base vectors is also in the Appendix for all plasmids used. In this way the full plasmid sequence for every plasmid used in these experiments is captured. Finally, the source of each component of all coding sequences are listed in the Appendix (Table E.1).

All plasmids were constructed through a combination of PCR, synthesis and restriction/ligation cloning. All PCR amplicons and coding sequences were sequence verified.

4.4.2 Cell culture

HeLa cells were a gift from Jonathan Choy (Simon Fraser University). MCF-7 cells were a gift from Gregg Morin (Canada's Michael Smith Genome Sciences Centre). 293T cells were a gift from Kieth Humphries (British Columbia Cancer Research Centre). YT-Indy and 721.221 cells were a gift from Judy Lieberman (Harvard University). YT cells were cultured in RPMI 1640 media, supplemented with 20% heat inactivated fetal bovine serum, 1X GlutaMAX, 1 mM sodium pyruvate, 10 mM HEPES, 0.1 mM beta-mercaptoethanol. All other cells were cultured in DMEM, supplemented with 10% heat inactivated fetal bovine serum and 1X GlutaMAX. All cell culture reagents were purchased from Thermo Fisher Scientific.

4.4.3 Transfection

HeLa cells

HeLa cells were transfected using Lipofectamine LTX. 2×10^5 cells were seeded in 2 mL in a well of a 6 well plate, the day prior to transfection. The day of transfection, DNA was diluted in OptiMEM, followed by the addition of PLUS reagent, followed by a 5 minute room temperature incubation. LTX

reagent was added, followed by a 20 minute room temperature incubation, during which time the cell growth media was replaced with fresh media. Finally the transfection mixture was added to the cell media, rocking the plate to mix the added reagents. The quantities used were: 1.25 µg plasmid DNA; 1.25 µl PLUS reagent; 3.75 µl LTX reagent; 500 µl OptiMEM. (All reagents from Thermo Fisher Scientific).

293T cells

293T cells were transfected using TransIT.

For GZB-DTA and GZB-HTK experiments 1/32 of a confluent 10 cm plate was seeded in 2 ml media in one well of a 6 well plate, the day prior to transfection. The day of transfection, DNA and TransIT were both diluted in OptiMEM, followed by a five minute room temperature incubation. The two mixtures were combined, vortexed and incubated for 30 minutes at room temperature. Finally, the transfection mixture was added to the cell media, rocking the plate to mix the added reagents. The quantities used were: 1.25 µg plasmid DNA; 7.25 µl TransIT reagent; 62.5 µl OptiMEM. (TransIT purchased from MirusBio and OptiMEM from Thermo Fisher Scientific).

For GZB-NFSA experiments 1.5×10^5 cells were seeded in 500 µL in a well of a 24 well plate, the day prior to transfection. The same protocol was used as above, with the following quantities: 0.5 µg plasmid DNA; 1.5 µl TransIT reagent; 25 µl OptiMEM.

MCF-7 cells

MCF-7 cells were transfected using Lipofectamine 3000. 1/2 of a confluent 10 cm plate was seeded in 10 mL in a 10 cm plate, the day prior to transfection. DNA and P3000 reagent were diluted in OptiMEM, as well as L3000 reagent separately. These mixtures were combined, incubated at room temperature for 10 minutes, and then added to the cell media, rocking the plate to mix the added reagents. The quantities used were: 14 µg plasmid DNA; 27.5 µl P3000 reagent; 41 µl L3000 reagent; 500 µl OptiMEM. (All reagents from Thermo Fisher Scientific).

YT cells

YT cells were electroporated using the Neon system (Thermo Fisher Scientific), using the 100 µl tip. 6×10^6 cells were washed once in PBS, and resuspended in Buffer R along with 20 µg plasmid DNA, in a final volume of 110 µl. The extra volume ensures no bubbles are generated in aspirating the cell mixture into the electroporation tip. Critically the plasmid DNA must be of a concentration of at least 1 µg/µl, and it must be prepared using an endotoxin free method. The quality of the plasmid prep greatly influences the electroporation efficiency as well as the post-electroporation viability. The apparatus was prepared as in the manufacturer's manual, using the E2 electrolytic buffer. The electroporation conditions were 3×10 ms pulses at 1250 V. The electroporated cells are then immediately added to 5 ml media spread across two wells of a 6 well plate.

4.4.4 Flow cytometry

Cells were harvested (via trypsinization if adherent), and resuspended in PBS supplemented with 10% complete media and 1 µg/ml of either propidium iodide (PI) (Thermo Fisher Scientific) or DAPI (Sigma) as viability stains. If cells were to be sorted, they were passed through a 35 µm nylon filter (BD Falcon). Cells were kept on ice and then analyzed on a BD Fortessa II, or sorted on either a BD Aria III or Fusion. For sorting, cells were sorted into complete media. In all flow cytometry experiments two initial gating steps were used. Debris was excluded by excluding cells at the bottom left corner of a PI vs FSC-A (forward scatter area) gate. Doublets were excluded using a hierarchical gating scheme: all cells with a wider pulse width signal were excluded first in FSC-W vs FSC-H (forward scatter width vs height) and then SSC-W vs SSC-H (side scatter width vs height).

4.4.5 Metabolic activity assay

Cell viability was assessed using PrestoBlue (Thermo Fisher Scientific cat # A13261), which is a resazurin based fluorescent metabolic activity assay. It is provided at a 10X concentration. Cell viability was measured at 1X PrestoBlue concentration in complete growth media in black walled, flat bottom, optically clear 96-well plates (BD Falcon cat # 353219). The fluorescence of each well was measured using a Tecan Safire2 plate reader. The data was acquired with the plate lid removed, with four reads per well, and a gain setting empirically determined by the instrument for each plate.

4.4.6 Direct toxin expression experiments

GZB-HTK in 293T cells

293T cells were transfected as above, and then the media was refreshed on all cells 24 hours after transfection. 5 µM ganciclovir was added at this time to appropriate wells, as was 70 ng/ml colcemid (Thermo Fisher Scientific cat # 15210-040) to control wells. 48 hours after transfection, cells were harvested via trypsinization, with all cells retained from supernatant, PBS wash and trypsinized cells. Cells were reseeded 1/8 in fresh media in 6 well plates and cultured for another 72 hours, with ganciclovir and colcemid supplemented where appropriate. 5 days after transfection, cells were harvested as above, and 1/80 of this cell suspension was reseeded in 100 µl media. These plates were incubated at 37 °C for 2 hours, and then 10 µl PrestoBlue was added to each well, followed by another incubation at 37 °C for 2 hours, followed by acquisition as above.

GZB-DTA in 293T cells

The GZB-DTA experiments were conducted as above, except that no ganciclovir was added, and the cells were assayed 48 hours after transfection, at which point the cells were harvested and prepared exactly as at the 5 day timepoint for the GZB-HTK experiments.

GZB-NFSA AND GZB-NFSB in 293T cells

The GZB-NFSA and GZB-NFSB experiments were conducted as for GZB-HTK, with the following differences. 48 hours after transfection, CB1954 was added at varying doses, with the DMSO concentration kept constant at 0.1%. 5 days after transfection, cells were harvested and prepared as above, with 1/10 of the cell suspension reseeded into 96 well plates for the viability assay.

GZB-NFSA and GZB-HTK in MCF-7 cells

MCF-7 cells were transfected as above, and 48 hours after transfection, GFP+ PI- cells were sorted directly into 100 μ l media in black walled, flat bottom, optically clear 96-well plates (BD Falcon). The media in the plates was supplemented with CB1954 or DMSO as a vehicle control prior to the sort. The DMSO concentration was kept constant at 0.1%. 72 hours after transfection the media was changed on all wells and replaced with fresh media. 9 days after transfection, 11 μ l PrestoBlue and 9 μ l media (20 μ l total volume) was added to all wells. (This was done to pipette a larger volume to minimize pipetting error.) The plates were incubated for one hour at 37 °C and then read as above.

GZB-HTK experiments were performed according to the same protocol, except that ganciclovir was used and the media was not changed at the 72 hour timepoint.

4.4.7 Cell labeling

Cells were fluorescently labeled with CFSE, Cell Proliferation Dye eFluor 450 (eF450), or Cell Proliferation Dye eFluor 670 (eF670) (all from eBioscience) following the manufacturer's protocol, except that only 1 PBS wash prior to labeling was done and only 1 media wash after labeling was done.

4.4.8 Cytotoxic lymphocyte co-culture experiments

Unless otherwise noted, 4×10^5 YT effector cells were combined with 1×10^5 target cells (either MCF-7 or 721) at a 4:1 effector:target (E:T) ratio in a final volume of 500 μ l YT media in 5 ml polystyrene round-bottom tubes (BD Falcon). The cell suspension was gently pelleted by spinning it at 200g for 15 seconds. The tubes were then incubated at 37 °C for 4 hours, and then prepared for flow cytometry or FACS sorting as above.

4.4.9 Isolation of YT Targeted cells using the FRET reporter

48 hours after transfection target cells expressing the granzyme B CY-FRET reporter were FACS sorted. The gating strategy was to first select YFP+ cells, and then select cells from a tight diagonal band in the CY-FRET (405 nm excitation; 525 nm emission) vs CFP gate. This latter step is to ensure a tight, clear, homogeneous FRET signal. These cells were then used in co-culture experiments with YT effectors. After the co-culture the cells were either analyzed for granzyme B status (via the FRET shift, as discussed in the main text) on a flow cytometer, or sorted and cultured for long-term survival experiments. In the latter case, target cells were identified as YFP+ (to differentiate from effectors). A

triangular FRET+ gate was set immediately below the diagonal line exhibited by targets cultured in the absence of effectors in the CY-FRET vs CFP gate (as shown in the main text). FRET+ cells are those that are shifted down into the triangular gate.

4.4.10 MCF-7 cell characterization

YT killing

MCF-7s expressing the FRET reporter were transfected, FACS sorted and co-cultured with YTs as above. After co-culture 2×10^3 PI-FRET+ or PI-FRET- cells were sorted into 100 μ l media in a well of a 96 well plate. These cells were cultured for 7 days, and then cell viability assessed using PrestoBlue as above.

YT E:T ratio dose response

MCF-7s were labeled with eF670 and then co-cultured with YTs as above. For all E:T ratios, 1×10^5 targets were used, and YT cell numbers were adjusted according to the E:T ratio. After the co-culture 1×10^3 viable targets (DAPI-eF670+) were sorted into 100 μ l media in a well of a 96 well plate. These cells were cultured for 7 days, and then cell viability assessed using PrestoBlue as above.

4.4.11 Lymphocyte resistance experiments

Small molecule inhibition

Target cells were CFSE labeled and then pretreated with DCI and Q-VD-OPh, both 10 μ M, for 90 minutes at 37 °C. The cells were then co-cultured with YT effectors, with the drug concentrations kept constant in the co-cultures. The cells were then either immediately analyzed by quantifying PI staining via flow cytometry, or isolated via FACS sorting. In the latter case 1×10^3 PI-FRET+ or PI-FRET- cells were sorted into 100 μ l media in a well of a 96 well plate. The media was presupplemented with the appropriate drug or vehicle control. These cells were cultured for 7 days, and then cell viability assessed using PrestoBlue as above.

Inhibitor of apoptosis gene overexpression

MCF-7s expressing the IAP-FRET constructs were transfected, FACS sorted and co-cultured with YTs above. After the co-culture 1×10^3 PI-FRET+ or PI-FRET- cells were sorted into 100 μ l media in a well of a 96 well plate. These cells were cultured for 7 days, and then cell viability assessed using PrestoBlue as above.

4.4.12 Enhancement of YT-Indy killing experiments

YTs were transfected as above, and enriched for PI-GFP+ cells via FACS sorting as above. Target MCF-7s were labeled with eF450 as above. Both cell populations were cultured separately overnight. Effector and target cell populations were then co-cultured as above. For all E:T ratios, 1×10^5 targets were used, and YT cell numbers were adjusted according to the E:T ratio. After the co-culture 1×10^3 viable targets (DAPI-GFP-eF670+) were sorted into 100 μ l media in a well of a 96 well plate. The media was presupplemented with either 10 μ M CB1954 or 0.1% DMSO as a vehicle control. The following day, the media was changed to drug free media for all wells. These cells were cultured for a further 6 days, and then cell viability assessed using PrestoBlue as above.

4.4.13 Drug reconstitution

All drugs were prepared fresh prior to experiments.

Ganciclovir

Ganciclovir was purchased from Invivogen (cat #sud-gcv). A 1000X stock solution was reconstituted by first combining ganciclovir powder in PBS, followed by vortexing for one minute. 10 μ l 12 M hydrochloric acid was then added per 1 ml PBS. The solution was then vortexed until the solid is completely dissolved. A vehicle control was also prepared in parallel.

CB1954

CB1954 was purchased from Sigma (cat # C2235). A 1000X stock solution was reconstituted in anhydrous DMSO.

Dichloroisocoumarin and Q-VD-Oph

Both reagents were purchased from Sigma (cat # D7910 and SMC0063 respectively), and reconstituted at 10 mM in anhydrous DMSO.

4.4.14 Statistical analysis

All analysis was done in R. For all experiments except those in Section 4.2.6, comparison of means was done using a single factor ANOVA (using the `aoV` command in R with the functional form `Mean ~ Treatment`) to test the null hypothesis that all means were sampled from the same distribution, followed by a post hoc Tukey's HSD test (`TukeyHSD` in R) to test the pairwise difference between all means in the set. The p-values annotated on all figures are the adjusted p-values resulting from the Tukey's HSD test.

The following approach was used to analyze the dose response curves of Section 4.2.6. Media background was first subtracted from raw fluorescent values, which were then normalized to readings from wells with no effectors. This created a dataset in which the MCF-7 viability ranged from 0

to 1. Subtracting these values from 1 converted the viabilities to cell death. The E:T ratios were then log transformed, giving a dataset of normalized viability vs. $\log(\text{E:T})$, for each of the 4 sample types. These sigmoid dose response curves were then fit with a logistic function of the form $y(x) = (1 + e^{m(EC_{50}-x)})^{-1}$, with y equal to the normalized MCF-7 cell death, and x equal to the log of the E:T ratio. The fit was done separately for each sample, using a nonlinear least squares method (the `nls` function in R). These fits produced estimates for EC_{50} for each sample and the standard error of these estimates. I then used a random sampling algorithm (specifically the `sim` values from the `predictNLS` function in the `propagate` package) to generate the 95% confidence bands for fitted dose response curves.

To test if the predicted EC_{50} s were significantly different I conducted a single factor ANOVA. I simulated observations of each EC_{50} by sampling a normal distribution with a mean of the EC_{50} and standard deviation of the standard error from the non-linear least squares fit— that is I sampled the sampling distribution of the mean. I sampled either $N = 2$ observations of each EC_{50} (the number of technical replicates for each dose response curve), or $N = 12$ (the degrees of freedom from the least squares fit). I then conducted a single factor ANOVA using these simulated observations of the EC_{50} s (using the `avov` command in R with the functional form `EC50 ~ Sample`). For both $N = 2$ and $N = 12$ observations, the null hypothesis that the EC_{50} s were sampled from the same distribution was rejected ($p = 3.6 \times 10^{-5}$ and $p = 2 \times 10^{-16}$ respectively). To be maximally conservative, I proceeded with the lower powered $N = 2$ set of simulated observations, and conducted a post hoc Tukey's HSD test (`TukeyHSD` in R) to test the pairwise difference between all EC_{50} s. These were all significantly different, with a minimum p-value of 0.011. It is these p-values that are reported in Table 4.1.

Finally to plot the dose response data and fits, I back transformed the $\log(\text{E:T})$ ratios, and the fitted EC_{50} s, by taking their exponential. To convert the standard error of the fitted parameters I used the following formula (derived using standard propagation of uncertainty: $\sigma_{EC_{50}} = |EC_{50} \log(EC_{50}) \sigma_{\log EC_{50}}|$

Chapter 5

Discussion, conclusions and future directions

Cellular therapeutics are likely to be a key medical intervention in the coming years. The ability to genetically modify cells raises the possibility of harnessing their diverse array of molecular function. To realize this potentially fully, efforts across a range of disciplines are underway to re-engineer biological systems into functional modules. Combining several of these modules in a cellular chassis has the possibility of offering incredibly useful cellular devices.

In this thesis I sought to contribute to this effort by repurposing the granzyme-perforin pathway as a delivery module for use in cellular therapeutics. This would enable cell-to-cell delivery of a therapeutic payload, from a prepositioned secretory lytic granule in a delivery lymphocyte, through perforin pores and into a target cell. When combined with other modules such as receptor targeting and prosthetic networks of control logic this could be used to deliver a large range of therapeutics throughout the body.

Towards this goal, in Chapter 2 I developed a computational biophysical model of the immunological synapse between a cytotoxic lymphocyte and its target. These results had one very important implication for this project: namely that simple diffusion is a perfectly plausible mechanism by which granzyme transits perforin pores. Receptor mediated endocytosis or granzyme-perforin interactions are not required, as have been suggested periodically over the last several decades. This is a crucial insight as it suggests that therapeutics that are present in the immunological synapse have the potential to transit perforin pores and enter the target cell. These results suggested that the problem of cell-to-cell delivery might be reducible to the problem of developing a method for releasing therapeutics into the synapse upon target cell recognition.

Two other biologically important insights were suggested by my model. First, the various adhesion and signaling molecules in the synapse create a crowded environment that is critical for lymphocyte killing of target cells, as without this crowding virtually all granzyme and perforin escape the synapse. Second, even in the presence of this crowding, substantial amounts of granzyme and perforin still

escape the synapse. I propose that the mechanism by which target cell specificity is maintained is that the requirement for high local concentrations of perforin (for pore formation) and granzyme (for cell entry) in the same place and time acts as a strong bimolecular filter. In the synapse, these conditions are met and granzyme enters the target cell; in the case of escaped granzyme or perforin, they both dilute rapidly and so those conditions are not met and thus granzyme does not enter bystander cells. These findings challenge the concept of a tight seal at the edges of the immunological synapse, which historically has been accepted as a given [256, 257], even though experimental evidence is mixed and conflicting [152, 183].

I then attempted to implement granzyme-perforin delivery experimentally in Chapter 3. To do this I focused on granzyme B as a molecular chaperone for inserting and trafficking a payload into and through the pathway. I designed a set of granzyme B derived tags and fused them to mCherry as a model payload. I screened these tags by assessing if they colocalized with lytic granules, using confocal microscopy. These experiments had interesting biological implications, in that they suggested that the domains of granzyme B that are responsible for directing it from the endoplasmic reticulum to the lytic granules are not contiguous in amino acid sequence space, and potentially are not localized to a single motif in the tertiary structure either. They also indicated that two candidate granzyme chaperones were promising. I moved on to test the capacity of these chaperones to fully transfer the payload from an effector to a target cell and found that full length granzyme fusions were indeed capable of this. I confirmed these results at the protein level.

Having demonstrated that granzyme-perforin mediated delivery was at least possible in principle, I attempted to use it to enhance lymphocyte killing via delivery of additional toxins to tumour cells in Chapter 4. I generated a variety of granzyme toxin fusions and found that all retained their activity, albeit with varying attenuation in potency. Efforts to develop truly lymphocyte resistant cells were ultimately unsuccessful. In light of these results, I attempted to still demonstrate enhanced killing using an effector dose response curve, with the aim of achieving a modest, but measurable increased, effect. Using an *E. coli* derived toxin/prodrug system, I did show that effector cells expressing the toxin fusion induced greater cell death than wild type effector cells. However the effect sizes were modest, and the relative contribution of the prodrug itself and the toxin fusions remains to be fully determined.

In the remainder of this chapter, I discuss some of the issues and outstanding problems encountered in this work, as well as ways these might be addressed. I close with some broader thoughts about engineering cellular therapeutics.

5.1 The utility of the granzyme-perforin pathway as a delivery system

The balance of evidence for the broad scale utility of granzyme-perforin mediated delivery presented in this thesis is equivocal. While I have confirmed that granzyme mCherry fusion proteins are transferred to target cells, it is only in a single effector and target cell line. Certainly the efficiency of this system is not perfect: in my proof-of-principle experiments I observed substantial background transfer of

unfused mCherry. However, there is always background in any biological system. In particular in this case the experiment was conducted in an artificial co-culture system in which there is ample time for spontaneous absorption of the mCherry protein from the media, which in turn might be present due to spontaneous effector cell lysis. When I attempted to use this system with a toxin payload, the resulting enhancement in cell killing was moderate, but measurable. I feel that the modest signal-to-noise ratio achieved in this work is not unreasonable: many mature technologies leveraging a new concept start out with poor efficiency. Subsequent iterative engineering can greatly improve the efficacy, robustness, broad scale applicability and so on.

A major unsolved problem in using the granzyme-perforin pathway as a delivery system is the requirement for the fusion proteins to be loaded into, and spend significant sequestered in, lytic granules. These granules are acidic, and contain a variety of proteases. Thus for every new protein payload, there is a chance that it will be substantially degraded, or at least decoupled from the granzyme chaperone. I have observed varying amounts of fusion protein breakdown for all fusion proteins tested in YT-Indys. In early experiments I observed considerable breakdown in another natural killer cell line (NK-92MI). This context dependency, which is unknowable for each protein until it is empirically tested, is a major drawback of this approach.

To address this problem, I have considered alternative methods for secretion into the immunological synapse. This approach is supported by the computational results presented here that indicate simple diffusion through perforin pores can transfer of a molecular species from the immunological synapse into the target cell. If this were true, and it has not yet been demonstrated experimentally, then the problem of developing a lymphocyte based delivery system becomes reduced to developing a secretory system that is activated by lymphocyte recognition of a target cell. In principle this could be achieved by placing the gene coding for the desired payload under control of a TCR response promoter, for example the Nuclear Factor of Activated T-cells (NFAT) promoter. If the gene contained a secretory signal sequence at its 5' end (for example the commonly used IL-2 or IgG sequences) TCR or CAR recognition of a target cell would initiate synthesis and then secretion of the payload. The problem with this approach is that by the time the protein was synthesized, the lymphocyte would have disengaged from the target cell. Thus, the secretion would not occur into the synapse: the cell-to-cell nature of the system would be lost. The result of this secretion would depend on the context. If the lymphocyte was in a target rich environment, it is possible that secretion would occur as the lymphocyte formed a synapse with another target, thus providing perforin pores the payload could access the target cell. In the case of a lymphocyte in a tumour this perhaps would be tolerable, since the secreted therapeutic would still target cells in a tumour. In other applications where cell-specific targeting is required, this could cause unacceptable adverse events.

This illustrates the main advantage of using the granzyme-perforin pathway: granules and their contents are already synthesized and prepositioned, ready for release. Therefore, to build a system that secretes a therapeutic into the synapse on the appropriate time scales, the therapeutic must be prepositioned as well. The mechanisms by which TCR signaling initiates lytic granule release are somewhat

understood, mainly involving MTOC polarization to the synapse followed by lytic granule polarization along microtubules. At the synapse the LGs first dock to the membrane, prime by undocking but remaining tethered, fuse to the membrane and finally exocytose their contents [129]. While it might be theoretically possible to insert a presynthesized payload into this chain of events at a late stage (and thereby avoid lytic granule degradation), there is certainly no obvious, tractable approach to interface with this complex machinery.

More broadly, a general consideration of the interaction between the payload and the delivery cell is important. As discussed above, lytic granule sequestration can be detrimental to the fusion protein integrity and activity. But the payload could easily be harmful to the cellular chassis as well. An obvious case is that of delivery of toxins. In my work here, I avoided this difficulty by using toxin prodrug systems, in which the delivery lymphocytes would not be affected until the prodrug were administered. Importantly, prodrug dosing would eliminate the delivery lymphocytes, which depending on the application might or might not be appropriate. But in the general case, such two component systems are unlikely to consistently be available. There are two ways in which this issue could be addressed or at least mitigated. The first is via sequestration: either in lytic granules, or another secretory vesicle. This could well be sufficient to prevent any major activity of the payload in the delivery cell. However, it reintroduces the problem of payload degradation. The second approach would be to engineer the delivery lymphocyte with additional components that protect it from the payload. For example, in the case of diphtheria toxin (which exerts its toxic effect by inhibiting elongation factor 2), overexpression of a mutant form of elongation factor 2 that is resistant to diphtheria toxin inhibition can protect cells from concurrent diphtheria toxin expression [258]. Again, this approach is certain to not be generally applicable, as a gene encodable factor that protects the delivery cell will frequently be unavailable. Furthermore, as increasing numbers of modifications are made to a cellular chassis, there is a significant metabolic load placed on the cell, and the potential for unexpected interactions between the engineered components increases with their number. This combinatorial increase in potential interactions highlights the need for orthogonal parts as cellular therapeutics increase in complexity.

Finally, a discussion of the cellular chassis itself is required. Exploiting the granzyme perforin pathway requires the use of a cytotoxic lymphocyte, as they are the cells that contain the pathway. While it is one of the main effector mechanisms of cytotoxic lymphocytes, it is not the only one. Death receptors and cytokines are other prominent examples. In order for this system to be generally applicable and deliver a range of payloads, and not simply toxins, these mechanisms would have to be eliminated. This is in principle achievable via either RNAi attenuation, or gene knockout. Two factors make this a significant challenge. The first is elucidating exactly which genes must be silenced in order to abrogate lymphocyte cytotoxicity. Regardless of their exact nature, it is likely to be more than five, possibly more than ten. This is certainly a substantial undertaking, but given the multiplex capabilities of the CRISPR/Cas9 system, it is not impossible. However, when considering the feasibility of this approach, it is important to also consider that in a clinical application, whatever modifications are made to the cellular chassis are likely to be made in primary cells, at least in the near term. (While the

use of immortalized cell lines has been reported clinically [259], their widespread use seems unlikely from a regulatory standpoint due to the risks of oncogenesis or otherwise uncontrolled proliferation. In the long term, standardized starting populations of heavily modified precursor cells that are then customized to a final personalized cellular product are conceivable, but this is a distant goal.) Germline modification of primary cells is a substantially greater challenge than for cell-lines, and indeed as-of-yet unproven: methods for simultaneously isolating, in a non-destructive manner, a population of cells that uniformly have a bi-allelic knockout of five to ten genes do not currently exist. Doing so and recovering meaningful, clinically useable numbers of cells would likely be harder still. A potential, speculative solution to this problem would be to reconstitute the granzyme-perforin pathway in a non-cytotoxic cellular chassis. While this sounds ambitious in the extreme, an analogous feat has been achieved for the initial steps of T-cell signaling in 293T cells [260].

It is clear that there are many large impediments to using the granzyme-perforin pathway and cytotoxic lymphocytes as a general delivery system. However, the core functionality that it provides, precise, cell-to-cell delivery, is sufficiently attractive, and my results so far sufficiently encouraging, that it is worth further pursuing this goal.

5.2 Further efforts to demonstrate the toxin-mediated enhancement of lymphocyte killing

Due to the limitations discussed above surrounding the cytotoxic nature of the delivery chassis, I view delivery of toxins to cancer cells in the context of adoptive cell therapy as the most likely potential near term application. This eliminates the need for chassis attenuation, and my results suggested that this approach may be feasible, but were not definitive. There are several avenues that could be pursued to increase the magnitude of the toxin-mediated enhancement of lymphocyte killing of target cells, in order to make such a determination.

5.2.1 Development of lymphocyte resistant target cells

First and foremost, a target cell population that is near or fully resistant to YF-Indy would greatly increase the potential observable enhancement of lymphocyte killing via toxin delivery. In the experiments presented in this thesis, the maximum theoretically attainable enhancement in killing was quite small, and was dependent on a bystander effect. Based on initial experiments it was known that any directly targeted cell would die, thus greatly increasing the baseline level of cell death, regardless of toxin transfer. If a target cell population were available in which YF-targeted cells survived, this would decrease that baseline, thereby increasing the maximum enhancement in killing that might be observed. Furthermore, it would expand the suite of potential toxins that could be employed, since the choice of toxins would no longer be confined to those that can produce a bystander effect.

I attempted to render cells apoptosis and lymphocyte resistant by inhibiting apoptotic cascades. A more direct approach would be to directly eliminate key mediators of that cascade. Using RNAi or

CRISPR/Cas9 to knock out the key executioner caspases 3 and 7 would be worth pursuing, especially in MCF-7s which are already caspase 3 deficient [240]. Additional inhibitor of apoptosis genes could also be tested, for example the BCL family [220], as well as the granzyme inhibitor PI-9 [232]. Finally, these approaches could be combined combinatorially. While this would be a substantial effort, it is also probably the most likely to succeed, as each component (caspase knockout or attenuation, caspase inhibition, and granzyme inhibition) would work in parallel, absorbing whatever leakage or overflow death signals that overcome the inhibition from the other components.

In the pursuit of fully lymphocyte cell lines, two points are worth considering. First, it is possible that such a hypothetical resistant cell population might also be resistant to additional toxins. However, this would be testable by direct expression of these toxins in the target cells. Second, if such a model system cannot be generated, it is worth considering that it may not be biologically relevant. Perhaps true lymphocyte resistance is not a characteristic of tumour cells. While there is a body of literature on the subject (Section 4.1), there is a far greater focus on the mechanisms by which tumours escape lymphocyte recognition and prevent lymphocyte adhesion to tumour cells [199]. Perhaps this is the greater issue. The immunosuppressive environment is also unquestionably an important factor in this regard [261], one which is not modeled in my current experimental design. Therefore, perhaps it is unrealistic to expect a target cell to survive an extended period of repeated attack in a completely isolated, artificial co-culture environment. Perhaps a more realistic model might be found in a tumour mouse model which might more fully capture the complexities of an actual tumour, including the tumour microenvironment. While this approach might allow for the observation of much more compelling effect sizes, it is hard to justify moving into an animal model system without convincing evidence of efficacy *ex vivo*, even if that evidence is harder to obtain. In closing, it is important to emphasize that this is not to say that lymphocyte delivery of toxins to tumour cells, if achieved, would not be therapeutically relevant. Rather, demonstrating its efficacy might require a more representative model system.

5.2.2 Improving fusion protein granule loading and delivery

Increasing enhancement of effector killing could also be achieved by addressing the other side of the problem: increasing the amount of toxins that are delivered to target cells. In this regard there are two strategies that would be worth immediately pursuing. The first is generating effector cells that express toxin fusion proteins from the endogenous granzyme locus. This could be done by using a CRISPR/Cas9 system to introduce a double stranded break at the 3' end of the granzyme B coding sequence, and supplying a suitable repair template that inserts a C-terminal toxin as a fusion protein. As is described in the Appendix, I have already done this with an mCherry fusion protein as a proof-of-principle, so I am confident that this step is feasible. A fusion protein that is expressed from the endogenous locus might increase the amount of fusion protein that is loaded into lytic granules, as compared to the amount loaded in cells where the fusion protein is being expressed ectopically. The main reason for this hypothesis is the possibility of a competition effect: namely that granzyme produced from its own locus is loaded more efficiently into lytic granules. This could be due to expression that is coordinated

with cell division or degranulation (when granules need to be populated or repopulated respectively). Or it could be that granzyme fusion proteins load less efficiently than does unfused granzyme. If this were the case, then expressing endogenous fusion proteins would essentially eliminate the competition posed by the unfused granzyme. In this case a biallelic modification of the granzyme locus would be necessary. Finally, in the case of ectopic overexpression, it is almost certain that granule loading is saturated. Excess fusion proteins are likely secreted, which *in vivo* could result in unwanted toxin activity at off-target sites. Expression from the endogenous granzyme promoter would likely greatly decrease this background.

A second approach worth considering would be the development of a minimal chaperone or tag that was still sufficient for delivery via perforin pores, but had improved stability and persistence in the harsh environment of lytic granules. Having observed appreciable fusion protein degradation across several fusion proteins and lymphocyte cell types, I initiated a series of mass spectrometry experiments to investigate if there were any regions in the fusion proteins that were more susceptible to breakdown. The data is shown in the Appendix, and indicates that the region surrounding and including the glycine-serine linker could be modified to minimize fusion protein degradation. In terms of developing a superior chaperone, efforts to map the critical domains of granzyme B would be informative, as would a similar investigation of other granzymes and lytic granule constituents. Finally, synthetic tags such as the glycosylation independent lysosomal tag (GILT) domain could also be investigated [262].

5.3 Future directions for granzyme-perforin delivery systems

The concept of granzyme-perforin mediated delivery is still in its infancy, and its practical utility remains to be robustly demonstrated. This makes any in depth discussion of its possible extensions, additions and applications of increased complexity premature. However, a certain amount of thought and attention has been invested in preparation for such an eventuality, and a brief summary is presented here.

Addition of receptor mediated targeting

For simplicity, natural killer cell lines with known target cell lines were used here as the most basic model system. As the approach is developed a return to primary T-cells would likely be appropriate. In both primary and immortalized cells, and in both T-cells and NK cells, targeting could be achieved via the addition of either a synthetic TCR, or a CAR, as discussed in the Introduction.

Addition of suicide switches

Especially given a recent series of fatalities in a CAR-T trial, the ability to rapidly and safely eliminate the administered cell product in the event of unforeseen toxicity will be a necessity moving forward. There are already clinically validated methods for doing this, such as the iCasp system [263]. Another promising approach is a dual CD20-CD34 marker which, when expressed on the surface of adminis-

tered CAR-T cells, allows for selective depletion of the administered cells using rituximab, a clinically approved anti-CD20 antibody [102]. However, in this latter case, whether the depletion is sufficiently rapid for clinical use, and its general clinical utility remain to be demonstrated in clinical trials.

Extension to non-protein payloads

All designs and approaches that have been presented here are limited to protein or peptide payloads that can be fused to a molecular chaperone that inserts the payload into the granzyme perforin pathway. However, it is possible to envision ways to extend this approach to nucleic acids and small molecules. In the case of small molecules, DNA aptamers against the protein chaperone (for example granzyme B), as well as the small molecule itself, could be generated using SELEX [264–266]. These could be joined using a DNA linker sequence, to create a bi-specific aptamer [267, 268]. In this way it would be in theory possible to couple the small molecule to the chaperone. In order to deliver large DNA constructs, a similar strategy could be pursued, using the bacterial plasmid segregation machinery [269]. The protein chaperone could be fused to a centromere binding protein (for example ParR). The DNA payload would then be cloned into a plasmid, which would contain a centromere-like DNA site to which the matching centromere binding protein would bind (for ParR this would be *parC*). Again this would couple the DNA payload to the protein chaperone. The first and most obvious unknown is whether these couplings would survive trafficking through the granzyme-perforin pathway, especially sequestration in the lytic granule. While these ideas are highly speculative, the delivery of whole nucleic acid circuits in particular is especially appealing, and perhaps eventually even entire synthetic genomes [270, 271].

Neutralization of the delivery chassis

For any application beyond the delivery of a payload to a target cell population for which death is the therapeutic objective, the cytotoxicity of the delivery lymphocyte chassis would need to be eliminated. Otherwise, whatever therapeutic or diagnostic effects that were achieved by payload delivery would be countered by the lymphocyte cytotoxicity. Here gene attenuation or knock out would both be relevant. Both an RNAi approach as well as CRISPR/Cas9 approaches, either actual knockout using Cas9, or knockdown using dCas9 fused to a repressor [272] could be considered. There would be two main challenges in this undertaking. First, it would be necessary to delineate exactly which genes would need to be targeted in order to eliminate lymphocyte cytotoxicity, and to what degree in the case of gene knockdown. A minimal set of genes would likely be granzymes A and B, Fas ligand and TNF- α , but others could well be required. Second, simultaneous, biallelic knockout of multiple genes is a challenging task. Furthermore, additional modifications of the cellular chassis are required for this application, such as the insertion of the payload. Finally, the efficiency of the entire process must be sufficiently high so as to recover a cell population of a viability, size and purity that is clinically usable. In this regard, attenuation is perhaps more feasible, since there is no requirement for isolation of cells with inactive target genes, but rather simply cells that have been successfully modified to express the at-

tenuating constructs (either RNAi based or Cas9 based). Using an immortalized cell line as the cellular chassis would make this undertaking slightly more feasible, since iterative modification steps would be possible. However, clinical use of immortalized cell lines would raise substantial regulatory concerns, although there is precedent in clinical trials [273]. Thus, it is likely that chassis neutralization will not be feasible without substantial improvements in genome editing methodology, which fortunately are not unlikely.

A tool for screening CARs for on-target off-tissue activity

A significant outstanding problem in the development pipeline of CARs is that of off-target activity. While *in vitro* cross-reactivity testing using ELISPOT is helpful, it cannot capture all eventualities [274, 275]. The payload delivery system presented here might be used as a unique screening tool for CAR off-target activity in animal models. I would envision generating delivery lymphocytes expressing a CAR under development, and a Cre recombinase payload. These cells would then be adoptively transferred to a mouse that had a transgenic *lacZ* Cre recombinase reporter germline modification [276]. Any tissue targeted by the CAR-T cells would also receive the Cre recombinase payload, triggering expression of the *lacZ* reporter. Following necropsy, standard X-gal staining would provide valuable information about which tissues were targeted, and these tissues could be further investigated. An important caveat to this approach is that the time for *lacZ* expression would have to be shorter than the time for target cell death, which is perhaps unlikely—making this another application for which a neutralized delivery cell chassis would be very useful. In this event, other payloads could be used which do not rely on target cell transcription for reporter activity. Epitope tags might be suitable payloads in this case, as tissue sections could be analyzed using standard immunohistochemistry techniques, to ascertain the same information. Using such a payload would not address the related issue of balancing the timing of necropsy such that the CAR-T cells are adequately distributed throughout the mouse, while ensuring that targeted cells have not progressed sufficiently in their apoptosis that they cannot be identified histologically. Finally, even were these challenges are overcome, it is important to note that the antigen profile of a mouse is clearly not entirely representative of that of a human. As various forms of humanized mouse models are developed, this discrepancy may decrease [277].

5.4 Broader insights into cellular therapeutics

Unlike modern computing or aerospace, which rely on parts built from the ground up in a step-by-step and hierarchical fashion, biology is not yet an engineering discipline. Natural development of molecules, pathways, and cells has not been guided by the principles of modularity, insulation, and robustness. Rather, biological systems evolved in an entirely opportunistic manner to occupy a niche that has a specific parameter set. Therefore, these systems can be highly unstable when perturbed or placed in an environment with a different parameter space. These intricate evolutionary origins have significant consequences for the systems we attempt to engineer. It is important to recognize that

progress will be impeded by the complexity of molecular and cell biology, which imposes a heavy burden of noise, unpredictability, and context dependency on the enterprise of engineering cellular therapeutics [278]. Successful examples are typically the result of trial and error, and contain many accessions to the messy realities of building functional biological systems. In these cases, countless design choices have been made, which are rational and empirical, explicit and implicit, deliberate and accidental. While the delivery system presented here cannot be classified as a successful cell-based therapeutic, several useful insights have been gleaned while attempting to make it so.

5.4.1 Context dependencies in cellular therapeutics

Biological molecules and pathways are highly evolved and networked [279], very sensitive to perturbation [280], and modifying them often results in cell death or unexpected failure modes [281]. The functions of individual molecules or pathways may be overlapping, integrated, redundant, and degenerate [282]. Rather than an insulated collection of pathways with specified interactions, the cell can exhibit characteristics of a single large network of multiplexed, interacting parts. This network may exhibit behavior that is probabilistic rather than deterministic, and in some cases may lack damping or input filtering to determine output responsiveness. Locally, the physical and biochemical niche, along with stromal cells and the extracellular matrix all provide critical stimuli and mechanical cues that influence cellular, tissue, and organ level differentiation and development.

This connectivity, and sensitivity injects a substantial amount of noise and unpredictability into engineered biological systems [283–285], and gives rise to what I would call context dependency. That is, the behavior, function, and stability of engineered components are entirely dependent on the molecular and cellular milieu in which they operate, and these characteristics may vary dramatically if this context is altered. Similarly, the connectivity and function of cellular pathways and the viability of the cellular chassis itself may also be highly sensitive to the introduction of a new component. While it may simply reflect our incomplete knowledge, the complexity inherent in these biological systems is such that their behavior can verge upon chaotic [286]. Predicting the effects of simple, small modifications within or between systems is in many cases impossible.

This has practical implications in the laboratory. Minor alterations to reagents or protocols can result in outsized effects, completely unexpected catastrophic failures, or worst of all, spurious results. Assays to test, debug, and validate cellular therapeutics are imperfect: the results are often ambiguous, and can frequently be misleading such that following insertion into the cell chassis, engineered components can display partial or incomplete failure modes that can be easy to miss. These issues are compounded by the sensitivity of the cellular chassis, and its tendency to initiate apoptosis upon perturbation. This is problematic, since the time window from the first signs of cellular dysfunction to cell death can be very short indeed, which makes troubleshooting and debugging engineered cells extremely challenging.

Thus, when we attempt to repurpose specialized molecules, interconnected pathways and sensitive cells, we should not be too surprised when they reveal themselves to be brittle and subject to failure.

After all, we are attempting to tweak highly evolved systems. However, evolution does not logically imply optimization, a crucial distinction that offers a way forward, and a window in which to work.

5.4.2 A framework for cell engineering

Where does this leave us? How can we succeed in this difficult environment, and manage the issues I have discussed above, while realizing the potential of repurposing biological systems? My solution has been to adopt empiricism and pragmatism as guiding principles. A set of rules has emerged from this approach, which my supervisor and I have proposed as a framework for building robust cell-based therapeutics in the complex, networked, and sensitive mammalian cell chassis [180].

Parsimony and simplicity

The design phase should always be guided by these two core principles. Unlike other engineering disciplines, in which efforts to consider and address edge cases may yield more robust function, in this field, these extra layers of design are more likely to have unforeseen negative consequences rather than improve system stability. Avoid the common desire to 'overbuild' systems by adding excessive features for downstream and long-term functionality that are, in reality, far more likely to result in unintended, and potentially very damaging and confounding, consequences.

Reuse

Pre-existing, validated designs and components should be used wherever possible. Unless novel or improved function is required, if a component has been shown to work sufficiently well for a given purpose, it should be used in place of any alternative, untested design, even if the latter is in theory superior. This concept should be extended all the way down to the level of nucleotide sequence, and the local sequence context should be maintained if possible.

Adaptation

If a molecule, pathway or cell with new or improved function is required, two options should be considered: (i) refinement and engineering of the existing part using either rational design or evolutionary methods, or (ii) sourcing a new component from other biological systems by testing those that have similar function, an approach that might be termed panning. Crucially, in either approach, the new and improved component should be considered as untested. Its behavior should not be inferred from that of the related component.

Step-wise testing

Always test, at the molecular level, each step in the production, maturation, and function of an engineered molecule, pathway or cell. While certainly useful during pilot work or as screening tools, proxy reporter constructs or assays should not be relied upon to definitively confirm the function and behavior

of engineered pathways. Similarly, investigators should make efforts to work in the intended cellular chassis, and as much as possible avoid reliance on model cell lines that are easy to work with, but less physiologically relevant. This should continue throughout the path to clinical use, with each modification tested in relevant animal models, and increasingly, in organ-on-a chip testbeds [287], which hopefully will continue to improve our ability to avoid adverse events that were missed in pre-clinical screening [275].

