

**Redirecting the cellular information flow
with programmable dCas9-based
chimeric receptors**



Toni A. Baeumler
St Hilda's College
University of Oxford

A thesis submitted for the degree of
Doctor of Philosophy
Hilary 2018

*“Potentially one of the most important concepts to arise in the history of mankind....
For the first time in all time a living creature understands its origin and can undertake
to design its future.”*

By Robert L. Sinsheimer
On human genetic modification
Published in American Scientist, 1969

*“But at times, I have been taken aback by just how intensely competitive
the study and use of CRISPR can be and how much it transformed
in a matter of years, becoming a global field that touches
virtually any researcher studying biology.”*

Jennifer A. Doudna & Samuel H. Sternberg
A Crack in Creation – Gene Editing And The Unthinkable Power
To Control Evolution, 2017

Redirecting the cellular information flow with programmable dCas9-based chimeric receptors

Toni A. Baeumler, DPhil in Medical Sciences, St Hilda's College,
University of Oxford, Hilary 2018

Abstract

Signal integration and transduction by cell-surface receptors is a complex, multi-layered process resulting in tight regulation of downstream mediators, which in turn elicit pre-defined native cellular responses. The modular architecture of transmembrane receptors provides a unique opportunity for engineering *de novo* sensor/effector circuits, enabling the development of custom cellular functions for research and therapeutic applications.

The signal transduction module of most existing chimeric receptors consists of either native intracellular domains or effectors domains fused to non-programmable DNA binding proteins. Therefore, these receptors can only engage in natural signalling pathways or drive the expression of artificial, pre-integrated transgenes. By harnessing the programmability of a nuclease deficient CRISPR/Cas9 (dCas9) signal transduction module and leveraging the evolutionarily optimised ligand-sensing capacity of native receptors, I have created a novel class of dCas9-based synthetic receptors (dCas9-synR).

I demonstrate that an optimised split dCas9-based core architecture and custom protease-based signal release mechanism can be standardised across multiple classes of extracellular domains to engineer receptor tyrosine kinase (RTK)-based and G-protein-coupled receptor (GPCR)-based chimeric receptors. dCas9-synRTK and dCas9-synGPCR integrate a broad variety of input signals (peptides, proteins, lipids, sugars) with highly specific and robust activation of any custom output transcriptional programme in an agonist dose-dependent manner. Finally, to showcase the therapeutic potential of dCas9-synRs, I used them to convert a pro-angiogenic signal into an anti-angiogenic response, deploy a chemokine/cytokine programme in response to tumour-enriched biomolecules, and induce insulin expression following glucose stimulation. The performance of dCas9-synRs and their unique versatility in redirecting the information flow makes them ideally suited to engineer designer cells capable of sensing specific disease markers and in turn drive various therapeutic programmes.

Declaration of Authorship

The thesis “Redirecting the cellular information flow with programmable dCas9-based chimeric receptors” I am submitting is entirely my own work except where otherwise indicated. It has not been submitted, either partially or in full, either for this Honour School or qualification or for another Honour School or qualification of this University, or for a qualification at any other institution. I have clearly indicated the presence of all material I have quoted from other sources, including any diagrams, charts, tables or graphs. I have clearly indicated the presence of all paraphrased material with appropriate references. I have acknowledged appropriately any assistance I have received in addition to that provided by my supervisor Professor Tudor A. Fulga. I have not copied from the work of any other candidate. I have not used the services of any agency providing specimen, model or ghostwritten work in the preparation of this thesis. All experiments were conducted in the Weatherall Institute of Molecular Medicine, University of Oxford, John Radcliffe Hospital, Headington, Oxford, OX3 9DS, UK.

This thesis is based on my work published in: Baeumler, T.A., Ahmed, A.A. & Fulga, T.A., 2017. Engineering synthetic signaling pathways with programmable dCas9-based chimeric receptors. *Cell Reports*, 20(11), pp.2639–2653. Most of the data, figures, tables, methods, and appendix of this thesis were published in this particular paper. The paper was published as an open access article under an Attribution 4.0 International (CC BY 4.0) license, permitting reuse of all content. It was written by myself and Tudor A. Fulga. Some illustrations/cartoons were adapted with kind permission from Quentin R.V. Ferry (Fulga Laboratory) and adapted from his publication: Ferry, Q.R.V., Lyutova, R. & Fulga, T.A., 2017. Rational design of inducible CRISPR guide RNAs for de novo assembly of transcriptional programs. *Nature Communications*, 8, p.14633.

David Knapp (Fulga lab) performed the analysis and generated the plots for Figure 4.8d-f and Figure 5.4e-f.

This thesis contains material that was presented at the conference “Genome Engineering: The CRISPR-Cas Revolution” from July 21-July 24, 2017 in Cold Spring Harbor, USA.

Lastly, a United Kingdom Patent (No. GB 1711470.3) was filed which contains data, figures, text, tables, methods, and appendix from the original paper and this thesis.

Toni A. Baeumler

Oxford, 7th March 2018

Acknowledgements

First and foremost, I would like to express my sincere gratitude to my DPhil supervisor Prof Tudor Fulga for accepting me as his graduate student. Throughout the last four years, he has been my mentor and showed a great deal of support, dedication, and resilience. He encouraged me to explore many uncharted waters in science under the motto “high risk, high reward” and guided me with scientific advice and knowledge. In the face of daring challenges, we both excelled as a team and pushed our work forward.

I would like to thank the Radcliffe Department of Medicine, the Medical Research Council, and St Hilda’s College for giving me the opportunity and the financial support to study at the University of Oxford.

I must express my appreciation to the members of my thesis advisory committee, Prof Jan Rehwinkel, Prof Persephone Borrow, and Dr Andrew Bassett.

I am beyond grateful to Dr Quentin Ferry, Dr Timothy Rajakumar and Dr Bruno Steinkraus for a great time in and outside of the lab, especially for teaming up for the OneStart Biotech Start-up competition 2014/15. I am very grateful to Dr Qianxin Wu, Dr Markus Toegel, Dr Aron Szabo, and Dr David Knapp for their continued support during the long journey on which the DPhil has taken me. I owe a debt of gratitude to Yale Michaels who has been a great source of advice for all areas of life and I hope I will be able to return the favour in the near future. I would also like to mention other members of the Fulga lab: Alice Lightowlers, Max Jamilly, Andrew Ramos, and Eunice Lee.

My thanks go to Prof Ahemd Ashour Ahemd and his lab, who have been very generous with help and support during my studies as well as Tatjana Sauka-Spengler’s lab and Thomas Milne’s lab for numerous helpful discussions about my research endeavours.

I would like to extend a special thanks to Dr Omer Habib and Prof Jin-Soo Kim from Seoul National University/Institute for Basic Science, South Korea for hosting me

as a visiting student for two months during my second DPhil year and for providing me with invaluable insights into the world of CRISPR.

Completing this work would have been all the more difficult if it were not for my very good friend Mohammad Karaminejadranjbar, who has always been there for me in good times but also in bad times. Without our infamous dinner in Goettingen, both of us would have never applied to Oxford, and the rest is history.

Last but not least, I am indebted to the unconditional support of my oldest and best friends and my family.

Publications

Toni A. Baeumler, Ahmed Ashour Ahmed and Tudor A. Fulga. Engineering synthetic signaling pathways with programmable dCas9-based chimeric receptors. **Cell Reports** 2017 Sep 12;20(11):2639-2653. doi: 10.1016/j.celrep.2017.08.044.

This work is covered and protected by a United Kingdom Patent (No. GB 1711470.3).

This study introduces the concept of dCas9-synRs, a nuclease deficient Cas9-based chimeric receptor scaffold that utilises endogenous receptor recognition domains to detect cognate ligands and trigger a user-defined transcriptional programme in response.

Qianxin Wu[#], Quentin RV. Ferry[#], **Toni A. Baeumler**, Yale S. Michaels, Dimitrios M. Vitsios, Omer Habib, Roland Arnold, Xiaowei Jiang, Stefano Maio, Bruno R. Steinkraus, Marta Tapia, Paolo Piazza, Ni Xu, Georg A. Holländer, Thomas A. Milne, Jin-Soo Kim, Anton J. Enright, Andrew R. Bassett and Tudor A. Fulga. *In situ* functional dissection of RNA *cis*-regulatory elements by multiplex CRISPR-Cas9 genome engineering. **Nature Communications** 2017 Dec 13;8(1):2109. doi 10.1038/s41467-017-00686-2

[#]These authors contributed equally to this work.

In this study, Dr Qianxin Wu and I utilised a CRISPR/Cas9-based multiplex screening approach to unbiasedly investigate the regulatory landscape encoded within the 3' untranslated region of mRNAs.

List of Abbreviations and Acronyms

In addition to standard abbreviations for metric units (e.g., ml), chemical symbols and formulas (e.g., HCl), amino acids (e.g., three letter code, Ser; or single letter code, S) and nucleic acid organic bases (e.g., G); the following abbreviations and acronyms listed below are used throughout this thesis.

3'LTR(Δ U3)	3'LTR without U3 region
4OHT	4-hydroxytamoxifen
5HT	serotonin/5-hydroxytryptamine
Ab	antibody
ABA	abscisic acid
AbCIDs	human antibody-based chemically induced dimerisers
Acr	phage-encoded anti-CRISPR proteins
AIDS	acquired immunodeficiency syndrome
APC	antigen presenting cells
APOBEC3	apolipoprotein B-editing catalytic polypeptide 3
ARRB2	β -arrestin2
ART	antiretroviral therapy
ASCL1	Achaete-scute homolog 1 gene
AsCpf1	<i>Acidaminococcus</i> sp. Cpf1
ATP	adenosine triphosphate
a.u.	arbitrary units
AVPR2	arginine vasopressin receptor 2
AZT	3-azido-3-deoxythymidine
BDKRB2	bradykinin receptor B2
bGH pA	bovine growth hormone polyadenylation signal
bp	base pairs
BSA	bovine serum albumin
CAR	chimeric antigen receptor
CAF	cancer-associated fibroblast
Cas	CRISPR-associated nuclease
Cascade	CRISPR-associated complex for antiviral defence
CBh	chicken β -actin hybrid promoter
CCR5	CC-chemokine receptor 5
CD3 ζ	CD3 zeta chain
CDS	coding DNA sequence
ChIP-seq	chromatin immunoprecipitation followed by deep DNA sequencing
CID	chemical inducer of dimerisation
CNO	clozapine-N-oxide
CPE	carboxypeptidase E
Cpf1	nuclease of type V CRISPR system
cPPT	central polypurine tract
CRISPR	clustered regularly interspaced short palindromic repeat
CRISPRa	CRISPR activation

CRISPRi	CRISPR interference
crRNA	CRISPR RNA
CRS	cytokine storm/cytokine release syndrome
Ct	cycle threshold
CTS	central termination sequence
CXCR4	C-X-C chemokine receptor type 4
DAPI	4',6-diamidino-2-phenylindole
DBD	DNA-binding domain
dCas9	CRISPR-associated, nuclease-deficient protein 9
dCas9(C)	C-terminal protein part of split dCas9
dCas9(N)	N-terminal protein part of split dCas9
DD	destabilisation domain
dEGFP	destabilised enhanced green fluorescent protein
DHFR	dihydrofolate reductase
DMEM	Dulbecco's modified Eagle's medium
DNMT3A	DNA methyl-transferase domain
Dox	doxycycline
DREADD	designer receptor exclusively activated by designer drug
EC ₅₀	half-maximal effective concentration
ECM	extracellular matrix
EDTA	ethylenediaminetetraacetic acid
EGFP	enhanced green fluorescent protein
EGFR	epidermal growth factor receptor
ER	oestrogen receptor
ER50	ER destabilisation domain
ERK	extracellular signal-regulated kinase
ERT2	hormone-binding domain of the oestrogen receptor
EYFP	enhanced yellow fluorescent protein
FACS	fluorescence-activated cell sorting
FADD	Fas-associated protein with death domain
FBS	foetal bovine serum
FDA	American Food and Drug Agency
FGFR1	fibroblast growth factor receptor 1
FKBP(12)	FK506-binding protein (12)
FRB	FKBP rapamycin binding
FSC	forward scatter
GA	gibberellin
Gab1	Grb2-associated binder protein 1
GABA	gamma-aminobutyric acid
GEF	guanine nucleotide exchange factor
GFP	green fluorescent protein
GLP1	glucagon-like peptide 1
gp120	envelope glycoprotein 120
GPCR	G-protein-coupled receptor
Grb2	growth factor receptor-bound protein 2
GRK	G-protein receptor kinase
HA	human influenza hemagglutinin
HBG1	haemoglobin subunit gamma-1