Uniqueness

Each combination of cellular chassis, genetic construct, experimental protocol, hardware apparatus, and so on, contains sufficient intricacies and permutations so as to render it unique. Extreme caution should be applied when mapping information from one experimental context to another, or between two types of biological molecule, pathway or cell. This concept extends to the level of the patient receiving the cellular therapeutic. The complex interplay between the immunogenicity of the cellular therapeutic and the patient's immune system has mostly necessitated cellular therapies that are either autologous or at least HLA matched. Even if these immune constraints are able to be overcome via innovation, it is likely that the complexity and immune network of each patient will necessitate a more personalized approach when administering cellular therapeutics, compared to historical small molecule therapy.

Empirical design

The interconnected web of biological context dependency remains opaque, especially at a molecular, mechanistic level. This makes predictive design of meaningful systems challenging. It is very difficult to predict how modifications to a given pathway component will affect other components in the pathway, the pathway itself, other pathways in the network, and the cell as a whole. Thus, in most cases theoretical predictions and designs are mainly useful at high level and preliminary stages of design, and require careful validation once implemented in the laboratory.

5.4.3 Grand challenges for cell engineering

The complexity and diversity of biological function is a double edged sword. On the one hand it provides a tantalizing set of parts from which to build cellular devices. On the other, its vagaries produce context dependencies that necessitate the empirical and pragmatic approach outlined in the previous section. This is not to minimize such an approach: the progress that has been made in this project has resulted from following such principles. I would venture the same is true for most other cellular therapies. Nor is it to suggest that biological knowledge is of secondary importance to a brute force, black box, empirical approach. Quite the opposite in fact: the intricacy of biological systems makes a deep and wide knowledge essential from design through development. Currently however, often this the knowledge is context dependent, making generalization difficult. This slows the progress

of developing specific cellular devices, since each effort is largely a new enterprise, rather than building on existing devices or technology. While this could be interpreted as being dismissive of cellular engineering, it is not meant to be, and indeed what has been achieved is remarkable. Rather it is simply a reflection of the early stage at which we find ourselves in developing the field. I suggest here two grand challenges that if realized would greatly accelerate its maturation.

Orthogonal systems

Development of biological modules or systems that operate independently of their surrounding environment. This would reduce or eliminate context dependencies of individual parts which would greatly increase the robustness and predictability of the systems which are built upon those parts, thus increasing the success rate of designs. Initially such requirements would likely necessitate a highly artificial and fragile system, which would be confined to use as a testbed, for prototyping or debugging. Even this would be of great benefit: consider a wind tunnel or a breadboard. In the long run it is possible that these systems could be sufficiently strengthened for practical use.

Orthogonal systems could be developed by following two main strategies: reconstitution, or *de novo* construction. Both are already underway. The most straightforward approach to the problem is to remove it: remove the cell. Cell free systems consisting of cellular extract have been available for some time, and are mature enough that commercial kits are available. However, more recently this approach is being used for rapid prototyping, and has been used to develop novel RNA circuitry [288] which could be used for RNA-based control logic. Immunological insight motivated the reconstitution of a T-cell receptor in a 293T model system, but it also provides inspiration for further reconstitution of a functional component in a non-native cellular chassis, which might allow for a greatly expanded range of function [260].

More ambitious still are the efforts underway to expand, adapt and create novel translational machinery. Synthetic ribosomes that recognize quadruplet codons, artificial amino acids and tRNAs, and *E. coli* strains that have redundant codons collapsed onto a single codon per amino acid have all been reported [289–291]. This opens the possibility of entirely parallel protein machineries operating in a cellular chassis: the endogenous machinery performing basic homeostatic functions, while the synthetic translational system produces the added functional components. These two systems could operate independently, free of cross-talk any the associated problems for both the viability host cell chassis and the performance of the added components. Even more far reaching is the complete chemical synthesis of the *Mycoplasma mycoides* genome [292], as well as the delineation of its minimal gene set for viability [271]. Intriguingly, approximately one third of this essential gene set could not be annotated with a known function, indicating the work yet to be done to realize a truly orthogonal minimal cell.

Obviously true orthogonality is likely impossible: there will always be context dependencies. But whatever insulation from these dependencies can be achieved will almost certainly bring substantial insight and biological knowledge, accelerate development cycles, and improve the robustness of cellular products. Perhaps most importantly, it would greatly improve the predictability of engineered systems,

perhaps releasing their design from the current empirically approach.

Predictive design

Development of a method for predicting the behavior of biological systems based on a minimal set of parameters. The application of quantitative, predictive models to the design and development of products is critical across all engineering disciplines. It allows for a streamlined, cheap, and rapid design process. It also massively increases the confidence that a design validated by the predictive theory is likely to work once constructed, albeit usually with modification.

In the fields of engineering, these models and theories are for the most part based on mathematical formalism built on top of physical and chemical principles. An extension of this approach to cellular and molecular biology would be transformative. Why has this not been achieved? Is it possible?

At first glance, it seems as though it should be straightforward. No one would question that ultimately biological systems are governed by physics and chemistry. Incredibly refined, mature and accurate theories exist for all of the molecular constituents of these systems. What is more, the last few decades have also seen the development of incredible computing power with which to implement these models. In my view the problem comes down to two fundamental challenges: complexity and parameterization.

For any molecular system of any size, we as a community have the appropriate mathematical and physical machinery with which to model it. Using the formalism of molecular dynamics, and the underlying electrostatics and quantum chemistry where needed, we understand in principle all of the interactions governing the molecular processes in biology. The system of equations could be written down, and in theory integrated forward in time. The problem is that we cannot even begin to apply the appropriate theories at the length and time scales that are necessary to usefully model biological systems. There would be too many constituent parts. Even initializing the simulation in a meaningful way would be near impossible. If such a feat were achieved, despite our vast computational resources, they would be massively insufficient for the task. This is the complexity problem. Solving it seems unlikely in the near term. Even a transformative advance in computational power (and accompanying algorithms for efficiently tracking and storing the particle trajectories, energy states and so on) would still leave the problem of initialization. It is unclear how this could be accomplished either manually or in an automated fashion.

As a result of these challenges, mathematical models of biology typically make use of supra-molecular methods. That is, for the most part, forces and interactions are modeled as homogenous bulk effects that are the same for all instances of a object. Even stochastic methods, such as those employed in Chapter 2 in which finite particles are tracked, still rely on descriptions of the forces and effects at supramolecular scales. This sufficiently reduces the complexity of these models such that they generally are tractable in terms of both initialization and computational requirements.

The cost of these simplifications is that these bulk effects are described by numerical parameters, and very often the models are incredibly sensitive to these parameters. Moreover, it is difficult to

appropriately assign values to these parameters that reflect the underlying biological process. This is widely acknowledged, and is usually addressed by exploring the parameter space in an unbiased manner, thus producing predictions for a range of parameter values selected to correspond to the range of possible biological extremes. This is a valid approach, but it presupposes the existence of an actual biological data set that the model can be compared to, so as to select the appropriate parameter range before making further predictions. These sets in general are rare. Thus there is a gap between the predictive models and the experimental biological implementation. This is the parameterization problem. That is, the amount of experimental work required to calibrate a predictive model by appropriately parameterizing it is sufficient to often undermine the model's predictive utility as an engineering tool. A related problem is that in many cases, the necessary experimental tools simply do not exist to generate a biological data set that correlates well with biophysical model parameters. Here it is important to note that the effort required to generate such a data set may be well worth it in the context of scientific discovery.

It is in resolving this parameterization problem that I think progress may be made. Improvements in microscopy technology may provide a wealth of experimental, time resolved data which may greatly increase the accuracy of empirical parameters for biophysical models, and provide a broader range of validation data sets. Another fruitful approach may be a systematic effort to characterize a host of biophysical parameters of certain model systems, for example a HeLa cell. This could provide quantitatively accurate models that could be tentatively extended to other systems, and incrementally adjusted as needed. Whatever approaches are used, the parameterization problem seems tractable, if there is a desire to pursue it.

Finally, I leave open the possibility for a new, non-reductionist theory of biology that is not based in mathematical physics or chemistry. This is not to say that such principles do not apply to biology, for they do. Rather, it seems at least possible that an alternative theory could be developed in parallel. It might be one better suited to the volume of constituent parts in a biological system, the multitude of different time and length scales that are relevant, the high number of different classes of objects, and the staggering complexity and heterogeneity of those classes. In general, scientific theory is so deeply rooted in some variation of reductionist mathematical physico-chemical approaches that it is hard to even conceive of what such a theory might look like. With no justification, I conjecture that an information centric approach might be a fruitful lamppost under which to look. It has an associated formalism, but is free from obvious physical or chemical constraints. However speculative this may seem, it seems equally rigid to not at least admit the possibility that an alternative, rigorous theory of biology might exist.

Of course, this but one vision of a way forward for cellular therapeutics. Perhaps both of those challenges are impossible. Perhaps they are unnecessary. New approaches may emerge, in which high throughput screening methods will render both biological knowledge and predictive design unnecessary. Or the payoff of cellular therapeutics will be sufficiently high that bespoke devices remain the norm. Or the whole field will remain a boutique cottage industry. I think the last outcome unlikely.

The power of encoding logic in cells equipped with a variety of biological functions will be impactful, even if the eventual realization is currently unclear.

5.4.4 The future of cellular therapeutics

As the cell is the basic unit of life, it seems evident that the ideal therapeutic modality would directly engage, cell-to-cell, with a diseased cellular target. Combined with every cell's capacity for molecularly encoded logic, this one-to-one interaction would allow for real-time therapeutic decision-making and target-cell discrimination at the site of active disease. Target cell surface molecules as well as environmental variables such as acidity or oxygen content could all be considered in selecting if, when, and which target cells to treat. The ability of a cell to sequester a therapeutic molecule intracellularly while actively trafficking to disease sites for molecular delivery would maximize the therapeutic molecule's biological activity, and minimize its off-target activity. Finally, disorders not amenable to small molecule or biologic therapy might respond to cell-to-cell therapy: engagement of multiple surface receptors or elimination via phagocytosis are both possible. The age of cell-based therapeutics has arrived, and their impact will continue to expand across medicine [177]. The array of functional parts, pathways, and cellular chassis already available within our cell repertoire represents an incredible resource for building these devices, one that will not be matched by ground up synthesis for a very long time. Instead, modifying and re-purposing biological systems will yield novel functional components and new waves of cell-based therapeutics.

Bibliography

- [1] Barton, F. B. *et al.* Improvement in Outcomes of Clinical Islet Transplantation: 1999-2010. *Diabetes Care* **35**, 1436–1445 (2012). DOI 10.2337/dc12-0063. → pages
- [2] Farney, A. C., Sutherland, D. E. R. & Opara, E. C. Evolution of Islet Transplantation for the Last 30 Years. *Pancreas* **45**, 8–20 (2016). DOI 10.1097/MPA.0000000000000391. → pages
- [3] Bruni, A., Gala-Lopez, B., Pepper, A. R., Abualhassan, N. S. & Shapiro, A. J. Islet cell transplantation for the treatment of type 1 diabetes: recent advances and future challenges. *Diabetes, metabolic syndrome and obesity : targets and therapy* **7**, 211–23 (2014). DOI 10.2147/DMSO.S50789. → pages
- [4] Kimbrel, E. A. & Lanza, R. Current status of pluripotent stem cells: moving the first therapies to the clinic. *Nature Reviews Drug Discovery* **14**, 681–692 (2015). DOI 10.1038/nrd4738. → pages
- [5] Takahashi, K. & Yamanaka, S. Induction of pluripotent stem cells from mouse embryonic and adult fibroblast cultures by defined factors. *Cell* **126**, 663–76 (2006). DOI 10.1016/j.cell.2006.07.024. → pages
- [6] Wu, S. M. & Hochedlinger, K. Harnessing the potential of induced pluripotent stem cells for regenerative medicine. *Nature Cell Biology* **13**, 497–505 (2011). DOI 10.1038/ncb0511-497. → pages
- [7] Takebe, T. *et al.* Vascularized and functional human liver from an iPSC-derived organ bud transplant. *Nature* **499**, 481–484 (2013). DOI 10.1038/nature12271. → pages
- [8] Lu, T.-Y. *et al.* Repopulation of decellularized mouse heart with human induced pluripotent stem cell-derived cardiovascular progenitor cells. *Nature communications* **4**, 2307 (2013). DOI 10.1038/ncomms3307. → pages
- [9] Tsukamoto, A. *et al.* Challenging Regeneration to Transform Medicine. *STEM CELLS Translational Medicine* **5**, 1–7 (2016). DOI 10.5966/sctm.2015-0180. → pages
- [10] Naldini, L. Gene therapy returns to centre stage. *Nature* **526**, 351–360 (2015). DOI 10.1038/nature15818. → pages
- [11] Little, M.-T. & Storb, R. History of haematopoietic stem-cell transplantation. *Nature Reviews Cancer* **2**, 231–238 (2002). DOI 10.1038/nrc748. → pages
- [12] Kaufmann, K. B., Büning, H., Galy, A., Schambach, A. & Grez, M. Gene therapy on the move. *EMBO Molecular Medicine* **5**, 1642–1661 (2013). DOI 10.1002/emmm.201202287. → pages

- [13] Hacein-Bey-Abina, S. *et al.* Sustained Correction of X-Linked Severe Combined Immunodeficiency by ex Vivo Gene Therapy. *New England Journal of Medicine* **346**, 1185–1193 (2002). DOI 10.1056/NEJMoa012616. → pages
- [14] Hacein-Bey-Abina, S. *et al.* Insertional oncogenesis in 4 patients after retrovirus-mediated gene therapy of SCID-X1. *The Journal of Clinical Investigation* **118** (2008). DOI 10.1172/JCI35700.3132. → pages
- [15] Biffi, A. *et al.* Lentiviral Hematopoietic Stem Cell Gene Therapy Benefits Metachromatic Leukodystrophy. *Science* **341** (2013). → pages
- [16] Cavazzana-Calvo, M. *et al.* Transfusion independence and HMGA2 activation after gene therapy of human β -thalassaemia. *Nature* **467**, 318–322 (2010). DOI 10.1038/nature09328. → pages
- [17] Laity, J. H., Lee, B. M. & Wright, P. E. Zinc finger proteins: new insights into structural and functional diversity. *Current opinion in structural biology* **11**, 39–46 (2001). → pages
- [18] Urnov, F. D., Rebar, E. J., Holmes, M. C., Zhang, H. S. & Gregory, P. D. Genome editing with engineered zinc finger nucleases. *Nature Reviews Genetics* **11**, 636–646 (2010). DOI 10.1038/nrg2842. → pages
- [19] Joung, J. K. & Sander, J. D. TALENs: a widely applicable technology for targeted genome editing. *Nature reviews. Molecular cell biology* **14**, 49–55 (2013). DOI 10.1038/nrm3486. → pages
- [20] Sander, J. D. & Joung, J. K. CRISPR-Cas systems for editing, regulating and targeting genomes. *Nature biotechnology* **32**, 347–55 (2014). DOI 10.1038/nbt.2842. → pages
- [21] Hsu, P. D., Lander, E. S. & Zhang, F. Development and Applications of CRISPR-Cas9 for Genome Engineering. (2014). NIHMS150003. → pages
- [22] Tsai, S. Q. & Joung, J. K. Defining and improving the genome-wide specificities of CRISPR-Cas9 nucleases. *Nature Reviews Genetics* **17**, 300–312 (2016). DOI 10.1038/nrg.2016.28. → pages
- [23] Chapman, M., Warren, E. H. & Wu, C. J. Applications of next-generation sequencing to blood and marrow transplantation. *Biology of blood and marrow transplantation : journal of the American Society for Blood and Marrow Transplantation* **18**, S151–60 (2012). DOI 10.1016/j.bbmt.2011.11.011. → pages
- [24] Lieber, M. R. The mechanism of double-strand DNA break repair by the nonhomologous DNA end-joining pathway. *Annual review of biochemistry* **79**, 181–211 (2010). DOI 10.1146/annurev.biochem.052308.093131. → pages
- [25] Liang, F., Han, M., Romanienko, P. J. & Jasin, M. Homology-directed repair is a major double-strand break repair pathway in mammalian cells. *Proceedings of the National Academy of Sciences of the United States of America* **95**, 5172–7 (1998). → pages
- [26] Cox, D. B. T., Platt, R. J. & Zhang, F. Therapeutic genome editing: prospects and challenges. *Nature medicine* **21**, 121–31 (2015). DOI 10.1038/nm.3793. → pages

- [27] Lombardo, A. *et al.* Site-specific integration and tailoring of cassette design for sustainable gene transfer. *Nature Methods* **8**, 861–869 (2011). DOI 10.1038/nmeth.1674. → pages
- [28] Genovese, P. *et al.* Targeted genome editing in human repopulating haematopoietic stem cells. *Nature* **510**, 235–240 (2014). DOI 10.1038/nature13420. → pages
- [29] Tebas, P. *et al.* Gene Editing of *CCR5* in Autologous CD4 T Cells of Persons Infected with HIV. *New England Journal of Medicine* **370**, 901–910 (2014). DOI 10.1056/NEJMoal300662. → pages
- [30] Fellmann, C., Gowen, B. G., Lin, P.-C., Doudna, J. A. & Corn, J. E. Cornerstones of CRISPR-Cas in drug discovery and therapy. *Nature Reviews Drug Discovery* **16**, 89–100 (2016). DOI 10.1038/nrd.2016.238. → pages
- [31] Tsai, S. Q. *et al.* GUIDE-seq enables genome-wide profiling of off-target cleavage by CRISPR-Cas nucleases. *Nature Biotechnology* **33**, 187–197 (2014). DOI 10.1038/nbt.3117. → pages
- [32] Yang, L. *et al.* Targeted and genome-wide sequencing reveal single nucleotide variations impacting specificity of Cas9 in human stem cells. *Nature Communications* **5**, 5507 (2014). DOI 10.1038/ncomms6507. → pages
- [33] Cho, S. W. *et al.* Analysis of off-target effects of CRISPR/Cas-derived RNA-guided endonucleases and nickases. *Genome research* **24**, 132–41 (2014). DOI 10.1101/gr.162339.113. → pages
- [34] Wei, X. *et al.* Mesenchymal stem cells: a new trend for cell therapy. *Acta Pharmacologica Sinica* **34**, 747–754 (2013). DOI 10.1038/aps.2013.50. → pages
- [35] Squillaro, T., Peluso, G. & Galderisi, U. Clinical Trials With Mesenchymal Stem Cells: An Update. *Cell Transplantation* **25**, 829–848 (2016). DOI 10.3727/096368915X689622. → pages
- [36] Cheng, Z. *et al.* Targeted Migration of Mesenchymal Stem Cells Modified With CXCR4 Gene to Infarcted Myocardium Improves Cardiac Performance. *Molecular Therapy* **16**, 571–579 (2008). DOI 10.1038/sj.mt.6300374. → pages
- [37] Suresh, S. C. *et al.* Thioredoxin-1 (Trx1) engineered mesenchymal stem cell therapy increased pro-angiogenic factors, reduced fibrosis and improved heart function in the infarcted rat myocardium. *International Journal of Cardiology* **201**, 517–528 (2015). DOI 10.1016/j.ijcard.2015.08.117. → pages
- [38] van Velthoven, C. T., Braccioli, L., Willemen, H. L., Kavelaars, A. & Heijnen, C. J. Therapeutic Potential of Genetically Modified Mesenchymal Stem Cells After Neonatal Hypoxic-Ischemic Brain Damage. *Molecular Therapy* **22**, 645–654 (2014). DOI 10.1038/mt.2013.260. → pages
- [39] Niess, H. *et al.* Selective Targeting of Genetically Engineered Mesenchymal Stem Cells to Tumor Stroma Microenvironments Using Tissue-Specific Suicide Gene Expression Suppresses Growth of Hepatocellular Carcinoma. *Annals of Surgery* **254**, 767–775 (2011). DOI 10.1097/SLA.0b013e3182368c4f. → pages

- [40] Sage, E. K. *et al.* Systemic but not topical TRAIL-expressing mesenchymal stem cells reduce tumour growth in malignant mesothelioma. *Thorax* **69**, 638–647 (2014). DOI 10.1136/thoraxjnl-2013-204110. → pages
- [41] Sage, E. K., Thakrar, R. M. & Janes, S. M. Genetically modified mesenchymal stromal cells in cancer therapy. *Cytotherapy* **18**, 1435–1445 (2016). DOI 10.1016/j.jcyt.2016.09.003. → pages
- [42] Corthay, A. Does the immune system naturally protect against cancer? *Frontiers in immunology* **5**, 197 (2014). DOI 10.3389/fimmu.2014.00197. → pages
- [43] Cramer, D. W. & Finn, O. J. Epidemiologic perspective on immune-surveillance in cancer. *Current opinion in immunology* **23**, 265–71 (2011). DOI 10.1016/j.coi.2011.01.002. → pages
- [44] Muul, L. M., Spiess, P. J., Director, E. P. & Rosenberg, S. A. Identification of specific cytolytic immune responses against autologous tumor in humans bearing malignant melanoma. *Journal of immunology (Baltimore, Md. : 1950)* **138**, 989–95 (1987). → pages
- [45] Rosenberg, S. A. *et al.* Use of Tumor-Infiltrating Lymphocytes and Interleukin-2 in the Immunotherapy of Patients with Metastatic Melanoma. *New England Journal of Medicine* **319**, 1676–1680 (1988). DOI 10.1056/NEJM19881223192527. → pages
- [46] Rosenberg, S. A. *et al.* Durable Complete Responses in Heavily Pretreated Patients with Metastatic Melanoma Using T-Cell Transfer Immunotherapy. *Clinical Cancer Research* **17**, 4550–4557 (2011). DOI 10.1158/1078-0432.CCR-11-0116. → pages
- [47] Brown, S. D. *et al.* Neo-antigens predicted by tumor genome meta-analysis correlate with increased patient survival. *Genome research* **24**, 743–50 (2014). DOI 10.1101/gr.165985.113. → pages
- [48] Rooney, M. S., Shukla, S. A., Wu, C. J., Getz, G. & Hacoen, N. Molecular and genetic properties of tumors associated with local immune cytolytic activity. *Cell* **160**, 48–61 (2015). DOI 10.1016/j.cell.2014.12.033. → pages
- [49] June, C. H. Principles of adoptive T cell cancer therapy. *The Journal of Clinical Investigation* **117** (2007). DOI 10.1172/JCI31446.1204. → pages
- [50] Dudley, M. E., Wunderlich, J. R., Shelton, T. E., Even, J. & Rosenberg, S. A. Generation of tumor-infiltrating lymphocyte cultures for use in adoptive transfer therapy for melanoma patients. *Journal of immunotherapy* **26**, 332–42 (2003). → pages
- [51] Jiang, Y., Li, Y. & Zhu, B. T-cell exhaustion in the tumor microenvironment. *Cell Death and Disease* **6**, e1792 (2015). DOI 10.1038/cddis.2015.162. → pages
- [52] Staveley-O’Carroll, K. *et al.* Induction of antigen-specific T cell anergy: An early event in the course of tumor progression. *Proceedings of the National Academy of Sciences of the United States of America* **95**, 1178–83 (1998). → pages
- [53] Ahmadzadeh, M. *et al.* Tumor antigen-specific CD8 T cells infiltrating the tumor express high levels of PD-1 and are functionally impaired. *Blood* **114**, 1537–44 (2009). DOI 10.1182/blood-2008-12-195792. → pages

- [54] Johnson, L. A. & June, C. H. Driving gene-engineered T cell immunotherapy of cancer. *Cell Research* **27**, 38–58 (2017). DOI 10.1038/cr.2016.154. → pages
- [55] Gros, A. *et al.* Prospective identification of neoantigen-specific lymphocytes in the peripheral blood of melanoma patients. *Nature Medicine* **22**, 433–438 (2016). DOI 10.1038/nm.4051. → pages
- [56] Morgan, R. A. *et al.* Cancer regression in patients after transfer of genetically engineered lymphocytes. *Science (New York, N.Y.)* **314**, 126–9 (2006). DOI 10.1126/science.1129003. → pages
- [57] Sharma, G. & Holt, R. A. T-cell epitope discovery technologies. *Human immunology* **75**, 514–9 (2014). DOI 10.1016/j.humimm.2014.03.003. → pages
- [58] Hicklin, D. J., Marincola, F. M. & Ferrone, S. HLA class I antigen downregulation in human cancers: T-cell immunotherapy revives an old story. *Molecular medicine today* **5**, 178–86 (1999). → pages
- [59] Bendle, G. M. *et al.* Lethal graft-versus-host disease in mouse models of T cell receptor gene therapy. *Nature medicine* **16**, 565–70, 1p following 570 (2010). DOI 10.1038/nm.2128. → pages
- [60] Rosenberg, S. A. Of Mice, Not Men: No Evidence for Graft-versus-Host Disease in Humans Receiving T-Cell Receptor-Transduced Autologous T Cells. *Molecular Therapy* **18**, 1744–1745 (2010). DOI 10.1038/mt.2010.195. → pages
- [61] Robbins, P. F. *et al.* A Pilot Trial Using Lymphocytes Genetically Engineered with an NY-ESO-1-Reactive T-cell Receptor: Long-term Follow-up and Correlates with Response. *Clinical Cancer Research* **21**, 1019–1027 (2015). DOI 10.1158/1078-0432.CCR-14-2708. → pages
- [62] Barrett, D. M., Singh, N., Porter, D. L., Grupp, S. a. & June, C. H. Chimeric antigen receptor therapy for cancer. *Annual review of medicine* **65**, 333–47 (2014). DOI 10.1146/annurev-med-060512-150254. → pages
- [63] Ahmad, Z. A. *et al.* scFv Antibody: Principles and Clinical Application. *Clinical and Developmental Immunology* **2012**, 1–15 (2012). DOI 10.1155/2012/980250. → pages
- [64] Dotti, G., Gottschalk, S., Savoldo, B. & Brenner, M. K. Design and development of therapies using chimeric antigen receptor-expressing T cells. *Immunological reviews* **257**, 107–26 (2014). DOI 10.1111/imr.12131. → pages
- [65] Gross, G., Waks, T. & Eshhar, Z. Expression of immunoglobulin-T-cell receptor chimeric molecules as functional receptors with antibody-type specificity. *Proceedings of the National Academy of Sciences of the United States of America* **86**, 10024–8 (1989). → pages
- [66] Sadelain, M., Brentjens, R. & Rivière, I. The Basic Principles of Chimeric Antigen Receptor Design. *Cancer Discovery* **3**, 388–398 (2013). DOI 10.1158/2159-8290.CD-12-0548. → pages
- [67] Jackson, H. J., Rafiq, S. & Brentjens, R. J. Driving CAR T-cells forward. *Nature Reviews Clinical Oncology* **13**, 370–383 (2016). DOI 10.1038/nrclinonc.2016.36. → pages

- [68] Vigneron, N. & Nathalie. Human Tumor Antigens and Cancer Immunotherapy. *BioMed Research International* **2015**, 1–17 (2015). DOI 10.1155/2015/948501. → pages
- [69] Hulsmeyer, M. *et al.* A Major Histocompatibility Complex{middle dot}Peptide-restricted Antibody and T Cell Receptor Molecules Recognize Their Target by Distinct Binding Modes: CRYSTAL STRUCTURE OF HUMAN LEUKOCYTE ANTIGEN (HLA)-A1{middle dot}MAGE-A1 IN COMPLEX WITH FAB-HYB3. *Journal of Biological Chemistry* **280**, 2972–2980 (2005). DOI 10.1074/jbc.M411323200. → pages
- [70] Maus, M. V. *et al.* An MHC-restricted antibody-based chimeric antigen receptor requires TCR-like affinity to maintain antigen specificity. *Molecular Therapy - Oncolytics* **3**, 16023 (2016). DOI 10.1038/mto.2016.23. → pages
- [71] Dai, H., Wang, Y., Lu, X. & Han, W. Chimeric Antigen Receptors Modified T-Cells for Cancer Therapy. *Journal of the National Cancer Institute* **108**, djv439 (2016). DOI 10.1093/jnci/djv439. → pages
- [72] Maude, S. L., Teachey, D. T., Porter, D. L. & Grupp, S. A. CD19-targeted chimeric antigen receptor T-cell therapy for acute lymphoblastic leukemia. *Blood* **125**, 4017–4023 (2015). DOI 10.1182/blood-2014-12-580068. → pages
- [73] Zhang, T. *et al.* Efficiency of CD19 chimeric antigen receptor-modified T cells for treatment of B cell malignancies in phase I clinical trials: a meta-analysis. *Oncotarget* **6**, 33961–71 (2015). DOI 10.18632/oncotarget.5582. → pages
- [74] Maude, S. L., Teachey, D. T., Porter, D. L. & Grupp, S. A. CD19-targeted chimeric antigen receptor T-cell therapy for acute lymphoblastic leukemia. *Blood* **125** (2015). → pages
- [75] Riddell, S. R. *et al.* Adoptive therapy with chimeric antigen receptor-modified T cells of defined subset composition. *Cancer journal (Sudbury, Mass.)* **20**, 141–4 (2014). DOI 10.1097/PPO.0000000000000036. → pages
- [76] Turtle, C. J. *et al.* CD19 CAR \AA T cells of defined CD4 + :CD8 + composition in adult B cell ALL patients. *Journal of Clinical Investigation* **126**, 2123–2138 (2016). DOI 10.1172/JCI85309. → pages
- [77] Sommermeyer, D. *et al.* Chimeric antigen receptor-modified T cells derived from defined CD8 + and CD4 + subsets confer superior antitumor reactivity in vivo. *Leukemia* **30**, 492–500 (2016). DOI 10.1038/leu.2015.247. → pages
- [78] Gattinoni, L. *et al.* A human memory T cell subset with stem cell \AA like properties. *Nature Medicine* **17**, 1290–1297 (2011). DOI 10.1038/nm.2446. → pages
- [79] Maude, S. L. *et al.* Chimeric Antigen Receptor T Cells for Sustained Remissions in Leukemia. *New England Journal of Medicine* **371**, 1507–1517 (2014). DOI 10.1056/NEJMoa1407222. → pages
- [80] Davila, M. L. *et al.* Efficacy and toxicity management of 19-28z CAR T cell therapy in B cell acute lymphoblastic leukemia. *Science translational medicine* **6**, 224ra25 (2014). DOI 10.1126/scitranslmed.3008226. → pages

- [81] Lee, D. W. *et al.* T cells expressing CD19 chimeric antigen receptors for acute lymphoblastic leukaemia in children and young adults: a phase 1 dose-escalation trial. *The Lancet* **385**, 517–528 (2015). DOI 10.1016/S0140-6736(14)61403-3. → pages
- [82] Hinrichs, C. S. & Restifo, N. P. Reassessing target antigens for adoptive T-cell therapy. *Nature Biotechnology* **31**, 999–1008 (2013). DOI 10.1038/nbt.2725. → pages
- [83] Newick, K., O'Brien, S., Moon, E. & Albelda, S. M. CAR T Cell Therapy for Solid Tumors. *Annual review of medicine* 1–14 (2016). DOI 10.1146/annurev-med-062315-120245. → pages
- [84] Bonifant, C. L., Jackson, H. J., Brentjens, R. J. & Curran, K. J. Toxicity and management in CAR T-cell therapy. *Molecular Therapy - Oncolytics* **3**, 16011 (2016). DOI 10.1038/mto.2016.11. → pages
- [85] Harlin, H. *et al.* Chemokine Expression in Melanoma Metastases Associated with CD8+ T-Cell Recruitment. *Cancer Research* **69**, 3077–3085 (2009). DOI 10.1158/0008-5472.CAN-08-2281. → pages
- [86] Wang, L.-C. S. *et al.* Targeting Fibroblast Activation Protein in Tumor Stroma with Chimeric Antigen Receptor T Cells Can Inhibit Tumor Growth and Augment Host Immunity without Severe Toxicity. *Cancer Immunology Research* **2**, 154–166 (2014). DOI 10.1158/2326-6066.CIR-13-0027. → pages
- [87] Caruana, I. *et al.* Heparanase promotes tumor infiltration and antitumor activity of CAR-redirected T lymphocytes. *Nature Medicine* **21**, 524–529 (2015). DOI 10.1038/nm.3833. → pages
- [88] Hatfield, S. M. *et al.* Immunological mechanisms of the antitumor effects of supplemental oxygenation. *Science Translational Medicine* **7**, 277ra30–277ra30 (2015). DOI 10.1126/scitranslmed.aaa1260. → pages
- [89] Fischer, K. *et al.* Inhibitory effect of tumor cell-derived lactic acid on human T cells. *Blood* **109**, 3812–3819 (2007). DOI 10.1182/blood-2006-07-035972. → pages
- [90] Jacobs, S. R. *et al.* Glucose uptake is limiting in T cell activation and requires CD28-mediated Akt-dependent and independent pathways. *Journal of immunology (Baltimore, Md. : 1950)* **180**, 4476–86 (2008). → pages
- [91] Ninomiya, S. *et al.* Tumor indoleamine 2,3-dioxygenase (IDO) inhibits CD19-CAR T cells and is downregulated by lymphodepleting drugs. *Blood* **125**, 3905–3916 (2015). DOI 10.1182/blood-2015-01-621474. → pages
- [92] Yong, C. S. M. *et al.* CAR T cell therapy of solid tumors. *Immunology and Cell Biology* 1–8 (2016). DOI 10.1038/icb.2016.128. → pages
- [93] Vignali, D. A. A., Collison, L. W. & Workman, C. J. How regulatory T cells work. *Nature reviews. Immunology* **8**, 523–32 (2008). DOI 10.1038/nri2343. → pages
- [94] Zhou, Q. *et al.* Program death-1 signaling and regulatory T cells collaborate to resist the function of adoptively transferred cytotoxic T lymphocytes in advanced acute myeloid leukemia. *Blood* **116**, 2484–2493 (2010). DOI 10.1182/blood-2010-03-275446. → pages

- [95] Liu, Y. *et al.* Inhibition of p300 impairs Foxp3+ T regulatory cell function and promotes antitumor immunity. *Nature Medicine* **19**, 1173–1177 (2013). DOI 10.1038/nm.3286. → pages
- [96] Muenst, S. *et al.* The immune system and cancer evasion strategies: therapeutic concepts. *Journal of Internal Medicine* **279**, 541–562 (2016). DOI 10.1111/joim.12470. → pages
- [97] Whiteside, T. L. Immune suppression in cancer: Effects on immune cells, mechanisms and future therapeutic intervention. *Seminars in Cancer Biology* **16**, 3–15 (2006). DOI 10.1016/j.semcancer.2005.07.008. → pages
- [98] Moon, E. K. *et al.* Multifactorial T-cell Hypofunction That Is Reversible Can Limit the Efficacy of Chimeric Antigen Receptor-Transduced Human T cells in Solid Tumors. *Clinical Cancer Research* **20**, 4262–4273 (2014). DOI 10.1158/1078-0432.CCR-13-2627. → pages
- [99] Pardoll, D. M. The blockade of immune checkpoints in cancer immunotherapy. *Nature Reviews Cancer* **12**, 252–264 (2012). DOI 10.1038/nrc3239. → pages
- [100] Ciceri, F. *et al.* Infusion of suicide-gene-engineered donor lymphocytes after family haploidentical haemopoietic stem-cell transplantation for leukaemia (the TK007 trial): a non-randomised phase I-II study. *The Lancet. Oncology* **10**, 489–500 (2009). DOI 10.1016/S1470-2045(09)70074-9. → pages
- [101] Di Stasi, A. *et al.* Inducible apoptosis as a safety switch for adoptive cell therapy. *The New England journal of medicine* **365**, 1673–83 (2011). DOI 10.1056/NEJMoa1106152. → pages
- [102] Philip, B. *et al.* A highly compact epitope-based marker/suicide gene for easier and safer T-cell therapy. *Blood* **124** (2014). → pages
- [103] Wu, C.-Y., Roybal, K. T., Puchner, E. M., Onuffer, J. & Lim, W. A. Remote control of therapeutic T cells through a small molecule-gated chimeric receptor. *Science (New York, N.Y.)* **350**, aab4077 (2015). DOI 10.1126/science.aab4077. → pages
- [104] Khalil, D. N., Smith, E. L., Brentjens, R. J. & Wolchok, J. D. The future of cancer treatment: immunomodulation, CARs and combination immunotherapy. *Nature Reviews Clinical Oncology* **13**, 273–290 (2016). DOI 10.1038/nrclinonc.2016.25. → pages
- [105] Newick, K., Moon, E. & Albelda, S. M. Chimeric antigen receptor T-cell therapy for solid tumors. *Molecular Therapy & Oncolytics* **3**, 16006 (2016). DOI 10.1038/mto.2016.6. → pages
- [106] Qasim, W. *et al.* Molecular remission of infant B-ALL after infusion of universal TALEN gene-edited CAR T cells. *Science Translational Medicine* **9**, eaaj2013 (2017). DOI 10.1126/scitranslmed.aaj2013. → pages
- [107] Ren, J. *et al.* Multiplex genome editing to generate universal CAR T cells resistant to PD1 inhibition. *Clinical Cancer Research* (2016). → pages
- [108] Themeli, M. *et al.* Generation of tumor-targeted human T lymphocytes from induced pluripotent stem cells for cancer therapy. *Nature Biotechnology* (2013). DOI 10.1038/nbt.2678. → pages

- [109] Kloss, C. C., Condomines, M., Cartellieri, M., Bachmann, M. & Sadelain, M. Combinatorial antigen recognition with balanced signaling promotes selective tumor eradication by engineered T cells. *Nature biotechnology* **31**, 71–5 (2013). DOI 10.1038/nbt.2459. → pages
- [110] Roybal, K. T. *et al.* Precision Tumor Recognition by T Cells With Combinatorial Antigen-Sensing Circuits. *Cell* **164**, 770–779 (2016). DOI 10.1016/j.cell.2016.01.011. → pages
- [111] Morsut, L. *et al.* Engineering Customized Cell Sensing and Response Behaviors Using Synthetic Notch Receptors. *Cell* **164**, 780–791 (2016). DOI 10.1016/j.cell.2016.01.012. → pages
- [112] Fedorov, V. D., Themeli, M. & Sadelain, M. PD-1- and CTLA-4-based inhibitory chimeric antigen receptors (iCARs) divert off-target immunotherapy responses. *Science translational medicine* **5**, 215ra172 (2013). DOI 10.1126/scitranslmed.3006597. → pages
- [113] Xie, Z., Wroblewska, L., Prochazka, L., Weiss, R. & Benenson, Y. Multi-input RNAi-based logic circuit for identification of specific cancer cells. *Science (New York, N.Y.)* **333**, 1307–11 (2011). DOI 10.1126/science.1205527. → pages
- [114] Bojar, D. & Fussenegger, M. The best of both worlds: reaping the benefits from mammalian and bacterial therapeutic circuits. *Current Opinion in Chemical Biology* **34**, 11–19 (2016). DOI 10.1016/j.cbpa.2016.05.012. → pages
- [115] Kemmer, C. *et al.* Self-sufficient control of urate homeostasis in mice by a synthetic circuit. *Nature Biotechnology* **28**, 355–60 (2010). DOI 10.1038/nbt.1617. → pages
- [116] Schukur, L., Geering, B., Charpin-El Hamri, G. & Fussenegger, M. Implantable synthetic cytokine converter cells with AND-gate logic treat experimental psoriasis. *Science Translational Medicine* **7**, 318ra201–318ra201 (2015). DOI 10.1126/scitranslmed.aac4964. → pages
- [117] Ye, H., Daoud-El Baba, M., Peng, R.-W. & Fussenegger, M. A synthetic optogenetic transcription device enhances blood-glucose homeostasis in mice. *Science (New York, N.Y.)* **332**, 1565–8 (2011). DOI 10.1126/science.1203535. → pages
- [118] Ye, H. *et al.* Pharmaceutically controlled designer circuit for the treatment of the metabolic syndrome. *Proceedings of the National Academy of Sciences of the United States of America* **110**, 141–6 (2013). DOI 10.1073/pnas.1216801110. → pages
- [119] Saxena, P., Charpin-El Hamri, G., Folcher, M., Zulewski, H. & Fussenegger, M. Synthetic gene network restoring endogenous pituitary-thyroid feedback control in experimental Graves' disease. *Proceedings of the National Academy of Sciences of the United States of America* **113**, 1244–9 (2016). DOI 10.1073/pnas.1514383113. → pages
- [120] Orive, G. *et al.* Cell encapsulation: Promise and progress. *Nature Medicine* **9**, 104–107 (2003). DOI 10.1038/nm0103-104. → pages
- [121] Barry, M. & Bleackley, R. C. Cytotoxic T lymphocytes: all roads lead to death. *Nature reviews. Immunology* **2**, 401–9 (2002). DOI 10.1038/nri819. → pages

- [122] Murphy, K., Travers, P. & Walport, M. *Janeway's Immunobiology* (Garland Science, New York, 2007), 7th edn. → pages
- [123] Caligiuri, M. A. Human natural killer cells. *Blood* **112**, 461–9 (2008). DOI 10.1182/blood-2007-09-077438. → pages
- [124] Lanier, L. L. Up on the tightrope: natural killer cell activation and inhibition. *Nature immunology* **9**, 495–502 (2008). DOI 10.1038/ni1581. → pages
- [125] Kruschinski, A. *et al.* Engineering antigen-specific primary human NK cells against HER-2 positive carcinomas. *Proceedings of the National Academy of Sciences of the United States of America* **105**, 17481–6 (2008). DOI 10.1073/pnas.0804788105. → pages
- [126] Vivier, E., Ugolini, S., Blaise, D., Chabannon, C. & Brossay, L. Targeting natural killer cells and natural killer T cells in cancer. *Nature reviews. Immunology* **12**, 239–52 (2012). DOI 10.1038/nri3174. → pages
- [127] Chowdhury, D. & Lieberman, J. Death by a thousand cuts: granzyme pathways of programmed cell death. *Annual review of immunology* **26**, 389–420 (2008). DOI 10.1146/annurev.immunol.26.021607.090404. → pages
- [128] Estébanez-Perpiña, E. *et al.* Crystal structure of the caspase activator human granzyme B, a proteinase highly specific for an Asp-P1 residue. *Biological chemistry* **381**, 1203–14 (2000). DOI 10.1515/BC.2000.148. → pages
- [129] de Saint Basile, G., Ménasché, G. & Fischer, A. Molecular mechanisms of biogenesis and exocytosis of cytotoxic granules. *Nature reviews. Immunology* **10**, 568–79 (2010). DOI 10.1038/nri2803. → pages
- [130] Vivier, E., Nunès, J. A. & Vély, F. Natural Killer Cell Signaling Pathways. *Science* **306**, 1517–1519 (2004). DOI 10.1126/science.1103478. → pages
- [131] Krzewski, K. & Coligan, J. E. Human NK cell lytic granules and regulation of their exocytosis. *Frontiers in Immunology* **3**, 335 (2012). DOI 10.3389/fimmu.2012.00335. → pages
- [132] Rudolph, M. G., Stanfield, R. L. & Wilson, I. a. How TCRs bind MHCs, peptides, and coreceptors. *Annual review of immunology* **24**, 419–66 (2006). DOI 10.1146/annurev.immunol.23.021704.115658. → pages
- [133] Dustin, M. L. & Long, E. O. Cytotoxic immunological synapses. *Immunological reviews* **235**, 24–34 (2010). → pages
- [134] Brownlie, R. J. & Zamoyska, R. T cell receptor signalling networks: branched, diversified and bounded. *Nature Reviews Immunology* **13**, 257–269 (2013). DOI 10.1038/nri3403. → pages
- [135] Trapani, J. A. & Smyth, M. J. Functional significance of the perforin/granzyme cell death pathway. *Nature reviews. Immunology* **2**, 735–47 (2002). DOI 10.1038/nri911. → pages
- [136] Lopez, J. A. *et al.* Perforin forms transient pores on the target cell plasma membrane to facilitate rapid access of granzymes during killer cell attack. *Blood* 2659–2668 (2013). DOI 10.1182/blood-2012-07-446146. → pages