HEK293(T)	human embryonic kidney cells 293 (large antigen T from SV40)
hESC	human embryonic stem cell
HGF	hepatocyte growth factor
HIV	human immunodeficiency virus
HIV-1 Ψ	human immunodeficiency virus 1 packaging signal
hM3/4D	human M3/4 muscarinic receptor DREADDs
hPSC	human pluripotent stem cell
hRO-i	human RASSL opioid G κ -coupled receptor
hRS-s	human RASSL based on serotonin 4 receptor
HSF1	human heat-shock factor 1
HSP90	heat shock protein 90
HSV-TK	herpes simplex virus thymidine kinase
HTLA	HEK293-derived cell line with β -arrestin2-TEV chimeric fusion gene
huEGFRt	truncated version of human EGFR polypeptide
iCasp9	inducible caspase 9
ICOS	inducible T cell co-stimulator
IFN γ	interferon gamma
Ig	immunoglobulin
IL1B	interleukin 1 beta
IL-2	interleukin 2
INS	insulin
iPSC	induced pluripotent stem cells
iSBH	inducible spacer-blocking hairpin
ITAM	immunoreceptor tyrosine-based activation
iTANGO2	inducible TANGO construct design 2
KORD	κ -opioid receptor DREADD
KRAB	Krüppel-associated box
LAT	linker for activation of T cell family member 1
LBD	ligand binding domain
lncRNA	long non-coding ribonucleic acid
LOV	light-oxygen-voltage
LPA	lysophosphatidic acid
LPAR	lysophosphatidic acid receptor
LSD1	lysine-specific histone demethylase 1A
LTR	long terminal repeat
MAPK	mitogen-activated protein kinase
MAPKKK	MAPK kinase kinase
MCP	MS2 bacteriophage coat protein
MESA	Modular extracellular sensor architecture
MET	see HGF
MHC	major histocompatibility complex
MIP-1 α	macrophage Inflammatory Protein 1 alpha
MSC	mesenchymal stem cell
mTOR	mammalian target of rapamycin
NES	nuclear export signal
NFAT	nuclear factor of activated T cells
NGF	nerve growth factor

NICD	Notch intracellular signalling domain
NLS	nuclear localisation signal
NSC	neural stem cell
nt	nucleotides
Opto-RTKs	optically controllable RTKs
PAM	protospacer adjacent motif
PBS	phosphate buffer saline
PC	protease chain
PC1/2/3	prohormone convertases
PCR	polymerase chain reaction
PD1	programmed cell death protein 1
PDGF	platelet-derived growth factor
PEI	polyethylenimine
P _{FWD}	forward primer
PI3K	phosphatidyl inositol 3-kinase
PKA/PKC	protein kinase A/C
PLC	phospholipase C
PLGF	placenta growth factor
P _{REV}	reverse primer
PTB	phosphotyrosine-binding domain
QF	transcription factor from <i>Neurospora crassa qa</i>
RASSL	receptor activated solely by a synthetic ligand
RET	rearranged during transfection receptor
Rta	Epstein-Barr virus-derived R transactivator
RTK	receptor tyrosine kinase
RT-qPCR	reverse transcription quantitative PCR
rtTA	reverse tetracycline-controlled transactivator
RVD	repeat variable diresidue
S1P	sphingosine-1-phosphate
SaCas9	<i>Streptococcus af</i> Cas9
SAM	synergistic activation mediator
SAMHD1	sterile alpha motif and histidine-aspartate domain-containing protein 1
scRNA	sgRNA scaffold
SFFV	spleen focus-forming virus
sgRNA	single guide RNA
SCAD	single-chain antibody domain
scFv	single chain variable fragment
SCR	scramble
SDS	sodium dodecyl sulfate
SDF-1 α	stromal cell-derived factor 1 alpha
SH2	Src homology-2 domain
SID4X	4 concatenated SID domains
SNTGV	chimeric Notch receptor with SCAD, Notch TMD, Gal4-VP16
SpCas9	<i>Streptococcus pyogenes</i> Cas9
SSC	side scatter
SV40	simian virus 40
synNotch	synthetic Notch receptor
synNQ	synthetic Notch receptor with QF

T1D	type 1 diabetes
T1Rs	taste receptors
T2D	type 2 diabetes
T7EI	T7 Endonuclease I
TALE	transcription activator-like effector
TANGO	transient interactions between two engineered protein components
TBS	Tris-buffered saline
TC	target chain
TCR	T cell receptor
TCS	TEV cleavage site
TDAG8	T-cell death-associated gene 8
TetR	tet repressor
TET1CD	catalytic domain of ten-eleven 1 hydroxylase
TEV	Nla tobacco etch virus
TEV(C)	C-terminal protein part of split TEV protease
TEV(N)	N-terminal protein part of split TEV protease
TF	transcription factor
TLS	tumour lysis syndrome
TM/TMD	transmembrane domain
TMP	trimethoprim
Tmpk	thymidylate kinase
TMt	transmembrane tether
TNF α	tumour necrosis factor alpha
TSP-1	thrombospondin 1
TR	transcriptional regulator domain
tracrRNA	trans-activating crRNA
Tris	tris(hydroxymethyl)aminomethane
TrkA	NGF/neurotrophin receptor
TRPV1	temperature-sensitive calcium channel V1
TSS	transcription start site
tTA	tetracycline-controlled transactivator
UAS	upstream activator sequence
V ₂ tail	AVPR2-originating C-terminal sequence
VEGF	vascular epidermal growth factor
VEGFR	vascular epidermal growth factor receptor
VP16	Herpes simplex virus protein vmw65
VPR	VP64, p65, Rta
ZAP70	ζ -associated protein of 70 kDa
ZFP	zinc finger protein

Table of Contents

Abstract.....	III
Declaration of Authorship.....	IV
Acknowledgments.....	VI
Publications.....	VIII
List of Abbreviations and Acronyms.....	IX
Table of Contents.....	XIV
List of Figures.....	XIX
List of Tables.....	XXII

Chapter 1 : Introduction.....	1
<i>1.1 Engineering artificial transcription factors.....</i>	<i>2</i>
1.1.1 Engineered non-programmable transcriptional regulators	2
1.1.2 Programmable, protein-based transcription factors	6
1.1.2.1 Zinc finger proteins (ZFPs).....	7
1.1.2.2 Transcription activator-like effectors (TALEs)	9
1.1.3 Programmable, RNA-guided dCas9-TF	10
1.1.3.1 Mechanism of clustered regularly interspaced short palindromic repeat (CRISPR)-associated nuclease 9 (Cas9).....	10
1.1.3.2 Harnessing CRISPR/dCas9 for inducible control of gene expression	12
1.1.3.3 CRISPR/dCas9 for transcriptional activation	14
1.1.3.4 CRISPR/dCas9 for transcriptional repression.....	15
1.1.3.5 Inducible dCas9-TF systems.....	17
<i>1.2 Signalling networks and receptor engineering</i>	<i>20</i>
1.2.1 Endogenous mammalian cell-surface receptors	22
1.2.1.1 G-protein-coupled receptors (GPCRs).....	22
1.2.1.2 Receptor tyrosine kinases (RTKs)	29
1.2.2 Prosthetic gene networks	33

1.2.3 Synthetic receptors	36
1.2.3.1 Chimeric antigen receptors (CARs)	37
1.2.3.2 Synthetic Notch receptor (synNotch)	43
1.2.3.3 Engineered G-protein-coupled receptors	47
1.2.3.3.1 Receptors Activated Solely by a Synthetic Ligand (RASSLs) and Designer Receptor Exclusively Activated by Designer Drugs (DREADDs).....	47
1.2.3.3.2 Non-programmable TF-based GPCRs (GPCR TANGO).....	51
1.2.3.4 Engineered receptor tyrosine kinases.....	53
1.2.3.5 Modular extracellular sensor architecture (MESA).....	55
1.3 Thesis Aims - Engineering a CRISPR/dCas9-based synthetic receptor scaffold.....	59
Chapter 2 : Materials and Methods	61
2.1 Reagents and kits.....	61
2.2 Plasmid constructs	62
2.2.1 EYFP reporter assay constructs	63
2.2.2 dCas9 ^{VP64} , Cas9, and associated constructs	63
2.2.3 TMt-dCas9 ^{VP64} , TMt-Cas9, and associated constructs	64
2.2.4 Split mono and dual TMt-dCas9 ^{VP64} and associated constructs	67
2.2.5 dCas9-synGPCR and associated constructs	70
2.2.6 dCas9-synRTK, MESA, and associated constructs	75
2.3 sgRNA sequences.....	77
2.4 Cell lines and culture conditions.....	78
2.5 Transient transfection protocols	78
2.6 Flow cytometry experiments and data analysis.....	81
2.7 Direct flow cytometry	82
2.8 RT-qPCR analysis.....	83
2.9 T7 endonuclease I (T7E1) assay	85
2.10 Confocal microscopy	86
2.11 Viral particle production and cell line generation.....	86

Chapter 3 : Establishing split transmembrane-tethered dCas9^{VP64} as signal transduction module.....	88
3.1 Chapter Aims.....	88
3.2 Results	88
3.2.1 Impact of localisation tags on the transcriptional potency of soluble dCas9 ^{VP64}	88
3.2.2 Engineering a conditional transmembrane-tethered dCas9 ^{VP64} (Tmt-dCas9 ^{VP64}) scaffold.....	96
3.2.3 A split Tmt-dCas9 ^{VP64} scaffold design for stringent OFF/ON state transitions	116
3.2.3.1 Split mono and dual Tmt-dCas9 ^{VP64}	122
3.2.3.2 Inducible split mono and dual Tmt-dCas9 ^{VP64}	125
3.3 Discussion	133
Chapter 4 : Engineering programmable chimeric dCas9-synGPCR scaffolds.....	135
4.1 Chapter Aims.....	135
4.2 Results	135
4.2.1 Chimeric, full length dCas9-based GPCRs are inadequate for attaining robust OFF/ON state transitions	135
4.2.2 Exploring design principles for synthetic split dCas9-based GPCRs (dCas9-synGPCRs)	139
4.2.3 Engineering an optimised split dual dCas9-synGPCR architecture with stringent OFF/ON state transition characteristics	147
4.2.3.1 Establishing and characterising a prototype split dual dCas9-synBDKRB2 receptor	147
4.2.3.2 Expanding the repertoire of ligand-activated dCas9-synGPCRs	153
4.2.4 Boolean AND logic gates for multi-factorial input processing	158
4.3 Discussion	164
4.3.1 Improving the scaffold of dCas9-synGPCRs.....	165
4.3.2 Considerations for the endogenous signalling pathways of GPCRs and their ligand specificities	166
4.3.3 Prospective Boolean AND logic gates with dCas9-synGPCRs.....	168

4.3.4 Alternative CRISPR-TF-based synthetic GPCRs.....	169
Chapter 5 : Engineering programmable chimeric dCas9-synVEGFRs	174
5.1 Chapter Aims.....	174
5.2 Results	174
5.2.1 Design and optimisation of dCas9-synVEGFR with stringent OFF/ON state transition characteristics	174
5.2.2 Boolean AND logic gates as inducible control switches.....	180
5.2.3 dCas9(N/C)-synVEGFR1/2 is superior to V-MESA-dCas9-TF	184
5.3 Discussion	186
5.3.1 Agonist repertoire for dCas9(N/C)-synVEGFR1/2	187
5.3.2 Alternative dCas9-TF-based synthetic receptor scaffold for VEGFA detection ...	188
Chapter 6 : Implementation of therapeutically-relevant transcriptional programmes with dCas9-synRs.....	190
6.1 Chapter Aims.....	190
6.2 Results	190
6.2.1 Angiogenic signal conversion mediated by dCas9(N/C)-synVEGFR1/2	191
6.2.2 Eliciting custom cytokine/chemokine programmes with dCas9-synGPCRs.....	193
6.2.3 Glucose-induced insulin expression in designer cells with dCas9(N/C)-synT1R3	194
6.3 Discussion	198
6.3.1 Converting various input signals into custom cytokine/chemokine output programmes	199
6.3.2 Engineering β -mimetic cells with dCas9-synT1R3	200
6.3.3 Choice of cell types for dCas9-synRs	205
Chapter 7 : General discussion and future perspectives	207
7.1 Designing the next generation of dCas9-synRs	208
7.1.1 Duration and reversibility of dCas9-synR-mediated signalling	209
7.1.2 Portability and refinement of dCas9-synGPCRs' mode of activation	212
7.1.3 A universal dCas9-synRTK for growth factor signal integration	216

7.1.4 Adapting the split dCas9 core architecture to engineer a novel class of synthetic Notch receptors (dCas9-synNotch).....	217
7.2 <i>Precautions and future applications of dCas9-synRs</i>	218
7.2.1 Immunological challenges of utilising dCas9-TFs	218
7.2.2 Multiplexing synthetic receptors for precision multi-input sensing, control, and response programmes	220
7.2.3 dCas9-synGPCRs as novel modality to fight HIV-1 infection.....	222
Chapter 8 : Bibliography	225
Appendix	A
<i>A: Amino acid sequence of constructs</i>	<i>A</i>
A:1 dCas9 ^{VP64} , Cas9, and associated constructs.....	A
A:2 TMt-dCas9 ^{VP64} , TMt-Cas9, and associated constructs	B
A:3 Split TMt-dCas9 ^{VP64} and associated constructs.....	G
A:4 dCas9-synGPCR and associated constructs.....	L
A:5 dCas9-synRTK, MESA, and associated constructs.....	S
<i>B: Transfection schematics for all experiments</i>	<i>V</i>