- [137] Law, R. H. P. *et al.* The structural basis for membrane binding and pore formation by lymphocyte perforin. *Nature* **468**, 447–51 (2010). DOI 10.1038/nature09518. → pages
- [138] Mouchacca, P., Schmitt-Verhulst, A.-M. & Boyer, C. Visualization of cytolytic T cell differentiation and granule exocytosis with T cells from mice expressing active fluorescent granzyme B. *PloS one* **8**, e67239 (2013). DOI 10.1371/journal.pone.0067239. → pages
- [139] Thiery, J. *et al.* Perforin pores in the endosomal membrane trigger the release of endocytosed granzyme B into the cytosol of target cells. *Nature immunology* **12**, 770–7 (2011). DOI 10.1038/ni.2050. → pages
- [140] Ormö, M. *et al.* Crystal structure of the *Aequorea victoria* green fluorescent protein. *Science* **273**, 1392–5 (1996). → pages
- [141] Mason-Osann, E., Hollevoet, K., Niederfellner, G., Pastan, I. & Sun, X. Quantification of recombinant immunotoxin delivery to solid tumors allows for direct comparison of in vivo and in vitro results. *Scientific Reports* **5**, 10832 (2015). DOI 10.1038/srep10832. → pages
- [142] Woodsworth, D. J., Dunsing, V. & Coombs, D. Design Parameters for Granzyme-Mediated Cytotoxic Lymphocyte Target-Cell Killing and Specificity. *Biophysical journal* **109**, 477–488 (2015). → pages
- [143] Voskoboinik, I., Smyth, M. J. & Trapani, J. A. Perforin-mediated target-cell death and immune homeostasis. *Nat. Rev. Immunol.* **6**, 940–52 (2006). DOI 10.1038/nri1983. → pages
- [144] Lieberman, J. Granzyme A activates another way to die. *Immunological reviews* **235**, 93–104 (2010). DOI 10.1111/j.0105-2896.2010.00902.x. → pages
- [145] Coombs, D. & Goldstein, B. Effects of the geometry of the immunological synapse on the delivery of effector molecules. *Biophysical journal* **87**, 2215–20 (2004). DOI 10.1529/biophysj.104.045674. → pages
- [146] Grakoui, A. *et al.* The immunological synapse: A molecular machine controlling T cell activation. *Science* **285**, 221–227 (1999). → pages
- [147] Waugh, S. M., Harris, J. L., Fletterick, R. & Craik, C. S. The structure of the pro-apoptotic protease granzyme B reveals the molecular determinants of its specificity. *Nature structural biology* **7**, 762–765 (2000). → pages
- [148] Burgoyne, R. D. & Morgan, A. Secretory granule exocytosis. *Physiol Rev* **83**, 581–632 (2003). DOI 10.1152/physrev.00031.2002. → pages
- [149] Ishiura, S. *et al.* Calcium is essential for both the membrane binding and lytic activity of pore-forming protein (perforin) from cytotoxic T-lymphocyte. *Molecular immunology* **27**, 803–7 (1990). → pages
- [150] Lopez, J. a. *et al.* Rapid and unidirectional perforin pore delivery at the cytotoxic immune synapse. *Journal of immunology* **191**, 2328–34 (2013). DOI 10.4049/jimmunol.1301205. → pages

- [151] Brown, A. C. N. *et al.* Remodelling of cortical actin where lytic granules dock at natural killer cell immune synapses revealed by super-resolution microscopy. *PLoS biology* **9**, e1001152 (2011). DOI 10.1371/journal.pbio.1001152. → pages
- [152] Stinchcombe, J. C., Bossi, G., Booth, S. & Griffiths, G. M. The immunological synapse of CTL contains a secretory domain and membrane bridges. *Immunity* **15**, 751–61 (2001). → pages
- [153] Elf, J., Don, A. & Ehrenberg, M. Mesoscopic reaction-diffusion in intracellular signaling. *SPIE: Fluctuations and noise in Biological, Biophysical and Biomedical Systems* **5110**, 114–125 (2003). → pages
- [154] Elf, J. & Ehrenberg, M. Spontaneous separation of bi-stable biochemical systems into spatial domains of opposite phases. *Systems biology*. **1**, 230–236 (2004). DOI 10.1049/sb. → pages
- [155] Gibson, M. a. & Bruck, J. Efficient Exact Stochastic Simulation of Chemical Systems with Many Species and Many Channels. *The Journal of Physical Chemistry A* **104**, 1876–1889 (2000). DOI 10.1021/jp993732q. → pages
- [156] Gillespie, D. T. A general method for numerically simulating the stochastic time evolution of coupled chemical reactions. *Journal of Computational Physics* **22**, 403–434 (1976). DOI 10.1016/0021-9991(76)90041-3. → pages
- [157] Saffman, P. G. & Delbruck, M. Brownian motion in biological membranes. *Proceedings of the National Academy of Sciences* **72**, 3111–3113 (1975). DOI 10.1073/pnas.72.8.3111. → pages
- [158] Gambin, Y. *et al.* Lateral mobility of proteins in liquid membranes revisited. *Proceedings of the National Academy of Sciences of the United States of America* **103**, 2098–102 (2006). DOI 10.1073/pnas.0511026103. → pages
- [159] Ramadurai, S. *et al.* Lateral diffusion of membrane proteins. *Journal of the American Chemical Society* **131**, 12650–6 (2009). DOI 10.1021/ja902853g. → pages
- [160] Lauffenberger, D. A. & Linderman, J. J. *Receptors. Models for binding, trafficking and signaling* (Oxford University Press, 1993). → pages
- [161] Jaspard, F., Nadi, M. & Rouane, A. Dielectric properties of blood: an investigation of haematocrit dependence. *Physiol. Meas.* **24**, 137–147 (2003). → pages
- [162] Maher, K. J., Klimas, N. G., Hurwitz, B., Fletcher, M. A. & Schiff, R. Quantitative Fluorescence Measures for Determination of Intracellular Perforin Content Quantitative Fluorescence Measures for Determination of Intracellular Perforin Content. *Clinical and Vaccine Immunology* **9**, 1248–1252 (2002). DOI 10.1128/CDLI.9.6.1248. → pages
- [163] Kelso, A. *et al.* The genes for perforin , granzymes A-C and IFN-g are differentially expressed in single CD8 + T cells during primary activation. *International Immunology* **14** (2002). → pages
- [164] Maier, T., Güell, M. & Serrano, L. Correlation of mRNA and protein in complex biological samples. *FEBS letters* **583**, 3966–73 (2009). DOI 10.1016/j.febslet.2009.10.036. → pages

- [165] Fischer, H., Polikarpov, I. & Craievich, A. F. Average protein density is a molecular-weight-dependent function 2825–2828 (2004). DOI 10.1110/ps.04688204.studies. → pages
- [166] Ellis, R. J. & Minton, A. P. Cell biology: join the crowd. *Nature* **425**, 27–8 (2003). DOI 10.1038/425027a. → pages
- [167] Ozbabacan, S. E. A., Engin, H. B., Gursoy, A. & Keskin, O. Transient protein-protein interactions. *Protein engineering, design & selection : PEDS* **24**, 635–48 (2011). DOI 10.1093/protein/gzr025. → pages
- [168] Tang, J. *et al.* Using two fluorescent probes to dissect the binding, insertion, and dimerization kinetics of a model membrane peptide. *Journal of the American Chemical Society* **131**, 3816–7 (2009). DOI 10.1021/ja809007f. → pages
- [169] Pipkin, M. E. & Lieberman, J. Delivering the kiss of death: progress on understanding how perforin works. *Current opinion in immunology* **19**, 301–8 (2007). DOI 10.1016/j.coi.2007.04.011. → pages
- [170] Haberman, R. *Elementary Applied Partial Differential Equations* (Prentice Hall, Upper Saddle River, NJ 07458, 1998), 3rd edn. → pages
- [171] Stinchcombe, J. C., Majorovits, E., Bossi, G., Fuller, S. & Griffiths, G. M. Centrosome polarization delivers secretory granules to the immunological synapse. *Nature* **443**, 462–465 (2006). DOI 10.1038/nature05071. → pages
- [172] Wiedemann, A., Depoil, D., Faroudi, M. & Valitutti, S. Cytotoxic T lymphocytes kill multiple targets simultaneously via spatiotemporal uncoupling of lytic and stimulatory synapses. *Proc. Natl. Acad. Sci. U S A* **103**, 10985–90 (2006). DOI 10.1073/pnas.0600651103. → pages
- [173] Dietrich, C., Yang, B., Fujiwara, T., Kusumi, A. & Jacobson, K. Relationship of lipid rafts to transient confinement zones detected by single particle tracking. *Biophys J* **82**, 274–84 (2002). DOI 10.1016/S0006-3495(02)75393-9. → pages
- [174] Kusumi, A., Ike, H., Nakada, C., Murase, K. & Fujiwara, T. Single-molecule tracking of membrane molecules: plasma membrane compartmentalization and dynamic assembly of raft-philic signaling molecules. *Semin Immunol* **17**, 3–21 (2005). DOI 10.1016/j.smim.2004.09.004. → pages
- [175] Lillemeier, B. F. & Klammt, C. How membrane structures control T cell signaling. *Frontiers in Immunology* **3** (2012). DOI 10.3389/fimmu.2012.00291. → pages
- [176] Couzin-Frankel, J. Cancer Immunotherapy. *Science* **342**, 1432–1433 (2013). → pages
- [177] Fischbach, M. A., Bluestone, J. A. & Lim, W. A. Cell-Based Therapeutics: The Next Pillar of Medicine. *Science translational medicine* **5**, 179ps7 (2013). → pages
- [178] D’souza, N. *et al.* Mesenchymal stem/stromal cells as a delivery platform in cell and gene therapies. *BMC Medicine* **13**, 186 (2015). DOI 10.1186/s12916-015-0426-0. → pages
- [179] Lim, W. A. Designing customized cell signalling circuits. *Nature Reviews Molecular Cell Biology* **11**, 393–403 (2010). DOI 10.1038/nrm2904. → pages

- [180] Woodsworth, D. J. & Holt, R. A. Cell-Based Therapeutics: Making a Faustian Pact with Biology. *Trends in Molecular Medicine* **23**, 104–115 (2017). DOI 10.1016/j.molmed.2016.12.004. → pages
- [181] Hinrichs, C. S. & Rosenberg, S. A. Exploiting the curative potential of adoptive T-cell therapy for cancer. *Immunological reviews* **257**, 56–71 (2014). DOI 10.1111/imr.12132. → pages
- [182] Voskoboinik, I., Whisstock, J. C. & Trapani, J. A. Perforin and granzymes: function, dysfunction and human pathology. (2015). → pages
- [183] Cartwright, A. N. R., Griggs, J. & Davis, D. M. The immune synapse clears and excludes molecules above a size threshold. *Nature Communications* **5**, 5479 (2014). DOI 10.1038/ncomms6479. → pages
- [184] Huang, L., Pike, D., Sleat, D. E., Nanda, V. & Lobel, P. Potential pitfalls and solutions for use of fluorescent fusion proteins to study the lysosome. (2014). → pages
- [185] Montel, A. H., Morse, P. A. & Brahmi, Z. Upregulation of B7 molecules by the Epstein-Barr virus enhances susceptibility to lysis by a human NK-like cell line. *Cellular immunology* **160**, 104–114 (1995). → pages
- [186] Shimizu, Y. & DeMars, R. Production of human cells expressing individual transferred HLA-A,-B,-C genes using an HLA-A,-B,-C null human cell line. *The Journal of Immunology* **142**, 3320–3328 (1989). → pages
- [187] Braulke, T. & Bonifacino, J. S. Sorting of lysosomal proteins. *Biochimica et Biophysica Acta (BBA) - Molecular Cell Research* **1793**, 605–614 (2009). DOI 10.1016/j.bbamcr.2008.10.016. → pages
- [188] Saftig, P. & Klumperman, J. Lysosome biogenesis and lysosomal membrane proteins: trafficking meets function. *Nature Reviews Molecular Cell Biology* **10**, 623–635 (2009). DOI 10.1038/nrm2745. → pages
- [189] Lopez, J. a. *et al.* Perforin forms transient pores on the target cell plasma membrane to facilitate rapid access of granzymes during killer cell attack. *Blood* (2013). DOI 10.1182/blood-2012-07-446146. → pages
- [190] Breitling, J. & Aebi, M. N-linked protein glycosylation in the endoplasmic reticulum. *Cold Spring Harbor perspectives in biology* **5**, a013359 (2013). DOI 10.1101/cshperspect.a013359. → pages
- [191] van Meel, E. *et al.* Multiple Domains of GlcNAc-1-phosphotransferase Mediate Recognition of Lysosomal Enzymes. *The Journal of biological chemistry* **291**, 8295–307 (2016). DOI 10.1074/jbc.M116.714568. → pages
- [192] Warner, J. B., Thalhauser, C., Tao, K. & Sahagian, G. G. Role of N-linked oligosaccharide flexibility in mannose phosphorylation of lysosomal enzyme cathepsin L. *The Journal of biological chemistry* **277**, 41897–905 (2002). DOI 10.1074/jbc.M203097200. → pages
- [193] Qian, Y. *et al.* Functions of the alpha, beta, and gamma subunits of UDP-GlcNAc:lysosomal enzyme N-acetylglucosamine-1-phosphotransferase. *The Journal of biological chemistry* **285**, 3360–70 (2010). DOI 10.1074/jbc.M109.068650. → pages

- [194] Corey, D. R. & Craik, C. S. An investigation into the minimum requirements for peptide hydrolysis by mutation of the catalytic triad of trypsin. *Journal of the American Chemical Society* **114**, 1784–1790 (1992). DOI 10.1021/ja00031a037. → pages
- [195] Gupta, R. & Brunak, S. Prediction of glycosylation across the human proteome and the correlation to protein function. *Pacific Symposium on Biocomputing. Pacific Symposium on Biocomputing* 310–22 (2002). → pages
- [196] Blázquez, M. & Shennan, K. I. J. Basic mechanisms of secretion : sorting into the regulated secretory pathway. *Biochemistry and Cell Biology* **191**, 181–191 (2000). → pages
- [197] Portales-Casamar, E. *et al.* A regulatory toolbox of MiniPromoters to drive selective expression in the brain. *Proceedings of the National Academy of Sciences of the United States of America* **107**, 16589–94 (2010). DOI 10.1073/pnas.1009158107. → pages
- [198] Igney, F. H. & Krammer, P. H. Immune escape of tumors: apoptosis resistance and tumor counterattack. *Journal of leukocyte biology* **71**, 907–20 (2002). → pages
- [199] Vinay, D. S. *et al.* Immune evasion in cancer: Mechanistic basis and therapeutic strategies. *Seminars in Cancer Biology* **35**, S185–S198 (2015). DOI 10.1016/j.semcancer.2015.03.004. → pages
- [200] Marcus, A. *et al.* Recognition of tumors by the innate immune system and natural killer cells. *Advances in immunology* **122**, 91–128 (2014). DOI 10.1016/B978-0-12-800267-4.00003-1. → pages
- [201] Brentjens, R. J. *et al.* CD19-Targeted T Cells Rapidly Induce Molecular Remissions in Adults with Chemotherapy-Refractory Acute Lymphoblastic Leukemia. *Science Translational Medicine* **5**, 177ra38–177ra38 (2013). DOI 10.1126/scitranslmed.3005930. → pages
- [202] Porter, D. L., Levine, B. L., Kalos, M., Bagg, A. & June, C. H. Chimeric antigen receptor-modified T cells in chronic lymphoid leukemia. *The New England journal of medicine* **365**, 725–33 (2011). DOI 10.1056/NEJMoa1103849. → pages
- [203] Fulda, S. Tumor resistance to apoptosis. *International journal of cancer* **124**, 511–5 (2009). DOI 10.1002/ijc.24064. → pages
- [204] Jensen, M. C. & Riddell, S. R. Design and implementation of adoptive therapy with chimeric antigen receptor-modified T cells. *Immunological reviews* **257**, 127–44 (2014). DOI 10.1111/imr.12139. → pages
- [205] Bonavida, B. Tumor Cell Resistance to Apoptosis by Infiltrating Cytotoxic Lymphocytes. In Yefenof, E. (ed.) *Innate and Adaptive Immunity in the Tumor Microenvironment*, 121–137 (Springer, 2008). → pages
- [206] Töpfer, K. *et al.* Tumor evasion from T cell surveillance. *Journal of biomedicine & biotechnology* **2011**, 918471 (2011). DOI 10.1155/2011/918471. → pages
- [207] Wong, R. S. Y. Apoptosis in cancer: from pathogenesis to treatment. *Journal of experimental & clinical cancer research : CR* **30**, 87 (2011). DOI 10.1186/1756-9966-30-87. → pages

- [208] Debatin, K.-M. Apoptosis pathways in cancer and cancer therapy. *Cancer immunology, immunotherapy : CII* **53**, 153–9 (2004). DOI 10.1007/s00262-003-0474-8. → pages
- [209] Lin, Y.-f. *et al.* Targeting the XIAP / caspase-7 complex selectively kills caspase-3 deficient malignancies. *The Journal of Clinical Investigation* **123**, 3861–3875 (2013). DOI 10.1172/JCI67951DS1. → pages
- [210] Devarajan, E. *et al.* Down-regulation of caspase 3 in breast cancer: a possible mechanism for chemoresistance. *Oncogene* **21**, 8843–51 (2002). DOI 10.1038/sj.onc.1206044. → pages
- [211] de Heer, P. *et al.* Caspase-3 activity predicts local recurrence in rectal cancer. *Clinical Cancer Research* **13**, 5810–5815 (2007). DOI 10.1158/1078-0432.CCR-07-0343. → pages
- [212] Huang, H. *et al.* Expression and prognostic significance of osteopontin and caspase-3 in hepatocellular carcinoma patients after curative resection. *Cancer science* **101**, 1314–9 (2010). DOI 10.1111/j.1349-7006.2010.01524.x. → pages
- [213] Iolascon, A. *et al.* Caspase 3 and 8 deficiency in human neuroblastoma. *Cancer genetics and cytogenetics* **146**, 41–7 (2003). → pages
- [214] Lakhani, S. A. *et al.* Caspases 3 and 7: key mediators of mitochondrial events of apoptosis. *Science (New York, N.Y.)* **311**, 847–51 (2006). DOI 10.1126/science.1115035. → pages
- [215] de Oca, J. *et al.* Caspase-3 activity, response to chemotherapy and clinical outcome in patients with colon cancer. *International Journal of Colorectal Disease* **23**, 21–27 (2008). DOI 10.1007/s00384-007-0362-3. → pages
- [216] Oudejans, J. J. *et al.* Absence of caspase 3 activation in neoplastic cells of nasopharyngeal carcinoma biopsies predicts rapid fatal outcome. *Modern pathology : an official journal of the United States and Canadian Academy of Pathology, Inc* **18**, 877–85 (2005). DOI 10.1038/modpathol.3800398. → pages
- [217] Sun, Y. *et al.* Differential caspase-3 expression in noncancerous, premalignant and cancer tissues of stomach and its clinical implication. *Cancer detection and prevention* **30**, 168–73 (2006). DOI 10.1016/j.cdp.2006.02.004. → pages
- [218] Winter, R. N., Kramer, A., Borkowski, A. & Kyprianou, N. Loss of caspase-1 and caspase-3 protein expression in human prostate cancer. *Cancer research* **61**, 1227–32 (2001). → pages
- [219] Small, S., Keerthivasan, G., Huang, Z., Gurbuxani, S. & Crispino, J. D. Overexpression of survivin initiates hematologic malignancies in vivo. *Leukemia* **24**, 1920–6 (2010). DOI 10.1038/leu.2010.198. → pages
- [220] Huber, C. *et al.* Inhibitors of apoptosis confer resistance to tumour suppression by adoptively transplanted cytotoxic T-lymphocytes in vitro and in vivo. *Cell death and differentiation* **12**, 317–25 (2005). DOI 10.1038/sj.cdd.4401563. → pages
- [221] Berezovskaya, O. *et al.* Increased Expression of Apoptosis Inhibitor Protein XIAP Contributes to Anoikis Resistance of Circulating Human Prostate Cancer Metastasis Precursor Cells Increased Expression of Apoptosis Inhibitor Protein XIAP Contributes to Anoikis Resistance of Circul. *Cancer Research* 2378–2386 (2005). → pages

- [222] Fulda, S. & Vucic, D. Targeting IAP proteins for therapeutic intervention in cancer. *Nature reviews. Drug discovery* **11**, 109–24 (2012). DOI 10.1038/nrd3627. → pages
- [223] Coumar, M. S., Tsai, F.-Y., Kanwar, J. R., Sarvagalla, S. & Cheung, C. H. A. Treat cancers by targeting survivin: just a dream or future reality? *Cancer treatment reviews* **39**, 802–11 (2013). DOI 10.1016/j.ctrv.2013.02.002. → pages
- [224] de Almagro, M. C. & Vucic, D. The inhibitor of apoptosis (IAP) proteins are critical regulators of signaling pathways and targets for anti-cancer therapy. *Experimental oncology* **34**, 200–11 (2012). → pages
- [225] LaCasse, E. C. Pulling the plug on a cancer cell by eliminating XIAP with AEG35156. *Cancer letters* **332**, 215–24 (2013). DOI 10.1016/j.canlet.2012.06.015. → pages
- [226] Obexer, P. & Ausserlechner, M. J. X-linked inhibitor of apoptosis protein - a critical death resistance regulator and therapeutic target for personalized cancer therapy. *Frontiers in oncology* **4**, 197 (2014). DOI 10.3389/fonc.2014.00197. → pages
- [227] Altieri, D. C. Survivin , cancer networks and pathway-directed drug discovery. *Nature reviews Cancer* **8**, 551–560 (2008). → pages
- [228] Bots, M. *et al.* SPI-CI and SPI-6 cooperate in the protection from effector cell mediated cytotoxicity. *Blood* **105**, 1153–1161 (2005). DOI 10.1182/blood-2004-03-0791.Supported. → pages
- [229] Bots, M., VAN Bostelen, L., Rademaker, M. T., Offringa, R. & Medema, J. P. Serpins prevent granzyme-induced death in a species-specific manner. *Immunology and cell biology* **84**, 79–86 (2006). DOI 10.1111/j.1440-1711.2005.01417.x. → pages
- [230] Ravi, R. *et al.* Resistance of cancers to immunologic cytotoxicity and adoptive immunotherapy via X-linked inhibitor of apoptosis protein expression and coexisting defects in mitochondrial death signaling. *Cancer research* **66**, 1730–9 (2006). DOI 10.1158/0008-5472.CAN-05-3377. → pages
- [231] Végran, F. *et al.* Survivin-3B potentiates immune escape in cancer but also inhibits the toxicity of cancer chemotherapy. *Cancer research* **73**, 5391–401 (2013). DOI 10.1158/0008-5472.CAN-13-0036. → pages
- [232] Medema, J. P. *et al.* Blockade of the granzyme B/perforin pathway through overexpression of the serine protease inhibitor PI-9/SPI-6 constitutes a mechanism for immune escape by tumors. *Proceedings of the National Academy of Sciences of the United States of America* **98**, 11515–20 (2001). DOI 10.1073/pnas.201398198. → pages
- [233] Stirpe, F. Ribosome-inactivating proteins. *Toxicon: official journal of the International Society on Toxinology* **44**, 371–83 (2004). DOI 10.1016/j.toxicon.2004.05.004. → pages
- [234] Pastan, I., Hassan, R., Fitzgerald, D. J. & Kreitman, R. J. Immunotoxin therapy of cancer. *Nature reviews. Cancer* **6**, 559–65 (2006). DOI 10.1038/nrc1891. → pages
- [235] Zhang, J., Kale, V. & Chen, M. Gene-directed enzyme prodrug therapy. *The AAPS journal* **17**, 102–10 (2015). DOI 10.1208/s12248-014-9675-7. → pages

- [236] Fillat, C., Carrió, M., Cascante, A. & Sangro, B. Suicide gene therapy mediated by the Herpes Simplex virus thymidine kinase gene/Ganciclovir system: fifteen years of application. *Current gene therapy* **3**, 13–26 (2003). → pages
- [237] Searle, P. F. *et al.* NITROREDUCTASE: A PRODRUG-ACTIVATING ENZYME FOR CANCER GENE THERAPY. *Clinical and Experimental Pharmacology and Physiology* **31**, 811–816 (2004). DOI 10.1111/j.1440-1681.2004.04085.x. → pages
- [238] Vass, S. O., Jarrom, D., Wilson, W. R., Hyde, E. I. & Searle, P. F. E. coli NfsA: an alternative nitroreductase for prodrug activation gene therapy in combination with CB1954. *British Journal of Cancer* **100**, 1903–1911 (2009). DOI 10.1038/sj.bjc.6605094. → pages
- [239] Kim, J. H. *et al.* High Cleavage Efficiency of a 2A Peptide Derived from Porcine Teschovirus-1 in Human Cell Lines, Zebrafish and Mice. *PLoS ONE* **6**, e18556 (2011). DOI 10.1371/journal.pone.0018556. → pages
- [240] Soule, H. D., Vazquez, J., Long, A., Albert, S. & Brennan, M. A human cell line from a pleural effusion derived from a breast carcinoma. *Journal of the National Cancer Institute* **51**, 1409–16 (1973). → pages
- [241] Baginska, J. *et al.* Granzyme B degradation by autophagy decreases tumor cell susceptibility to natural killer-mediated lysis under hypoxia. *Proceedings of the National Academy of Sciences of the United States of America* **110**, 17450–5 (2013). DOI 10.1073/pnas.1304790110. → pages
- [242] Mahrus, S. & Craik, C. S. Selective chemical functional probes of granzymes A and B reveal granzyme B is a major effector of natural killer cell-mediated lysis of target cells. *Chemistry & biology* **12**, 567–77 (2005). DOI 10.1016/j.chembiol.2005.03.006. → pages
- [243] Deveraux, Q. L. & Reed, J. C. IAP family proteins—suppressors of apoptosis. *Genes & development* **13**, 239–52 (1999). → pages
- [244] Kalos, M. *et al.* T cells with chimeric antigen receptors have potent antitumor effects and can establish memory in patients with advanced leukemia. *Science translational medicine* **3**, 95ra73 (2011). DOI 10.1126/scitranslmed.3002842. → pages
- [245] Greco, R. *et al.* Improving the safety of cell therapy with the TK-suicide gene. *Frontiers in pharmacology* **6**, 95 (2015). DOI 10.3389/fphar.2015.00095. → pages
- [246] Fesnak, A. D., June, C. H. & Levine, B. L. Engineered T cells: the promise and challenges of cancer immunotherapy. (2016). → pages
- [247] Peng, W. *et al.* Loss of PTEN Promotes Resistance to T Cell-Mediated Immunotherapy. *Cancer Discovery* **6** (2016). → pages
- [248] Amantana, A., London, C. A., Iversen, P. L. & Devi, G. R. X-linked inhibitor of apoptosis protein inhibition induces apoptosis and enhances chemotherapy sensitivity in human prostate cancer cells. *Molecular cancer therapeutics* **3**, 699–707 (2004). → pages
- [249] Flanagan, L. *et al.* High levels of X-linked Inhibitor-of-Apoptosis Protein (XIAP) are indicative of radio chemotherapy resistance in rectal cancer. *Radiation Oncology* **10**, 131 (2015). DOI 10.1186/s13014-015-0437-1. → pages

- [250] Garg, H., Suri, P., Gupta, J. C., Talwar, G. P. & Dubey, S. Survivin: a unique target for tumor therapy. *Cancer Cell International* **16**, 49 (2016). DOI 10.1186/s12935-016-0326-1. → pages
- [251] Turley, S. J., Cremasco, V. & Astarita, J. L. Immunological hallmarks of stromal cells in the tumour microenvironment. *Nature Reviews Immunology* **15**, 669–682 (2015). DOI 10.1038/nri3902. → pages
- [252] Wang, C. *et al.* High throughput sequencing reveals a complex pattern of dynamic interrelationships among human T cell subsets. *Proceedings of the National Academy of Sciences of the United States of America* **107**, 1518–1523 (2010). DOI 10.1073/pnas.0913939107. → pages
- [253] Li, M. *et al.* Optimal promoter usage for lentiviral vector-mediated transduction of cultured central nervous system cells. *Journal of Neuroscience Methods* **189**, 56–64 (2010). DOI 10.1016/j.jneumeth.2010.03.019. → pages
- [254] Astrakhan, A. *et al.* Ubiquitous high-level gene expression in hematopoietic lineages provides effective lentiviral gene therapy of murine Wiskott-Aldrich syndrome. *Blood* **119**, 4395–4407 (2012). DOI 10.1182/blood-2011-03-340711. → pages
- [255] Choi, P. J. & Mitchison, T. J. Imaging burst kinetics and spatial coordination during serial killing by single natural killer cells. *Proceedings of the National Academy of Sciences of the United States of America* **110**, 6488–93 (2013). DOI 10.1073/pnas.1221312110. → pages
- [256] Lieberman, J. The ABCs of granule-mediated cytotoxicity: new weapons in the arsenal. *Nature reviews. Immunology* **3**, 361–70 (2003). DOI 10.1038/nri1083. → pages
- [257] Griffiths, G. M., Tsun, A. & Stinchcombe, J. C. The immunological synapse: a focal point for endocytosis and exocytosis. *The Journal of Cell Biology* **189**, 399–406 (2010). DOI 10.1083/jcb.201002027. → pages
- [258] Wang, Z. *et al.* Development of a nonintegrating Rev-dependent lentiviral vector carrying diphtheria toxin A chain and human TRAF6 to target HIV reservoirs. *Gene therapy* **17**, 1063–76 (2010). DOI 10.1038/gt.2010.53. → pages
- [259] Klingemann, H., Boissel, L. & Toneguzzo, F. Natural Killer Cells for Immunotherapy - Advantages of the NK-92 Cell Line over Blood NK Cells. *Frontiers in immunology* **7**, 91 (2016). DOI 10.3389/fimmu.2016.00091. → pages
- [260] James, J. R. & Vale, R. D. Biophysical mechanism of T-cell receptor triggering in a reconstituted system. *Nature* **487**, 64–9 (2012). DOI 10.1038/nature11220. → pages
- [261] Rabinovich, G. A., Gabrilovich, D. & Sotomayor, E. M. Immunosuppressive Strategies that are Mediated by Tumor Cells. *Annual Review of Immunology* **25**, 267–296 (2007). DOI 10.1146/annurev.immunol.25.022106.141609. → pages
- [262] Lebowitz, J. H. *et al.* Glycosylation-independent targeting enhances enzyme delivery to lysosomes and decreases storage in mucopolysaccharidosis type VII mice (2003). → pages
- [263] Di Stasi, A. *et al.* Inducible Apoptosis as a Safety Switch for Adoptive Cell Therapy. *New England Journal of Medicine* **365**, 1673–1683 (2011). DOI 10.1056/NEJMoal1106152. → pages

- [264] Ellington, A. D. & Szostak, J. W. In vitro selection of RNA molecules that bind specific ligands. *Nature* **346**, 818–822 (1990). DOI 10.1038/346818a0. → pages
- [265] Zhu, Q., Liu, G. & Kai, M. DNA Aptamers in the Diagnosis and Treatment of Human Diseases. *Molecules* **20**, 20979–20997 (2015). DOI 10.3390/molecules201219739. → pages
- [266] Kimoto, M., Yamashige, R., Matsunaga, K.-i., Yokoyama, S. & Hirao, I. Generation of high-affinity DNA aptamers using an expanded genetic alphabet. *Nature Biotechnology* **31**, 453–457 (2013). DOI 10.1038/nbt.2556. → pages
- [267] Merkle, T., Holder, I. T. & Hartig, J. S. The dual aptamer approach: rational design of a high-affinity FAD aptamer. *Org. Biomol. Chem.* **14**, 447–450 (2016). DOI 10.1039/C5OB02026C. → pages
- [268] Zhou, J. & Rossi, J. Aptamers as targeted therapeutics: current potential and challenges. *Nature Reviews Drug Discovery* (2016). DOI 10.1038/nrd.2016.199. → pages
- [269] Schumacher, M. A. Bacterial plasmid partition machinery: a minimalist approach to survival. *Current opinion in structural biology* **22**, 72–9 (2012). DOI 10.1016/j.sbi.2011.11.001. → pages
- [270] Annaluru, N. *et al.* Total Synthesis of a Functional Designer Eukaryotic Chromosome. *Science* **344**, 55–58 (2014). DOI 10.1126/science.1249252. → pages
- [271] Hutchison, C. A. *et al.* Design and synthesis of a minimal bacterial genome. *Science* **351**, aad6253–aad6253 (2016). DOI 10.1126/science.aad6253. → pages
- [272] Qi, L. S. *et al.* Repurposing CRISPR as an RNA-guided platform for sequence-specific control of gene expression. *Cell* **152**, 1173–83 (2013). DOI 10.1016/j.cell.2013.02.022. → pages
- [273] Suck, G. *et al.* NK-92: an ‘off-the-shelf therapeutic’ for adoptive natural killer cell-based cancer immunotherapy. *Cancer Immunology, Immunotherapy* **65**, 485–492 (2016). DOI 10.1007/s00262-015-1761-x. → pages
- [274] Linette, G. P. *et al.* Cardiovascular toxicity and titin cross-reactivity of affinity-enhanced T cells in myeloma and melanoma. *Blood* **122**, 863–71 (2013). DOI 10.1182/blood-2013-03-490565. → pages
- [275] Cameron, B. J. *et al.* Identification of a Titin-Derived HLA-A1-Presented Peptide as a Cross-Reactive Target for Engineered MAGE A3-Directed T Cells. *Science Translational Medicine* **5**, 197ra103–197ra103 (2013). DOI 10.1126/scitranslmed.3006034. → pages
- [276] Soriano, P. Generalized lacZ expression with the ROSA26 Cre reporter strain. *Nature genetics* **21**, 70–1 (1999). DOI 10.1038/5007. → pages
- [277] Ito, R., Takahashi, T., Katano, I. & Ito, M. Current advances in humanized mouse models. *Cellular & molecular immunology* **9**, 208–14 (2012). DOI 10.1038/cmi.2012.2. → pages
- [278] Kwok, R. Five hard truths for synthetic biology. *Nature* **463**, 288–90 (2010). DOI 10.1038/463288a. → pages

- [279] Barabási, A.-L. & Oltvai, Z. N. Network biology: understanding the cell's functional organization. *Nature Reviews Genetics* **5**, 101–113 (2004). DOI 10.1038/nrg1272. → pages
- [280] Hornung, G. & Barkai, N. Noise Propagation and Signaling Sensitivity in Biological Networks: A Role for Positive Feedback. *PLoS Computational Biology* **4**, e8 (2008). DOI 10.1371/journal.pcbi.0040008. → pages
- [281] Elmore, S. Apoptosis: a review of programmed cell death. *Toxicologic pathology* **35**, 495–516 (2007). DOI 10.1080/01926230701320337. → pages
- [282] Edelman, G. M. & Gally, J. A. Degeneracy and complexity in biological systems. *Proceedings of the National Academy of Sciences* **98**, 13763–13768 (2001). DOI 10.1073/pnas.231499798. → pages
- [283] Del Vecchio, D. Modularity, context-dependence, and insulation in engineered biological circuits. *Trends in biotechnology* **33**, 111–9 (2015). DOI 10.1016/j.tibtech.2014.11.009. → pages
- [284] Tawfik, D. S. Messy biology and the origins of evolutionary innovations. *Nature Chemical Biology* **6**, 692–696 (2010). DOI 10.1038/nchembio.441. → pages
- [285] Tsimring, L. S. Noise in biology. *Reports on progress in physics. Physical Society (Great Britain)* **77**, 026601 (2014). DOI 10.1088/0034-4885/77/2/026601. → pages
- [286] Uversky, V. N. Dancing Protein Clouds: The Strange Biology and Chaotic Physics of Intrinsically Disordered Proteins. *Journal of Biological Chemistry* **291**, 6681–6688 (2016). DOI 10.1074/jbc.R115.685859. → pages
- [287] Esch, E. W., Bahinski, A. & Huh, D. Organs-on-chips at the frontiers of drug discovery. *Nature Reviews Drug Discovery* **14**, 248–260 (2015). DOI 10.1038/nrd4539. → pages
- [288] Takahashi, M. K. *et al.* Rapidly Characterizing the Fast Dynamics of RNA Genetic Circuitry with Cell-Free Transcription–Translation (TX-TL) Systems. *ACS Synthetic Biology* **4**, 503–515 (2015). DOI 10.1021/sb400206c. → pages
- [289] Chin, J. W. Expanding and Reprogramming the Genetic Code of Cells and Animals. *Annual Review of Biochemistry* **83**, 379–408 (2014). DOI 10.1146/annurev-biochem-060713-035737. → pages
- [290] Lajoie, M. J. *et al.* Genomically Recoded Organisms Expand Biological Functions. *Science* **342** (2013). → pages
- [291] Ostrov, N. *et al.* Design, synthesis, and testing toward a 57-codon genome. *Science* **353** (2016). → pages
- [292] Gibson, D. G., Smith, H. O., Iii, C. A. H., Venter, J. C. & Merryman, C. chemical synthesis of the mouse mitochondrial genome **7** (2010). DOI 10.1038/nMeth.1515. → pages

Appendix A

Generating granzyme B mCherry fusion proteins expressed from genomic granzyme B locus using CRISPR/Cas9

To try to improve granzyme B payload fusion protein loading into lytic granules, and hence potentially transfer to target cells and fusion protein stability, I generated YTs that express granzyme B mCherry fusion proteins from the genomic granzyme B locus. (See Section 5.2.2 for further discussion and motivation.) Specifically I used the CRISPR/Cas9 system to insert a sequence consisting of a glycine serine linker followed by the crmCherry coding sequence immediately preceding the granzyme B stop codon.

I first designed two guide RNAs that had predicted cut sites within ten base pairs of the stop codon (see below for details of gRNA sequence and design). I cloned these gRNAs into pX330 (a plasmid expressing spCas9 and gRNA from CBh and U6 promoters respectively, addgene cat #42230). I also cloned a small fragment of the genomic granzyme B sequence immediately surrounding the stop codon (putative cut sites) into a plasmid that consists of GFP followed by RFP out of frame. The small fragment was inserted between the two fluorescent proteins. Thus if the gRNA cuts in that region, the NHEJ repair will result in around 30% in frame RFP. I tested the two guide RNAs using this plasmid (data not shown) and found that gRNA-1 (with sequence CATGAAACGCTACTA ACTAC) had a much higher cutting efficiency.

I then designed a donor template, consisting of an insert of a glycine serine linker followed by crmCherry, and left and right homology arms for regions immediately 5' and 3' of the stop codon in the genomic granzyme B locus (see below for details of template design, homology arm PCR from genomic DNA, and template assembly and cloning). The homology arms were 1 kb, and were amplified from genomic DNA extracted from the cell line NK-92MI. This donor was initially assembled in a TOPO vector (using tdTomato as a payload), and then subcloned into an MND vector. For the results presented here, the donor was further subcloned into a modified pdL vector that had the CAG promoter deleted

(see Appendix E for details on the pdL vector) and the tdTomato payload was replaced by an mCherry payload. This final donor plasmid was named pDN_GZB_E5-MCH (see Section A.2 for plasmid map and sequence).

I electroporated the pX330 (gRNA-1) and pDN_GZB_E5-MCH plasmids into YT-Indys using the Neon electroporation system (Thermo Fisher Scientific), using mainly the manufacturers recommendations. For details see the Methods sections of Chapters 3 and 4. Here I note any relevant details to this particular experiment.

- YTs were passaged 1/2 24 hours prior.
- Plasmids and DNA amounts. The negative controls were pX330 or pDN_GZB_E5-MCH alone, 20 ug. The positive control was 10 ug of pdL_MCH an mCherry expression plasmid. Finally, the experimental condition was both pX330 and pDN_GZB_E5-MCH at 1:1 molar ratio, 20 ug DNA.
- 100 ul tip using buffer R, final volume 110 ul

2 weeks after electroporation I FACS sorted cells for RFP signal, and then lysed 2×10^5 cells in Laemmli sample buffer. These lysates were size separated by gel electrophoresis transferred to blots and probed for mCherry and granzyme B (see Methods sections of Chapter 3 for details on sample preparation, Western blotting and antibodies).

I also extracted genomic DNA from 8×10^5 cells using DNAzol (Thermo Fisher Scientific) following the manufacturer's instructions. I used this genomic DNA as template for a PCR in which the forward primer (ACAGCTGCTCACTGTTGGGG) annealed 5' of the 5' end of the left homology arm (that is in the granzyme B gene, but not in the donor template) and the reverse primer (TTG-TACAGCTCGTCCATGCC) annealed in the mCherry insert (that is only in the donor template and nowhere in the granzyme B gene), so therefore the desired amplicon (2.8 kb) should only amplify from genomic DNA with the insert at the correct granzyme B locus. The PCR was conducted using Taq PCR supermix (Thermo Fisher Scientific) with 0.25 ul gDNA as template, 1 ul each of 10 uM primers in a 50 ul reaction. The cycling conditions were 57C annealing temperature, 3 minute extension time, for 30 cycles.

The results of the FACS sort, Western blot, and PCR screen are shown in Figure A.1 and demonstrate locus specific insertion of a granzyme B mCherry fusion protein expressed from the native granzyme B promoter. The fusion protein is functional (FACS and microscopy) and stable (I monitored expression for 6 weeks via FACS and microscopy). Important given the purpose of these experiments, is that the observed degradation of these fusion proteins is similar to what is observed when the fusion proteins are transiently overexpressed (data not shown).

I repeated these experiments exactly as above, using a different natural killer cell line (NK-92MI). NK92s had previously been used extensively in this project and had been shown to degrade granzyme payload fusion proteins to a very large extent. I was interested to see if expression from the endogenous

locus decreased this degradation. The experimental protocols were identical to those for YTs. The results are shown in Figure A.1. Again, the degradation is substantial and similar to what was previously observed for fusions transiently overexpressed (data not shown).

This data demonstrates that functional fusion proteins can be produced from the genomic granzyme B locus, but that the degradation of the fusion proteins is not substantially altered.