List of Figures

Figure 1.1 Programmable artificial transcription factors (TFs).....	8
Figure 1.2 dCas9-TF for transcriptional activation and repression.....	13
Figure 1.3 Inducible dCas9-TF for controlled endogenous gene regulation.....	18
Figure 1.4 Endogenous cell signalling.....	21
Figure 1.5 G-protein-coupled receptor signalling.	24
Figure 1.6 Classical arrestin-mediated signalling by GPCRs.....	27
Figure 1.7 A representative of RTKs: vascular endothelial growth factor receptors (VEGFRs).....	30
Figure 1.8 Chimeric antigen receptor (CAR) design inspired by native T cell receptor (TCR) architecture.....	38
Figure 1.9 Synthetic Notch receptor (synNotch) design inspired by native Notch receptor architecture.....	44
Figure 1.10 Engineering GPCR signalling capabilities.....	48
Figure 1.11 Modular extracellular sensor architecture (MESA).....	56
Figure 1.12 Conceptual overview of dCas9-based synthetic receptors (dCas9-synRs). 60	
Figure 3.1 Subcellular localisation of soluble dCas9 ^{VP64} variants.....	90
Figure 3.2 EYFP reporter assay and quantification.....	92
Figure 3.3 Transcriptional activation potential of differently, subcellular-localised soluble dCas9 ^{VP64}	94
Figure 3.4 Conceptual overview of transmembrane-tethered dCas9 ^{VP64} (TMt-dCas9 ^{VP64}).	97
Figure 3.5 Subcellular localisation of transmembrane-tethered dCas9 ^{VP64}	99
Figure 3.6 Control antibody staining for anti-HA and anti-myc.	101
Figure 3.7 Transcriptional activation potential of conditionally transmembrane-tethered dCas9 ^{VP64} (TMt-dCas9 ^{VP64}).	102

Figure 3.8 TEV protease specifically releases transmembrane-bound NLS-dCas9 ^{VP64} from the membrane.	104
Figure 3.9 TMt-NLS-dCas9 ^{VP64} transcriptional activation is independent of TEV protease.	105
Figure 3.10 TMt ^{NES} -dCas9 ^{VP64} transcriptional OFF/ON transition is dependent on TEV protease.	107
Figure 3.11 TEV independent reporter activation of lowly expressed TMt-NLS-dCas9 ^{VP64[DOX]}	108
Figure 3.12 Subcellular localisation of transmembrane-tethered Cas9-mCherry variants.	110
Figure 3.13 Establishing HEK293T_dEGFP for a genomic dEGFP disruption assay...	112
Figure 3.14 TMt-Cas9-mCherry variants diminish dEGFP fluorescence in HEK293T_dEGFP cells.	113
Figure 3.15 TMt-Cas9-mCherry variants disrupt dEGFP in the genome of HEK293T_dEGFP cells independent of TEV protease-mediated liberation.	115
Figure 3.16 Genomic dEGFP disruption is mediated by sgEGFP-guided TMt-Cas9-mCherry variants.	117
Figure 3.17 Screen to identify split dCas9 ^{VP64} configurations with improved OFF/ON state transition characteristics.	119
Figure 3.18 Rapamycin abrogates OFF/ON state transition in split dCas9 FKBP/FRB pair configurations.	121
Figure 3.19 Conceptual overview of split mono and dual TMt-dCas9 ^{VP64}	123
Figure 3.20 dCas9(C) ^{VP64} reassembles with liberated dCas9(N) fragment. (a).....	124
Figure 3.21 Inducible transcriptional activation with soluble split TEV for split mono and dual TMt-dCas9 ^{VP64} (TMt-dCas9-iTEV ^{soluble}).	126
Figure 3.22 Inducible transcriptional activation of an endogenous gene with split dual TMt-dCas9-iTEV ^{soluble}	128

Figure 3.23 Concept of an engrafted split TEV in split mono and dual TMt-dCas9 ^{VP64} (TMt-dCas9-iTEV ^{engrafted}).	130
Figure 3.24 Inducible transcriptional activation with split mono and dual TMt-dCas9-iTEV ^{engrafted} .	131
Figure 3.25 Inducible transcriptional activation of an endogenous gene with split dual TMt-dCas9-iTEV ^{engrafted} .	132
Figure 4.1 Concept of GPCR-coupled dCas9 ^{VP64} as novel modality for the rewiring of signal pathways.	137
Figure 4.2 BDKRB2-dCas9 ^{VP64} recapitulates poor OFF/ON state transition behaviour of TMt-dCas9 ^{VP64} .	138
Figure 4.3 Scaffold optimisation for split mono dCas9(N)-synBDKRB2.	141
Figure 4.4 Extracellular domain optimisation for split mono dCas9(N)-synBDKRB2.	143
Figure 4.5 Implementation of the split mono dCas9-synGPCR designs for AVPR2.	145
Figure 4.6 Endogenous gene upregulation with split mono dCas9-synBDKRB2 and split mono dCas9-synAVPR2.	146
Figure 4.7 Design and functional validation of a non-optimised split dual dCas9-synBDKRB2 architecture.	148
Figure 4.8 Optimisation of split dual dCas9-synBDKRB2 designs.	150
Figure 4.9 Characterisation of dCas9(N/C)-synBDKRB2.	152
Figure 4.10 Vasopressin and SDF-1 responsive dCas9-synGPCRs.	154
Figure 4.11 Lipid responsive dCas9(N/C)-synLPARs.	156
Figure 4.12 D-glucose responsive dCas9(N/C)-synT1Rs.	157
Figure 4.13 2-input Boolean AND logic gates with heterologous dCas9-synGPCRs.	159
Figure 4.14 2-input Boolean AND logic gates with dCas9(N/C)-synBDKRB2 and an inert drug.	161
Figure 4.15 3-input Boolean AND logic gates with two dCas9-synGPCRs and one inert drug.	163
Figure 4.16 Alternative CRISPR-TF-based GPCR architectures.	170

Figure 5.1 Identification of optimal VEGFR combinations for dCas9-synVEGFR.	176
Figure 5.2 Identification of an optimal TCS combination in dCas9(N/C)-VEGFR1/2.....	178
Figure 5.3 dCas9(N/C)-VEGFR1/2 is expressed on the cell-surface.	179
Figure 5.4 Functional characterisation of dCas9(N/C)-VEGFR1/2.....	181
Figure 5.5 2-input Boolean AND logic gate strategies for dCas9(N/C)-synVEGFR1/2 with an inert drug.	183
Figure 5.6 Comparing dCas9(N/C)-synVEGFR1/2 with MESA-dCas9-TF receptors....	185
Figure 6.1 Angiogenic signal conversion mediated by dCas9(N/C)-synVEGFR1/2.	192
Figure 6.2 Cytokine signal conversion mediated by dCas9(N/C)-synCXCR4.	195
Figure 6.3 Tumour biomarker evoked multi-factorial cytokine/chemokine response programme mediated by dCas9(N/C)-synLPAR1.	196
Figure 6.4 Glucose-induced insulin expression in β -mimetic cells mediated by dCas9(N/C)-synT1R3.	197
Figure 7.1 Next-generation dCas9-synR instalments.....	215

List of Tables

Table 1 List of reagents, including chemicals, enzymes, and standard molecular biology kits.	62
Table 2 List of spacer sequences for targeting endogenous genes with dCas9-TF.....	77
Table 3 List of forward and reverse primers used for qPCR.	85

Chapter 1 : Introduction

The purpose of the field of synthetic biology is perhaps best described by the late theoretical physicist Richard P. Feynman (1918-1988) who is believed to have written on a blackboard "What I cannot create, I do not understand." (Caltech archives, <http://archives.caltech.edu/pictures/1.10-29.jpg>). This statement does not necessarily imply that if we managed to create something, we would understand it. On the contrary, understanding is necessary to create something but not sufficient. One could also say that biology, in principle, is about understanding life, while synthetic biology aims to engineer life by assembling its basic components (Elowitz & W. A. Lim 2010; W. A. Lim 2010).

As a general blueprint, synthetic biologists have exploited existing elements (DNA, RNA, protein) from all the three domain systems (bacteria, archaea, and eukaryota) (Woese et al. 1990) to create artificial signalling pathways and gene networks in bacteria, yeast, and more recently in more complex mammalian cells (Weber & Fussenegger 2009; Kiel et al. 2010; Weber & Fussenegger 2010; T. K. Lu et al. 2009; Wieland & Fussenegger 2012; S. Ausländer & Fussenegger 2013; Lienert et al. 2014; Schukur & Fussenegger 2016; Roybal & W. A. Lim 2017; Kitada et al. 2018). The conceptualisation of synthetic mammalian gene circuits has also become an attractive prospect for novel therapeutic applications in modern days (T. K. Lu et al. 2009; Schukur & Fussenegger 2016; Weber & Fussenegger 2009). Important milestones in therapeutic mammalian synthetic biology were the design of chimeric antigen receptors (CARs) to fight cancer (Fesnak et al. 2016; W. A. Lim & June 2017; Tran et al. 2017) and prosthetic gene networks to treat metabolic and autoimmune diseases (S. Ausländer & Fussenegger 2013; Teixeira & Fussenegger 2017). There are many more emerging technological advances with the potential to catalyse the development of new generations of therapeutic designer cells.

In the first part of this introduction, I discuss non-programmable and programmable artificial transcription factors that have been used to regulate the expression of transgenes and endogenous genes, respectively. In the second half, I explore prosthetic gene networks and current synthetic receptor designs to redirect the information flow in cells. By highlighting inventive steps and discussing shortcomings and drawbacks of such engineered receptors, I hope to evoke an appreciation for the laborious, iterative processes in receptor engineering and to demonstrate the need for novel approaches in chimeric receptor development.

1.1 Engineering artificial transcription factors

1.1.1 Engineered non-programmable transcriptional regulators

Transcription factors (TFs) are components of the cell that regulate gene expression at the level of DNA and thereby direct diverse cellular processes (e.g., differentiation, proliferation) (Rosenfeld et al. 2005). They actively recruit co-activators to initiate (ON switch) or co-repressors to prevent (OFF switch) transcription from a target gene (Ptashne 1986).

DNA-binding domains

Generally, a TF is a two-part entity consisting of a DNA-binding domain (DBD) and a transcriptional regulator domain (TR) that can either activate or repress the expression of a gene. One of the earliest characterised TF is Gal4 from *Saccharomyces cerevisiae*. By binding to its specific 17bp upstream activator sequence (UAS) as a homodimer, it induces gene transcription (Guarente et al. 1982; West et al. 1984; Johnston & R. W. Davis 1984; Giniger et al. 1985). This intrinsic property translates perfectly to mammalian cells. In 1988, it was shown that Gal4 could recognise and bind to its cognate UAS sequence, engrafted as a synthetic operator element upstream of a

minimal promoter, to robustly initiate the expression of a transgene (Kakidani & Ptashne 1988; N. Webster et al. 1988).

The GAL4 DBD can be readily separated from the rest of the protein, rendering it into a sole DBD (amino acids 1-147) (Brent & Ptashne 1985; Keegan et al. 1986; J. Ma & Ptashne 1987) with only minimal remaining transcriptional activity (Kakidani & Ptashne 1988; Shi et al. 1991). The splitting of functional domains highlights two important features, one being the general modular structure of a TF and the other one being an opportunity to exchange endogenous domains with heterologous domains thereby conferring new properties.

Transcriptional regulator domains

Potential chimeric transcription factors could be equipped with novel activation domains. VP16 is a 65 kD polypeptide that functions as a transactivator for the herpes simplex virus-1 genes. VP16 activates gene expression by interacting either directly or indirectly (via co-activators and mediators) with the core transcriptional machinery of mammalian cells, eventually leading to the recruitment of RNA polymerase II (Hampsey 1998; Mittler et al. 2003; Vojnic et al. 2011; F.J. Yang et al. 2004). The minimal activation domain of VP16 is described by the eleven amino acid sequence DALDDFDLML (Triezenberg et al. 1988; Seipel et al. 1992). Since the VP16 core sequence acts synergistically when recruited to the same synthetic operator (Carey et al. 1990; Emami & Carey 1992; Y.-S. Lin et al. 1990), a duplication (Seipel et al. 1992) and eventually a quadruplication (Emami & Carey 1992; Beerli et al. 1998), termed VP64, was conceived for more potent transcriptional gene activation. Other examples include p65, which is a subunit of the C-terminal part of the human NF- κ B transcription factor (residues 288–548) (Ruben et al. 1991; P. A. Moore et al. 1993; Schmitz & Baeuerle 1991) and the C-terminal activation domain of human heat-shock factor 1 (HSF-1) (Green et al. 1995).

In contrast, there are also transcriptional repressor domains, such as the Krüppel-associated box (KRAB) domain, in particular, the residues 1-75 of KOX1 (Margolin et al. 1994).