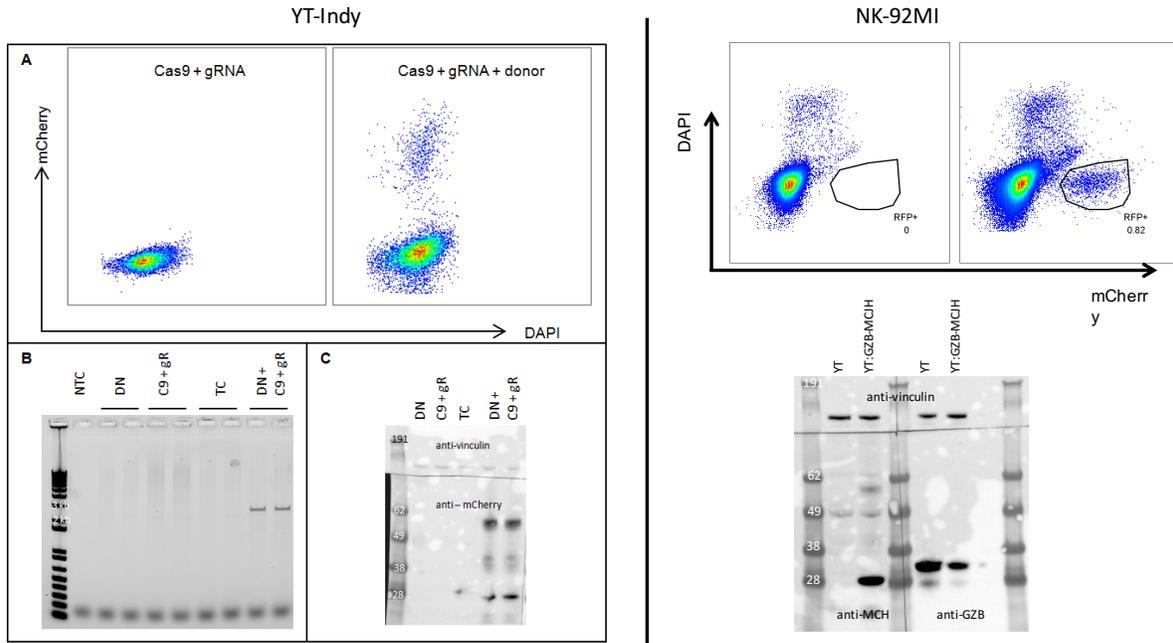


Figure A.1: Molecular characterization of genomically expressed granzyme B mCherry fusion proteins. CRISPR-cas9 generation of YT-Indy cells expressing granzyme-B-mCherry fusion proteins from the endogenous granzyme B locus. **Left panel.** A) FACS data showing that cells transfected with both cas9, gRNA and donor plasmids had a clear RFP+ population (right), while those transfected with cas9 and gRNA alone did not (left). B) Agarose gel of PCR reactions using genomic DNA from transfected YT-Indys as template, a forward primer that anneals at the 5' end of the granzyme B coding sequence, and a reverse primer that anneals at the 3' end of mCherry. Amplicons consistent with mCherry insertion at the endogenous granzyme B locus are observed in samples from cells that were transfected with cas9, gRNA and donor plasmid (far right, DN + C9 + gR), but not in any other sample. C) Western blot of whole cell lysates from transfected YT-Indys. Blot was probed with anti-mCherry antibody. Bands consistent with granzyme B-mCherry fusion proteins are observed in samples from cells that were transfected with cas9, gRNA and donor plasmid (far right, DN + C9 + gR), but not in any other sample. **Right panel** As for left panel, but using NK-92MI cells. Abbreviations: DN: donor plasmid; C9: cas9; gR: guide RNA; NTC: no template control; TC: transfection control.

A.1 Design of guide RNAs and donor template

TEMPLATE DESIGN

Design of donor vector for cas9 mediated fusion of payload to genomic GZMB locus. Need homology arms surrounding GZMB stop codon, flanking desired insert. Want this whole construct to be embedded inside LTRs of viral vector so that it may be packaged into virus.

Donor Template: LHA—GS_LINKER—PAYLOAD—2A—GFP—STOP—RHA

LHA: 1 kb NK-92MI genomic sequence immediately upstream of GZMB STOP codon

RHA: 1 kb NK-92MI genomic sequence immediately downstream of GZMB STOP codon

PAYLOAD: In assembly process will be tdTomato



Assembly strategy:

1. PCR amplify homology arms with tailed primers. Primer tails include homology overlaps to adjacent components for Gibson assembly, as well as NotI sites to liberate homology arms from sequencing vector.
2. TOPO clone amplicons and sequence verify.
3. Liberate INSERT (GS_LINKER—PAYLOAD—2A—GFP—STOP) from pMND:GZB-TDT_GFP (dHL_0355) via restriction digest (NotI/XhoI).
4. Open pMND (vHL_0021) by cutting out promoter and GFP reporter (Clal/Sall) immediately inside the LTRs) to function as BACKBONE.
5. Liberate RHA, LHA from TOPO vectors by NotI restriction digest.
6. 4 part Gibson Assembly:
BACKBONE—LHA—INSERT—RHA—BACKBONE
Homology for Gibson assembly provided by homology arm amplicons.
**NOTE: This assembly strategy requires 3'→5' exonuclease activity of the polymerase used. The standard NEB Gibson kit uses a polymerase that has fairly nonrobust activity. Thus it will be better to use either their NEB HiFi builder kit, or follow original Gibson protocol directly.

Primer design: 5'—NOTI—GIBSON_OVERLAP—RE_SITE—ANNEAL—3'

1. Gibson overlap design:
Tm>48C; length>20bp (since doing more than 3 part assembly)
Choose to do 25 bp overlaps, all of which had Tm high enough.
For restriction digest liberated components, there remains a few bases of the old RE sequence. To completely eliminate old RE and replace with new one (see below) move Gibson overlap further into vector. This will result in small 3' mismatches, which requires 3'→5' exonuclease activity.

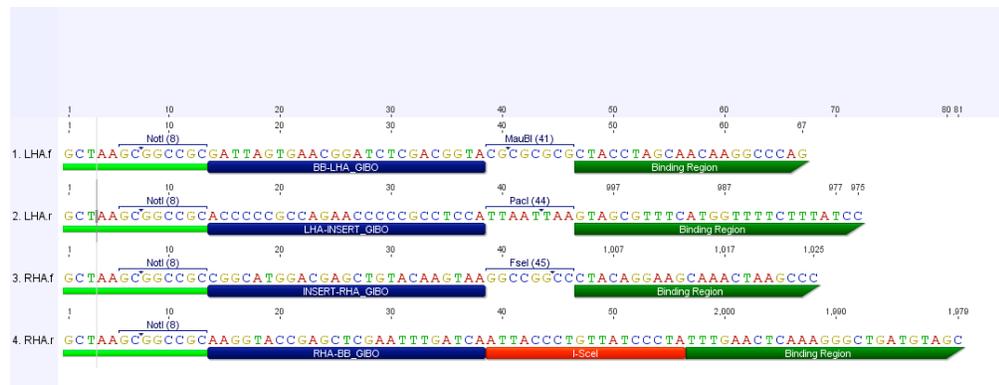
BACKBONE—LHA: GATTAGTGAACGGATCTCGACGGTA
 LHA—INSERT: TGGAGCGGGGGTCTGGCGGGGGT (RC: ACCCCCGCCAGAACCCCGCCTCCA)
 INSERT—RHA: CGGCATGGACGAGCTGTACAAGTAA
 RHA—BACKBONE: TGATCAAATTCGAGCTCGGTACCTT (RC: AAGGTACCGAGCTCGAATTTGATCA)

2. RE site selection:

LHA.f: Use *MauBI* to replace *Clal*, since 8 cutter more unique. CGCGCGCG
 LHA.r: Use *PacI* to replace *NotI*, which is not unique in MND vector (*NotI* cuts in backbone).
 TTAATTAA
 RHA.f: Use *FseI* to replace *XhoI*, since 8 cutter more unique. GGCCGGCC
 RHA.r: Use *I-SceI* to replace *Sall*. Choose homing endonuclease so guaranteed is always unique
 so that vector can be linearized for transfection. TAGGGATAACAGGGTAAT. (RC:
 ATTACCCTGTTATCCCTA)

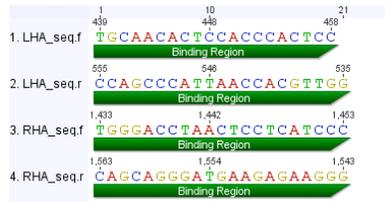
3. Annealing region design:

Standard 4 requirements: GC content roughly 50%, length ~ 20 bp, Tm ~50-60C, GC clamp.
 LHA.f: CTACCTAGCAACAAGGCCAG
 LHA.r: GTAGCGTTTCATGGTTTCTTTATCC
 RHA.f: CTACAGGAAGCAAAGTAAAGCC
 RHA.r: TTTGAACTCAAAGGGCTGATGTAGC



Sequencing primers:

Designed one forward and one reverse in the middle of each homology arm.



GUIDE DESIGN

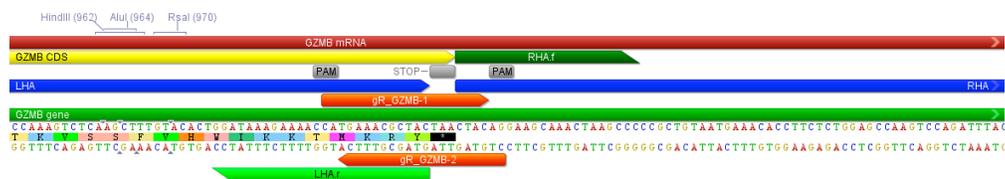
Took 70 bp on either side of GZMB stop codon, input into Zhang lab guide design tool:

<http://crispr.mit.edu/>.

all guides
scored by inverse likelihood of offtarget binding
mouse over for details ... show legend

	score	sequence	
Guide #1	88	CTGTAGTTAGTACGCTTCA	TGG
Guide #2	85	CATGAAACGCTACTAAGTAC	AGG
Guide #3	84	AAGGTGTTTCATTACAGCG	GGG
Guide #4	83	GAAAGTGTTCATTACAGCG	GGG
Guide #5	81	ACTTTGGTGCAGGCTGTTGG	AGG
Guide #6	78	AAGTCTCAAGCTTTGTACAG	TGG
Guide #7	75	GAGACTTTGGTGCAGGCTGG	TGG
Guide #8	70	GTGTACAAAGCTTGGAGCTT	TGG
Guide #9	69	AAAGCTTGCAGGCTTGGTGG	AGG
Guide #10	69	AGAAGGTGTTTCATTACAGC	GGG
Guide #11	62	GAGAAAGTGTTCATTACAGC	GGG
Guide #12	48	CTGTAAAGAAACACCTTCTC	TGG
Guide #13	48	TCTGGACTTGGCTCCAGAGA	AGG

Picked top two guides; all others have cut sites >10bp away from stop codon (e.g. homology junctions).

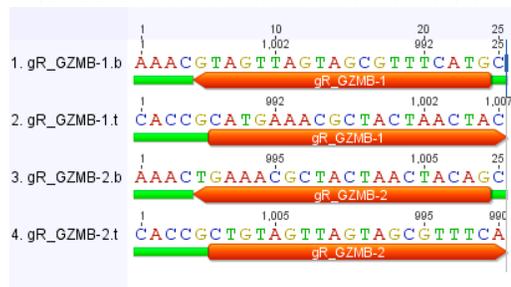


Ordered top and bottom oligos for annealing with overhangs for ligation into Zhang lab plasmids.

Example oligo design: Note that the NGG PAM is **not** included in the designed oligos.

Genomic Sequence 5' ... GACCACAGTCTGATCAGTTTTTCCTTGGGCTGCAA... 3'
3' ... CTGGTGTGACTAGTCAAAAGGAAACCCGACGTT... 5'

Oligo 1 → 5' - CACCGCAGTCTGATCAGTTTTTCCTT - 3'
3' - CTCAGACTAGTCAAAAGGAAACAAA - 5' ← Oligo 2



A.2 pDN_GZB_E5-MCH plasmid

A.2.1 Plasmid map

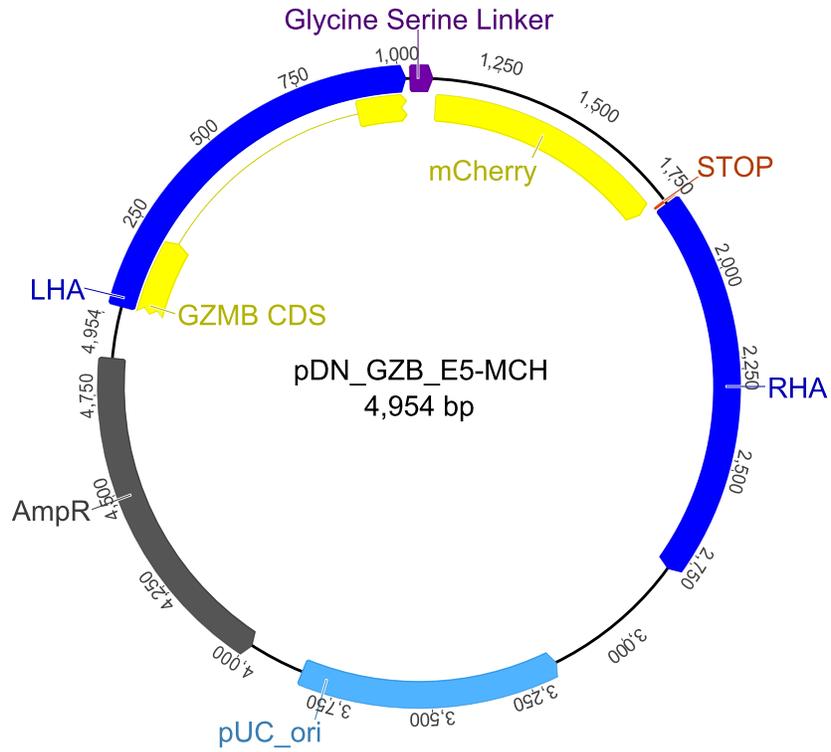


Figure A.2: pDN_GZB_E5-MCH plasmid. The annotations correspond to those in the genbank file.

A.2.2 Plasmid sequence

LOCUS pDN_GZB_E5-MCH 4954 bp DNA circular UNA 23-MAR-2017
DEFINITION Concatenation of 2 sequences.
ACCESSION urn:local...1490302855544.11
VERSION urn:local...1490302855544.11
KEYWORDS .
SOURCE
ORGANISM

FEATURES Location/Qualifiers
source 23..1022
/label="LHA"
CDS join(<23..238,882..>1022)
/gene="GZMB"
/note="Derived by automated computational analysis using
gene prediction method: BestRefSeq."
/codon_start=1
/product="granzyme B precursor"
/protein_id="NP_004122.2"
/db_xref="GI:221625528"
/db_xref="CCDS:CCDS9633.1"
/db_xref="GeneID:3002"
/db_xref="HGNC:HGNC:4709"
/db_xref="MIM:123910"
/translation="MQPILLLLAFLLLPRADAGEIIGGHEAKPHSRPYMAYLMIWDQK
SLKRCGGFLIRDDFVLTAHCWGSSINVTLGAHNIKEQEPTQQFIPVKRPIHPAYNP
KNFSDIMLLQLERKAKRTRAVQPLRLPSNKAQVKPGQTCSVAGWGQTAPLGKHSHTL
QEVKMTVQEDRKCESDLRHYDSTIELCVGDPEIKKTSFKGDSGGPLVCNKVAQGIVS
YGRNINGMPPRACTKVSSFVHWIKTKMKRY"
/label="GZMB CDS"
misc_feature 1032..1091
/note="Geneious type: polylinker"
/label="GS_Linkers_Rand"
CDS 1101..1775
/label="mCherry"
terminator 1782..1787
/label="STOP"
source 1794..2793
/label="RHA"
misc_feature complement(3170..3840)
/note="Geneious type: Origin of Replication"
/label="pUC_ori"
misc_feature complement(3985..4845)
/note="Geneious type: Marker"
/label="AmpR"

ORIGIN

```
1 ccccgtagc gagataatc atctacatg caacaaggcc caggatgaag cagggcagac
61 atgcagtggt gccggctggg ggcagacggc ccccttgga aaacactcac acacactaca
121 agaggtgaag atgacagtgc aggaagatcg aaagtgcgaa tctgacttac gccattaita
181 cgacagtacc attgagttgt gcgtggggga cccagagatt aaaaagactt cctttaaggt
241 aagactatgc acctgcctgg attggctctt gggagaaaga tgtttgggga atatctgaga
301 cctggagact caagtagtgg gggactcctt caccactag actgtgatat ttctctctgg
361 aaagagaaga ggggactaga ctgagctggg gagaattag ggcctctgca aactaccag
```

421 gaggcttatg gtggatggtg cttctttgga aggatgaatt tgcaacactc cacccactcc
481 aggtcacaga tattaggaaa ctgtgccac tgggggtgca gtaattataa ccagggtgtg
541 cttcagaggc tggtagccaa cgtggttaat gggctgggcc tccatggtgg acatcagccc
601 tccttgccca cttctgggtc cttaaacagc caacgggcc acatacctcc gatctcagga
661 tctgggggac atgacggagg ctggcccctg ggatgaggg aagcagtaac aatgtccagg
721 gccagagctt ggcagctggg ggcaccagc ggctgccct gccctctggt ctcccacatg
781 taggctgtgc aagtggcct ttctaaaag ggggctgag atggaagaga gggcaggacc
841 cggaggagca tcagctcagt ccttccactc tctattcaca gggggactct ggaggccctc
901 ttgtgtgtaa caaggtggcc cagggcattg tctcctatgg acgaaacaat ggcatgcctc
961 cacgagcctg caccaaagtc tcaagctttg tacactggat aaagaaaacc atgaaacgct
1021 acgcgcccg cggaggcggg ggttctggcg ggggtggatc aggggtgga ggctccggtg
1081 gaggcggtc gggcgcccc atcatcaagg agttcatgcg cttaagggtg cacatggagg
1141 gctccgtgaa cggccacgag ttcgagatcg agggcgagg cgagggccgc ccctacgagg
1201 gcaccagac cgccaagctg aaggtgacca agggggccc cctgcccttc gcctgggaca
1261 tcctgtcccc tcagttcatg tacggctcca aggcctact gaagcaccgc gccgacatcc
1321 ccgactactt gaagctgtcc ttccccgagg gcttcaagt ggagcgcgtg atgaacttcg
1381 aggacggcgg cgtgtgacc gtgaccagg actcctcct gcaggacggc gaggttatct
1441 acaaggtaga gctgcggcgc accaacttcc cctccgacgg ccccgtaatg cagaagaaga
1501 ccatgggctg ggaggcctcc tccgagcggg tgtaccccga ggacggcggc ctgaaggcgg
1561 agatcaagca gaggtgaag ctgaaggacg gcgccacta cgacgctgag gtcaagacca
1621 cctacaaggc caagaagccc gtgcagctgc cggcgcccta caacgtcaac atcaagttgg
1681 acatcacctc ccacaacgag gactacacca tctggaaca gtacgaacgc gccgagggcc
1741 gccactccac cggcgcatg gacgagctgt acaaggaatt ctaatagctc gagctacagg
1801 aagcaaaacta agcccccgct gtaatgaaac accttctctg gagccaagtc cagatttaca
1861 ctgggagagg tgccagcaac tgaataaata cctcttagct gagggtgaaa gctggtttct
1921 tgittattca ttgacctca ttctcaggca ccacatctgc gctatgcagg ccaatgacac
1981 aatittgctg ttttctgctt tctcctctcc cctcaccctc tgccacctcc ccaaaccctc
2041 acatgaagct gatactcagc tcttctctat ccacaccagt ttctccagg cctgcccttc
2101 tgccaaggct gaagctgagc accatcagga gacaacatgg accactttgg tctggggct
2161 ttgggtaaac ttcttaacct cttctccagt gttacatgac agagaaaaaa gggataatac
2221 catgggacct aactcctcat cccccactgg ggctcctcat tctcccctgg gcttagtttc
2281 tctaccctcc tctgagctca aggctcagct cgtcctccag cctcttggct gcccttctc
2341 ttcacccctg ctgagtgctc tcagaatcca ccaactcttg tctctccag accacactga
2401 tctgatctgg cccctccctc atatctacc acctaagata cccagagacc catgtggttc
2461 cataagggcc ttgccactga gacgccagcc catctcatgc cctggcagag aggggcctca
2521 gaaaaaccag gcctgtgtgg caaccaggta agacctatgg aggacaaggc tggcacgggtc
2581 tctctccaac ccttggctcc atctctcccc taggtagggc cagctcaacc cctcccattc
2641 agcccagtgct cctcccatac actcaaggtt cactgccacc ctgggcagtc agcaggctga
2701 gcccctttaa acctgttctt cttggctact gctggcctct aggctaagat tccctgctag
2761 caacctgggc tacatcagcc ctttgagttc aaagtgcacc tctagctaga gcttggcgtg
2821 atcatggtca tagctgtttc ctgtgtgaaa ttggtatccg ctcaaatc cacacaacat
2881 acgagccgga agcataaagt gtaaagcctg ggggcctaa tgagttagct aactcacatt
2941 aattgcgttg cgctcactgc ccgctttcca gtcgggaaac ctgtcgtgcc agctgcatta
3001 atgaatcggc caacgcggcg ggagaggcgg ttgctgtatt gggcgtcttt ccgcttctc
3061 gctcactgac tgcgtcgcct cggctcgttc gctgcggcga gcggtatcag ctactcaaa
3121 ggcggtaata cggttatcca cagaatcagg ggataacgca ggaagaaca tgtgagcaaa
3181 aggccagcaa aaggccagga accgtaaaaa ggccgcgttg ctggcgtttt tccataggct
3241 ccgccccctc gacgagcatc acaaaaatcg acgctcaagt cagagggtgc gaaaccgac
3301 aggactataa agataccagg cgtttcccc ttggaagctcc ctctgtcgtc ctctgttcc
3361 gacctgccc cttaccgat accgttccgc ctttctcct tccgggaagg tggcgcttc
3421 tcatagctca cgctgtaggt atctcagttc ggtgtaggtc gttcgtcca agctgggctg
3481 tgtgcacgaa cccccgttc agcccgaccg ctgcgcctta tccggtaact atcgtctga
3541 gtccaaccgc gtaagacacg acttatcgcc actggcagca gccactggta acaggattag
3601 cagagcgagg tatgtaggcg gtgctacaga gttcttgaag ttggtggccta actacggcta

3661 cactagaaga acagtatttg gtatctgccc tctgctgaag ccagttacct tcggaaaaag
3721 agttggtagc tcttgatccg gcaaaaaaac caccgctggt agcgggtttt ttgtttgcaa
3781 gcagcagatt acgcgagaa aaaaaggatc tcaagaagat cctttgatct tttctacggg
3841 gtctgacgct cagtggaacg aaaactcacg ttaagggatt ttggtcatga gattatcaaa
3901 aaggatcttc acctagatcc ttttaaaatta aaaatgaagt tttaaatcaa tctaaagtat
3961 atatgagtaa acttggctctg acagttacca atgcttaatc agtgaggcac ctatctcagc
4021 gatctgtcta tttcgttcat ccatagtgc ctgactcccc gtcgtgtaga taactacgat
4081 acgggagggc ttaccatctg gccccagtgc tgcaatgata ccgagagacc cacgctcacc
4141 ggctccagat ttatcagcaa taaccagcc agccggaagg gccgagcgca gaagtgggcc
4201 tgcaacttta tccgctcca tccagtctat taattgttgc cgggaagcta gagtaagtag
4261 ttcgccagtt aatagtttgc gcaacgttgt tgccattgct acaggcatcg tgggtgcacg
4321 ctctgctgtt ggtatggctt cattcagctc cggttcccaa cgatcaaggc gagttacatg
4381 atccccatg ttgtgcaaaa aagcggttag ctctctcggc cctccgatcg ttgtcagaag
4441 taagttggcc gcagtgttat cactcatggt tatggcagca ctgcataatt ctcttactgt
4501 catgccatcc gtaagatgct tttctgtgac tgggtgagtac tcaaccaagt cattctgaga
4561 atagtgtatg cggcgaccga gttgctcttg cccggcgta atacgggata ataccgcgcc
4621 acatagcaga actttaaaag tgctcatcat tggaaaacgt tcttcggggc gaaaactctc
4681 aaggatctta ccgctgttga gatccagttc gatgtaacc actcgtgcac ccaactgatc
4741 ttcagcatct tttactttca ccagcgttcc tgggtgagca aaaacaggaa ggcaaaatgc
4801 cgcaaaaaag ggaataaggc cgacacggaa atgttgaata ctcatactct tccttttca
4861 atattattga agcatttata agggttattg tctcatgagc ggatacatat ttgaatgtat
4921 ttagaaaaat aaacaaatag gggttccgag agct

//

Appendix B

Mass spectrometry based investigation of granzyme B mCherry fusion protein regions of instability

To try to decrease the amount of degradation of the granzyme B payload fusion proteins—and hence potentially increase the efficiency of payload transfer to target cells, and expand the range of possible delivery cell chassis—I conducted a mass spectrometry experiment to investigate if any regions in the granzyme B mCherry fusion protein were more susceptible to breakdown. (See Section 5.2.2 for further discussion and motivation.) Specifically I expressed the fusion proteins in YT-Indy cells, ran the whole cell lysate through a shotgun mass spectrometry pipeline, and then realigned the peptides to the fusion proteins sequence, looking for regions of instability.

I electroporated the pdL plasmids coding for mCherry (MCH) and granzyme B fused to MCH via a glycine serine linker (GZB-MCH) into YTs using the Neon electroporation system (Thermo Fisher Scientific), using mainly the manufacturers recommendations. For details see the Methods sections of Chapters 3 and 4. The plasmids were pdL_MCH and pdL_GZB-MCH respectively (see Appendix E for plasmid sequences and maps) and 20 ug of DNA was used. 48 hours after transfection I FACS sorted viable RFP+ cells. I then lysed these samples in Laemmli sample buffer, and size separate 8×10^5 cell equivalents via gel electrophoresis, transferred the samples to blots and probed for mCherry (see Methods sections of Chapter 3 for details on sample preparation, Western blotting and antibodies). Using an identical aliquot of the cell lysate sample, and loading equivalent cell amounts, I ran a second gel using the same running conditions and stained this gel with a non-specific Coomassie blue stain to identify all proteins in the gel.

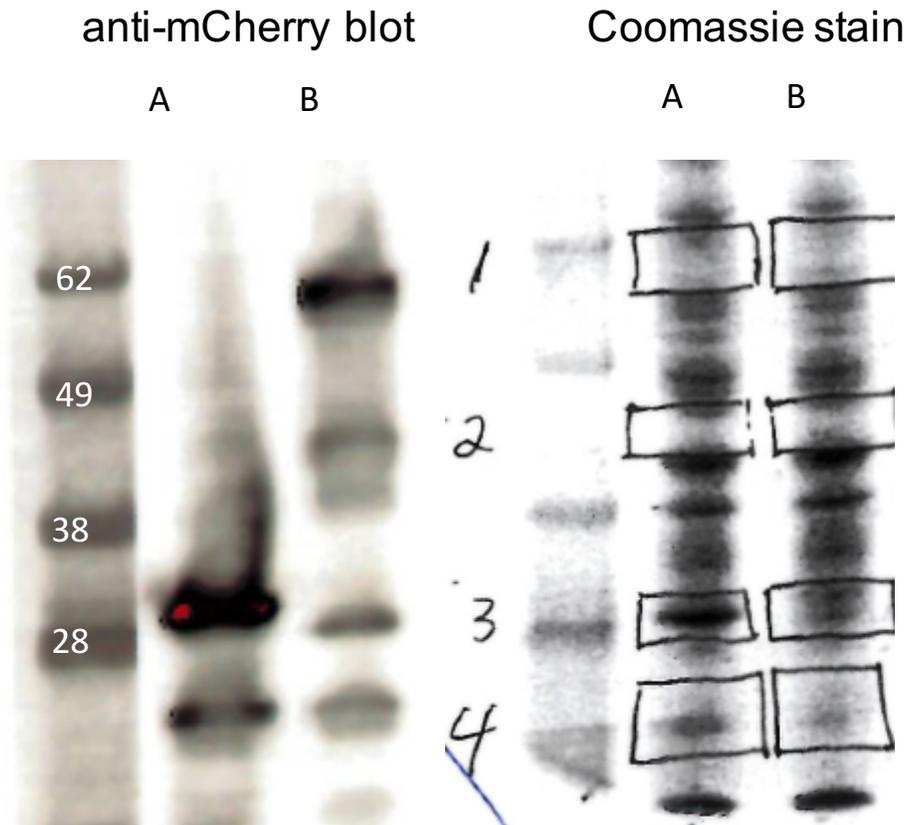


Figure B.1: Western blot and gel of samples used for mass spectrometry experiments. The images come from two separate gels, loaded with identical aliquots of the same sample. A = YTs expressing MCH; B = YTs expressing GZB-MCH. Handwritten numbers in middle are bands that were excised (approximate location hand drawn drawn on Coomassie stain) for mass spectrometry investigation. The letters A and B, and numbers 1-4 correspond to those in fig.... Numbers at far left are sizes in kDa of molecular marker.

The blot and gel images are shown in Figure B.1. For each sample, I identified four bands of interest, labeled 1-4, and consisting of (in order):

1. Full length fusion protein
2. Prominent breakdown product of full length fusion protein
3. Full length mCherry and/or fusion protein cleaved or decoupled approximate at the linker (since the sizes of granzyme B and crmCherry are both approximately 30 kDa)
4. Prominent breakdown product of mCherry

The Coomassie stained gel was submitted to the GSC Proteomics core facility. The bands were excised, the proteins in-gel digested with trypsin, and run on an Orbitrap Fusion mass spectrometer.

The resulting mass-charge ratio spectra were compared to the theoretical peptide spectra of GZB-MCH using SeqQuest, generating a list of peptides. These peptides were then aligned to the GZB-MCH amino acid sequence using a custom script written in R. The data is split into two figures Figure B.2 (for tryptic peptides) and Figure B.3 (non-tryptic peptides).

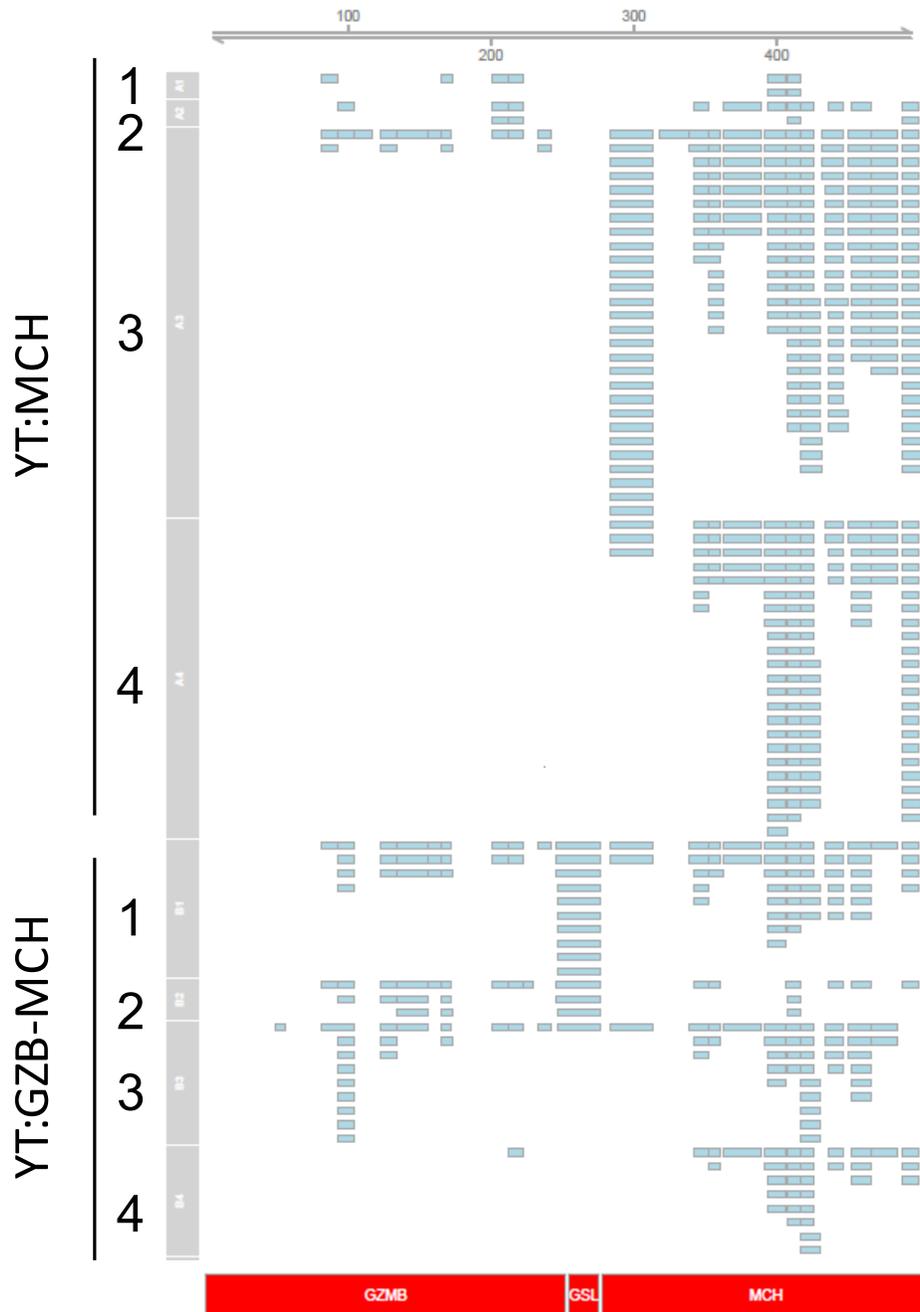


Figure B.2: Mass spectrometry identified tryptic peptides from whole cell lysates. Numbers at top and red schematic at bottom are reference GZB-MCH sequence . Each rectangle is individual peptide, mapped to its location in the reference.

Together this data suggests two tentative findings. First, there appears to be some semi-conserved pattern of N-terminal degradation of mCherry, as can be seen from comparing the coverage going from A3 to A4 in Figure B.2 (recall A = YT:MCH, and 3 = 30kda expected MCH size, 4 = 20kda breakdown region). This results in a substantial loss of coverage in the N-terminal region of Cherry, not just a loss of depth (as I would expect if the proteins in this region were just random breakdown products). This conclusion is further supported by the data in Figure B.3, which shows a significant density of non-tryptic peptides near the N-terminus of the protein. Notably this degradation seems somewhat dependent on N-terminal exposure, as these peptides are far more common in the unfused samples (which have a free N-terminus) than in the fusion protein samples.

Second, it appears that there is substantial 'breakage' in the peri-linker region of GZB-MCH. My argument here is the same: going from B1 to B3 (i.e. from the region I would expect full length fusion protein to the region where GZB or MCH would run independently) in Figure B.2, you see a substantial loss of coverage, mainly focused around the linker (with an almost complete loss of linker peptides). My assessment of this pattern is that it is more of an 'explosion' than a clean break. That is, if the protein just snapped in two, then I would expect both halves to migrate at 30 kda, and thus retain all of the peptides found in the full length fusion protein (losing at most one unique peptide, if the break point were in the middle of that peptide, but then I would potentially expect a non-tryptic peptide). However, if the middle of the protein is cleaved or somehow otherwise degraded in a localized but not amino acid restricted fashion, then all of the small fragments would migrate off the bottom of the gel, and thus generating the loss of peptides we see. One possible mechanism for is that if there are have two core globular proteins with hanging fragments of unstructured polypeptides, these are degraded away (by exopeptidases, mechanical shear etc). So I think a region of fragility, with frequent breakage, followed by further degradation from the break point to the core peptides, could explain these observations.

Based on this data I think it would be worth considering re-engineering the fusion protein. Since the domains necessary for granzyme B to act as a suitable chaperone are not fully delineated, it would be inadvisable to modify it. However, a different linker could be tested. This would be a judicious first modification to make, based on the model of peri-linker instability described above. One possible alternative is a shorter glycine serine linker that has been used in granzyme B tdTomato fusion proteins in a mouse model system [138] (with nucleotide and amino acid sequences GGCGGGTCTGGCG-GTGGGGGATCGGCCAACGGATCC and GGSGGGSANGS respectively).

The crmCherry protein used here is a variant of mCherry with the 12 N-terminal amino acids of mCherry deleted [184]. This was done in an effort to improve the stability of mCherry in lysosomes, after substantial N-terminal degradation was observed. It is possible that further elimination of N-terminal residues could improve the stability even more. This could be guided by the constraint of avoiding disrupting the beta-barrel or fluorophore containing alpha helix of the mCherry protein.

Finally, this pipeline provides a template for investigating the characteristics of fusion protein breakdown. This could be employed in attempts to extend this system to other cytotoxic lymphocyte chassis or when developing new payloads for delivery.

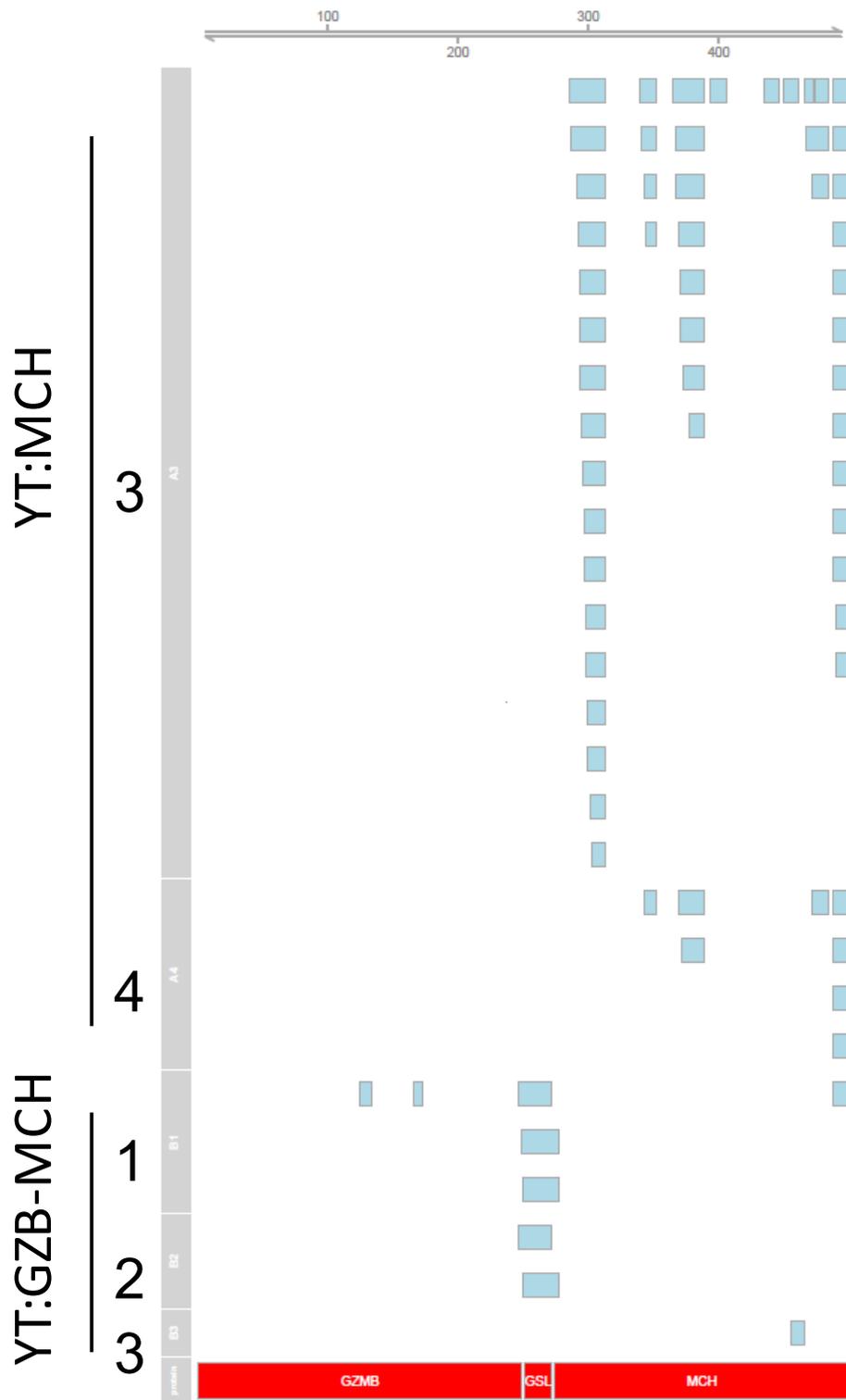


Figure B.3: Mass spectrometry identified non-tryptic peptides from whole cell lysates. Numbers at top and red schematic at bottom are GZB-MCH sequence. Each rectangle is individual peptide, mapped to its location in the reference.

Appendix C

Generation of a mutant EEF2 gene which protects cells from DTA when the two are co-expressed

The diphtheria toxin A fragment (DTA) was shown to be potent as a C-terminal granzyme fusion (GZB-DTA) and would be an attractive payload to deliver to tumour cells Section 4.2.1. However, for lymphocyte delivery, it is likely that some mechanism would be required to protect the delivery lymphocyte from GZB-DTA autotoxicity. Fortunately, a mutant form of EEF2—recall DTA inhibits protein synthesis by ribosylating EEF2 (Section 4.1)—has been reported, the overexpression of which renders cells resistant to DTA toxicity [252]. I set out to confirm this result and test if co-expression of this mEEF2 with GZB-DTA would protect cells from toxicity.

I first constructed a mEEF2 clone as follows. The wild type gene was obtained from the MGC collection (BC126259). The required mutation is G717R (GGA to CGA). To introduce this mutation I conducted a two step PCR reaction the first to insert the mutation, which would generate two overlapping fragments, and then a second fusion PCR with only external primers.

The annealing regions of the 4 primers were:

1. EEF2_G717R_IN.r: GATCTGGCCCCCTCGGCGGTGGATG
2. EEF2_G717R_IN.f: CATCCACCGCCGAGGGGGCCAGATC
3. EEF2_Ex.r: CAATTTGTCCAGGAAGTTGTCCAG
4. EEF2_Ex.f: ATGGTGAACCTCACGGTAGACCAG

The first PCR reaction used 0.5 ul Accuprime PFX (Thermo Fisher Scientific), 5 ul 10X reaction mix, 1.5 ul each primer (1 with 4 and 2 with 3, two separate reactions) and 30 ng EEF2 cDNA in a final volume of 50 ul. The cycling parameters were 55C annealing temperature and a 2m30s extension time,

for 30 cycles. The amplicons were size separated by gel electrophoresis and 2 and 0.4 kb bands were excised and column purified. The second PCR was conducted using 25 ng of the larger fragment and 5 ng of the smaller fragment (this to ensure equimolar amounts of each template). Only the external primers were used (3 and 4), otherwise the PCR reaction was as above. The cycling conditions were 64C annealing temperature, a 3m10s extension time for 27 cycles. The amplicons were size separated by gel electrophoresis and a 2.4 kb band was excised and column purified. This amplicon was ligated into a TOPO vector and sequence verified.

I then constructed three separate plasmids coding for: (i) GFP; (ii) GZB-DTA-2A-GFP; (iii) mEEF2-2A-GZB-DTA-2A-GFP, where 2A denotes a 2A peptide which results in ribosomal skipping. Thus these coding sequences are expressed as a single transcript, but result 1, 2 and 3 separate proteins respectively (see Section 4.2.1 and reference [239] for details). These coding sequences were expressed from the MND promoter in a MND lentiviral transfer plasmid (see Section 4.4 for details and references on the MND vector, and Section E.2 for plasmid maps and sequences).