Small molecule-inducible chimeric transcription factors

Most of the earliest chimeric TF were derived from operons, which are rather complex prokaryotic regulatory elements (Yarranton 1992; Gossen et al. 1993), such as the tetracycline-controlled transactivator (tTA) (Gossen & Bujard 1992). tTA consists of the tet repressor (TetR) that is fused to VP16 (Gossen & Bujard 1992). TetR originates from the tetracycline-resistance operon Tn10 in *E.coli*. To harness the ability of tTA to initiate transgene expression, its cognate operator repeat sequences were incorporated upstream of a minimal promoter sequence (Gossen & Bujard 1992). One important property of tTA is its conditional DNA binding activity. In the presence of the small molecule tetracycline, tTA cannot bind its target sequence, preventing transcriptional activation (Gossen & Bujard 1992). A subsequent mutagenesis screen identified a reverse Tet repressor (rTetR) that requires tetracycline for cognate operon sequence binding (Hecht et al. 1993), leading to the development of a reverse tetracycline-controlled transactivator (rtTA) (Gossen et al. 1995). This example describes one of the first instances of a controllable artificial gene expression system and an OFF/ON switch relying on the presence or absence of an externally-delivered, inert stimulus (Allgood & Eastman 1997).

The creation of prokaryotic synthetic gene-switch TFs illustrate several important lessons: (i) synthetic TFs can be composed of several distinct domains that are interchangeable; (ii) the property of the DBD is maintained in the milieu of mammalian cells; (iii) cognate operator repeat sequences can be cloned upstream of minimal promoter sequences to initiate transgene expression (Yarranton 1992; Gossen et al. 1993).

Prokaryotic-derived chimeric OFF/ON switch TFs led to the development of various ligand-controllable gene circuits in mammalian cells. Examples include switches responsive to metabolites (e.g., phloretin (Gitzinger et al. 2009)), L-arginine (Hartenbach et al. 2007), and licenced food additives (e.g., benzoate and vanillate (Gitzinger et al. 2012; M. Xie et al. 2014)).

Primitive chimeric intracellular receptors

The successful adaptation of ligand-controllable gene-switch TFs sparked an exploration into other regulatory units that can be externally controlled by co-delivered inducer or repressor molecules. This search led to the creation of chimeric intracellular receptors that are composed of three domains: DBD, TR and a separate ligand binding domain (LBD) (Allgood & Eastman 1997).

This novel class of chemically inducible gene switches can be divided into two categories in which the LBD is based on either dimerisation subunits or nuclear receptors. Chemical inducer of dimerisation (CID)-regulated gene switches rely on homo-/heterodimerisation domains. One half is fused to a DBD, the other to the TR. Only in the presence of their cognate small molecule, these two domains undergo homo-/heterodimerisation, thereby bringing DBD and TA in proximity and reconstituting a functional TF. Inducers for CID-based chimeric intracellular receptors include rapamycin (Rivera et al. 1996; Belshaw et al. 1996; Ho et al. 1996; Pollock & Clackson 2002; Liberles et al. 1997), plant hormone abscisic acid (ABA) (Liang et al. 2011) and blue light (Xue Wang et al. 2012; Kennedy et al. 2010).

Nuclear receptor-based chimeric intracellular receptors contain the LBD of steroid receptors. The oestrogen receptor (ER)-based chimeric intracellular receptors contain the ER domain that sequesters the TF in the cytoplasm. In the presence of oestrogen or its derivatives, the complex translocates into the nucleus where it exerts

its transcriptional function (N. J. Webster et al. 1988; N. Webster et al. 1988; Roemer & Friedmann 1993; Whelan & N. Miller 1996; Superti-Furga et al. 1991; Braselmann et al. 1993).

Progesterone receptor-based chimeric intracellular receptors can be activated by progesterone, or after LBD mutation by the antagonist antiprogestins (Vegeto et al. 1992; Delort & Capecchi 1996; Y Wang et al. 1994). Improved OFF/ON state transitions of such chimeric TFs were achieved by reflecting on the arrangement of the different domains within the TF and on different linker length between the domains, highlighting the importance of construct optimisation studies (Y Wang et al. 1997).

A plethora of artificial transcription factors has been developed enabling the manipulation of mammalian gene expression at the transcription level. However, using CIDs or intracellular LBDs limit the inducibility to cell-permeable input signals (e.g., diffusible small molecules, light). Rendering cells responsive to extracellular cues, however, can be achieved by providing a link between the extracellular and the intracellular space, for example via the LBD of cell-surface receptors (Chapter 1.2).

All presented DBDs are inherently restricted to their cognate DNA binding sites making them non-programmable. As a consequence, only target genes that carry pre-defined sequences in an artificial operator region on a plasmid or as a transgene in the genome can be targeted. To gain control of endogenous gene expression with artificial TFs requires the deployment of programmable TFs, which I discuss next in Chapters 1.1.2 and 1.1.3.

1.1.2 Programmable, protein-based transcription factors

With the discovery of customisable and programmable DNA binding proteins, zinc finger proteins (ZFPs) (J. Miller et al. 1985; Brown et al. 1985; Vincent 1986; Berg 1988) and transcription activator-like effectors (TALEs) (Bonas et al. 1989; Römer et al. 2007;

Kay et al. 2007), user-defined targeting of virtually any stretch of DNA in the genome became a reality.

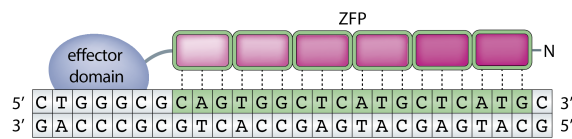
1.1.2.1 Zinc finger proteins (ZFPs)

ZFP are DNA-binding proteins that consist of several Cys2-His2 zinc finger domains (Figure 1.1). The cysteine and histidine residues in each domain coordinate a zinc ion which confers protein stability (Pavletich & Pabo 1991). Each finger consists of approximately 30 amino acids that form a $\beta\beta\alpha$ structure in which the α -domain recognises the major groove of the double helix (M. S. Lee et al. 1989; Pavletich & Pabo 1991; Elrod-Erickson et al. 1996). This mechanism guarantees the sequence-specific recognition of 3 base pairs (bp) in the DNA. Individual finger domains were subsequently rearranged to create artificial ZFP that can virtually bind to any desired DNA sequence (Wright et al. 2006; Carroll et al. 2006; Sander et al. 2011).