I transfected these three plasmids into 293T cells using TransIT (Mirus Bio). (See Section 4.4 for details on 293T cells and transfection). 48 hours after transfection I imaged the cells using a fluorescence microscope and then ran them on a flow cytometer. The results are shown in Figure C.1, and show that: (i) GZB-DTA nearly completely abrogates GFP expression and (ii) mEEF2 co-expressed with GZB-DTA simultaneously restores GFP expression, albeit at lower a lower expression level, and in a lower fraction of cells.

This data suggests that DTA could be delivered to target cells as a granzyme B fusion protein, if mEEF2 were co-expressed along with the fusion protein, which could be done using 2A peptides as was done here. Importantly, since the mature mEEF2 and GZB-DTA proteins are separated, mEEF2 would not be loaded into lytic granules nor delivered to target cells, thus both preserving the integrity of the delivery lymphocyte, and avoiding counteracting the toxic effect of the delivered GZB-DTA in the target cells.

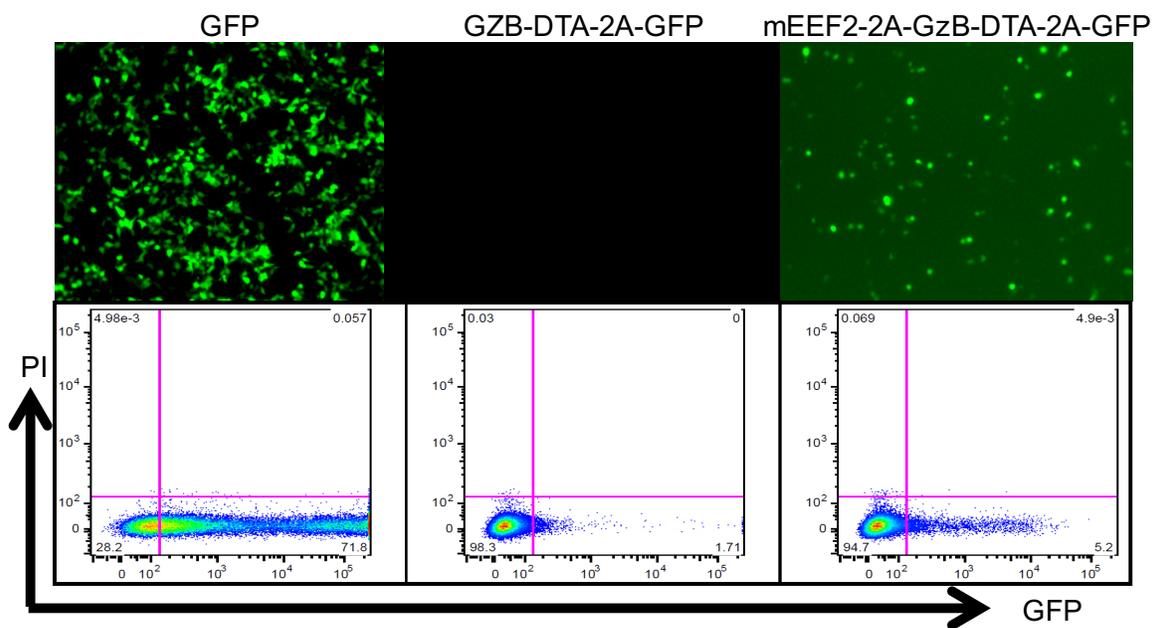


Figure C.1: Evidence that mEEF2 co-expression with GZB-DTA restores protein synthesis function. Top and bottom panels are us the same samples: cells were imaged and then analyzed via flow cytometry.

Appendix D

Code

D.1 Core MATLAB scripts implementing computational biophysical model (Chapter 2)

In this section I include a minimal working example of the stochastic simulation algorithm implementation. This consists of several MATLAB source code files, and a parameter file. For the work presented in the main text, this core set of code was compiled into a standalone executable (using MATLAB's built in capacity for this) on Westgrid. The result is a binary executable and an execution shell script. A custom python wrapper script was used to initialize the simulations across the parameter ranges, and submit them to the scheduler. Finally a small bash script was used to concatenate the various output files, and a MATLAB script to summarize this output. Since these are 'computational' steps, which would not change the actual data, and are fairly specific to Westgrid, they are not included.

D.1.1 IMS_Ex.m

```
%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
%
%Date:2015/07/12
%
%Written by: Daniel Woodsworth (daniel.woodsworth@gmail.com)
%
%
%Summary: Simulates time evolution of granzyme and perforin molecules
%within immunological synapse, from molecule's release from lytic granule
%through synapse diffusion, perforin insertion, pore formation and granzyme
%internalization.
%
%Units: (unless otherwise noted) microns, seconds.
%
%Methods: Uses spatial stochastic simulation algorithm (adapted
%from Elf, 2004). See Elf, Systems Biology, 2004 for further details
%regarding this method.
%
```

```

%Input: single parameter, an integer (simulation number) of class char.
%
%Requires:
%1. Helper scripts, within same directory, included with this file.
%2. Parameter file named 'params.txt', in same directory, included.
%
%Output:
%1. summary.txt: Most human readable, has simulation endpoint data.
%2. simVarSummary.csv: Has more endpoint data, csv format for data
%analysis.
%3. xx_timeCourse.csv: (xx is number provided as input) Has time evolution
%data of species numbers, csv formatted.
%
%
%For further details see paper: Woodsworth DJ et al., Biophysical Journal,
%2015.
%
%Free to use, please provide attribution.
%
%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%

function IMS_Ex(sim_num)

tic %Start clock to calculate total CPU time

%-----Parameter Input-----%

fid = fopen('params.txt','r');

%Parameter file must have format: name of parameter=value
%Read parameter file into cell array: first column is parameter names
%second is parameter values. All strings at this stage
C = textscan(fid,'%s %s','delimiter ','=');

%Convert to hash table, with names as keys, and values as values
params = containers.Map(C{1},C{2});

%Initialize local variables from hash table

%Geometry
R_syn = str2num(params('Radius_of_synapse'));
h = str2num(params('Height_of_synapse'));
r_LG = str2num(params('Radius_of_lytic_granule'));
l = str2num(params('Mesh_cell_size'));

%Molecule Sizes
r_GzB = str2num(params('Radius_of_granzyme_B'));
r_Pfn_XA = str2num(params('Cross-sectional_perforin_radius'));

%Pore values
n_pore = str2num(params('Number_of_perforin_monomers_per_pore'));
r_pore = str2num(params('Luminal_radius_of_perforin_pore'));

%Biophysical Constants

```

```

etaLipid = str2num(params('Viscosity_of_cell_membrane_(lipid)'));
etaH2O = str2num(params('Viscosity_of_water'));

Kb = str2num(params('Boltzmann_constant'));
T37 = str2num(params('Temperature_in_Kelvin'));

K_pfn_ins = str2num(params('Rate_of_perforin_insertion_into_membrane'));

%Hindered diffusion parameter, between 0 and 1.
%This allows for systematic exploration of space
%To use biophysical calculation of this parameter, set it to -1

hdc = str2num(params('Hindered_diffusion_coeffecient'));

%Intrasynaptic parameters
V_avg_mol = str2num(params('Average_volume_of_molecule_in_synapse'));
frac_occ = str2num(params('Fraction_of_synapse_occupied_by_molecules'));
Kd = str2num(params('Average_dissociation_constant_for_non-specific_binding'));

%Molecule numbers
N_PFN = str2num(params('Number_of_perforin_released_into_synapse'));
N_GZB = str2num(params('Number_of_granzyme_B_released_into_synapse'));

%Global Simulation_parameters
T = str2num(params('Maximum_simulation_time'));
outIncr = str2num(params('Time_interval_for_data_output'));
pVal = str2num(params('Current_Varied_Parameter_Value'));

%-----Shuffle random number generator-----%

%Ensure unique random number stream for simulation.

rng('shuffle');
sd = int32(rand*1e8 + str2num(sim_num));
rng(sd);

%-----Spatial Discretization-----%

%Discretize circular synapse. Inscribe circle inside square.
%Define origin in lower left of square. Mesh location
%defined by centre coordinates of mesh.
%Each cell indexed by rastering from bottom left to top right
%Mesh cell size must be chosen to be larger than perforin pore diameter

n_cells_per_side = ceil(2*R_syn/l); % number of mesh cells to go across square
n_cells = n_cells_per_side^2; % number of cells in system

%-----Calculate Derivative Constants-----%

r_Pfn = 2*r_Pfn_XA; %volume averaged pfn radius

```

```

%radius of PFN oligo is average of long and short dimensions of chain
r = @(i) r_Pfn_XA*(i+1)/2;

%Diffusivities for solution particles from Einstein Smoulchowski relation
D_GzB = Kb*T37/(6*pi*etaH20*r_GzB);
D_Pfn = Kb*T37/(6*pi*etaH20*r_Pfn);

%Diffusivities for membrane inserted perforin oligo is calculated using
%results from Saffman-Delbruck

gamma = 0.5772; %Eulers constant
h = 0.01; % Thickness of plasma membrane
D_Pfn_jmer = @(j) ((Kb*T37)/(4*pi*etaLipid*h))*(log(etaLipid*h/(etaH20*r(j))))-gamma);

%-----Hindered diffusion & macromolecular crowding-----%

%Assume that molecules diffusing through synapse slowed down by both
%occupied space of synapse due to other synaptic molcules (e.g. ICAM/LFA,
%CD45, TCR:MHC) and non-specific binding with these molcules

%So D_eff = (Kd/(R+Kd))*(1-f)D_free, where Kd is dissociation constant, R is
%density of nonspecific binders, and f is volume fraction occupied by all
%molcules.
%R = f/(Avg. Molecule in Syanpse)

%Switch: set hdc to -1 to use biophysical calculation; otherwise can use
%lumped hdc parameter to explore effects of varying hindered diffusion
if(hdc < 0)

    %last term to convert to Molar (NOTE this term accounts for nm vs um!!)
    R_binder = (frac_occ/V_avg_mol)*(6/10);

    %Effective diffusivities
    D_GzB_eff = (Kd/(R_binder+Kd))*(1-frac_occ)*D_GzB;
    D_Pfn_eff = (Kd/(R_binder+Kd))*(1-frac_occ)*D_Pfn;

else

    D_GzB_eff = hdc*D_GzB;
    D_Pfn_eff = hdc*D_Pfn;

end

%-----System Size-----%

%Number of species in system: inserted perforin jmers 1-18, synaptic
%perforin and granzyme
N_species = n_pore+2;

%State matrix N, each row is a subvolume
%Order of species in state matrix N: P1,...,P18,Psyn,Gsyn
N = zeros(n_cells, N_species);

```

```

%-----Set Reactions-----%
%
%We have following reactions:
%
%
%                               Kij
%Perforin oligomerization: Pj+Pi -> Pi+j, j=1,2,3; i = j,...,18-j
%
%                               K1
%Bulk perforin insertion into membrane: Pb -> P1
%
%                               Kg
%Granzyme translocation through perforin pore: G + P18 -> Gi + P18
%
%Important to note only allow 1,2,3mers join to form higher oligos, >4mer
%cannot join with each other, only 1,2,3mer.

%Number of interactions:
%monomer to jmer (17)
%dimer to jmer (15)
%trimer to jmer (13)
%perforin insertion and granzyme translocation (2)
n_rxns = 17+15+13+2;

%Stoichiometry matrix S.
%Each row corresponds to a a given reaction in the system, row order must
%match column order of A/ai (see below).
%Each column corresponds to a species in the system, column order must
%match that of N above.
%Sij is the change in number of molecules of the jth species that occurs
%when the jth reaction occurs.

%Construct S in 4 parts, one each for each 1,2,3mer reactions, finally for
%bulk species.

S = [];

%cycle through 1,2,3mers

for j=1:3

    %Can only have reactions up to n_pore (18) - j. E.g. if have 3mer, can
    %only have 3mer joining a 15mer to form an 18mer, whereas with 1mer can
    %have 1mer joining 17mer to form 18mer, so more reactions. This is the
    %interpretation of the n_pore-j below: number of reactions for a jmer.

    subS=zeros(n_pore-j, N_species);
    for i=j:(n_pore-j)

        %Exception for two identical oligos combining to form multimer
        %In this case lose 2 of one species, manually set appropriately
        if(i == j)
            subS(i,j) = -2;
            subS(i,i+j) = 1;
        end
    end
end

```

```

        continue;
    end

    %For all other oligo interactions , Lose imer and jmer , gain i+jmer
    subS(i, j) = -1;
    subS(i, i) = -1;
    subS(i, i+j) = 1;
end

%Add this subsection of matrix to total stoichiometry matrix
%Since indexing starts at j , will have empty rows of subS from 1 to j-1
%So only append rows from j onwards
S = [S;subS(j:end,:)];
end

%Last part for pfn insertion and granzyme translocation
subS = zeros(2, N_species);

%Pfn insertion
subS(1, n_pore+1) = -1;
subS(1, 1) = 1;

%Granzyme translocation
subS(2, n_pore+2) = -1;

%append
S = [S;subS];

%-----Set rate constants-----

%Perforin oligo formation (Kij)
%Rate constant for creation of perforin multimer from two perforin oligos

%Use result from Lauffenburger & Linderman for particle finding a trap
% $k = 2\pi D * K_{on} / (2\pi D + K_{on} * \ln(b/s))$ ,  $K_{on}$  is the chemical rate
% $b = \sqrt{SA / (\pi * P_j)}$ ,  $P_j$  is the number of PFN oligo jmers finding the
%larger PFN imer trap,  $D$  is sum of diffusivities
% $s$  is reaction radius, take this to be sum radius of PFN imer and jmer
%Take diffusion limited case  $\rightarrow K = 2\pi D / \ln(b/s)$ 
%This  $k$  has units of  $L^2/s$ , so need to divide by  $SA$  of mesh cell  $l^2$ 
%So  $k_i = 2\pi D / ((l^2) * \ln(b/s))$ . Note that  $s$  is a function of  $i$ , and  $b$  is
%a function of  $P_1$ 

%Since these values change over time , need to define them internally to
%simulation.

b = @(Pj) sqrt(l^2/(pi*Pj));
s = @(i, j) r(i) + r(j);
K = @(i, j, Pj) 2*pi*(D_Pfn_jmer(i) + D_Pfn_jmer(j))/((l^2)*log(b(Pj)/s(i, j)));

%Rate of granzyme finding pore
%Assume granzyme in cuboid of height  $h$ , whose projection onto membrane
%surface is a mesh cell

```

```

%Then rate of diffusive movement across cuboid is (3D / h^2), solely along z
%axis. This is in competition with diffusive jumps out of cuboid.
%Assume that this rate is modulated by A_p/A_tot, e.g. the fraction of the
%pore size compared to the mesh cell size
%So rate is (3 D A_p) / (L^2 h^2)
Kg = 3*D_GzB_eff*(pi*r_pore^2)/((L^2)*(h^2));

%----- Initiliaziation -----%

%First discretize synapse into square cells, and calculate which cells
%correspond to central region where LGs are released. Track if cell is
%external, internal, or internal to LG.

%Raster from bottom left to top right. For each cell check if its centre is
%within radius of synapse, and within radius of lytic granule.
%For lytic granule, assign indices to vector so that molecules can be
%initially distributed there

%BinMat matrix below defines if mesh cell is inside (1) or outside(0)
%synapse, or inside LG (2)
BinMat = zeros(n_cells_per_side);

%Boolean vector, determines if cell is internal or external, according to
%cell index
IsInt = zeros(n_cells,1);

%Cell index. Starts from bottom left, increases L→R, Bot→Top
Ci = 1;

%For given subvolume, give x,y coordinates
SVcoords = zeros(n_cells,2);

LGInd = []; %vector with indices of all cells within LG radius

%Raster over discretized space
for ix = 1:n_cells_per_side
    for iy = 1:n_cells_per_side

        %spatial coords of current cell (defined centrally)
        %Note -1 is 'one', others are lowercase L
        x = (ix-1)*l + l/2;
        y = (iy-1)*l + l/2;
        SVcoords(Ci,:) = [x y];

        %Check if displacement from centre synapse less than radius of
        %synapse minus mesh cell length, to ensure at least 1 cell is
        %external all the way round
        if ((x-R_syn)^2 + (y-R_syn)^2 <= (R_syn-l)^2)
            BinMat(ix,iy) = 1;
            IsInt(Ci) = 1;
        end
    end
end

```

```

    %Check if displacement from centre synapse less than LG radius
    if ((x-R_syn)^2 + (y-R_syn)^2 <= r_LG^2)
        BinMat(ix, iy) = 2;
        LGInd = [LGInd, Ci];
    end

    Ci = Ci+1;

end
end

%Now generate initial distributions. Only need to update those cells that
%have species in them initially.

%Required matrices

%Rate matrix R. Row is a subvolume, columns are sum (within that subvolume)
%of interaction propensities, diffusion propensities and all propensities.

R = zeros(n_cells,3);

%Propensity matrix A. Each row corresponds to a subvolume, each column is a
%interaction (e.g. a 'chemical' reaction)
A = zeros(n_cells, n_rxns);

%Diffusion matrix D. Each row corresponds to subvolume, each column is
%diffusive propensity for a species to jump out of subvolume. Column order
%matches state matrix N
D = zeros(n_cells, N_species);

%Initialization list for next event time for each cell used to build binary
%tree
%Since most mesh cells have no species in them, their next event time is
%infinite
%First row is index of mesh cell, second is next event time

InitList = [[1:n_cells]' inf(n_cells,1)];
Qarray = [1:n_cells]';

%Initialize cells that have nonzero species numbers
%1. Distribute species
%2. Calculate diffusive and reaction propensities, update matrices
%3. Calculate next event time for these cells
%All other cells have 0 species; hence no chemical or diffusion events

%Since dividing total number of molecules across lytic granules, round to
%nearest integer; cannot have decimal molecules.
N_PFN_per_LG_cell = round(N_PFN/length(LGInd));
N_GZB_per_LG_cell = round(N_GZB/length(LGInd));

%Adjust total number of molecules in system accordingly for internal
%consistency.
N_PFN = N_PFN_per_LG_cell*length(LGInd);

```

```

N_GZB = N_GZB_per_LG_cell*length(LGInd);

for i = LGInd

    %Note: all calculations in this loop are for a given cell

    %Add perforin and granzyme to state matrix
    N(i,n_pore+1) = N_PFN_per_LG_cell;
    N(i,n_pore+2) = N_GZB_per_LG_cell;

    %Calculate interaction propensities for current cell
    ai = calcIntrxnPropens(N(i,:),n_pore,K, K_pfn_ins,Kg);

    %Insert this nonzero row into propensity matrix
    A(i,:) = ai;

    %now recalculate a0
    a0 = sum(ai);

    %Calculate diffusion propensities for current cell
    di = calcDiffPropens(N(i,:), n_pore, D_Pfn_jmer, D_Pfn_eff, D_GzB_eff, l);

    %Insert this nonzero row into propensity matrix
    D(i,:) = di;

    %recalculate d0
    d0 = sum(di);

    %update total propensisty
    s0 = a0 + d0;

    %Update rate matrix for this SV
    R(i,:) = [a0 d0 s0];

    %Calculate next event time for this cell, which will now be less than
    %infinite since have molecules in cell
    tau = -log(rand)/s0;

    %Put this into the initialization list. Note that the constructor for
    %the binary tree accepts an arbitrary unordered list
    InitList(i,2) = tau;

end

%Build event Q and Qarray
[Q,Qarray] = buildQ(InitList,Qarray);

%-----Simulation-----%

%Current simulation time
t = 0;

```

```

%Counting variables
N_GzB_Int = 0; %Internalized GzB
Lost = [0 0]; %Lost PFN and GzB
T_pore_form = []; %Time when each pore forms
Num_pore_formed = 0; %Total number of pores formed
maxmer = 0; %Maximum PFN oligomer (18 if pore forms)
stuckCount = 0; %Boolean: 1 if system is stuck, 0 otherwise.

%outMat is matrix that tracks number of species in system over time
%Initialize with initial values
outMat = [t,N_PFN,N_GZB, Lost(1),Lost(2),Num_pore_formed,(Num_pore_formed > 0),N_GzB_Int,pVal];

%tOut is next time at which to record state of system, occurs at intervals
%of outIncr
tOut = outIncr;

while(t < T)

    %First check if system is 'stuck', e.g. no 1,2,3mers, no pores, no
    %free solution molecules. In this case nothing can happen in system,
    %but higher order PFN jmers will diffuse forever. To avoid this useless
    %simulation, terminate.

    monomer = sum(N(:,1));
    dimer = sum(N(:,2));
    trimer = sum(N(:,3));
    num_free = sum(N(:,n_pore+1)) + sum(N(:,n_pore+2));

    if(num_free == 0 && monomer == 0 && dimer == 0 && trimer == 0)

        stuckCount = 1;
        break;
    end

    %Otherwise, proceed according to SSSA

    %Next reaction will be in subvolume at top of event tree
    sv = Q(1,1);
    %Will occur at time t from subvolume at top of event tree
    t = Q(1,2);

    %State vector for that subvolume
    Nsv = N(sv,:);

    %Total interaction propensities. Calculated as above.
    a0 = R(sv,1);

    %Total propensity
    s0 = R(sv,3);

    %Choose if event is reaction or diffusion
    if(rand < a0/s0)
        %Reaction

```

```

%Propensity vector for that subvolume (it was calculated when an event
%last occurred that influenced the subvolume). Or, in the case of
%ICs, it was calculated during initialization.

a_sv = A(sv,:);

    num_pore_prior = Nsv(n_pore);

%Determine reaction that occurs
rxn_ind = find(cumsum(a_sv) > rand*a0, 1, 'first');

%Track internalized GzB
if (rxn_ind == n_rxns)
    N_GzB_Int = N_GzB_Int + 1;
end

%Update species numbers, and update state matrix
Nsv = Nsv + S(rxn_ind,:);
N(sv,:) = Nsv;

    num_pore_post = Nsv(n_pore);

%Track maxmer
[Temp,pfn_oli_prod] = max(S(rxn_ind,1:n_pore));
if (pfn_oli_prod > maxmer)
    maxmer = pfn_oli_prod;
end

%Recalculate propensities for this subvolume
a_sv = calcIntrxnPropens(Nsv,n_pore,K, K_pfn_ins,Kg);
d_sv = calcDiffPropens(Nsv, n_pore, D_Pfn_jmer, D_Pfn_eff, D_GzB_eff, l);

%Update propensity matrices
A(sv,:) = a_sv;
D(sv,:) = d_sv;

%Update rate matrix
a0 = sum(a_sv);
d0 = sum(d_sv);
s0 = a0 + d0;
R(sv,:) = [a0 d0 s0];

%Track time of pore formation
if ((num_pore_post - num_pore_prior) > 0)
    T_pore_form = [T_pore_form t];
    Num_pore_formed = Num_pore_formed + 1;
end

%Calculate next event time for this cell based on updated state
tau = -log(rand)/s0;

%Working in absolute time, so next event time is added to time that

```

```

%already occurred
tnext = t + tau;

%Update event tree
[Q, Qarray] = updateQ(1,tnext,Q,Qarray);

else
    %Diffusion

    %Diffusive propensity for subvolume
    d_sv = D(sv,:);

    %Total diffusive propensities
    d0 = R(sv,2);

    %species that diffuses
    species_ind = find(cumsum(d_sv) > rand*d0, 1, 'first ');

    %Determine which adjacent subvolume particle jumps to (N,S,E,W,
    %with equal probability)
    rn = rand;

    if(rn < 0.25)
        %North
        %SV above is current cell volume index + k, where k is number
        %of cells on side length of square
        svd = sv + n_cells_per_side;

    elseif(rn >= 0.25 && rn < 0.5)
        %East
        %SV left of current cell is sv - 1
        svd = sv - 1;

    elseif(rn >= 0.5 && rn < 0.75)
        %West
        %SV right of current cell is sv + 1
        svd = sv + 1;
    else
        %South
        %SV above is current cell volume index - l, where l is number
        %of cells on side length of square
        svd = sv - n_cells_per_side;
    end

    %Check if destination subvolume is external to synapse

    if( !sInt(svd) == 0)

        %Bulk species are allowed to escape
        if(species_ind > n_pore)

            %In this case remove species from origin subvolume, update
            %state matrix
            Nsv(species_ind) = Nsv(species_ind)-1;

```

```

N(sv,:) = Nsv;

%DO NOT add species to destination subvolume, it is lost
%Instead update lost particle counter
Lost(species_ind - n_pore) = Lost(species_ind - n_pore) + 1;

%Recalculate propensities for this subvolume
a_sv = calcIntrxnPropens(Nsv,n_pore,K, K_pfn_ins,Kg);
d_sv = calcDiffPropens(Nsv, n_pore, D_Pfn_jmer, D_Pfn_eff, D_GzB_eff, l);

%Update propensity matrices
A(sv,:) = a_sv;
D(sv,:) = d_sv;

%Update rate matrix
a0 = sum(a_sv);
d0 = sum(d_sv);
s0 = a0 + d0;
R(sv,:) = [a0 d0 s0];

%Calculate next event time for this cell based on updated state
tau = -log(rand)/s0;

%Working in absolute time, so next event time is added to time that
%already occurred
tnext = t + tau;

%Update event tree
[Q, Qarray] = updateQ(1,tnext,Q,Qarray);

end

%Membrane bound pfn oligos not allowed to escape hence they are
%kept in current volume. Therefore nothing changes, so no need
%to update Q, propensities etc.

else
%If jumping to internal subvolume, update both cells

%Destination subvolume state vector
Nsvd = N(svd,:);

%Update origin and destination subvolume state vectors
Nsv(species_ind) = Nsv(species_ind)-1;
Nsvd(species_ind) = Nsvd(species_ind) + 1;

N(sv,:) = Nsv;
N(svd,:) = Nsvd;

%Recalculate propensities for both subvolumes
a_svor = calcIntrxnPropens(Nsv,n_pore,K, K_pfn_ins,Kg);
d_svor = calcDiffPropens(Nsv, n_pore, D_Pfn_jmer, D_Pfn_eff, D_GzB_eff, l);

a_svddest = calcIntrxnPropens(Nsvd,n_pore,K, K_pfn_ins,Kg);

```

```

d_svdest = calcDiffPropens(Nsvd, n_pore, D_Pfn_jmer, D_Pfn_eff, D_GzB_eff, l);

%Update propensity matrices
A(sv,:) = a_svor;
D(sv,:) = d_svor;
A(svd,:) = a_svdest;
D(svd,:) = d_svdest;

%Update rate matrix
a0o = sum(a_svor);
d0o = sum(d_svor);
s0o = a0o + d0o;
R(sv,:) = [a0o d0o s0o];

a0d = sum(a_svdest);
d0d = sum(d_svdest);
s0d = a0d + d0d;
R(svd,:) = [a0d d0d s0d];

%Calculate next event time for subvolumes
tau_o = -log(rand)/s0o;
tau_d = -log(rand)/s0d;

%Working in absolute time, so next event time is added to time that
%already occurred
tnext_o = t + tau_o;
tnext_d = t + tau_d;

%Update event tree
%Origin volume is by definition in head node
[Q, Qarray] = updateQ(1,tnext_o,Q,Qarray);

%Destination volume found from Qarray
Qpos = Qarray(svd);
[Q, Qarray] = updateQ(Qpos,tnext_d,Q,Qarray);

end
end

%Having moved forward in time, check if far enough to require
%outputting data

if (t>tOut)

    outMat = [outMat;[tOut,sum(N(:,n_pore+1)),sum(N(:,n_pore+2)),Lost(1),Lost(2),
        Num_pore_formed,(Num_pore_formed > 0),N_GzB_Int,pVal]];
    tOut = tOut + outIncr;

end

end
end

```

```

%Write time course summary
csvwrite( strcat( sim_num, '_timeCourse.csv' ), outMat );

%Save workspace
save( 'workspace.mat' );

%Print comma seperated summary line
varVect = [N_PFN N_GZB Lost(1) Lost(2) maxmer Num_pore_formed N_GzB_Int D_GzB D_GzB_eff D_Pfn
           D_Pfn_eff K_pfn_ins Kd Kg V_avg_mol frac_occ str2num(sim_num) pVal hdc];
csvwrite( 'simVarSummary.csv', varVect );

%Print endpoint summary data file
FID = fopen( 'summary.txt', 'w' );

%TimeStamp
fprintf( FID, '%s\n', datestr( clock ) );

%JobID
fprintf( FID, 'JobID:%s\n', sim_num );

%Working directory
fprintf( FID, 'WD:%s\n', pwd );

%RNG seed
fprintf( FID, 'RNG seed:%f\n', sd );

%Total length of simulation (biophysical time)
fprintf( FID, 'Total Biophysical Sim Time = %f\n', t );

%System Stuck
fprintf( FID, 'System stuck? %d\n', stuckCount );

%Maximum PFN oligo
fprintf( FID, 'Maxmer = %f\n', maxmer );

%Number of pores formed
fprintf( FID, 'Total Num Pores = %f\n', Num_pore_formed );

%Time of pore formation (0 if none form)
t = 0;
for t = T_pore_form
    fprintf( FID, 'Pore formed at time = %f\n', t );
end

%Lost Molecules
fprintf( FID, 'Number of molecules escaped synapse:\n' );
fprintf( FID, 'Perforin = %d/%d\n', Lost(1), N_PFN );
fprintf( FID, 'Granzyme = %d/%d\n', Lost(2), N_GZB );

%Amount of Granzyme Internalized
fprintf( FID, 'Number of granzyme molecules internalized = %d\n', N_GzB_Int );

%Calculate positions of pores
for i = 1:n_cells

```

```

        if (N(i,n_pore) > 0)
            fprintf(FID, 'Location of pores:\n');
            fprintf(FID, 'XY Coords: (%f,%f). Number: %d\n', SVcoords(i,1), SVcoords(i,2), N(i,n_pore));
        end
    end
end

elapsedTime = toc;

fprintf(FID, 'Elapsed CPU time = %f\n', elapsedTime);
fprintf(FID, '\n');
fclose(FID);

end

```

D.1.2 buildQ.m

```

function [Q, Qarray] = buildQ(initList, oldQarray)

Q = initList;
n = length(Q);
i = floor(n);

if (mod(n,2) == 0)
    Q(n+1,1) = n+1;
    Q(n+1,2) = inf;
    oldQarray(n+1) = n+1;
end

while (i > 0)

    [Q, oldQarray] = percDown(i, Q, oldQarray);
    i = i-1;
end

Qarray = oldQarray;
end

```

D.1.3 calcDiffPropens.m

```

function D = calcDiffPropens(N, n_pore, D_Pfn_jmer, D_Pfn, D_GzB, l)

%Calculate total propensities for diffusive jump for each species in a
%given subvolume.
%This is defined as  $d_i = (D_i/l^2)*X_i$ , where  $D_i$  is the diffusivity of
%the species,  $l$  is the length of the mesh cell and  $X_i$  is the number of
%species in the subvolume

%Diffusion propensity vector, column order is same as state matrix N
D = [];

%First do all membrane bound pfn oligos

```

```

for i = 1:n_pore

    di = N(i)*D_Pfn_jmer(i);
    D = [D di];

end

%Free perforin
di_pfn = N(n_pore+1)*D_Pfn;

%Free perforin
di_gzb = N(n_pore+2)*D_GzB;

D = [D di_pfn di_gzb];

D = 4*D/(l^2);

end

```

D.1.4 calcInterxnpromens.m

```

function A = calcIntrxnPropens(N, n_pore, K, K_pfn_ins, Kg)

%Calculate total propensities for each interaction in subvolume
%This is the intrinsic probability of the rxn (c_i) times the number of
%possible reactions (the product of the number of reactants)
%a0 is the sum of all probabilities

%Propensity vector
A = [];

%Construct propensities for pfn oligomerization first
%Do in three iterations, one for each 1,2,3mer that can add to chain
for j=1:3
    for i=j:(n_pore-j)

        %Exception in the case of two identical oligos combining
        %If only 1 oligo, then propensity will be non-zero because
        %c = Ni*Ni = Ni^2 = 1
        %But cannot have single oligo combine with itself.
        %Manually set propensity to zero
        if(i == j && N(i) == 1)
            subA = 0;
            A = [A;subA];
            continue;
        end

        %Calculate propensity: ai = ci*hi; ci is intrinsic rate
        %with units of s^-1; hi is combinatorial number of ways this
        %reaction can occur, e.g. product of two species numbers
        c = K(i,j,N(j));
        h = N(j)*N(i);
        subA = c*h;
    end
end

```

```

        A = [A;subA]; % append
    end

end

%Propensity for pfn insertion
subA = K_pfn_ins*N(n_pore+1);
A = [A;subA];

%Propensity for Granzyme translocation
subA = Kg*N(n_pore)*N(n_pore+2);
A = [A;subA];

A = A';

end

```

D.1.5 percDown.m

```

function [newList, newQarray] = percDown(i, list, oldQarray)

Lchild = @(i) 2*i;
Rchild = @(i) 2*i+1;

while(2*i <= length(list)) %E.g. while current node still has child

    lc = Lchild(i);
    rc = Rchild(i);

    if(list(lc,2) < list(rc,2))
        mc = lc;
    else
        mc = rc;
    end

    if(list(mc,2) < list(i,2))

        [list, oldQarray] = swap(i,mc,list,oldQarray);
    else
        break;
    end

    i = mc;

end

newList = list;
newQarray = oldQarray;

end

```

D.1.6 Qsort.m

```
function [Q, Qarray] = Qsort(i,oldQ, oldQarray)

Lchild = @(i) 2*i;
Rchild = @(i) 2*i+1;
parent = @(i) floor(i/2);

Q = oldQ;
Qarray = oldQarray;

p = parent(i);
if (p<1)
    p = 1;
end

if (Q(i,2) < Q(p,2))

    [Q, Qarray] = swap(i ,p,Q, Qarray);
    [Q, Qarray] = Qsort(p,Q, Qarray);

elseif(2*i <= length(Q) && Q(i,2) > min(Q(Lchild(i),2),Q(Rchild(i),2)))

    [Q, Qarray] = percDown(i ,Q, Qarray);
end

end
```

D.1.7 swap.m

```
function [newlist, newQarray] = swap(i,j,list, Qarray)

Qarray(list(i,1)) = j;
Qarray(list(j,1)) = i;
newQarray = Qarray;

temp = list(i,:);
list(i,:) = list(j,:);
list(j,:) = temp;
newlist = list;

end
```

D.1.8 updateQ.m

```
function [Q, Qarray] = updateQ(i, val, oldQ, oldQarray)

Q = oldQ;
Q(i,2) = val;
```

```
[Q, Qarray] = Qsort(i,Q,oldQarray);
```

```
end
```

D.2 Image filtration and colocalization analysis: MATLAB source code (Chapter 3)

This section contains the MATLAB source code that filters raw images. The output images from this filtering were then used as input for a second script which quantified the colocalization, but that is simple, standard code using standard algorithms, and is not included.

```
%-----%
%—Name: LAGfilter_batch.m
%—Author: Daniel Woodsworth
%—Date: March 3, 2017
%
%LAG = local and global filter
%
%Basic idea is to account for local variations in background signal, but
%also to achieve robust filtering of noise (so local and global filtering)
%
%To do this, first calculate local background for each pixel, subtract this
%background from original image.
%
%Then calculate global median absolute deviation. Use this as estimate of
%variance of pixel noise, and define threshold as some multiple of this.
%Define all pixels below this threshold as noise, and set to zero.

%Local pixel intensity idea comes from Dunn et al 2011
%Median absolute deviation idea comes from Josh Scuri

clear all
close all
clc

#####PARAM SET HERE#####
%Set paths here for input and output directories
inPath = '/Volumes/DAN/coloc/RAW/';
outPath = '/Volumes/DAN/coloc/FILT/';

#####PARAM SET HERE#####
%length on either side of current pixel that to make box for median
%calculation
L = 12;

#####PARAM SET HERE#####
%Number of standard deviations above MAD to consider as above background
%These are empirical
C0_threshold = 3;
C1_threshold = 6;

#####PARAM SET HERE#####
```

```

%Regex pattern for extracting unique identifier of sample image (e.g.
%within given sample, whatever IDs the actual image files)
CaptureNumberPattern = '^Capture (\d+)_';

%Get directories in basepath
tmp = dir(inPath);
dirs = {tmp.name};
dirs([1,2]) = []; %Delete . and ..

for dirindex = 1:length(dirs)

    %Assume each directory name is sample name
    SampleName = dirs{dirindex};
    path = fullfile(inPath,dirs{dirindex});
    writePath = fullfile(outPath,dirs{dirindex});

    %Get all image file names, assuming tiff extension
    im_files = dir(fullfile(path,'*.tiff'));
    fnames = {im_files.name};

    %Get all unique captures. Each capture has separate tiff for each
    %channel. Assume format of Name_Cx.tiff, where x = 0,1 is channel id.
    %Strip channel id to get actual name, then strip duplicates.
    [temp, basenames, extensions] = cellfun(@fileparts,fnames, 'UniformOutput', false);
    ImageNames = cellfun(@(x) x(1:end-3), basenames, 'UniformOutput', false);
    ImageNames = unique(ImageNames);

    mkdir(writePath);

    SampleName

    for i = 1:length(ImageNames)

        %Get unique identifier of image file name
        [tokens,matches] = regexp(ImageNames{i},CaptureNumberPattern,'tokens','match');
        CaptureNumber = cell2mat(tokens{1});

        CaptureNumber

        C0_img_name = fullfile(path, strcat(ImageNames{i},'_C0',extensions{1}));
        C1_img_name = fullfile(path, strcat(ImageNames{i},'_C1',extensions{1}));

        %% Read image
        imf = imfinfo(C0_img_name);
        C0_img = zeros(imf(1).Height, imf(1).Width, 'single');
        C0_img(:, :) = imread(C0_img_name);

        imf = imfinfo(C1_img_name);
        C1_img = zeros(imf(1).Height, imf(1).Width, 'single');
        C1_img(:, :) = imread(C1_img_name);

        %——C0 = green channel = lytic granule filter——%

```

```

%Useful for punctuate type objects

%min and max pixels to consider, to account for edge cases
min_pix = L+1;
max_pix = size(C0_img,1)-L;

%Initialize background matrix to maximum pixel intensity
bg = repmat(max(C0_img(:)),size(C0_img,1),size(C0_img,1));

%Calculate localized background pixel intensity for whole image
for k = min_pix:max_pix
    for j = min_pix:max_pix

        sub = C0_img((k-L):(k+L),(j-L):(j+L));

        bg(k,j) = median(sub(:));

    end
end

%Subtract local background from image
C0_im_bgfilt = C0_img-bg;

%Set any negative pixels to 0
C0_im_bgfilt(C0_im_bgfilt < 0) = 0;

%Now subtract MAD from image to filter pixel noise
%See below for more information
%Since so many pixels are 0, use all nonzero pixels to calculate MAD

%I don't know why using the original image works better, but it cleans up
%pixel noise better (essentially a higher threshold)

C0_pos = C0_img(C0_img > 0);
mad_C0 = median(abs(C0_pos(:) - median(C0_pos(:)))));
noise_std_C0 = 1.4826 * mad_C0; %See below for explanation

%Noise filter. Set all pixels below threshold to 0
C0_mad_sub = C0_im_bgfilt;

C0_mad_sub(C0_mad_sub < noise_std_C0 * C0_threshold) = 0;

%C1 = RED = mCherry channel
%——Good for homogenous structures———%
%
%Just do MAD threshold, but with 6sigma

C1_pos = C1_img(C1_img > 0);

```

```

mad_C1 = median(abs(C1_pos(:) - median(C1_pos(:)))));
noise_std_C1 = 1.4826 * mad_C1; %See below for explanation

%Noise filter. Set all pixels below 6 sigma threshold to 0
C1_mad_sub = C1_img;

C1_mad_sub(C1_mad_sub < noise_std_C1 * C1_threshold) = 0;

%Write images
C0_OUT = fullfile(writePath, strcat(ImageNames{i}, '_filt_C0', extensions{1}));
C1_OUT = fullfile(writePath, strcat(ImageNames{i}, '_filt_C1', extensions{1}));

imwrite(uint16(C0_mad_sub), C0_OUT);
imwrite(uint16(C1_mad_sub), C1_OUT);

end

end

%—————Background on MAD derivation etc—————%

% Convert the median absolute deviation to something more resembling a
% standard deviation. This is used instead of calculating the standard
% deviation directly because the median absolute deviation ignores outliers
% (and actual spots), giving a result after normalisation closer to the
% standard deviation of noise alone, with other image features ignored.
% (See http://en.wikipedia.org/wiki/Median\_absolute\_deviation#Relation\_to\_standard\_deviation
% for the conversion factor)

%calculate the absolute median deviation
%This is mad = median(abs(Intensity_i - median(Intensity)))
%In other words: calculate median intensity of images using all pixels.
%Subtract this from each pixel.
%Take absolute value of these values
%Calculate median of these values
%
%Sigma ~1.4mad (see wikipedia article on mad)
%To be very sure that retain only signal, take all pixels that have
%intensity that is 6*sigma above 1.4*mad

```

Appendix E

Plasmids

E.1 Granzyme B derived chaperone-mCherry fusion protein plasmids (Chapter 3)

E.1.1 Base pdL plasmid map

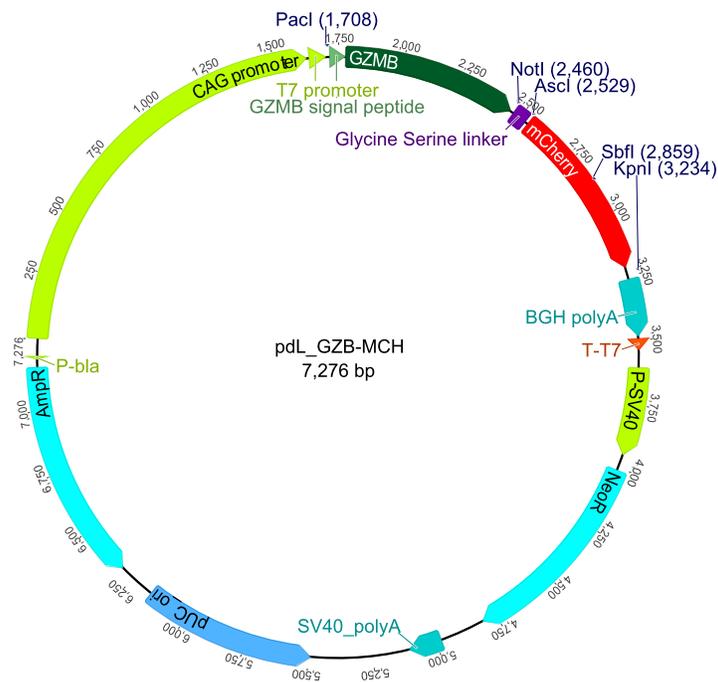


Figure E.1: Base pdL vector. In this example the insert coding sequence is the granzyme B-crmCherry fusion protein (GZB-MCH). The annotations correspond to those in the genbank file.

E.1.2 Base pdL plasmid sequence

LOCUS pdL_GZB-MCH 7276 bp DNA circular UNA 21-FEB-2017
DEFINITION .
ACCESSION .
VERSION .
KEYWORDS .
SOURCE
ORGANISM

FEATURES Location/Qualifiers
promoter 2..1628
/label="CAG promoter"
promoter 1635..1702
/label="T7 promoter"
sig_peptide 1717..1776
/gene="GZMB"
/label="GZMB signal peptide"
mat_peptide 1777..2457
/gene="GZMB"
/product="granzyme B"
/label="GZMB"
misc_feature 2467..2526
/note="Geneious type: polylinker"
/label="Glycine Serine linker"
CDS 2536..3216
/label="mCherry"
polyA_signal 3262..3486
/label="BGH polyA"
terminator 3495..3538
/label="T-T7"
promoter 3606..3949
/label="P-SV40"
misc_feature 4011..4805
/note="Geneious type: Marker"
/label="NeoR"
polyA_signal 4979..5109
/label="SV40_polyA"
misc_feature complement(5492..6162)
/note="Geneious type: Origin of Replication"
/label="pUC_ori"
misc_feature complement(6307..7167)
/note="Geneious type: Marker"
/label="AmpR"
promoter complement(7202..7208)
/label="P-bla"

ORIGIN

```
1 ctgcaggcgt tacataactt acggtaaatg gcccgctgg ctgaccgccc aacgaccccc
61 gccattgac gtcaataatg acgtatgttc ccatagtaac gccaataggg actttccatt
121 gacgtcaatg ggtggagtat ttacggtaaa ctgcccactt ggcagtacat caagtgtatc
181 atatgccaag tacgccccct attgacgtca atgacggtaa atggcccgcc tggcattatg
241 cccagtacat gaccttatgg gactttccta cttggcagta catctacgta ttagtcatcg
301 ctattaacat ggtcgagggt agccccacgt tctgcttcac tctccccatc tcccccccct
361 cccaccccc aattttgtat ttatttattt tttaattatt ttgtgcagcg atgggggagg
```

421 gggggggggg ggggcgcgcg ccaggcgggg cggggcgggg cgaggggagg ggcggggcga
481 ggcggagagg tgcggcgca gccaatcaga ggcggcgct ccgaaagttt cctttatgg
541 cgaggcgcg ggcggcgcg ccctataaaa agcgaagcgc ggcggggggg gggagtcgct
601 ggcagctgc cttcgcccc tgccccctc cgcccgccgc tcgcccgcgc cgccccggct
661 ctgactgacc gcgttactcc cacaggtgag cggcggggac ggcccttctc ctccgggctg
721 taattagcgc ttggttfaat gacggctgtt tcttttctg tggctgcgtg aaagcctga
781 ggggctccgg gagggccctt tggcggggg gagcggctcg gggggtgctg gcgtgtgtgt
841 gtgcgtgggg agcggccgct gcgctccgc gctgcccgcc ggctgtgagc gctcggggcg
901 cggcggggg cttgtgcgc tccgagtggt gcgcagggg agcggcgccg gggcggtgc
961 ccccggtgac ggggggggct gcgaggggaa caaaggctgc gtgcggggtg tgtcgtggg
1021 ggggtgagca gggggtgtgg gcgctcggt cgggctgcaa cccccctgc accccctcc
1081 ccgagttgct gacacggcc cggcttcggg tgcggggctc cgtacggggc gtggcgggg
1141 gctcgccgtg ccgggcccc ggtggcgca ggtgggggtg ccgggccccg cggggccgc
1201 tggggccggg gagggctcgg gggagggggc cggcgcccc cggagcgccg gcggtgtcg
1261 aggcgcgcg agccgagcc attgccttt atggtatcg tgcgagaggc cgcagggact
1321 tctttgttc caaatctgtg cggagccgaa atctggagg cgcccgccga cccccctag
1381 cggcgcggg gcgaagcgtt ggcggcgccg caggaaggaa atgggccccg agggcctcg
1441 tgcgtgcgg cgccgcccgc ccttctccc tctccagcct cggggtgtc cgcgggggga
1501 cggctgcctt cgggggggac gggcagggc ggggttcggc ttctgcgtg tgaccggcg
1561 ctctagacaa tigtactaac ctctctctt tctctctct gacaggttgg tgtacagtag
1621 cttccaccac cggttaatac gactcactat aggctagcat ttaggtagaca ctatagaata
1681 caagctactt gttcttttt cattaattaa gccgccatgc aaccaatcct gcttctgctg
1741 gccttctcc tgcgtcccag ggcagatgca ggggagatca tgggggaca tgaggccaag
1801 cccactccc gccctacat ggcttatctt atgatctggg atcagaagtc tctgaagagg
1861 tgcggtggtt tctgatacg agacgacttc gtgctgacag ctgctactg ttggggaagc
1921 tccataaatg tcacctggg ggcaccacat atcaagaac aggagccgac ccagcagttt
1981 atccctgtga aaagaccat cccccatca gcctataatc ctaagaactt ctccaacgac
2041 atcatgtctc tgcagctgga gagaaaggcc aagcggacca gagctgtgca gccctcagg
2101 ctacctagca acaaggccca ggtgaagcca ggcagacat gcagtggtgc cggctggggg
2161 cagacggccc cctgggaaa aactcacac aactacaag aggtgaagat gacagtgag
2221 gaagatcgaa agtgcgaatc tgaactacgc cattattacg acagtaccat tgagttgtc
2281 gtgggggacc cagagattaa aaagacttcc ttaaggggg actctggagg ccctctgtg
2341 tgaacaagg tggcccaggc cattgtctcc tatggagcaa acaatggcat gcctccacga
2401 gcctgacca aagtctcaag cttgtacac tggataaaga aaacctgaa acgctacgg
2461 gccgctggag gtgggggttc tggcgggggt ggatcagggg gtggaggttc cggtgagggt
2521 gggctggggc gcgccatcat caaggagttc atgcgcttca aggtgcacat ggagggtcc
2581 gtgaaccggc acgagttcga gatcgagggc gaggcgagg ggcggcccta cgagggcacc
2641 cagaccgcca agctgaaggt gaccaagggt ggccccctgc ccttcgctg ggacatcctg
2701 tccccctagt tcatgtacgg ctccaaggcc tacgtgaagc acccgccga catccccgac
2761 tacttgaagc tctcttccc cgagggttc aagtgaggc gcgtgatgaa ctgcaggac
2821 ggcggtgtg tgaccgtgac ccaggactcc tccctgcagg acggcgagtt catctacaag
2881 gtgaagctgc gcggaccaca ctccccctc gacggccccg taatgcagaa gaagacctg
2941 ggctgggagg cctcctccga gcggatgtac cccgaggagc gcgccctgaa gggcgagatc
3001 aagcagaggc tgaagctgaa ggacggcgcc cactacgagc ctgaggtaa gaccacctac
3061 aaggccaaga agcccgtgca gctgccggc gcctacaacg tcaacatcaa gttggacatc
3121 acctcccaca acgaggacta caccatcgtg gaacagtacg aacgcgccga gggcgccac
3181 tccaccggcg gcatggacga gctgtacaag gaattctaat agctcgaggg taccactagt
3241 tagcccgtg atcagcctcg actgtgcctt ctagttgcca gccatctgtt gttgcccct
3301 cccccgtgcc ttccttgacc ctggaagggt ccactcccac tgcctttcc taataaaatg
3361 aggaatgtc atcgcaattg ctgagtaggt gtcattctat tctgggggtt ggggtggggc
3421 aggacagcaa gggggaggat tgggaagaca atagcaggca tgcgtgggat gcggtgggct
3481 ctatgttaac gcgtataacc cctggggcc tctaaacggg tcttgagggg tttttgact
3541 cgagcgattt cggcctattg gttaaaaat gagctgattt acaaaaaat taacgcaat
3601 taattctgtg gaatgtgtg cagttagggt gtggaaagtc cccaggctcc ccagcaggca

3661 gaagtatgca aagcatgcat ctcaattagt cagcaaccag gtgtggaag tccccaggct
3721 ccccagcagg cagaagtatg caaagcatgc atctcaatta gtcagcaacc atagtcccg
3781 ccctaactcc gcccatccc cccctaactc cgcccagttc cgcccattct cggcccatg
3841 gctgactaat ttttttatt tatgcagagg ccgaggccgc ctctgcctct gagctattcc
3901 agaagtagtg aggaggcttt ttggaggcc taggctttt caaaaagctc cgggagctt
3961 gtataccat ttccgatct gatcaagaga caggatgagg atcgtttcgc atgattgaac
4021 aagatggatt gcacgcagg tctccggcc cttgggtgga gaggctattc ggctatgact
4081 gggcacaaca gacaatcggc tgctctgat cgcctgtt cggctgtca gcgcaggggc
4141 gcccgttct tttgtcaag accgacctgt ccggtgccct gaatgaactg caggacgagg
4201 cagcgcggct atcgtggctg gccacgacgg gcgttccttg cgcagctgtg ctgcacgttg
4261 tactgaagc ggaaggac tggtgctat tggcggaagt gccgggagc gatctcctgt
4321 catctcacct tgctcctgcc gagaaagtat ccatcatggc tgatgcaatg cggcggctgc
4381 atacgcttga tccggctacc tgcccattcg accaccaagc gaaacatcgc atcgagcgag
4441 cacgtactcg gatggaagcc ggtcttctg atcaggatga tctggacgaa gagcatcagg
4501 ggctcgcgcc agccgaactg ttccgacggc tcaaggcgcg catgcccgcg ggcgaggatc
4561 tcgtcgtgac ccatggcgat gcctgcttgc cgaatatcat ggtggaaaat ggccgctttt
4621 ctggattcat cgactgtggc cggctgggtg tggcggaccg ctatcaggac atagcgttgg
4681 ctaccctgat tattgctgaa gagcttggcg gcgaatgggc tgaccgcttc ctctgcttt
4741 acggtatcgc cgctcccgat tcgcagcga tcgccttcta tcgccttctt gacgagttct
4801 ctgagcggg actctggggt tcgaaatgac cgaccaagcg acgccaacc tgccatcacg
4861 agatttcgat tccaccgcc ccttctatga aagggtgggc ttcggaatcg tttccggga
4921 cgcggctg atgatctcc agcgcgggga tctcatgctg gagttcttcg cccaccccaa
4981 cttgtttatt gcagcttata atggttaca ataaagcaat agcatcaca atttcacaaa
5041 taaagcattt tttcactgc attctagtgg tggttgtcc aaactcatca atgtatctta
5101 tcatgtctgt ataccgtcga cctctagcta gagcttggcg taatcatggt catagctgtt
5161 tcctgtgtga aattgttata cgctcacaat tccacacaac atacgagccg gaagcataaa
5221 gtgtaaagcc tggggtgctt aatgagttag ctaactcaca ttaattgcgt tgcgctcact
5281 gcccgccttc cagtccggaa acctgtcgtg ccagctgcat taatgaatcg gccaacgcgc
5341 ggggagagcc ggtttcgta ttgggcgctc ttccgcttcc tcgctcactg actcgtcgcg
5401 ctcggtcgtt cggctcggcg gagcggatc agctcactca aaggcggtaa tacggttata
5461 cacagaatca ggggataaac caggaaagaa catgtgagca aaaggccagc aaaaggccag
5521 gaaccgtaaa aaggccgcgt tgcctggcgtt ttccatagg ctccgcccc ctgacgagca
5581 tcacaaaaat cgacgctcaa gtcagaggtg gcgaaaccgg acaggactat aaagatacca
5641 ggcgtttccc cctggaagct ccctcgtgcg ctctcctgtt cccaccctgc cgcttaccgg
5701 atacctgtcc gcctttctcc ctccgggaag cgtggcgctt tctcatagct cacgctgtag
5761 gtatctcagt tcggtgtagg tcgttcgctc caagctgggc tegtgtcacg aacccccgt
5821 tcagcccgcg cgctgcgctt tatccggtaa ctatcgtctt gagtccaacc cggtagaca
5881 cgacttatac cactggcag cagccactgg taacaggatt agcagagcga ggtatgtagg
5941 cgggtctaca gatttctga agtgggtggc taactacggc tacactagaa gaacagtatt
6001 tggatctgc gctctgctga agccagttac ctccgaaaa agagtggta gctcttgatc
6061 cggcaaaaa accaccgctg gtacgggtt tttgtttgc aagcagcaga ttacgocgag
6121 aaaaaaagga tctcaagaag atcctttgat ctttctacg ggtctgacg ctactggaa
6181 cgaaaactca cgttaaggga tttgtgcat gagattatca aaaaggatct tcacctagat
6241 ccttttaaat taaaaatgaa gttttaaatac aatctaaagt atatatgagt aaacttggtc
6301 tgacagttac caatgcttaa tcagtgaggc acctatctca gcgatctgtc tattcgttc
6361 atccatagtt gcctgactcc ccgtcgtgta gataactacg atacgggagg gcttaccatc
6421 tggccccagt gctgcaatga taccgcgaga cccacgctca cggctccag atttatcagc
6481 aataaacccg ccagccggaa gggccgagcg cagaagtggc cctgcaactt tatccgcctc
6541 catccagctc attaattgtt gccgggaagc tagagtaagt agttccagc ttaaatagttt
6601 gcgcaacggt gttgccattg ctacaggcat cgtgggtgca cgctcgtctg ttggtatggc
6661 ttcattcagc tccggttccc aacgatcaag gcgagttaca tgatcccca tgttgtcaa
6721 aaaaagcgtt agctccttcg gtctccgat cgttgtcaga agtaagtgg cgcaggtgtt
6781 atcactcatg gttatggcag cactgcataa ttctcttact gtcatgcat ccgtaagatg
6841 cttttctgtg actggtgagt actcaaccaa gtcattctga gaatagtga tgcggcgacc

6901 gagttgctct tgcccggcgt caatacggga taataccgcg cccatatgca gaactttaa
6961 agtgctcadc attgaaaac gttcttggg gcgaaaactc tcaaggatct taccgctggt
7021 gagatccagt tcgatgtaac ccaactcgtc acccaactga tcttcagcat cttttacttt
7081 caccagcgtt tctgggtgag caaaaacagg aaggcaaat gccgcaaaa agggataag
7141 ggcgacacgg aatgttgaa tactcatact cttcctttt caatattatt gaagcattta
7201 tcagggttat tgtctcatga gcggatacat attgaaatgt atttagaaaa ataacaat
7261 aggggttccg cgagct

//

E.1.3 Coding sequence inserts

Coding sequences listed below extend from the PacI site to KpnI site of the pdL vector.

>GZB-MCH

TTAATTAAGCCGCCATGCAACCAATCCTGCTTCTGCTGGCCTTCTCCTGCTGCCAGGGCAGATGCAGGGGAGATCATC
GGGGGACATGAGGCCAAGCCCCACTCCCGCCCTACATGGCTTATCTTATGATCTGGGATCAGAAGTCTCTGAAGAGGTG
CGGTGGCTTCTGATACGAGACGACTTCGTGCTGACAGCTGCTCACTGTTGGGGAAGCTCCATAAATGTCACCTTGGGGG
CCCACAATATCAAGAACAGGAGCCGACCCAGCAGTTTATCCCTGTGAAAAGACCCATCCCCATCCAGCCTATAATCCT
AAGAAGTCTTCCAACGACATCATGCTACTGACGCTGGAGAGAAAGGCCAAGCGGACCCAGAGCTGTGCAGCCCCCTCAGGCT
ACCTAGCAACAAGGCCCAGGTGAAGCCAGGCAGACATGCAGTGTGGCCGGCTGGGGGCAGACGGCCCCCTGGGAAAAC
ACTCACACACTACAAGAGGTGAAGATGACAGTGCAGGAAGATCGAAAGTGCGAATCTGACTTACGCCATTATTACGAC
AGTACCATTGAGTTGTCCGTGGGGACCCAGAGATTA AAAAGACTTCTTTAAGGGGACTCTGGAGGCCCTCTTGTGTG
TAACAAGGTGGCCAGGGCATTGTCTCTATGACGAAACAATGGCATGCCTCCACGAGCCTGCACCAAAGTCTCAAGT
TTGTACACTGGATAAAGAAAACCATGAAAACGCTACGCGGCCCTGGAGGTGGGGTTCCTGGCGGGGTGGATCAGGGGGT
GGAGGTTCCGGTGGAGGTGGGTCCGGCGCCCATCATCAAGGAGTTCATGCGCTTCAAGGTGCACATGGAGGGCTCCGT
GAACCGCCACGAGTTCCGAGATCGAGGGCAGGGGCGAGGGCCGCCCTACGAGGGCACCCAGACCGCCAAAGCTGAAGGTGA
CCAAGGTGGCCCCCTGCCCTTCGCTGGGACATCCTGTCCCTCAGTTCATGTACGGCTCCAAGGCCTACGTGAAGCAC
CCCGCCGACATCCCGACTACTTGAAGCTGTCTTCCCGAGGGCTTCAAGTGGGAGCGGTGATGAACCTCGAGGAOCCG
CGGGTGGTGAOCGTGACCCAGGACTCTCCCTGCAGGACGGGAGTTCATCTACAAGGTGAAGCTGCGCGGCCAACCT
TCCCTCCGAGGCCCGTAATGCAGAAGAAGACCATGGGCTGGGAGGCCCTCCGAGCGGATGTACCCCGAGGAGCCG
GCCCTGAAGGGGAGATCAAGCAGAGGCTGAAGCTGAAGGAOCCGCGGOCCTACGAGCTGAGGTCAAGACACCTACAA
GGCCAAAGAACCCGCTGCAGCTGCCCGGCGCTACAACGTCAACATCAAGTTGGACATCACCTCCACAAOCGAGGACTACA
CCATCGTGAACAGTACGAACGCGCCGAGGGCCGCACTCCACCGGGCCATGGACGAGCTGTACAAGGAATTCTAATAG
CTCGAGGGTACC

>GZBSS-MCH

TTAATTAAGCCGCCATGCAACCGATCTTGTGCTGCTGGCTTTTCTGCTGTTGCCAAGGGCAGACGCTGGAGAGGGCGCG
CCCATCATCAAGGAGTTCATGCGCTTCAAGGTGCACATGGAGGGCTCCGTGAACGGCCACGAGTTCCGAGATCGAGGGCGA
GGGCGAGGGCCGCCCTACGAGGGCACCCAGACCGCCAAAGCTGAAGGTGAACAAAGGGTGGCCCCCTGCCCTTCGCTGGG
ACATCCTGTCCCTCAGTTCATGTACGGCTCCAAGGCCTACGTGAAGCACCCCCGCCGACATCCCCGACTACTTGAAGCTG
TCTTCCCGAGGGCTTCAAGTGGGAGCGGTGATGAACCTCGAGGAOCCGCGGTGGTGAOCGTGACCCAGGACTCCTC
CCTGCAGGACGGCGAGTTCATCTACAAGGTGAAGCTGCGCGGCCAACCTTCCCTCCGAGCGGCCCGTAATGCAGAAGA
AGACCATGGGCTGGGAGGCCCTCCGAGCGGATGTACCCCGAGGAGCGGCCCTGAAGGGGAGATCAAGCAGAGGCTG
AAGCTGAAGGACGGCGGOCCTACGAGCTGAGGTCAAGACACCTACAAGGCCAAGAAGCCCGTGCAGCTGCCCGGCGC
CTACAACGTCAACATCAAGTTGGACATCACCTCCACAAOCGAGGACTACACCATCGTGAACAGTACGAACGCGCGGAGG
GCCGCACTCCACCGCGCATGGACGAGCTGTACAAGGAATTCTAATAGCTCGAGGGTACC

>GZBSS-MCH-GZBSM

TTAATTAAGCCGCCATGCAACCGATCTTGTGCTGCTGGCTTTTCTGCTGTTGCCAAGGGCAGACGCTGGAGAGGGCGCG
CCCATCATCAAGGAGTTCATGCGCTTCAAGGTGCACATGGAGGGCTCCGTGAACGGCCACGAGTTCCGAGATCGAGGGCGA
GGGCGAGGGCCGCCCTACGAGGGCACCCAGACCGCCAAAGCTGAAGGTGAACAAAGGGTGGCCCCCTGCCCTTCGCTGGG
ACATCCTGTCCCTCAGTTCATGTACGGCTCCAAGGCCTACGTGAAGCACCCCCGCCGACATCCCCGACTACTTGAAGCTG
TCTTCCCGAGGGCTTCAAGTGGGAGCGGTGATGAACCTCGAGGAOCCGCGGTGGTGAOCGTGACCCAGGACTCCTC
CCTGCAGGACGGCGAGTTCATCTACAAGGTGAAGCTGCGCGGCCAACCTTCCCTCCGAGCGGCCCGTAATGCAGAAGA
AGACCATGGGCTGGGAGGCCCTCCGAGCGGATGTACCCCGAGGAGCGGCCCTGAAGGGGAGATCAAGCAGAGGCTG
AAGCTGAAGGACGGCGGOCCTACGAGCTGAGGTCAAGACACCTACAAGGCCAAGAAGCCCGTGCAGCTGCCCGGCGC

```

CTACAACGTCAACATCAAGTTGGACATCACCTCCACAAACGAGGACTACACCATCGTGGAAACAGTACGAACGOGCCGAGG
GCOGCCACTCCACGGCGGCATGGACGAGCTGTACAAGGAATTCACAGCTGCTCACTGTTGGGAAGCTCCATAAATGTC
ACCTTGGGGGCCACAATATCAAAGAACAGGAGCCGACCCAGCAGTTTATCCCTGTGAAAAGACCCATCCCCATCCAGC
CTATAATCCTAAGAACTTCTCCAACGACATCATGCTACTGCAGCTGGAGAGAAAAGGGTACC
>MCH
TTAATTAAGCCGCCATGATCATCAAGGAGTTCATGCGCTTCAAGGTGCACATGGAGGGCTCCGTGAACGCCACGAGTTC
GAGATCGAGGGCGAGGGCGAGGGCCGCCCTACGAGGGCACCCAGACCGCCAAGCTGAAGGTGAOCCAAGGGTGGCCCOCT
GCCCTTCGCTGGGACATCCTGTCCCTCAGTTTCATGTACGGCTCCAAGGCCTACGTGAAGCACCCGCGGACATCCCGG
ACTACTTGAAGCTGTCCTTCCCGAGGGGCTCAAGTGGGAGCGCGTGATGAACCTCGAGGACGGCGGCTGGTGACCGTG
ACCCAGGACTCCTCCCTGCAGGACGGCGAGTTCATCTACAAGGTGAAGCTGCGCGGCACCAACTTCCCTCCGACGGGCC
CGTAATGCAGAAGAAGACCATGGGCTGGGAGGCCTCCTCCGAGCGGATGTACCCCGAGGACGGCGCCCTGAAGGGCGAGA
TCAAGCAGAGGCTGAAGCTGAAGGAACGGCGGCGACTACGACGCTGAGGTCAAGACCCTACAAGGCCAAGAAGCCCGTG
CAGCTGCCCGGGCGCTACAACGTCAACATCAAGTTGGACATCACCTCCACAAACGAGGACTACACCATCGTGAACAGTA
CGAACGCGCGGAGGGCCGCCACTCCACCGGCGGCATGGAACGAGCTGTACAAGGAATTCATAAGCTCGAGGGTACC

```

E.2 Granzyme B-toxin fusion protein plasmids and IAP-FRET plasmids (Chapter 4)

E.2.1 Base pdL plasmid map

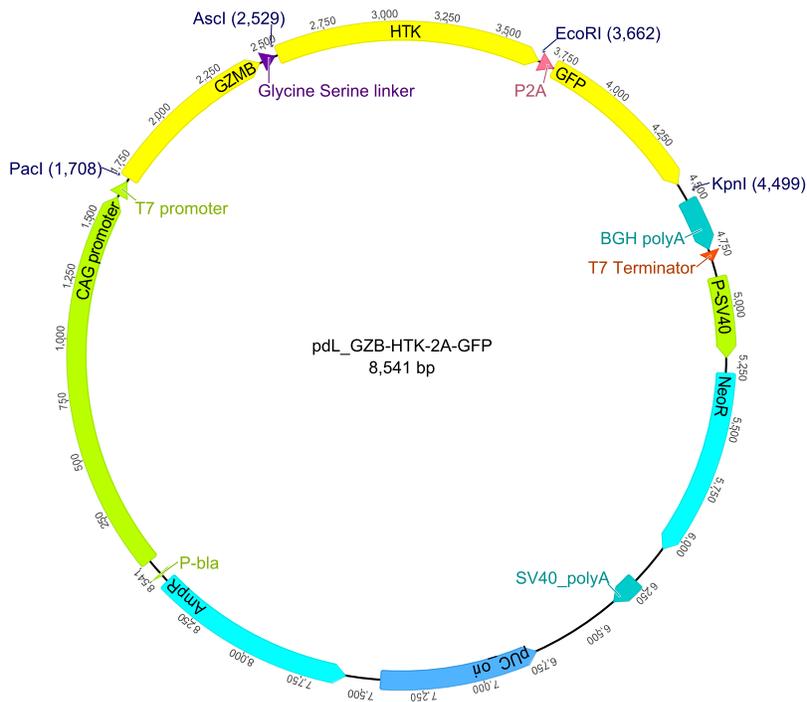


Figure E.2: Base pdL vector. In this example the insert coding sequence is the granzyme B-thymidine kinase fusion protein, followed by a 2A-GFP protein. The annotations correspond to those in the genbank file.

E.2.2 Base pdL plasmid sequence

LOCUS pdL_GZB-HTK-2A-GFP 8541 bp DNA circular UNA 21-FEB-2017
DEFINITION .
ACCESSION .
VERSION .
KEYWORDS .
SOURCE
ORGANISM

FEATURES Location/Qualifiers
promoter 2..1628
/label="CAG promoter"
promoter 1635..1702
/label="T7 promoter"
CDS 1717..2457
/label="GZMB"
misc_feature 2467..2526
/note="Geneious type: polylinker"
/label="Glycine Serine linker"
CDS 2536..3661
/label="HTK"
misc_signal 3667..3732
/label="P2A"
CDS 3733..4452
/label="GFP"
polyA_signal 4527..4751
/label="BGH polyA"
terminator 4760..4803
/label="T7 Terminator"
promoter 4871..5214
/label="P-SV40"
misc_feature 5276..6070
/note="Geneious type: Marker"
/label="NeoR"
polyA_signal 6244..6374
/label="SV40_polyA"
misc_feature complement(6757..7427)
/note="Geneious type: Origin of Replication"
/label="pUC_ori"
misc_feature complement(7572..8432)
/note="Geneious type: Marker"
/label="AmpR"
promoter complement(8467..8473)
/label="P-bla"

ORIGIN

```
1 ctgcaggcgt tacataactt acggtaaatg gcccgctgg ctgaccgccc aacgaccccc
61 gccattgac gtcaataatg acgtatgttc ccatagtaac gccaataggg actttccatt
121 gacgtcaatg ggtggagtat ttacggtaaa ctgcccactt ggcagtacat caagtgtatc
181 atatgccaag tacgccccct attgacgtca atgacggtaa atggcccgcc tggcattatg
241 cccagtacat gaccttatgg gactttccta cttggcagta catctacgta ttagtcatcg
301 ctattaacat ggtcgagggt agccccacgt tctgcttcac tctccccatc tcccccccct
361 cccaccccc aattttgtat ttattttatt tttattatt ttgtgcagcg atgggggagg
421 gggggggggg ggggcgcgcg ccaggcgggg cggggcgggg cgaggggagg ggcggggcga
```

481 ggcgagaggg tgcggcggca gccaatcaga gcggcgcgct ccgaaagttt ccttttatgg
541 cgaggcggcg gcggcggcgg ccctataaaa agcgaagcgc gcggcggggc gggagtcgct
601 ggcagcgtgc cttcgccccg tgccccgctc cgccgccgcc tcgccccgcc cgccccggct
661 ctgactgacc gcgttactcc cacaggtgag cgggcgggac ggcccttctc ctccgggctg
721 taattagcgc ttggtttaat gacggcttgt ttctttctg tggctgcgtg aaagccttga
781 ggggctccgg gagggccctt tgtcggggg gagcggctcg gggggtgcgt gcgtgtgtgt
841 gtgcgtgggg agcggccgct gcggctccgc gctgccggc ggctgtgagc gctcggggcg
901 cggcgggggg cttgtgcgc tccgcagtgt gcgcgagggg agcggcgccg gggcggtgc
961 ccccggtgac ggggggggct gcgaggggaa caaaggctgc gtgcggggtg tgtcgtggg
1021 ggggtgagca ggggggtgtg gcgcgtcggg cgggctgcaa cccccctgc accccctcc
1081 ccgagttgct gacacggccc cggcttcggg tcgggggctc cgtacggggc gtggcgggg
1141 gctcgcctg ccgggccccg ggtggcggca ggtgggggtg ccgggccccg cggggcggc
1201 tcgggccccg gagggctcgg gggagggggc cggcggcccc cggagcgccg gcggctgtc
1261 aggcgcggcg agccgcagcc attgcctttt atggtaatcg tcgagagggg cgcagggact
1321 tcctttgtcc caaatctgtg cggagccgaa atctggagg cgccgccga cccccctag
1381 cggcgcgggg gcgaagcggg gcggcgccg caggaaggaa atgggccccg agggccttcg
1441 tgcgtgcgg cgccgcccgc cccttctcc tctccagcct cggggtgtc cgcgggggga
1501 cggctgcctt cgggggggac ggggcagggc ggggttcggc ttctggcgtg tgaccggcg
1561 ctctagacaa ttgtactaac ctcttctct tctctctct gacaggttgg tftacagtag
1621 cttccaccac cggttaatac gactactat aggttagcat ttaggtagaca ctatagaata
1681 caagctactt gttctttttg cattaattaa gccgccatgc aaccaatcct cttctgtctg
1741 gccttctcc tgcgtcccag ggcagatgca ggggagatca tcgggggaca tgaggccaag
1801 ccccaactcc gccctacat ggcttatctt atgatctggg atcagaagtc tctgaagagg
1861 tgcggtggtt tctgatacg agacgacttc gtgtgacag ctgctcactg ttggggaagc
1921 tccataaatg tcacctggg ggcccacaat atcaaagaac aggagccgac ccagcagttt
1981 atccctgtga aaagacccat cccccatcca gcctataatc ctaagaactt ctccaacgac
2041 atcatgtctc tgcagctgga gagaaaggcc aagcggacca gagctgtgca gccctcagg
2101 ctacctgaca acaaggccca ggtgaagcca gggcagacat gcagtggtgc cggctggggg
2161 cagacggccc ccctgggaaa aactcacac aactacaag aggtgaagat gacagtcag
2221 gaagatgaa agtgcgaatc tgcactacgc cattattacg acagtaccat tgagtgtgc
2281 gtgggggacc cagagattaa aaagacttcc ttaaggggg actctggagg ccctctgtg
2341 tgaacaagg tgcccaggg cattgtctcc tatggacgaa acaatggcat gcctccacga
2401 gcctgacca aagtctcaag cttgtacac tggataaaga aaaccatgaa acgctacgg
2461 gccgtggag gcggggggtt tggcgggggt ggatcagggg gtggaggctc cggtgaggc
2521 gggtcggggc cgcccgttc gtacccctgc catcaacacg cgtctgcgtt cgaccaggct
2581 gcgcgttctc gggccatag caaccgact acggcgttgc gccctcgccg gcagcaagaa
2641 gccacggaag tccgctgga gcagaaaatg cccacgctac tgcgggttta tatagacggt
2701 cctcacggga tggggaaac caccaccag caactgctgg tggccctggg ttgcgcgac
2761 gatctgtct acgtaccga gccgatgact tactggcagg tgcggggg cttccagaca
2821 atcggaaca tctacaccac acaacaccgc ctgaccagg gtgagatc ggccggggac
2881 gcggcggtg taatgacaag cggccagata acaatgggca tgccttatgc cgtgaccgac
2941 gccgttctg ctctcatal cgggggggag gctgggagct cacatgccc gccccggcc
3001 ctaccctca tcttcgacc ccatcccac gccgccctc tgtgtacc ggcgcgcgca
3061 taccttatg gcagcatgac ccccaggcc gtgctggcgt tctgtggcct catcccgg
3121 acctgcccg gcacaaacat cgtgtgggg gccctccgg aggacagaca catcgaccg
3181 ctggcctaac gccagcggc cggcagcgg cttgacctg ctatgctggc cgcgattcgc
3241 cgcgttacg ggctgctgc caatacggg cggatctgc agggcgggc gctgtggcg
3301 gaggattgg gacagcttc ggggacggc gtgcccccc aggggtccga gccccagagc
3361 aacgccccg cacgaccca tctcggggac acgttattha ccctgttctg ggcccccgag
3421 ttgctggccc ccaacggcga cctgtacaac gtgtttgctt gggccttggc cgtcttgcc
3481 aaacgcctcc gtccatgca cgtctttatc ctggattacg accaatcgc cggcggtgc
3541 cgggacggc tgcgtcaact tacctcggg atgttcaga cccacgtac caccggggc
3601 tccataccga cgtctgca cctggcggc acgtttgcc gggagatgg ggaggctaac
3661 gaattcgaa gcggagctac taactcagc ctgctgaagc aggtgggaga cgtggaggag

3721 aacctggac ctatggtgag caagggcgag gagctgttca cccgggtggt gcccatcctg
3781 gtcgagctgg acggcgacgt aaacggccac aagttcagcg tgtccggcga gggcgagggc
3841 gatgccacct acggcaagct gacctgaag ttcactgca ccaccggcaa gctgcccgtg
3901 ccctggccca ccctcgtgac caccctgacc tacggcgtgc agtgctttag ccgctacccc
3961 gaccacatga agcagcacga ctcttcaag tccgccatgc ccgaaggcta cgtccaggag
4021 cgcaccatct tctcaagga cgacggcaac tacaagacc gcgccgaggt gaagttcgag
4081 ggcgacacc ttgtgaaccg catcgagctg aagggcatcg actcaagga ggacggcaac
4141 atcctggggc acaagctgga gtacaactac aacagccaca acgtctatat catggccgac
4201 aagcagaaga acggcatcaa ggtgaacttc aagatccgcc acaacatcga ggacggcagc
4261 gtgcagctcg ccgaccacta ccagcagaac acccccatcg gcgacggccc cgtgctgctg
4321 cccgacaacc actacctgag caccagctcc gccctgagca aagaccccaa cgagaagcgc
4381 gatcacatgg tctgctgga gttcgtgacc gccgcccggga tcaactcgg catggacgag
4441 ctgtacaagt aactcgagag atccccggg gtcgactgat caaatcgag ctcggtacca
4501 ctagttagcc cgctgatcag cctcgactgt gccttctagt tgcagccat ctgttgtttg
4561 cccctcccc gtgccttctc tgaccctgga aggtgccact cccactgtcc tttcctaata
4621 aaatgaggaa attgcatcgc attgtctgag taggtgctat tctattctgg ggggtggggt
4681 gggcgaggac agcaaggggg aggattggga agacaatagc aggcattgct gggatgcggt
4741 gggctctatg gtaacgcgta taacccttg gggcctctaa acgggtcttg aggggttttt
4801 tgactgagc gatttcggcc tattgttaa aaaatgagct gatttaacaa aaatttaacg
4861 cgaattaatt ctgtggaatg tgtgtagtt aggggtgga aagtccccag gctccccagc
4921 aggcagaagt atgcaaagca tgcactcaa ttagtcagca accagggtg gaaagtcccc
4981 aggtcccca gcaggcagaa gtatgcaaag catgcatctc aattagtcag caaccatagt
5041 cccgccccta actccgccc tcccgccct aactccgcc agtccgcc attctccgcc
5101 ccatggctga ctatatttt ttatttatgc agaggccgag gccgcctctg cctctgagct
5161 attccagaag tagtgaggag gcttttttg aggcctagcc ttttgcaaaa agctcccggg
5221 agcttgata tccattttc gatctgatca agagacagga tgaggatcgt ttcgatgat
5281 tgaacaagat ggattgcacg caggttctcc ggccgcttg gtggagaggc tattcggcta
5341 tgactgggca caacagacaa tcggctgctc tgatgccgcc gtgtccggc tgcagcgca
5401 gggcgcccc gttcttttt tcaagaccga cctgtccggt gccctgaat aactgcagga
5461 cgaggcagcg cggctatcgt ggctggccac gacggcgctt cctgcgag ctgtgctcga
5521 cgtgtcact gaagcggaa gggactggct gctattgggc gaagtgcgg ggcaggatct
5581 cctgtcatct caccttgctc ctgcccagaa agtatccatc atggctgatg caatgcccgg
5641 gctgcatacg ctgtatccg ctacctgcc attcgaccac caagcgaac atcgcatcga
5701 gcgagcacgt actcggatgg aagccggct tgcgatcag gatgatctgg acgaagagca
5761 tcaggggctc gcgccagccg aactgttcgc caggctcaag gcgcgatgc ccgacggcga
5821 ggatctcgtc gtgaccatg gcgatgcctg ctgcccgaat atcatggtgg aaaatggccg
5881 ctttctgga ttcactgact gtggccggct ggggtggtgc gaccgctatc aggacatagc
5941 gttggctacc cgtgatattg ctgaagagct tggcggcga tggctgacc gcttctcgt
6001 gctttacggt atcgccgctc ccgattcgca gcgcatcgc ttctatcgcc ttcttgacga
6061 gttcttctga gcgggactct ggggttggaa atgaccgacc aagcagcgc caacctgcca
6121 tcacgagatt tcgattccac cgcgccttc tatgaaaggt tgggctcgg aatcgtttc
6181 cgggacgccc gctggatgat cctccagcgc ggggatccta tgcctggatt ctccgcccac
6241 cccaacttgt ttattgcagc ttataatggt tacaataaaa gcaatagcat cacaaattc
6301 acaataaag cattttttc actgcattct agttgtggt tgcctaaact catcaatgta
6361 tcttatcatg tctgtatacc gtcgacctct agctagagct tggcgtaatc atggtcatag
6421 ctgtttcctg tgtgaaattg ttatccgctc acaattccac acaacatacg agccggaagc
6481 ataaagtgta aagcctgggg tcctaatga gtgagctaac tcacattaat tgcgttgccg
6541 tcaactcccc ctttccagtc gggaaacctg tctgtccagc tgcattaatg aatcgccaa
6601 cgcgcgggga gaggcggttt gcgtattggg cgctctccg ctctctcgt cactgactcg
6661 ctgcccctgg tcttccgct gcggcgagcg gtatcagctc actcaaaggc ggtaatacgg
6721 ttatccacag aatcagggga taacgcagga aagaacatgt gagcaaaagg ccagcaaaag
6781 gccagggaacc gtaaaaaggc cgcgttgctg gcgttttcc ataggctccg ccccctgac
6841 gagcatcaca aaaaatcgac ctcaagtcag aggtggcgaa acccgacagg actataaaga
6901 taccaggcgt ttccccctg aagctccctc gtgcgctctc ctgttccgac cctgcccgtt

6961 accggatacc tgtccgcctt tctcccttcg ggaagcgtgg cgctttctca tagctcacgc
7021 ttaggtatc tcagttcggg ttaggtcgtt cgctccaagc tgggctgtgt gcacgaacct
7081 cccgttcagc ccgaccgctg cgccttatcc ggtaactatc gtcttgagtc caaccggta
7141 agacacgact tatcgccact ggcagcagcc actggaaca ggattagcag agcgaggat
7201 gtaggcgggt ctacagagtt ctgaagtg tggcctaact acggctacac tagaagaaca
7261 gtatttgga tctgcgctct gctgaagcca gttacctcg gaaaagagt tggtagctct
7321 tgatcggca aacaaaccac cgctggtagc ggttttttg ttgcaagca gcagattacg
7381 cgagaaaaa aaggatctca agaagatcct ttgatcttt ctacggggtc tgacgctcag
7441 tgaacgaaa actcacgta agggattttg gtcatgagat tatcaaaaag gatcttcacc
7501 tagatccttt taaattaa atgaagttt aaatcaatc aaagtatata tgagtaaact
7561 tggctgaca gttaccaatg cttaatcagt gaggcaccta tctcagcgat ctgtctatft
7621 cgttcacca tagttgcctg actccccgctc gtgtagataa ctacgatacg ggagggcta
7681 ccatctggcc ccagtgtctc aatgataccg cgagaccac gtcaccggc tccagattta
7741 tcagcaataa accagccagc cggaggggcc gagcgcagaa gtggtcctgc aactttatcc
7801 gcctccatcc agtctattaa ttgtgcccgg gaagctagag taagtagtcc gccagttaat
7861 agtttgcgca acgittgtgc cattgctaca ggcatcgtg tgcacgctc gtcgtttggt
7921 atggcttcat tcagctccgg ttccaacga tcaaggcgag ttacatgac cccatgttg
7981 tgcaaaaaa cggttagctc cttcggctct ccgatcgtt tcagaagtaa gttggccgca
8041 gtgttatcac tcatggttat ggcagcactg cataattctc ttactgtcat gccatccgta
8101 agatgctttt ctgtgactgg tgagtactca accaagtcat tctgagaata gtgtatcgg
8161 cgaccgagtt gctcttgccc ggcgtcaata cgggataata ccgcccaca tagcagaact
8221 ttaaaagtc tcatcattg aaaacgttct tggggcgaa aactctcaag gatcttaccg
8281 ctgttgagat ccagttcgat gtaaccact cgtgcacca actgatcttc agcatcttt
8341 actttacca gcgtttctgg gtgagcaaaa acaggaaggc aaaatgccgc aaaaaagga
8401 ataagggcga cacggaaatg ttgaatactc atactcttc tttttcaata ttattgaagc
8461 atttatcagg gttattgtct catgagcgga tacatattg aatgtattta gaaaataaa
8521 caaatagggg ttccgcgagc t

//

E.2.3 Base pMND plasmid map

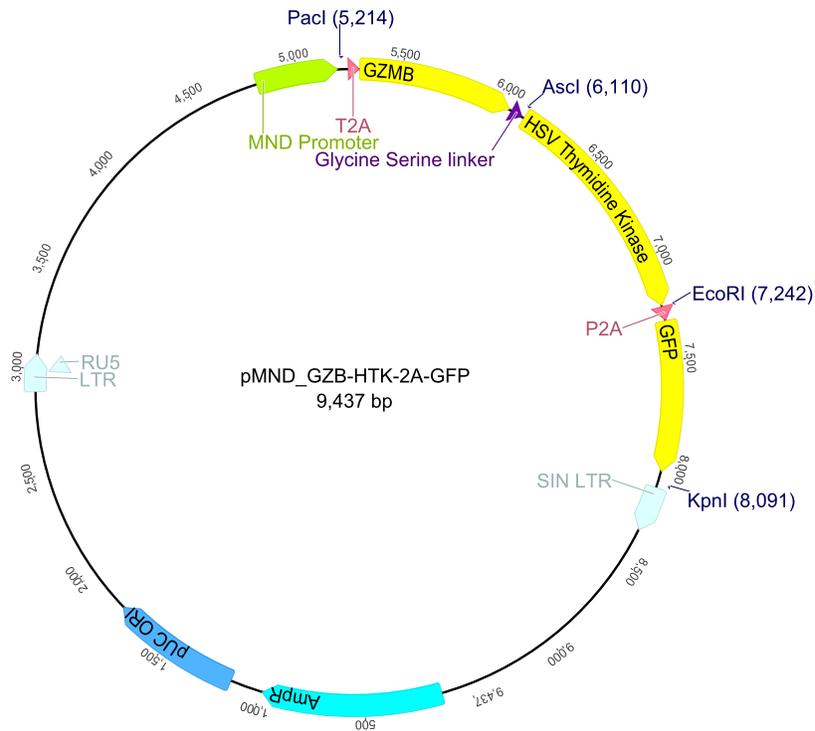


Figure E.3: Base pMND vector. In this example the insert coding sequence is the granzyme B-thymidine kinase fusion protein, followed by a 2A-GFP protein. The annotations correspond to those in the genbank file.

E.2.4 Base pMND plasmid sequence

```

LOCUS      pMND_GZB-HTK-2A-GFP          9437 bp    DNA     circular UNA 26-JAN-2017
DEFINITION .
ACCESSION .
VERSION   .
KEYWORDS  .
SOURCE
  ORGANISM .

FEATURES             Location/Qualifiers
     misc_feature    132..992
                     /note="Geneious type: Marker"
                     /label="AmpR"
     misc_feature    1147..1766
                     /note="Geneious type: Origin of Replication"
                     /label="pUC ORI"
     LTR              2891..3071
                     /label="LTR"
     LTR              2990..3072
                     /label="RU5"
     promoter        4803..5194
                     /label="MND Promoter"
     misc_signal     5244..5297
                     /label="T2A"
     CDS              5298..6038
                     /gene="GZMB"
                     /EC_number="3.4.21.79"
                     /note="fragmentin 2; cathepsin G-like 1; T-cell serine
                     protease 1-3E; cytotoxic serine protease B; C11; CTLA-1;
                     fragmentin-2; human lymphocyte protein; cytotoxic
                     T-lymphocyte proteinase 2"
                     /codon_start=1
                     /product="granzyme B precursor"
                     /protein_id="NP_004122.2"
                     /db_xref="GI:221625528"
                     /db_xref="CCDS:CCDS9633.1"
                     /db_xref="GeneID:3002"
                     /db_xref="HGNC:4709"
                     /db_xref="MIM:123910"
                     /translation="MQPILLLLAFLLLPADAGEIIGGHEAKPHSRPYMAYLMIWDQK
                     SLKRCGGFLIRDDFVLTAAHCWGSSINVTLGAHNIKEQEPTQQFIPVKRPIHPAYNP
                     KNFSNDIMLLQLERKAKRTRAVQPLRLPSNKAQVKPGQTCSVAGWGQTAPLGKHSHTL
                     QEVKMTVQEDRKCESDLRHYYDSTIELCVGDPEIKKTSFKGDSGGPLVCNKVAQGIVS
                     YGRNNGMPPRACTKVSSFVHWIKTKMKRY"
                     /label="GZMB"
     misc_feature    6048..6107
                     /note="Geneious type: polylinker"
                     /label="Glycine Serine linker"
     CDS              6117..7240
                     /label="HSV Thymidine Kinase"
     misc_signal     7247..7312
                     /label="P2A"
     CDS              7313..8032

```

```

                                /label="GFP"
LTR                            8112..8329
                                /label="SIN LTR"
ORIGIN
  1  caggTggcac tttcgggga aatgtgcgcg gaaccctat ttgtttatt ttctaaatac
 61  attcaaatat gtatccgctc atgagacaat aaccctgata aatgcttcaa taatattgaa
121  aaaggaagag tatgagtatt caacatttcc gtgtcgcctt tattcccttt ttgcgggcat
181  ttgccccttc tgtttttgct caccagaaa cgctggtgaa agtaaaagat gctgaagatc
241  agttgggtgc acgagtggtt tacatcgaac tggatctcaa cagcggtaag atccttgaga
301  gtttcgccc cgaagaacgt ttccaatga tgagcacttt taaagttctg ctatgtggcg
361  cggattattc ccgtattgac gccgggcaag agcaactcgg tcgccgata cactatttct
421  agaatgactt ggttgagtac tcaccagtca cagaaaagca tcttacggat ggcatgacag
481  taagagaatt atgcagtgtc gccataacca tgagtataa cactgcggcc aacttacttc
541  tgacaacgat cggaggaccg aaggagctaa ccgctttttt gcacaacatg ggggatcatg
601  taactgcctt tgatcgttgg gaaccggagc tgaatgaagc cataccaacg gacgagcgtg
661  acaccacgat gcctgtagca atggcaacaa cgttgcgcaa actattaact ggcgactac
721  ttactctagc ttcccggcaa caattaatag actggatgga ggcgataaaa gttgcaggac
781  cacttctgct ctcggccctt ccggctggct ggtttattgc tgataaatct ggagccggtg
841  agcgtgggtc tcgcggtatc atgcagcac tggggccaga tgtaagccc tcccgtatcg
901  tagttatcta cacgacgggg agtcaggcaa ctatggatga acgaaataga cagatcgtcg
961  agataggtgc ctactgatt aagcattggt aactgtcaga ccaagtttac tcatatatac
1021 tttagattga tttaaaactt ctttttaat ttaaaggat ctaggatgaag atcctttttg
1081 ataatctcat gacaaaaatc ccttaacgtg agttttcgtt ccactgagcg tcagaccccg
1141 tagaaaagat caaaggatct tcttgagatc ctttttttct gcgcgtaatc tgctgcttgc
1201 aaacaaaaaa accaccgcta ccagcgggtg tttgtttgcc ggatcaagag ctaccaactc
1261 tttttccgaa ggtaactggc ttcagcagag cgcagatacc aaatactgtc cttctagtgt
1321 agccgtagtt agggcaccac ttcaagaact ctgtagcacc gcctacatac ctgctctgct
1381 taatcctgtt accagtggtc gctgccagtg gcgataagtc gtgtcttacc gggttggact
1441 caagacgata gttaccggat aaggcgcagc ggtcgggctg aacggggggt tcgtgcacac
1501 agcccagctt ggagcgaacg acctacaccg aactgagata cctacagcgt gagctatgag
1561 aaagcggcac gcttccgaa gggagaaagg cggacaggta tccgtaagc ggcagggtcg
1621 gaacaggaga gcgcacgagg gagcttccag ggggaaacgc ctggtatctt tatagtcctg
1681 tcgggtttcg ccacctctga ctgagcgtc gatTTTTgtg atgctcgtca ggggggcgga
1741 gcctatggaa aaacgccagc aacgcggcct tttacggtt cctggccttt tgctggcctt
1801 ttgctcacat gttctttcct gcgttatccc ctgattctgt ggataaccgt attaccgctt
1861 ttgagtgagc tgataccgct cgccgcagcc gaacgaccga gcgcagcggg tcagtgagcg
1921 aggaagcgga agagcgccca atacgcaaac cgcctctccc cgcgcgttgg ccgattcatt
1981 aatgcagctg gcacgacagg ttcccgact ggaagcggg cagtgagcgc aacgcaatta
2041 atgtgagtta gctcactcat taggcacccc aggccttaca ctttatgctt ccggctcgta
2101 tgttgtgtgg aattgtgagc ggataacaat ttcacacagg aaacagctat gaccatgatt
2161 acgccaagcg cgcaattaac ctcactaaa gggacaaaaa gctggagctg caagcttggc
2221 cattgcatac gttgtatcca tatcataata tgtacattta tattggctca tgtccaacat
2281 taccgccatg ttgacattga ttattgacta gttattaata gtaatcaatt acggggtcat
2341 tagttcatag cccatataat gagttccgcg ttacataact tacggtaaat ggcccgcctg
2401 gctgaccgcc caacgacccc cgcccattga cgtcaataat gacgtatgtt cccatagtaa
2461 cgccaatagg gactttccat tgacgtcaat ggggtggagta ttacggtaa actgcccact
2521 tggcagtaca tcaagtgtat catatgcca gtacgcccc tattgacgtc aatgacggta
2581 aatggcccgc ctggcattat gcccagtaca tgacctatg ggactttcct acttggcagt
2641 acatctactg attagtcatc gctattacca ttggtatgcg gttttggcag tacatcaatg
2701 ggcgtggata gcggtttgac tcacggggat ttccaagtct ccacccatt gacgtcaatg
2761 ggagtttgtt ttggcaccia aatcaacggg actttccaaa atgtcgtaac aactccgccc
2821 cattgacgca aatggcggtt aggcgtgtac ggtgggaggt ctatataagc agagctcgtt
2881 tagtgaaccg gggctctctt ggtagacca gatctgagcc tgggagctct ctggctaact
2941 agggaacca ctgcttaagc ctcaataaag cttgccttga gtgcttcaag tagtgtgtgc

```

3001 ccgtctgttg tgtgactctg gtaactagag atccctcaga cccttttagt cagtgtggaa
3061 aatctctagc agtggcgccc gaacagggac ctgaaagcga aagggaaacc agaggagctc
3121 tctcgacgca ggactcggct tgctgaagcg cgcacggcaa gagggcaggg gcggcgactg
3181 gtgagtacgc caaaaatttt gactagcgga ggctagaagg agagagatgg gtgcgagagc
3241 gtcagtatta agcgggggag aattagatcg cgatgggaaa aaattcggft aaggccaggg
3301 ggaaagaaaa aatataaatt aaaacatata gtatgggcaa gcagggagct agaacgattc
3361 gcagttaatc ctggcctggt agaacaatca gaaggctgta gacaaatact gggacagcta
3421 caaccatccc ttcagacagg atcagaagaa cttagatcat tatataatac agtagcaacc
3481 ctctattgtg tgcatacaag gatagagata aaagacacca aggaagcttt agacaagata
3541 gaggaagagc aaaacaaaag taagaccacc gcacagcaag cggccgctga tcttcagacc
3601 tggaggagga gatatgaggg acaattggag aagtgaatta tataaatata aagtagtaaa
3661 aattgaacca ttaggagtag caccaccaa ggcaagaga agagtggtc agagagaaaa
3721 aagagcagtg ggaataggag ctttgttctt tgggttcttg ggagcagcag gaagcactat
3781 gggcgacgcc tcaatgacgc tgacggtaca ggccagacaa ttattgtctg gtatagtga
3841 gcagcagaac aatttgctga gggctattga ggcgcaacag catctgttgc aactcacagt
3901 ctggggcagc aagcagctcc aggcaagaat cctggctgtg gaaagatacc taaaggatca
3961 acagctcctg gggatttggg gttgctctgg aaaactcatt tgcaccactg ctgtgccttg
4021 gaatgctagt tggagtaata aatctctgga acagattgga atcacacgac ctggatggag
4081 tgggacagag aaattaacaa ttacacaagc ttaatacact ccttaattga agaatcgcaa
4141 aaccagcaag aaaagaatga acaagaatta ttggaattag ataaatgggc aagtttggg
4201 aattggttta acataacaaa ttggctgtg tatataaaat tattcataat gatagtagga
4261 ggcttggtag gtttaagaat agttttgct gtactttcta tagtgaatag agttaggcag
4321 ggatattcac cattatcgtt tcagaccacc ctccaaccc cgaggggacc cgacaggccc
4381 gaaggaatag aagaagaagg tggagagaga gacagagaca gatccattcg attagtgaac
4441 ggatctcgac ggtatcgata agctaattca caaatggcag tattcatcca caattttaa
4501 agaaaagggg ggattggggg gtacagtga ggggaaagaa tagtagacat aatagcaaca
4561 gacatacaaa ctaaagaatt acaaaaacaa attacaaaa ttcaaaattt tggggtttat
4621 tacagggaca gcagagatcc agtttgggaa ttgcttgat cgattagtcc aatttgttaa
4681 agacaggata tcagtgttcc aggcctctagt ttgactcaa caatatacc agctgaagcc
4741 tatagagtac gagccataga tagaataaaa gattttatt agtctccaga aaaagggggg
4801 aatgaaagac cccacctgta ggtttggcaa gctaggatca aggttaggaa cagagagaca
4861 gcagaatatg ggccaacag gatattctgt gtaagcagtt cctgcccgg ctcagggcca
4921 agaacagttg gaacagcaga atatgggcca aacaggatat ctgtggtaa cagttcctgc
4981 cccggctcag ggccaagaac agatggtccc cagatgcggt cccgccctca gcagtttcta
5041 gagaaccatc agatgtttcc aggggtcccc aaggacctga aatgacctg tgccttattt
5101 gaactaacca atcagttcgc ttctcgcttc tgttcgctgc ctctgtctcc ccgagctcaa
5161 taaaagagcc cacaaccctc cactcggcgc gatctagatc tcgaatcgtt aattaaattg
5221 ccgccatggg atccggaagc ggagagggca gaggaagctt gtaaacatgc ggtgacgtcg
5281 aggagaatcc tggacctatg caaccaatcc tgcctctgct ggcttctcct ctgctgccc
5341 gggcagatgc aggggagatc atcgggggac atgaggccaa gcccactcc cggccctaca
5401 tggcttattc tatgatctgg gatcagaagt ctctgaagag gtgcggtggc ttctgatac
5461 gagacgactt cgtgctgaca gctgctcact gttgggaag ctccataaat gtcaccttgg
5521 gggcccacaa tatcaaagaa caggagccga cccagcagtt tatccctgtg aaaagaccca
5581 tccccatcc agcctataat cctaagaact tctccaacga catcatgcta ctgcagctgg
5641 agagaaagcc caagcggacc agagctgtgc agcccctcag gctacctagc aacaaggccc
5701 aggtgaagcc agggcagaca tgcagtgtgg ccggctgggg gcagacggcc cccctgggaa
5761 aacactcaca cacactacaa gaggtgaaga tgacagtga ggaagatcga aagtgcgaat
5821 ctgacttacg ccattattac gacagtacca ttgagttgtg cgtgggggac ccagagatta
5881 aaaagacttc cttaagggg gactctggag gccctctgt gtgtaacaag gtggcccagg
5941 gcattgtctc ctatggacga aacaatggca tgcctccac agcctgcacc aaagtctcaa
6001 gctttgtaca ctggataaag aaaacatga aacgctacgc ggccgctgga ggcgggggtt
6061 ctggcggggg tggatcaggg ggtggaggct ccgggtggag cgggtcgggc gcgcatggc
6121 ttcgtacccc tgccatcaac acgcgtctgc gttcgaccag gctgcgctt ctgcgggcca
6181 tagcaaccga cgtacggcgt tgcgccctcg ccggcagcaa gaagccaggg aagtccgcct

6241 ggagcagaaa atgcccacgc tactgcgggt ttatatagac ggtcctcacg ggatggggaa
6301 aaccaccacc acgcaactgc tgggtggcct gggttcgcgc gacgatatcg tctacgtacc
6361 cgagccgatg acttactggc aggtgctggg ggcttccgag acaatcgga acatctacac
6421 cacacaacac cgcctcgacc agggtgagat atcgccggg gacgcccggg tggtaatgac
6481 aagcgccagc ataacaatgg gcatgcctta tgcctgacc gacgcccgtc tggctcctca
6541 tatcgggggg gaggtgagg gctcacatgc cccgccccg gccctcacc tcatcttca
6601 ccgcatccc atcgccgcc tcctgtgcta cccggccgag cgataccta tgggcagcat
6661 gacccccag gccgtgctgg cgttcgtggc cctcatccc cggacctgc ccggcacaaa
6721 catcgtgttg ggggcccttc cggaggacag acacatcgac cgcctggcca aacgccagcg
6781 cccggcgag cggcttgacc tggctatgct ggccgcgatt cggcgcttt acgggctgct
6841 tgccaatacg gtgcggatc tgcagggcgg cgggtcgtgg cgggaggatt ggggacagct
6901 ttcggggagc gccgtgccg cccaggggag cgagccccag agcaacgcgg gccccagacc
6961 ccatatcggg gacacgttat ttaccctgtt tcgggcccc gagttgctgg cccccaacgg
7021 cgacctgtac aacgtgtttg cctgggcctt ggacgtctt gccaaacgcc tccgtccat
7081 gcacgtcttt atcctggatt acgaccaatc gcccgccggc tgcggggagc ccctgtgca
7141 acttacctcc gggatgttcc agaccacgt caccacccc ggctccatac cgacgatcg
7201 cgacctggcg cgcacgtttg cccgggagat gggggaggcc gaattcgaa ggggagctac
7261 taacttcagc ctgctgaagc aggttgaga cgtggaggag aacctggac ctatggtgag
7321 caaggcgag aagctgttca cgggggtggt gccatcctg gtcgagctgg acggcgactg
7381 aaacggccac aagttcagcg tgtccggcga gggcgaggcc gatgccact acggcaagct
7441 gacctgaag ttcactgca ccaccggcaa gctgccctg ccctggccca ccctcgtgac
7501 caccctgacc tacggcgtgc agtgcttca cgcctacccc gaccacatga agcagcacga
7561 cttctcaag tccgccatgc ccgaaggcta cgtccaggag cgcaccatct tctcaagga
7621 cgacggcaac tacaagacc gcgcccaggt gaagttcag ggcgacacc tggtaaccg
7681 catcgagctg aagggcacg acttcaagga ggacggcaac atcctggggc acaagctgga
7741 gtacaactac aacagccaca acgtctatat catggccgac aagcagaaga acggcatca
7801 ggtgaacttc aagatccgcc acaacatcga ggacggcagc gtcagctcg ccgaccacta
7861 ccagcagaac accccatcg gcgacggccc cgtgctgctg cccgacaacc actacctgag
7921 caccagtcg gccctgagca aagaccccaa cgagaagcgc gatcacatgg tctgctgga
7981 gttcgtgacc gccgcccggg tactctcgg catggacgag ctgtacaagt aactcgagag
8041 atctctcgag agatccccc gggctgactg atcaaatcg agctcgttac cttaagacc
8101 aatgacttac aaggcagctg tagatcttag ccactttta aaagaaaagg ggggactgga
8161 agggctaatt cactccaac gaagacaaga tctgctttt gcttgactg ggtctctg
8221 gttagaccag atctgacct gggagctctc tggctaacta ggaacccac tgcctaagcc
8281 tcaataaagc ttgcctttag tgcctcaagt agtgtgtgcc cgtctgtgtg gtgactctgg
8341 taactagaga tccctcagac ccttttagtc agtgtgaaa atctctagca gtagtagttc
8401 atgtcatctt attattcagt atttataact tgcaaagaaa tgaatatcag agagttagag
8461 gaactgtttt attgcagctt ataatggta caaataaagc aatagcatca caaattcac
8521 aaataaagca ttttttca c tgcattctag ttgtggtttg tccaaactca tcaatgtatc
8581 ttatcatgtc tggctctagc tatcccggcc ctaactccgc ccatcccggc ctaactccg
8641 cccagttccg cccattctcc gccccatggc tgactaat tttttat tta tgcagaggcc
8701 gaggcgcgct cggcctctga gctattccag aagtagtgag gaggctttt tggaggccta
8761 ggcctttgcg tcgagacgta cccaattcgc cctatagtga gtcgtattac gcgcgctcac
8821 tggccgtcgt tttaacaagc cgtgactggg aaaaccctgg cgttacccaa cttaatcgcc
8881 ttgcagcaca tcccccttc gccagctggc gtaatagcga agaggcccgc accgatcgcc
8941 ctccaaca gttgcgcagc ctgaatggcg aatggcgcga cgcgcctgt agcggcgcat
9001 taagcggcgc ggggtgtggg gttacgcgca gcgtgaccgc tacacttgcc agcgcctag
9061 cggccgctcc tttcgtcttc ttcccttctt ttctcgccac gttcggcggc tttcccgtc
9121 aagctctaaa tcgggggctc cctttagggt tccgatttag tgccttaccg cacctcgacc
9181 ccaaaaaact tgattagggt gatggttcac gtagtgggcc atcgccctga tagacggtt
9241 ttgccctttt gacgtggag tccacgttct ttaatagtgg actctgttc caaactggaa
9301 caaacctcaa ccctatctcg gtctattctt ttgatttata agggattttg ccgattcgg
9361 cctattggtt aaaaaatgag ctgatttaac aaaaatttaa cgcgaatttt aaaaaatat
9421 taacgtttac aatttcc

//

E.2.5 Source of component parts

Table E.1: Source of component parts of pdL vectors.

Symbol	Component Name	Source
GZB	granzyme B	Synthesized. Sequence from NCBI RefSeq Gene database (Accession: NG_028340.1)
GSL	glycine serine linker	Synthesized
T2A and P2A	Ribosomal skipping peptides	MND vector series
DTA	Diphtheria toxin A fragment	PGKdtapA (Addgene plasmid 13440)
PEA	Pseudomonas exotoxin A fragment	pRB1101 (a gift from Ira Pastan)
HTK	Herpes simplex virus thymidine kinase	pAL119-TK (Addgene plasmid 21911)
NFSA and NFSB	<i>nfsA</i> and <i>nfsB</i> nitroreductases	DH5-alpha <i>E. coli</i> genomic DNA
XIAP	X-linked inhibitor of apoptosis protein	Mammalian Gene Consortium cDNA (Clone ID: 5532247)
SURV	Survivin	Synthesized. Sequence from NCBI RefSeq gene database (Accession: NM_001168.2)

E.2.6 Coding sequence inserts

Coding sequences listed below extend from the end of the mammalian promoter (CAG for pdL and MND for pMND) to the BGH polyA (for pdL) or the LTR (for pMND). Sequences are given in fasta format, with the header line having the format:

```
>plasmid_backbone | coding_sequence
```

```
>pMND_|_GZB-HTK-2A-GFP
```

```
TAGATCTCGAATCGTTAATTAATTAATTCGCCCATGGGATCCGGAAGCGGAGAGGGCAGAGGAAGTCTGCTAACATGCGGTG  
ACGTCGAGGAGAATCCTGGACCTATGCAACCAATCCTGCTTCTGCTGGCCTTCTCCTGCTGCCAGGGCAGATGCAGGG  
GAGATCATCGGGGACATGAGGCCAAGCCCACTCCCGCCCTACATGGCTTATCTTATGATCTGGGATCAGAAGTCTCT  
GAAGAGGTGCGGTGGCTTCTGATACGAGACGACTTCGTGCTGACAGCTGCTCACTGTTGGGAAGCTCCATAAATGTCA  
CCTTGGGGGCCACAATATCAAAGAACAGGAGCCGACCCAGCAGTTTATCCCTGTGAAAAGACCCATCCCCATCCAGCC  
TATAATCCTAAGAACTTCTCCAACGACATCATGCTACTGCAGCTGGAGAGAAAGGCCAAGCGGACCCAGAGCTGTGCAGCC  
CCTCAGGCTACCTAGCAACAAGGCCAGGTGAAGCCAGGGCAGACATGCAGTGTGCGCCGCTGGGGGCAGACGGCCCCC  
TGGGAAAACACTCACACACACTACAAGAGGTGAAGATGACAGTGCAGGAAGATCGAAAGTGCGAATCTGACTTACGCCAT  
TATTAACGACAGTACCATTGAGTTGTGCGTGGGGGACCCAGAGATTAATAAGACTTCCCTTAAAGGGGACTCTGGAGGCC  
TCTTGTGTGTAACAAGGTGGCCAGGGCATTGTCTCCTATGGACGAAACAATGGCATGCCCTCCACGAGCCTGCACCAAAG  
TCTCAAGCTTTGTACTGGATAAAGAAAACCATGAAACGCTAOCGGCCGCTGGAGGCGGGGTTCTGGCGGGGTGGA  
TCAGGGGTGGAGGCTCOGTTGGAGGCGGGTGGGGGCGCATGGCTTCGTACCCCTGCCATCAACACGCGTCTGCGTTC  
GACCAGGCTGCGGTTCTCGGGGCATAGCAACCCAGCTAOCGGGTTGCGCCCTGCGGGCAGCAAGAAGCCAOGGAAGT  
COGCTGGAGCAGAAAATGCCACGCTACTGCGGGTTTATATAGACGGTCTCAOCGGGATGGGAAAACCAACCAACGC  
AACTGCTGGTGGCCCTGGGTTGCGCGCAGATATCGTCTAOCGTAOCAGCCGATGACTTACTGGCAGGTGCTGGGGCT  
TCGAGACAATCGGAACATCTACACCACACAACCCGCTCGACAGGGTGAATATCGGCGGGGACGCGCGGTGTT  
AATGACAAGCGCCAGATAACAATGGGCATGCCCTTATGCGGTGACCGACCGCTTCTGGCTCCTCATATCGGGGGGAGG  
CTGGGAGCTCACATGCCCGCCCGCCCTCACCCCTCATCTTCGACCCGCATCCATCGCGCCCTCCTGTGCTAOCGG  
GCCCGCGATACCTTATGGGCAGCATGACCCCGCAGCCGCTGCTGGCGTTGCTGGCCCTCATCCCGCCGACCTTGCCCGG  
CACAAACATCGTGTGGGGGCCCTTCGGAGGACAGACACATCGACCGCTGGCCAAACGCCAGCGCCCGCGGAGCGGG  
TTGACCTGGCTATGCTGGCCGCGATTGCGCGCTTTACGGGCTGCTTGCCAATACGGTGCAGTATCTGCAGGGCGCGGG  
TCGTGGCGGAGGATTGGGACAGCTTTTCGGGACGGCGTGCOCGCCAGGGTGCOCGAGCCAGAGCAACGCGGGGCC  
ACGACCCATATCGGGGACAGCTTATTTACCCGTTTTCGGGCCCCGAGTTGCTGGCCCCAACGGGACCTGTACAACG  
TGTTTGCTGGGCTTGAGCGTCTGGCCAAACGCTCCGTCATGCAGCTTCTTATCCTGGATTACGACCAATCGCCC
```

GCOGGCTGOCGGGACGCOCTGCTGCAACTTACCTCOGGGATGGTCCAGAACCCAGTCAACCACCCCGGCTCCATAOCGAC
GATCTGOGACCTGGOCGCGACGTTTGOOOGGAGATGGGGGAGGCOGAATTOGGAAGOGGAGCTACTAACTTCAGCCTGC
TGAAGCAGGCTGGAGAGCTGGAGGAGAAOCTGGACCTATGGTGAACAAGGGOGAGGAGCTGTTCAOOGGGTGGTGOCC
ATOCTGGTGOAGCTGGACGGCAGCTAAACGGCCACAAGTTCAAGCTGTCCGGCAGGGOGAGGGCGATGCCAOCTAOGG
CAAGCTGACCCCTGAAGTTTCACTCTGCACACCCGGCAAGCTGCCCGTCCCTGGCCACCCCTCGTGACCACCCCTGACCTAOC
GCGTGCAGTGTCTCAGCCGCTACCCCGACACATGAAGCAGCAGACTTCTTCAAGTCCGCCATGCCCGAAGGCTACGTC
CAGGAGCGCACCATCTTCTTCAAGGACGACGGCAACTACAAGACCCCGCCGAGGTGAAGTTGAGGGGACACCCCTGGT
GAACCCGATCGAGCTGAAGGGCATCGACTTCAAGGAGGACGGCAACATCCTGGGGCACAAGCTGGAGTACAACCTACAACA
GCCACAACGCTTATATCATGGCCGACAAAGCAGAAGAACGGCATCAAGGTGAACCTCAAGATCCGCCACAACATCGAGGAC
GGCAGCGTGCAGCTOCCGACCACTACCAGCAGAACACCCCATCGGCGACGGCCCGGTGCTGCTGCCGACAACCCACTA
CCTGAGCACCCAGTCCGCCCTGAGCAAAGACCCCAAGGAGAAGCGCGATCACATGGTCTGCTGGAGTTGTTGACCGCCG
CCGGGATCACTCTGGCATGGACGAGCTGTACAAGTAACTCGAGAGATCTCTCGAGAGATCCCGGGGTGACTGATCA
AATTCGAGCTCGGTACCTTTAAGACCAATGACTTACA

>pMND_1_FRET

TAGATCTCGAATCGAATTTAATTAATTAATTCOCCCATGGGATCCATGTTGGTGTCTGTTGCCCTGATGAAGTCTCAGGCCCTT
GAGCAGCTTGAGAGTATAATCAACTTTGAAAACTGACTGAATGGACAGTTCTAATGTTATGGAAGAGAGGAAGATCAA
AGTGGAAATTCGGAAGCGGAGCTACTAACTTCAGCCTGCTGAAGCAGGCTGGAGACGTGGAGGAGAACCCTGGACCTGGCC
GGCCTATGGTGAAGCAGGGGAGGAGCTGTTCAOOGGGTGGTGOCCATCCTCGTGGTGAAGCTGGACGGGACGTAAACGGC
CACAAGTTCAAGCTGTCCGGCAGGGGCGAGGGGCGATGCCAOCTACGGCAAGCTGACCTGAAGTTTCACTGCACCAOCCGG
CAAGCTGCCCGTGGCCCGCCCTCGTGACCAOCCCTGAOCTGGGGCGTGCAGTGTCTCGCCCGCTACCCCGACCCACA
TGAAGCAGCAGACTTCTTCAAGTCCGCCATGCCCGAAGGCTACGTCAGGAGCGCACCATCTTCTTCAAGGACGACGGC
AACTACAAGACCCCGCCGAGGTGAAGTTGAGGGGACACCCCTGGTGAACCCGATCGAGCTGAAGGGCATOGACTTCAA
GGAGGACGGCAACATCCTGGGGCACAAGCTGGAGTACAACGCCATCAGCGACAACGTCTATATCACCGCCGACAAGCAGA
AGAAGCGCATCAAGGCCAATTTCAAGATCCGCCACAACATCGAGGACGGCAGCGTGCAGCTGGOOAGCCACTACCAGCAG
AACACCCCATCGGGGACGGCCCGGTGCTGCTGCCGACAACCCACTACCTGAGCAOCCAGTCCAAGCTGAGCAAAGACCC
CAOAGAGAAGCGCGATCACATGGTCTGCTGGAGTTGTTGACCGCCCGGGATCACTCTGGCATGGACGAGCTGTACA
AGTTAACTAGTGTGGCCCGACTTCCGGCAGGATOGATATGGTGAACAAGGGGAGGAGCTGTTCAOOGGGTGGTGOCC
ATOCTGGTGOAGCTGGACGGCAGCTAAACGGCCACAAGTTCAAGCTGTCCGGCAGGGOGAGGGCGATGCCAOCTAOGG
CAAGCTGACCCCTGAAGTTTCACTCTGCACACCCGGCAAGCTGCCCGTCCCTGGCCACCCCTCGTGACCACCTTCGGCTAOC
GCTGCAGTGTCTCGCCGCTACCCCGACACATGAAGCAGCAGACTTCTTCAAGTCCGCCATGCCCGAAGGCTACGTC
CAGGAGCGCACCATCTTCTTCAAGGACGACGGCAACTACAAGACCCCGCCGAGGTGAAGTTGAGGGGACACCCCTGGT
GAACCCGATCGAGCTGAAGGGCATCGACTTCAAGGAGGACGGCAACATCCTGGGGCACAAGCTGGAGTACAACCTACAACA
GCCACAACGCTTATATCATGGCCGACAAAGCAGAAGAACGGCATCAAGGTGAACCTCAAGATCCGCCACAACATCGAGGAC
GGCAGCGTGCAGCTOCCGACCACTACCAGCAGAACACCCCATCGGCGACGGCCCGGTGCTGCTGCCGACAACCCACTA
CCTGAGCTACCAGTCCGCCCTGAGCAAAGACCCCAAGGAGAAGCGCGATCACATGGTCTGCTGGAGTTGTTGACCGCCG
CCGGGATCACTCTGGCATGGACGAGCTGTACAAGTAAAGGGCGCCCTCGAGAGATCCCGGGGTGACTGATCAAATT
CGAGCTCGGTACCTTTAAGACCAATGACTTACA

>pdL_1_GFP

ACCGGTTAATACGACTCACTATAGGCTAGCATTAGGTGACACTATAGAATACAAGCTACTTGTCTTTTTGCACGCGCG
CGAGGAGGGCCACCATGGTGAAGCAAGGGOGAGGAGCTGTTCAOOGGGTGGTGOCCATCCTGGTGOAGCTGGACGGOGAC
GTAAACGGCCACAAGTTCAAGCTGTCCGGCAGGGGCGAGGGCGATGCCAOCTACGGCAAGCTGACCCCTGAAGTTCACTG
CACCACCGGCAAGCTGCCCGTGGCCCGCCCTCGTGACCAOCCCTGAOCTACGGCGTGCAGTGTCTCAGCCGCTAOC
CCGACACATGAAGCAGCAGACTTCTTCAAGTCCGCCATGCCCGAAGGCTACGTCAGGAGCGCACCATCTTCTTCAAG
GACGACGGCAACTACAAGACCCCGCCGAGGTGAAGTTGAGGGGACACCCCTGGTGAACCGCATOGAGCTGAAGGGCAT
CGACTTCAAGGAGGACGGCAACATCCTGGGGCACAAGCTGGAGTACAACCTACAACAGCCACAACGTCTATATCATGGCCG
ACAAGCAGAAGAAGCGCATCAAGGTGAACCTTCAAGATCCGCCACAACATOGAGGACGGCAGCGTGCAGCTGCGGACCCAC
TACCAGCAGAACACCCCATCGGCGACGGCCCGGTGCTGCTGCCGACAACCCACTACCTGAGCACCCAGTCCCGCCCTGAG
CAAAGACCCCAAGGAGAAGCGCGATCACATGGTCTGCTGGAGTTGTTGACCGCCCGGGATCACTCTGGCATGGACG
AGCTGTACAAGTAAAGCGCCCGGACTCTAGATCATAATCAGCCATACCACATTTGTAGAGGTTTTACTTGTCTTAAAAA
ACCTCCACACCTCCCGTTAATTAACCGGATTAGCCTCTGCTTAAACTAGTTAGCCCGCTGATCAGCCTCGA

>pdL_1_XIAP-2A-FRET

ACCGGTTAATACGACTCACTATAGGCTAGCATTAGGTGACACTATAGAATACAAGCTACTTGTCTTTTTGCATTAATT
AAGCCGCCATGACTTTTTAAGGATTTGAAGGATCTAAAACCTTGTGTACCTGCAGACATCAATAAGGAAGAAGAAATTTGTA
GAAGAGTTTAAATAGATTAATAAATCTTTGCTAATTTTCAAGTGGTGTGCTGTTTCAAGCATCAACACTGGCACGAGCAGG

GTTCCTTTACTGGTGAAGGAGATACCGTGC GGTCCTTGTGTCATGCAGCTGTAGATAGATGGCAATATGGAGACT
CAGCAGTTGGAAGACACAGGAAAGTATCCCAAATTCAGATTTATCAACGGCTTTTATCTTGAAAATAGTGCCACGCAG
TCTACAAATCTGGTATCCAGAATGGTCAGTACAAGTTGAAAATCTCTGGGAAGCAGAGATCATTTCCTTAGACAG
GCCATCTGAGACACATGCAGACTATCTTTGAGAACTGGGCAGGTTGTAGATATATCAGACCCATATACCCGAGGAACC
CTGCCATGTATAGTGAAGAAGCTAGATTAAGTCCCTTCAGAACTGGCCAGACTATGCTCACCTAACCCCAAGAGATTA
GCAAGTCTGGACTCTACTACACAGGTATTGGTGACCAAGTGCAGTGCCTTTGTTGTGGTGGAAAACGAAAATTTGGGA
ACCTTGTGATCGTGCCCTGGTCAGAACACAGGCCACACTTTCCCTAATTGCTTCTTTGTTTGGGCCGGAATCTTAATATTC
GAAGTGAATCTGATGCTGTGAGTTCTGATAGGAATTCACAAATTCACAAATCTTCCAAGAAAATCCATCCATGGCAGAT
TATGAAGCACGGATCTTTACTTTTGGGACATGGATATACTCAGTTAACAAAGGAGCAGCTTCCAAGAGCTGGATTTTATGC
TTTAGTGAAGGTGATAAAGTAAAGTGCCTTCACTGTGGAGGAGGCTAACTGATTGGAAGCCCAGTGAAGACCCCTGGG
AACAAATGCTAAATGGTATCCAGGGTGC AAATATCTGTTAGAACAGAAGGGACAAGAATATATAACAATATTCATTTA
ACTCATTCACTTGGAGGTGTCTGGTAAGAACTACTGAGAAAACACCATCACTAACTAGAAGAAATGATGATACCATCTT
CCAAAATCCTATGGTACAAGAAGCTATACGAATGGGGTTCAGTTTCAAGGACATTAAGAAAATAATGGAGGAAAAAATTC
AGATATCTGGGAGCAACTATAAATCACTTGAAGTTCCTGGTTGCAGATCTAGTGAATGCTCAGAAAAGACAGTATGCAAGAT
GAGTCAAGTCAAGCTTACTACAGAAAAGAGATTAGTACTGAAGAGCAGCTAAGGCGCTGCAAGAGGAGAAGCTTTGCAA
AATCTGTATGGATAGAAATATTGCTATCGTTTTGTTCCCTTGGGACATCTAGTCACTTGTAAACAATGTGCTGAAGCAG
TTGACAAGTGTCCATGTGCTACACAGTCACTTTCAAGCAAAAAATTTTATGTCTGAATTCGGAAGCGGAGCTACT
AAGTTCAGCCTGCTGAGCAAGGCTGAGAGCTGAGAGGAAACCCCTGGAACCTGCGCGCTATGTTGAGCAAGGGCGAGGA
GCTGTTCAACGGGGTGGTGCCATCCTGGTGGAGCTGGAACGGGACGTTAAACGGCCACAAGTTCAAGCGTGTCCGGCGAGG
GCGAGGGCGATGCCACCTACCGCAAGCTGACCTGAAGTTTCACTTGCACCCACCGGCAAGCTGCCCGTGCCCTGGCCACC
CTCGTGACCACCTGACCTGGGGCTGCAGTGCCTGCGCGCTACCCCGACCCATGAAGCAGCACGACTTCTTCAAGTC
CGCCATGCCGAAGGCTACGTCCAGGAGCGCACCATCTTCTTCAAGGACGACGCAACTACAAGACCCGCGCGAGGTGA
AGTTGAGGGGACACCCCTGGTGAACCGCATCGAGCTGAAGGGCATCGACTTCAAGGAGGACGGCAACATCCTGGGGCAC
AAGCTGGAGTACAACGCCATCAGCGACAACGCTATATCACCGCCGACAAGCAGAAGAAACGGCATCAAGGCCAATTCAA
GATCGCCACAACATOGAGGACGGCAGGCTGCAGCTGCGGACCACTACCCAGCAGAACACCCCATCGGCGAOGGCCCCG
TGCTGCTGCCGACAACTACCTGAGCACCCAGTCCAAGCTGAGCAAAGACCCCAACGAGAAGCGCGATCACATGGTC
CTGCTGGAGTTGCTGACCGCGCGGGGATCACTCTCGGCATGGACGAGCTGTACAAGTTAACTAGTGTGGGCCCGACTT
CGGCAGGATOGATATGGTGAAGGGCGAGGAGCTGTTCAACGGGGTGGTGCCATCCTGGTGGAGCTGGAOGGGGACG
TAAACGGCCACAAGTTCAAGCGTGTCCGGGAGGGCGAGGGGATGCCACCTACGGCAAGCTGACCCGAAAGTTCACTGTC
ACCACCGGCAAGCTGCCGCTGCCCTGGCCACCCCTCGTGACCACCTTCGGCTACGGCTGCAGTGCCTGCCCGCTACCC
CGACCACATGAAGCAGCACGACTTCTTCAAGTCCGCCATGCCGAAGGCTACGTCCAGGAGCCACCATCTTCTTCAAGG
ACGACGGCAACTACAAGACCCGCGCGAGGTGAAGTTGAGGGGACACCCCTGGTGAACCGCATCGAGCTGAAGGGCATC
GACTTCAAGGAGGACGGCAACATCCTGGGGCACAAGCTGGAGTACAACACTACAACAGCCACAACGCTTATATCATGGCCGA
CAAGCAGAAGAACGGCATCAAGGTGAACCTTCAAGATCCGCCACAACATCGAGGAOGGCAGCGTGCAGCTGCGCGACCACT
ACCAGCAGAACACCCCATCGGGCAGCGCCCCGCTGCTGCTGCCGACAACCACTACCTGAGCTACCAGTCCGCCCCGAGC
AAAGACCCCAACGAGAAGCGCGATCACATGGTCTGCTGGAGTTGCTGACCGCGCGGGGATCACTCTCGGCATGGACGA
GCTGTACAAGTAAGCGCGCCCTOGAGAGATCCCCCGGGTGGACTGATCAAAATTOGAGCTCGGTACCACTAGTTAGCCC
GCTGATCAGCCTCGA

>pdL_|_GZB-DTA

ACCGGTTAATACGACTCACTATAGGCTAGCATTTAGGTGACACTATAGAATACAAGCTACTTGTCTTTTGCACGCGCG
CGAGGAGGGCCACCATGCAACCAATCCTGCTTCTGCTGGCTTCCCTCCTGCTGCCAGGGCAGATGCAGGGGAGATCATC
GGGGGACATGAGGCCAAGCCCACTCCCGCCCTACATGGCTTATCTTATGATCTGGGATCAGAAGTCTCTGAAGAGGTG
CGGTGGCTTCTGATACGAGACGACTTCGTGCTGACAGCTGCTCACTGTTGGGAAGCTCCATAAATGTCACTTGGGGG
CCACAATATCAAAGAACAGGAGCGCACCCAGCAGTTTATCCCTGTGAAAAGACCCATCCCCATCCAGCCTATAATCCT
AAGAACTTCTCAACGACATCATGCTACTGCAGCTGGAGAGAAAAGGCCAAGCGGACACAGAGCTGTGCAGCCCCCTCAGGCT
ACCTAGCAACAAGGCCAGGTGAAGCCAGGGCAGACATGCAGTGTGGCCGGCTGGGGCAGACGGCCCCCTGGGAAAAC
ACTCACACACACTACAAGAGGTGAAGATGACAGTGCAGGAAGATOGAAAGTGGGAATCTGACTTACGCCATTATTACGAC
AGTACCATTGAGTTGTGCGTGGGGACCCAGAGATTA AAAAGACTTCTTTAAGGGGACTCTGGAGGCCCTCTTGTGTG
TAACAAGGTGGCCAGGGCATTGTCTCTATGGACGAAACAATGGCATGCCTCCACGAGCTGCACCAAAGTCTCAAGCT
TTGTACTGATAAAGAAAACCATGAAACGCTACGGCGCGCTGGAGGCGGGGTTCTGGCGGGGTGGATCAGGGGGT
GGAGGCTCCGGTGGAGCGGGTGGGGCGCCGATCCTGATGATGTTGTTGATTCTTCAAATCTTTTGTGATGGAAAA
CTTTTCTTCTGATACCAGGGACTAAACCTGGTTATGTAGATTCCATTCAAAGGTATACAAAAGCCAAAATCTGGTACAC
AAGGAAATATGACGATGATTGGAAGGGTTTTATAGTACCAGACAATAAATACGACGCTGCGGGTACTCTGTAGATAAT
GAAAACCCGCTCTCTGAAAAGCTGGAGGCGTGGTCAAAGTGAAGTATCCAGGACTGACGAAGGTTCTCCCACTAAAAGT

GGATAATGCCGAAACTATTAAGAAAGAGTTAGGTTTAAAGTCTCACTGAACCGTTGATGGAGCAAGTCGGAACGGAAGAGT
TTATCAAAGGTTCCGTGATGGTGTCTTCGCGTGTAGTGTCTCAGCCTTCCCTTCGCTGAGGGGAGTTCTAGCGTTGAATAT
ATTAATAACTGGGAACAGCGGAAAGCGTTAAGCGTAGAACTTGAGATTAATTTGAAACCCGTTGAAAAACGTTGCCAAGA
TGCGATGTATGAGTATATGGCTCAAGCCTGTGCAGGAAATCGTGTCCAGGCGATCTCTTTGTGAAGGAACCTTACTTCTGT
GGTGTGACATAAATGGACAACTACCTACAGAGATTTAAAGCTCTAACTTAATTAACGCGATTAGCCTCTGCTTAAACTA
GTTAGCCCCGCTGATCAGCCTCGA

>pdL_|_GZB-NFSB-2A-GFP

ACCGGTTAATACGACTCACTATAGGCTAGCATTAGGTTGACACTATAGAATACAAGCTACTTGTCTTTTTGCATTAATT
AAGCCGCCATGCAACCAATCCTGCTTCTGCTGGCCTTCCCTGCTGCCAGGGCAGATGCAGGGGAGATCATCGGGGA
CATGAGGCCAAGCCCCACTCCCGCCCTACATGGCTTATCTTATGATCTGGGATCAGAAGTCTCTGAAGAGGTGCGGTGG
CTTCTGATACGAGACGACTTCGTGCTGACAGCTGCTCACTGTTGGGGAAGCTCCATAAATGTCACTTGGGGGCCACA
ATATCAAAGAACAGGAGCCGACCCAGCAGTTTATCCCTGTGAAAAGACCCATCCCCATCCAGCCTATAATCCTAAGAAC
TTCTOCAACGACATCATGCTACTGACAGCTGGAGAGAAAGGCAAGCGGACCAGAGCTGTGCAGCCCTCAGGCTACCTAG
CAACAAGGCCAGGTGAAGCCAGGGCAGACATGCAGTGTGCOGGCTGGGGGCAGACGGCCCCCTGGGAAAACACTCAC
ACACACTACAAGAGGTGAAGATGACAGTGCAGGAAGATCGAAAGTGCAGTCTGACTTACGCCATTATACGACAGTACC
ATTGAGTTGTGCGTGGGGGACCCAGAGATTAAGAACTTCCCTTAAAGGGGACTCTGGAGGCCCTCTTGTGTAAACAA
GGTGGCCAGGGCATTGTCTCTATGACGAAACAATGGCATGCCTCCACGAGCTGCACCAAAAGTCTCAAGCTTTGTAC
ACTGGATAAAAGAAACCTGAAACGCTACCGCCGCTGGAGGCGGGGTTCTGGCGGGGTGATCAGGGGTGGAGGC
TCCGTTGGAGGCGGTTCCGGCCGCCCGATATCATTCTGTGCCCTTAAAGCGTCACTTCCACTAAGGCATTGTATGCCAG
CAAAAACTTACCCCGGAACAGGCGCAGCAGATCAAAACGCTACTGCAATACAGCCCATCCAGCACCAACTCCAGCCGT
GGCATTTTATTGTTGCCAGCACGGAAGAAGTAAAGCGCGTGTGCCAAATCCGCTGCCGGTAATTACGTGTTCAACGAG
CGTAAATGCTTGTGCTCGCAGTGTGTTCTGTGCAAAAACCGCATGGACGATGCTGGCTGAAGCTGGTGT
TGACCAGGAAGATGCCGATGGCCGCTTTGCCACGCGGAAGCGAAAGCGGCAACGATAAAGTCCGCAAGTTCTTCGCTG
ATATGCACCGTAAAGATCTGCATGATGATGCAGAGTGGATGGCAAAACAGGTTTATCTCAACGTCCGTAACCTCCTGCTC
GGGTGGCGGCTCTGGGTCTGGACGCGGTACCCATCGAAGTTTTGACGCGCCATCCTCGATGCAGAAATTTGGTCTGAA
AGAGAAAGGCTACACCAGTCTGGTGGTGTTCGGTAGGTATCACAGCGTTGAAGATTTAAAGCTACGCTGCCGAAAT
CTGCTGCCGCAAAACATCACCTTAAACGAAAGTGAATTCGGAAGCGGAGCTACTAATTCAGCCTGCTGAAGCAGGCT
GGAGAGTGGAGGAGAACCCCTGGAOCTATGGTGAACAAGGGGAGGAGCTGTTCAACGGGGTGGTCCCATCCTGGTGA
GCTGGACGGGACGTAACCGCCACAAGTTGAGCGTGTCCGGGAGGGGAGGCGATGCCACCTACGGCAAGCTGACCC
TGAAGTTGATCTGCACCAACCGCAAGCTGCCCGTCCCTGGCCACCCCTCGTGACCACCTGACCTACGGCGTGCAGTGC
TTCAGCCGCTACCCCGACACATGAAGCAGCACGACTTCTTCAAGTCCGCCATGCCGGAAGGCTACGTCCAGGAGCGCAC
CATCTTCTCAAGGACGACGGCAACTACAAGACCCCGCCGAGGTGAAGTTCGAGGGCGACCCCTGGTGAACCGCATCG
AGCTGAAGGGCATCGACTTCAAGGAGGACGGCAACATCCTGGGGCACAAGCTGGAGTACAACACTACAACAGCCACAAGCTC
TATATCATGGCCGACAAGCAGAAGAACGGCATCAAGGTGAACCTCAAGATCCGCCACAACATCGAGGACGGCAGCGTGA
GCTCGCCGACCACTACCAGCAGAACACCCCATCGGGACGGCCCGGTGCTGCTGCCGACAACCACTACCTGAGCAACC
AGTCCGCCCTGAGCAAAGACCCCAACGAGAAGCGGATCACATGGTCTGCTGGAGTTGTTGACCCGCGCGGGATCACT
CTCGCATGGACGAGCTGTACAAGTAACTCGAGAGATCCCCGGGGTGCAGTGTCAAATTCGAGCTCGGTACCACTAGT
TAGCCCGCTGATCAGCCTCGA

>pdL_|_NFSB-2A-GFP

ACCGGTTAATACGACTCACTATAGGCTAGCATTAGGTTGACACTATAGAATACAAGCTACTTGTCTTTTTGCATTAATT
AAGCCGCCATGACGCCAACCTTGAACCTTATTTGTGCCATCGCTCCATTCCGCAATTCAGTGAACCCATTTCCGAA
GCGCAGCGTGAGCGATTATTAACAGCGCCGTCGACGTCAGTCCAGTTCCAGTTTTTTGACAGTGCAGTAGCATTATTCGAT
TACCGACAAAGCGTTACGTGAAGAACTGGTGAACGCTGACCGCGGGCAAAAACACGTAGCGCAAGCGCGGAGTTCTGGG
TGTTCTGTGCCGACTTAAACCGCATTTACAGATCTGTCCGGATGCTCAGCTGGCCCTGGCGGAACAACTGTTGCTGGT
GTGTTGATACGGCAATGATGGCGCAGAATGCATTAATCGCAGCGGAATCGCTGGGATTGGGCGGGTATATATCGCCGG
CCTGCGCAATAATATTGAAGCGGTGACGAACTGCTTAAATTAACCGCAGCATGTTCTGCCGCTGTTTGGGCTGTGCTTG
GCTGGCCTGCGGATAATCCGGATCTTAAAGCGGTTTAAACCGCCTCCATTTGGTGCATGAAAACAGCTATCAACCGCTG
GATAAAGGCGCACTGGCGCAGTATGACGAGCAACTGGGGAAATATTACCTCACCCGTGGCAGCAATAATGCCGGGATAC
CTGGAGGATCATATCCGCCAACAATCATTAAAGAAAGCGCCCATTTATCTGATTATTTGCACAAACAGGGTTGGG
CGACGCGGAATTCGGAAGCGGAGCTACTAATCTCAGCCTGCTGAAGCAGGCTGGAGAGCTGGAGGAGAACCCTGGACCT
ATGGTGAACAAGGGGAGGAGCTGTTCAACGGGGTGGTCCCATCCTGGTGCAGCTGGAAGCGGAGCTAAACGGCCACAA
GTTGAGCGTGTCCGGGAGGGGAGGGCGATGCCACCTACGGCAAGCTGACCCGTAAGTTTCACTGACCCACCGGCAAGC
TGCCCGTGCCTGGCCACCCCTCGTGAACCCCTGACCTACGGCGTGCAGTGTCTTACGCCGCTACCCCGACCATGAAG
CAGCAGACTTCTTCAAGTCCGCCATGCCGGAAGGCTACGTCCAGGAGCGCACCATCTTCTTCAAGGACGACGGCAACTA

CAAGACCCGCGCAGGTGAAGTTGAGGGGACACCCCTGGTGAACCGCATOGAGCTGAAGGGCATOGACTTCAAGGAGG
ACGGCAACATCCTGGGGCACAAGCTGGAGTACAACAGCCACAACGCTATATCATGGCCGACAAGCAGAAGAAC
GGCATCAAGGTGAACCTCAAGATCOGCCACAACATOGAGGACGGCAGCGTGCAGCTOCCOACCCTACCAGCAGAACAC
CCCCATGGCGACGGCCCCGCTGCTGCTGCCCGACAACCCTACCTGAGCACCCAGTCCGCCCTGAGCAAAGACCCCAACG
AGAAGCGGATCACATGGTCTGCTGGAGTTGCTGACCCGCCCGGGGATCACTCTCGGCATGGAACGAGCTGTACAAGTAA
CTCGAGAGATCCCCCGGGTGCAGTGAATTCGAGCTCGGTACCCTAGTTAGCCCCGCTGATCAGCCTCGA

>pdL_ |_SURV-2A-FRET

ACCGGTTAATACGACTCACTATAGGCTAGCATTAGGTGACACTATAGAATACAAGCTACTTGTCTTTTTGCATTAATT
AAGCCGCCATGGGTGCCCGACGTTGCCCCCTGCCTGGCAGCCCTTTCTCAAGGACCACCGCATCTCTACATCAAGAAC
TGCCCTTCTTGAGGGCTGCGCCTGCACCCCGGAGCGGATGGCCGAGGCTGGCTTATCCACTGCCCCACTGAGAACGA
GCCAGACTTGGCCAGTGTCTTCTGCTTCAAGGAGCTGGAAGGCTGGGAGCCAGATGACGACCCCATAGAGGAACATA
AAAAGCATTGTCGGTTGCGCTTCTTCTGCTCAAGAAGCAGTTTGAAGAATTAACCCCTGGTGAATTTTTGAAACTG
GACAGAGAAAGAGCCAAGAACAATTCACAAAGGAAACCAACAATAAGAAGAAAGAAATTTGAGGAACTGCGGAGAAAGT
GCGCGTGCATCGAGCAGCTGGCTGCCATGGATGAATTOGGAAGCGGAGCTACTAATTCAGCCTGCTGAAGCAGGCTG
GAGAGTGGAGGAGAACCCTGGACCTGGCCCGCCTATGGTGAGCAAGGGGAGGAGCTGTTCAOCCGGGTGGTGCCCATC
CTGGTCGAGCTGAGCGCGACGTAACCGGCCACAAGTTCAAGCTGTCCCGGAGGGGAGGGCGATGCCACCTAOCGCCAA
CTGACCTGAAAGTTCACTGCCACCCGCAAGCTGCCCGTCCCTGGCCACCCCTGCTGACCCACTGACTGCGGGCG
TGCAGTGTCTCCCGCCCTACCCGACCCACATGAAGCAGCACGACTTCTTCAAGTCCGCCATGCCGAAGGCTACGTCOCAG
GAGCGCACCTTCTTCAAGGACGACGGCAACTACAAGAACCCCGCCGAGGTGAAGTTGAGGCGACACCCCTGGTGAA
CCGCATCGAGCTGAAGGGCATCGACTTCAAGGAGACGGCAACATCCTGGGGCACAAGCTGAGTACAACGCCATCAGCG
ACAAGCTTATATCACCGCCGACAAGCAGAAGAAGGCATCAAGGCCAATTCAGATCCGCCACAACATCGAGGACGGC
AGCGTGCAGCTCGCCGACCACTACCAGCAGAACACCCCATCGCGACGGCCCGCTGCTGCTGCCGACAACCCTACCT
GAGCACCCAGTCCAAGCTGAGCAAAGACCCCAACGAGAAGCGGATCACATGGTCTGCTGGAGTTCGTGACCCCGCCG
GGATCACTCTCGGCATGAGCAGCTGTACAAGTTAAGTGTGGCCCGACTTCGGCAGGATOGATATGGTGAGCAAG
GGGAGGAGCTGTTCAOCCGGGTGGTGCCATCCTGGTGGAGCTGGACGGGACGTAACCGCCACAAGTTCAAGGTGTC
CGGGAGGGGAGGGGATGCCACCTAOCGCCAAGCTGACCTGAAGTTCACTGCAOCCOCCGCAAGCTGCCCGTCCCT
GGCCACCCCTGTCGACCCCTTCGGCTACGGCTGCAGTCTTCCCGGCTACCCCGACCATGAAGCAGCAGACTTC
TTCAAGTCCGCATGCCGAAGGCTACGTCCAGGAGCGCACCATCTTCTTCAAGGACGACGGCAACTACAAGACCCGCGC
CGAGGTGAAGTTGAGGGCGACACCCCTGGTGAACCGCATOGAGCTGAAGGGCATOGACTTCAAGGAGGACGGCAACATCC
TGGGGCACAAGCTGGAGTACAACACTACAACAGCCACAACGTCTATATCATGGCCGACAAGCAGAAGAAGCGCATCAAGGTG
AACTTCAAGATCCCGCACACATCGAGGACGGCAGCGTGCAGCTCGCCGACCACTACCAGCAGAACACCCCATCCGCGA
CGCCCCCGTGTCTGCTGCCCGACAACCCTACCTGAGCTACCAGTCCGCCCTGAGCAAAGACCCCAACGAGAAGCGCGATC
ACATGGTCTGCTGGAGTTGCTGACCCCGCCGGATCACTCTCGGCATGGACGAGCTGTACAAGTAAGCCGCGCCCTCG
AGAGATCCCCCGGGTGCAGTGAATTCGAGCTCGGTACCCTAGTTAGCCCCGCTGATCAGCCTCGA

>pdL_ |_GZB-PEA

ACCGGTTAATACGACTCACTATAGGCTAGCATTAGGTGACACTATAGAATACAAGCTACTTGTCTTTTTGCACGCGCG
CGAGGAGGGCCACCATGCAACCAATCCTGCTTCTGCTGGCTTCTCCTGCTGCCAGGGCAGATGCAGGGGAGATCATC
GGGGGACATGAGGCCAAGCCCACTCCCGCCCTACATGGCTTATCTTATGATCTGGGATCAGAAGTCTCTGAAGAGGTG
CGGTGGCTTCTGATACGAGACGACTTCGTGCTGACAGCTGCTCACTGTTGGGAAGCTCCATAAATGTCACCTTGGGGG
CCACAATATCAAAGAACAGGAGCCGACCCAGCAGTTTATCCCTGTGAAAAGACCCATCCCCATCCAGCCTATAATCCT
AAGAACTTCTCAACGACATCATGCTACTGCACTGGAGAGAAAGGCCAAGCGACAGAGCTGTGACGCCCTCAGGCT
ACCTAGCAACAAGGCCAGGTGAAGCCAGGGCAGACATGCACTGTGGCCGGCTGGGGGACAGCGCCCGCTGGGAAAAC
ACTCACACACTACAAGAGGTGAAGATGACAGTGCAGGAAGATCGAAAGTGCGAATCTGACTTACGCCATTATTACGAC
AGTACCATTGAGTTGTGCGTGGGGGACCCAGAGATTAAGAAAGACTTCCCTTAAAGGGGACTCTGGAGGCCCTCTTGTG
TAACAAGGTGGCCAGGCATTGTCTCTATGGACGAAAACATGGCATGCCTCCACGACCTGCACCAAAGTCTCAAGCT
TTGTACTGATGATAAAGAAAACCATGAAACGCTACGCGCCCGCTGGAGGGGGGTTCTGGGGGGGTGGATCAGGGGGT
GGAGGCTCCGCTGGAGGGGGTGGGGGGCCCAACCGGTGCCGAGTTCTGGGGGACGGTGGCGATGTGCTTTAG
CACCCGTGGTACCAGAAGCTGACGGTGTAGAGCGCTGCTGCAGGCACATGCTGAGCTGGAAGAGCGTGGCTATGTATCG
TTGGCTACCACGGCACTTTCTGGAAGCAGCTCAGTCCATCGTGTGGTGGTGTCCGTGCCCGTTCTCAAGACTGGAT
GCGATTTGGCGTGGTTCTACATTGCAGGCGATCCAGCGCTGCATACGGTTATGCGCAGGACCCAGGAACCGGATGCTCG
CGGTGCCATTOGTAATGGTGGCTGCTGCGGTATATGTGCCGGTTCAGCCTGCCGGCTTCTACCCACTAGCCTGA
CCCTGGCCGCGCCGAGGGCGGGTGAAGTGAACGCTCTGATTGGTCACTCTGCCCTCGCCCTGGATGCCATCAC
GGCCAGAGGAGGAGGGCGGTGCTGGAACCAATCTGGGCTGGCCGCTGCTGAACGTACGGTCTTATTCCGAGCGC
GATTCTACCAGTCTCTGTAACGTTGGCGGGCATCTGGACCCATCTTCTATTCCAGATAAGGAGCAGGCAATCTCCGCGC

TGCCGATTATGCAAGCCAAACCGGGTAAACCACCTCGTGAAGATCTGAAATAACTCGAGCCTGCCGGCCTTAATTAACG
CGATTAGCCTCTGCTTAAACTAGTTAGCCCGCTGATCAGCCTCGA

>pdL_ | _NFSB-2A-GFP

ACCGGTTAATACGACTCACTATAGGCTAGCATTAGGTGACACTATAGAATACAAGCTACTTGTCTTTTTGCATTAATT
AAGCCGCCATGGATATCATTCTGTGCGCTTAAAGCGTCATTCCACTAAGGCATTTGATGCCAGCAAAAACTTACCCCG
GAACAGGCCGAGCAGATCAAAACGCTACTGCAATACAGCCATCCAGCACCAACTCCAGCCGTGGCATTATTATTGTTGC
CAGCACGGAAGAAGGTAAAGCGCGTGTGCCAAATCCGCTGCCGGTAATTACGTGTTCAACGAGCGTAAAAATGCTTGATG
CCTCGCACGTGCTGGTGTCTGTGCAAAAAACCGCATGGACGATGCTGGCTGAAGCTGGTTGTTGACCAGGAAGATGCC
GATGGCCGCTTTGCCACGCCGAAGCGAAAGCCGCGAACGATAAAGGTCGCAAGTTCTTCGCTGATATGCACCGTAAAGA
TCTGCATGATGATGCAGAGTGGATGGCAAAAACAGGTTTATCTCAACGTCGGTAACCTCCGCTCGCGGTGGCGGCTCTGG
GTCTGGACGCGGTACCCATCGAAGGTTTTGACGCCGCCATCCGATGCAGAAATTTGGTCTGAAAAGAGAAAGGCTACACC
AGTCTGGTGGTTGTTCCGGTAGGTCATCACAGCGTTGAAGATTTTAAACGCTACGCTGCCGAAATCTCGTCTGCCGAAAA
CATCACCTTAAACGAAGTGAATTCGGAAGCGGAGCTACTAATTCAGCCTGCTGAAGCAGGCTGGAGACGTGGAGGAGA
AOCCTGGACCTATGGTGAGCAAGGGGOGAGGAGCTGTTACOCGGGGTGGTCCCATCCTGGTOGAGCTGGACGGOGACGTA
AACGGCCACAAGTTACGGTGTCCGGOGAGGGOGAGGGOGATGCCACCTAOCGCAAGCTGAOCCCTGAAGTTTCATCTGCAC
CACCGGCAAGCTGCCGTGCCCTGGCCACCCCTGTAACACCCCTGACCTAOCGGGTGCAGTCTTACGCCGCTACCCCG
ACCACCTGAAGTGAACGACGACTTCTTCAAGTCCGCCATGCCGGAAGGCTACGTCCAGGAGCGCACCATCTTCTTCAAGGAC
GACCGCTAACAGACCCCGCGAGGTGAAGTTGAGGGGACACCCCTGGTGAACCGCATCGAGCTGAAGGGCATCGA
CTTCAAGGAGGACGGCAACATCCTGGGGCACAAAGCTGGAGTACAACACTACAACAGCCACAACGCTCTATATCATGGCCGACA
AGCAGAAGAACCGCATCAAGGTGAACCTTCAAGATCCGCCACAACATCGAGGACGGCAGCGTGCAGCTCGCCGACCACTAC
CAGCAGAACACCCCATCGCGCAGCGCCCGTGTCTGCTGCCGACAACCACTACCTGAGCACCCAGTCCGCCCTGAGCAA
AGACCCCAACGAGAAGCGCATCACATGGTCTGCTGGAGTTGCTGACCGCCCGGGATCACTCTCGGCATGGACGAGC
TGTACAAGTAACTCGAGAGATCCCCCGGGTCCGACTGATCAAATTCGAGCTCGGTACCACACTAGTTAGCCCGCTGATCAGC
CTCGA

>pMND_ | _GZB-2A-GFP

TAGATCTCGAATCGTTAATTAATTTGCGGCCATGGGATCCGGAAGCGGAGAGGGCAGAGGAAGTCTGCTAACATGCGGTG
ACGTCGAGGAGAATCCTGGACCTATGCAACCAATCCTGCTTCTGCTGGCCTTCTCCTGCTGCCAGGGCAGATGCAGGG
GAGATCATCGGGGACATGAGGCCAAGCCCATCCCGCCCTACATGGCTTATCTTATGATCTGGGATCAGAAGTCTCT
GAAGAGGTGCGGTGCTTCTGATACGAGACGACTTCTGCTGACAGCTGCTCACTGTTGGGAAGCTCCATAAATGTCA
CCTTGGGGGCCACAATATCAAAGAACAGGACCGCACCCAGCAGTTTATCCCTGTGAAAAGACCCATCCCCATCCAGCC
TATAATCCTAAGAATCTTCCAACGACATCATGCTACTGCAGCTGGAGAGAAAGGCCAAGCGGACCAGAGCTGTGACGCC
CCTCAGGCTACCTAGCAACAAGGCCAGGTGAAGCCAGGGCAGACATGCAGTGTGCCCGCTGGGGGCAGACGGCCCGCC
TGGGAAAACACTCACACACTACAAGAGGTGAAGATGACAGTGCAGGAAGATCGAAAGTGCGAATCTGACTTACGCCAT
TATTAACGACAGTACCATTGAGTTGTGCGTGGGGACCCAGAGATTAATAAGACTTCTTTAAGGGGGACTCTGGAGGCC
TCTTGTGTGTAACAAGGTGGCCAGGGCATTGTCTCCTATGGACGAAACAATGGCATGCCCTCCACGAGCCTGCACCAAG
TCTCAAGCTTTGTACTCTGGATAAAGAAAACCATGAAACGCTACGAAATTCGGAAGCGGAGCTACTAACTTACGCTGCTG
AAGCAGGCTGGAGACGTGGAGGAGAACCCTGGACCTATGGTGAGCAAGGGGAGGAGCTGTTACCGGGGTGGTGOCCAT
CCTGGTOGAGCTGGACGGGACGTAACCGGCCACAAGTTACGCGTGTCCGGOGAGGGOGAGGGOGATGCCACCTAOCGCA
AGCTGAOCCCTGAAGTTTCATCTGCACCACCGGCAAGCTGCCCGTCCCTGGCCACCCCTGTAOCCACCTGACCTAOCGC
GTGCAAGTCTTACCGGCTACCCCGACCATGAAGCAGCACGACTTCTTCAAGTCCGCCATGCCGGAAGGCTACGTCCA
GGAGCGCACCATCTTCTTCAAGGACGACGGCAACTACAAGACCCCGGAGGTGAAGTTGAGGGCGACACCCCTGGTGA
ACCGCATCGAGCTGAAGGGCATCGACTTCAAGGAGGACGGCAACATCCTGGGGCACAAAGCTGGAGTACAACACTACAACAGC
CACAACTCTATATCATGGCCGACAAGCAGAAGAACGGCATCAAGGTGAACCTTCAAGATCCGCCACAACATCGAGGACGG
CAGCGTGCAGCTGCGGACCACTACCAGCAGAACACCCCATCGGGCAGCGCCCGTGTCTGCTGCCGACAACCACTACC
TGAGACCCAGTCCGCCCTGAGCAAAGACCCCAACGAGAAGCGCATCACATGGTCTGCTGGAGTTGCTGACCGCGCC
GGGATCACTCTCGCATGGACGAGCTGTACAAGTAACTCGAGAGATCTCTOGAGAGATCCCCCGGGTCCGACTGATCAA
TTCGAGCTCGGTACCTTTAAGACCAATGACTTACA

>pdL_ | _GZB-HTK-2A-GFP

ACCGGTTAATACGACTCACTATAGGCTAGCATTAGGTGACACTATAGAATACAAGCTACTTGTCTTTTTGCATTAATT
AAGCCGCCATGCAACCAATCCTGCTTCTGCTGGCCTTCTCCTGCTGCCAGGGCAGATGCAGGGGAGATCATCGGGGA
CATGAGGCCAAGCCCACTCCCGCCCTACATGGCTTATCTTATGATCTGGGATCAGAAGTCTCTGAAGAGGTGCGGTGG
CTTCTGATACGAGACGACTTCTGCTGACAGCTGCTCACTGTTGGGAAGCTCCATAAATGTACCTTGGGGGCCACA
ATATCAAAGAACAGGAGCCGACCCAGCAGTTTATCCCTGTGAAAAGACCCATCCCCATCCAGCCTATAATCCTAAGAAC
TTCTCAACGACATCATGCTACTGCAGCTGGAGAGAAAGGCCAAGCGGACAGAGCTGTGCAGCCCTCAGGCTACCTAG

CAACAAGGCCAGGTGAAGCCAGGGCAGACATGCAGTGTGGCCGGCTGGGGCAGAACGGCCCCCTGGGAAACACTCAC
ACACACTACAAGAGGTGAAGATGACAGTGCAGGAAGATCGAAAGTGCAGTCTGACTTACGCCATTATTACGACAGTACC
ATTGAGTTGTGCGTGGGGGACCCAGAGATTAAGAAAGACTTCCCTTTAAGGGGGACTCTGGAGGCCCTCTTGTGTAAACA
GGTGGCCAGGGCATTGTCTCTATGGACGAAACAATGGCATGCCTCCACGAGCCTGCACCAAAGTCTCAAGCTTTGTAC
ACTGATAAAGAAAACCATGAAACGCTACGGGCCCTGGAGGGGGGGTTCTGGGGGGGTGGATCAGGGGGTGGAGGC
TCCGGTGGAGGGGGTGGGGCGGGCCCTTCTGACCCCTGCCATCAACACGGCTCTGGTTCGACCCAGGCTGGGGTTT
TCCGGGCCATAGCAACCGACGTACGGCGTTGGCCCTCGCCGCGCAGCAAGAAGCCACGGAAAGTCCGCCCTGGAGCAGAAAA
TGCCACCGCTACTGCGGGTTTATATAGACGGTCTCACGGGATGGGAAAAACACCACCGCAACTGCTGGTGGCCCTG
GGTTCGCGCAGCATATGCTACTGACCCGAGCCGATGACTTACTGGCAGGTGCTGGGGGCTTCGAGACAATCGCGAA
CATCTACACCACACAACCCGCCCTCGACCAGGTGAGATATCGCCGGGACCGGGCGGTGTAATGACAAGGCCCCAGA
TAACAATGGGCATGCCCTTATGCCGTGACCGACCGCTTCTGGCTCCTCATATCGGGGGGAGGCTGGGAGCTCACATGCC
CCGCCCGGCCCTCACCTCATCTTCCAGCCCATCCCATCGCCCTCCTGTGCTACCCGGCCGGCGGATACTTAT
GGGACGATGACCCCCAGGCGGTGCTGGGGTTCGTCGCTCATCCCGGACCTTCCCGGCACAAACATCGTGTGG
GGGCCCTTCGGAGGACAGACATGACCGCTGGCCAAAGCCAGCGCCCGGGGAGGGCTTGAACCTGGCTATGCTG
GCCGGATTCCGGCGTTTACGGCTGCTTGCATAAAGGTGGGTATCTGCAGGGGGGGGGTGGTGGCGGGAGGATTG
GGGACAGCTTTGGGAAAGCGCGTGGCCCGCCAGGTGGGAGCCCGAGCAAGCGGGCCACGACCCCATATGGGG
ACACGTTATTTACCGTGTTCGGCCCGCCGAGTTGCTGGCCCGCAAGCGGACCTGTACAACGTTTTCCTGGCCCTTG
GACGCTTGGCCAAAGCCCTCCGTCACGCTCTTTATCTGAGTACGACCAATCGCCCGCGGCTGCCGGGACGC
CCTGCTGCAACTTACTCCGGGATGTCAGACCCACGTCACCCACCCCGGCTCCATACCGACGATCTGCGACCTGGCGC
GCAAGTTTGGCCGGGAGATGGGGGAGGCTAACGAATTCGGAAGCGGAGCTACTAATTCAGCCTGCTGAAGCAGGCTGGA
GAGTGGAGGAGAACCTGGACCTATGTTGAGCAAGGGCGAGGAGCTGTTACCCGGGGTGGTCCCATCCTGTTGAGCT
GGACGGCAGCTAAACGGCCACAAGTTACGGTGTCCGGCGAGGGCGAGGGCGATGCCACCTAAGCAAGCTGACCTGA
AGTTTCATCTGCACCCAGGCAAGCTGCCCGTCCCTGGCCACCCCTCGTGAACCCCTGACCTACGGCGTGCAGTGTTC
AGCCGTACCCCGACACATGAAGCAGCAGGACTTCTTCAAGTCCGCCATGCCGAAGGCTACGTCAGGAGCGCACCAT
CTTCTTCAAGGACGACGGCAACTACAAGACCCCGCGGAGGTGAAGTTTCGAGGGGACACCCCTGGTGAACCGCATCGAGC
TGAAGGGCATGACTTCAAGGAGGACGGCAACATCTGGGGCACAAGCTGGAGTACAACACTACAACAGCCACAACGCTCTAT
ATCATGGCCGACAAGCAGAAGAAGGCATCAAGGTGAATTCAGATCCGCCACAACATCGAGGACGGCAGCGTGCAGCT
CGCCGACCACTACCAGCAGAACCCCATCGGCGACGGCCCGTGTGCTGCCGACAACCACTACCTGAGCAACCGT
CCGCCCTGAGCAAGACCCCAACGAGAAGCGGATCACATGGTCTGCTGGAGTTGCTGACCGCCCGGGGATCACTCTC
GGCATGGACGAGCTGTACAAGTAACTCGAGAGATCCCCGGGGTGGACTGATCAAATTCGAGCTCGGTACCACTAGTTAG
CCCCTGATCAGCTCGA

>pMND_1_GZB-DTA-2A-GFP

TAGATCTCGAATCGTTAATTAATTTGCCGCCATGGGATCCGGAAGCGGAGAGGGCAGAGGAAGTCTGCTAACATGCGGTG
ACGTGAGGAGAATCCTGGACCTATGCAACCAATCCTGCTTCTGCTGGCCTTCTCCTGCTGCCAGGGCAGATGCAGGG
GAGATCATCGGGGGACATGAGGCCAAGCCCCACTCCCGCCCTACATGGCTTATCTTATGATCTGGGATCAGAAGTCTCT
GAAGAGGTGCGGTGGCTTCTGATACGAGACGACTTCTGCTGACAGCTGCTCACTGTTGGGGAAGCTCCATAAATGTCA
CCTTGGGGGCCACAATATCAAAGAACAGGACGCCAGCCAGCAGTTTATCCCTGTGAAAAGACCCATCCCCATCCAGCC
TATAATCCTAAGAACTTCTCAACGACATCATGCTACTGCAGCTGGAGAGAAAAGGCCAAGCGGACAGAGCTGTGCAGCC
CCTCAGGCTAOCCTAGCAACAAGGCCAGGTGAAGCCAGGGCAGACATGCAGTGTGGCCGGCTGGGGCAGAGGGCCCC
TGGGAAAACACTCACACACTACAAGAGGTGAAGATGACAGTGCAGGAAGATCGAAAGTGCAGTCTGACTTACGCCAT
TATTACGACGATACCATTGAGTTGTGCGTGGGGGACCCAGAGATTAAGAAAGACTTCCCTTTAAGGGGGACTCTGGAGGCC
TCTTGTGTGTAACAAGGTGGCCAGGGCATTGTCTCTATGACGAAACAATGGCATGCCCTCCAGAGCCTGCACCAAAG
TCTCAAGCTTTGTACTGATAAAGAAAACCATGAAACGCTACCGCGCCGCTGGAGGGGGGGTCTGGCGGGGGTGG
TCAGGGGGTGGAGGCTCCGGTGGAGGGGGTGGGGCGCCCGATCCTGATGATGTTGTTGATTCTTCTAATCTTTTGT
GATGAAAACTTTTCTCGTACCACGGGACTAAACCTGGTTATGTAGATTCCATTCAAAAAGGTATACAAAAGCCAAAAT
CTGGTACACAAGGAAATATGACGATGATTGAAAAGGGTTTATAGTACCGACAATAAATACGACGCTGCGGGATACTCT
GTAGATAATGAAAACCGCTCTCTGGAAAAGCTGGAGGCGTGGTCAAAGTACGATATCCAGGACTGACGAAAGTTCTGCG
ACTAAAAGTGGATAATGCCGAACTATTAAGAAAAGGTTAGGTTAAGTCTCACTGAACCGTTGATGGAGCAAGTCGGAA
CGGAAGAGTTTATCAAAGGTTCCGGTGTGGTCTCGCGTGTAGTCTCAGCCTTCCCTCGCTGAGGGGAGTTCTAGC
GTTGAATATATTAATACTGGGAACAGGGCAAGCGTTAAGCGTAGAACTTGAATTAATTTGAAAACCGTGGAAAACG
TGGCCAAGATGCGATGATGAGTATGGCTCAAGCCTGTGCAGGAAATCGTGTGAGGGGATCTCTTTGTGAAGGAACT
TACTTCTGTGGTGTGACATAATTTGACAAAACCTACCTACAGAGATTTAAAGCTCGAATTCGGAAGCGGAGCTACTAATTC
AGCCTGCTGAAGCAGGCTGGAGAGGTGGAGGAGAACCCCTGGACCTATGTTGAGCAAGGGCGAGGAGCTGTTACCCGGGGT
GGTGGCCATCCTGTTGAGCTGGAAGCGGACGTAACCGGCCACAAGTTGAGCGTGTCCGGCGAGGGCGAGGGCGATGCCA

CCTACGGCAAGCTGACCCTGAAGTTCATCTGCACCACCGGCAAGCTGCOOCTGCCCTGGCCACCOCTCGTGACCACCOCTG
AOCTAOGGCGTGCACTGCTTCAGCOGCTAOCOCGACCACATGAAGCAGCAGACTTCTTCAAGTCCGOCATGCCCCAAGG
CTAOGTCCAGGAGCGCACCATCTTCTTCAAGGACGACGGCAACTACAAGACCCGCGCCGAGGTGAAGTTCGAGGGGACACA
COCTGGTGAACCCATCGAGCTGAAGGGCATCGACTTCAAGGAGGACGGCAACATCCTGGGGCACAAGCTGGAGTACAAC
TACAACAGCCACAACGTCTATATCATGGCCGACAAGCAGAAGACCGCATCAAGGTGAACCTCAAGATCCGCCACAACAT
CGAGGACGGCAGCGTGCAGCTGCGCGACCCTACCAGCAGAACACCCCATCGCGCAGGGCCOCTGCTGCTGCCGACA
ACCCTAOCCTGAGCACCAGTCCGCCCTGAGCAAAGACCCCAACGAGAAGCGCGATCACATGGTCTGCTGGAGTTCGCTG
ACCGCCGCCGGATCACTCTCGGCATGGACGAGCTGTACAAGTAACTCGAGAGATCTCTCGAGAGATCCCCCGGGGTGCA
CTGATCAAATTCGAGCTCGGTACCTTTAAGACCAATGACTTACA

>pMND_1_GFP

TAGATCTCGAATCGTTAATTAATGCCCCTGGAATTCGGAAGCGGAGCTACTAATTCAGCCTGCTGAAGCAGGCTG
GAGACGTGGAGGAGAACCCCTGGACCTATGGTGAAGGCGGAGGAGCTGTTCAOCCGGGTGGTGCOCATCCTGGTGGAG
CTGGAOCCGGOAGTAAACGGCCACAAGTTCAGOGTGTCCGGGAGGGGAGGGGATGCCAOCCTACGGCAAGCTGACCOCT
GAAGTTCATCTGCACCACCGGCAAGCTGCOOCTGCCCTGGCCACCOCTCGTGACCACCOCTGACCTAOCGGGTGCAGTGTCT
TCAGCOGCTAOCOCGACCACATGAAGCAGCAGACTTCTTCAAGTCCGCCATGCCOAGGCTAOCCTCAGGAGCGCAACC
ATCTTCTTCAAGGACGACGGCAACTACAAGACCCGCGCCGAGGTGAAGTTCGAGGGGACACCOCTGGTGAACCCGATCGA
GCTGAAGGGCATCGACTTCAAGGAGGACGGCAACATCCTGGGGCACAAGCTGGAGTACAACAGCCACAACGTCT
ATATCATGGCCGACAAAGCAGAAACCGGCATCAAGGTGAACCTCAAGATCCGCCACAACATCGAGGACGGCAGCGTGCAG
CTCGCCGACCACTACCAGCAGAACACCCCATCGCGCAGGGCCOCTGCTGCTGCCGACAACCACTAOCCTGAGCAOCCA
GTCCGCCCTGAGCAAAGACCCCAACGAGAAGCGCGATCACATGGTCTGCTGGAGTTCGTAACCCGCCGGGATCACTC
TCGGCATGGAAGCTGTACAAGTAACTCGAGAGATCTCTCGAGAGATCCCCCGGGGTGACTGATCAAATTCGAGCTCG
GTACCTTTAAGACCAATGACTTACA

>pdL_1_GZB-NFSA-2A-GFP

ACCGTTAATACGACTCACTATAGGCTAGCATTAGGTGACACTATAGAATACAAGCTACTTGTCTTTTTGCATTAATT
AAGCCGCCATGCAACCAATCCTGCTTCTGCTGGCCTTCTCTGCTGCCAGGGCAGATGCAGGGGAGATCATCGGGGA
CATGAGGCCAAGCCCACTCCCGCCCTACATGGCTTATCTTATGATCTGGGATCAGAAGTCTCTGAAGAGGTGCGGTGG
CTTCTGATACGAGACGACTTCGTGCTGACAGCTGCTCACTGTTGGGAAAGCTCCATAAATGTCAOCTTGGGGGCCACA
ATATCAAAGAACAGGAGCCGACCCAGCAGTTTATCCCTGTGAAAAGACCCATCCCCATCCAGCCTATAATCCTAAGAAC
TTCOCAACGACATCATGCTACTGCAGCTGGAGAGAAAGGCCAAGCGGACAGAGCTGTGCAGCCOCTCAGGCTACCTAG
CAACAAGGCCAGGTGAAGCCAGGGCAGACATGCAGTGTGGCCGGCTGGGGCAGACGCCOCCOCTGGGAAAACACTCAC
ACACACTACAAGAGGTGAAGATGACAGTGCAGGAAGATCGAAAAGTGCGAATCTGACTTACGCCATTATTACGACAGTACC
ATTGAGTTGTGCGTGGGGGACCCAGAGATTAAGAAAGACTTCCCTTAAGGGGGACTCTGGAGGCCCTCTGTGTGAACAA
GGTGGCCAGGGCATTGTCTCTATGGACGAAACAATGGCATGCCTCCACAGCCCTGCACAAAAGTCTCAAGCTTTGTAC
ACTGGATAAAGAAAACATGAACCGCTAOCGGCCCGCTGGAGGCGGGGTTCTGGCGGGGTGGATCAGGGGGTGGAGGC
TCCGGTGGAGGCGGGTCCGGCCGCCACGCCAACCATTGAACCTTATTTGTGGCCATCGCTCCATTGCCATTTCACTGA
TGAACCCATTTCCGAAGCGCAGCGTGGAGGCGATTATTAACAGCGCCCGTGGACGCTCCAGTTCAGTTTTTGTGAGTGA
GTAGCATTATTCGATTACCGACAAGCGTTACGTGAAGAACTGGTGAAGCTGACCGGGCGGCAAAAACAGTAGCGCAA
GCGGGCGAGTTCGGGTGTTCTGTGCCACTTTAACCGCCATTTACAGATCTGTCCGATGCTCAGCTCGGCCTGGCGGA
ACAACCTGTTGCTCGGTGCTGTTGATACGGCAATGATGGCGCAGAATGCATTAAATCGCAGCGGAATCGCTGGGATTGGCG
GGGTATATACGGCGGCTGCGCAATAATTTAAGCGGTGACGAACTGCTTAAATACCGCAGCATGTTCTGCCGCTG
TTTGGCTGTGCTTGGCTGGCTGCGGATAATCCGATCTTAAGCCGCTTTACCGCCTCCATTTTGGTGCATGAAAA
CAGCTATCAACCGCTGGATAAAGGGCGCACTGGCGCAGTATGACGAGCAACTGGCGGAATATTACCTCACCCGTGGCAGCA
ATAATCGCCGGATACCTGGAGCGATCATATCCGCGAACAATCATTAAAGAAAGCCGCCATTTATTCTGATTATTTG
CACAAAACAGGGTTGGCGACGCGGAATTCGGAAGCGGAGCTACTAATTCAGCCTGCTGAAGCAGGCTGGAGACGTGGA
GGAGAACCCCTGGACCTATGTTGAGCAAGGGCGAGGAGCTGTTACCGGGGTGGTGCOCATCCTGGTGCAGCTGGACGGCG
ACGTAACCGCCACAAGTTCAGCGTGTCCGGCAGGGGAGGGGATGCCACCTACGGCAAGCTGACCCCTGAAGTTCATC
TGACCACCGGCAAGCTGCOOCTGCCCTGGCCACCOCTCGTGACCACCOCTGACCTAOCGGGTGCAGTGTCTCAGCOGCTA
CCCCGACCACATGAAGCAGCAOACTTCTTCAAGTCCGCCATGCCOAGGCTACGTCCAGGAGCGCACCATCTTCTTCA
AGGACGACGGCAACTACAAGACCCGCGCGAGGTGAAGTTCGAGGGGACACCOCTGGTGAACCGCATCGAGCTGAAGGGC
ATCGACTTCAAGGAGGACGGCAACATCCTGGGGCACAAGCTGGAGTACAACACTACAACAGCCACAACGTCTATATCATGGC
CGACAAGCAGAAGAACGGCATCAAGGTGAACCTCAAGATCCGCCACAACATCGAGGACGGCAGCGTGCAGCTGCGCGAAC
ACTACCAGCAGAACCCOCCATCGCGCAGCGCCOCTGCTGCTGCCGACAACCACTACCTGAGCACCAGTCCGCCOCTG
AGCAAAGACCCCAACGAGAAGCGCGATCACATGGTCTGCTGGAGTTCGTAACCGCCGCCGGGATCACTCTCGGCATGGA
CGAGCTGTACAAGTAACTCGAGAGATCCCCCGGGGTGACTGATCAAATTCGAGCTCGGTACCACTAGTTAGCCCGCTGA

TCAGCCTCGA