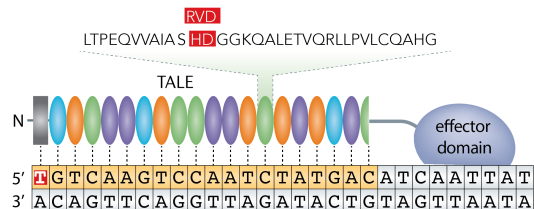
Programmable ZFP fused to effector domains (ZFP transcription factors, ZFP-TFs) offered the unique opportunity to transcriptionally regulate target-specific endogenous loci *in situ* (Sera 2009; Klug 2010). For gene activation, ZFP-TF were created by fusing effector domains such as p65 or VP16 derivatives to ZFP (Q. Liu et al. 1997; P. Q. Liu et al. 2001; L. Zhang et al. 2000; Bae et al. 2003; Beerli, Dreier, et al. 2000; Beltran et al. 2007; Rebar et al. 2002; Y. Lu et al. 2008; Passananti et al. 2010). For gene repression, effector domains such as DNA methyltransferase (Dnm1) (G. L. Xu & Bestor 1997) and KRAB repressor domain (Beerli, Dreier, et al. 2000; Papworth et al. 2003; Reynolds et al. 2003), were deployed.

Furthermore, ZFP-TF were engineered to contain LBD domains of nuclear hormone receptors to render them responsive to exogenously-delivered compounds, creating chemically inducible gene switches (Beerli, Schopfer, et al. 2000; L. Xu et al. 2001; Dent et al. 2007; Magnenat et al. 2008; Schwimmer et al. 2012).

a Zinc finger protein transcription factor (ZFP-TF)



b Transcription activator-like effector transcription factor (TALE-TF)



c Clustered regularly interspaced short palindromic repeat (CRISPR)-associated, nuclease-deficient protein 9 (dCas9) transcription factor (dCas9-TF)

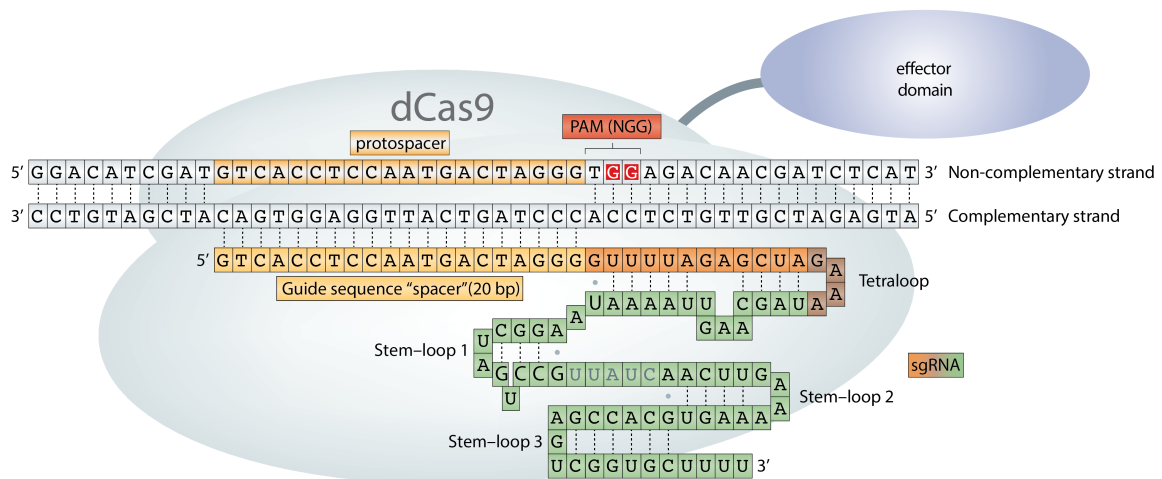


Figure 1.1 | Programmable artificial transcription factors (TFs). (a) Schematic representation of a zinc finger protein transcription factor (ZFP-TF). The ZFP is fused to a transcriptional effector domain. Each finger recognises 3 bp. Typically, when using a single ZFP, 6 fingers are connected, recognising 18 bp (Beerli & Barbas 2002). (b) Schematic representation of an engineered transcription activator-like effector transcription factor (TALE-TF). Each TALE repeat recognises one base pair with its repeat variable diresidue (RVD; shown in red) at position 12 and 13. Target sequence depicted is 20 bp. (c) Schematic representation of CRISPR-associated, nuclease-deficient protein 9 (dCas9) fused to an effector domain, constituting an artificial transcription factor (dCas9-TF). dCas9-TF recognises a genomic locus via the chimeric single guide RNA (sgRNA). PAM, protospacer adjacent motif. Modified from and reprinted by permission from Macmillan Publishers Ltd: [Nat Rev Genet] (Kim, H. & Kim, J.-S., 2014. A guide to genome engineering with programmable nucleases. Nature Reviews Genetics, 15(5), pp.321–334.), copyright (2014). Licence #4250930537596.

1.1.2.2 Transcription activator-like effectors (TALEs)

Transcription activator-like effectors (TALEs) originate from the plant pathogen *Xanthomonas sp.* (Bonas et al. 1989; Boch & Bonas 2010). The central DBD of TALEs from *Xanthomonas sp.* consist of between 1.5 and 33.5 tandem domains (54-909 amino acids) with 17.5 being the most abundant (Boch & Bonas 2010). Each tandem domain is composed of 34-35 amino acids, specifically recognizing a particular DNA base pair (Figure 1.1) (Boch et al. 2009; Moscou & Bogdanove 2009). DNA nucleobase specificity is conferred by two adjacent amino acids at position 12 and 13 in each repeat (Boch et al. 2009; Moscou & Bogdanove 2009). This variable region is referred to as repeat variable diresidues (RVDs) in which amino acid residues HD, NI, NK, and NG specifically and reliably bind to C, A, G, and T nucleotides, respectively. The elucidated simple RVD code enabled the creation of customised, sequence-specific TALEs (Morbitzer et al. 2011; F. Zhang et al. 2011; Reyon et al. 2012; Briggs et al. 2012; Schmid-Burgk et al. 2013; Y. Kim et al. 2013).

TALE-TFs were created either by deploying VP16 or p65 (J. C. Miller et al. 2011; Geissler et al. 2011; Y. Li et al. 2012) to induce gene upregulation or by coupling TALEs to KRAB (Y. Li et al. 2012) and DNMT3 domains (Amabile et al. 2016) for gene repression.

Inducible TALE-TFs were engineered to gain exogenous control over effector function either by exploiting oestrogen or progesterone receptor LBD (Mercer et al. 2014) or by deploying light-induced heterodimerisation domains (Konermann et al. 2013).

Taken together, engineered ZFP-TFs and TALE-TFs were the first adapted programmable TFs. However, customising DNA binding sequences requires time-consuming molecular biology work and laborious validation, especially for ZFPs (Wright et al. 2006; Carroll et al. 2006; Ramirez et al. 2008; Morbitzer et al. 2011; F. Zhang et

al. 2011; Reyon et al. 2012; Briggs et al. 2012; Schmid-Burgk et al. 2013; Y. Kim et al. 2013).

1.1.3 Programmable, RNA-guided dCas9-TF

Harnessing the power of clustered regularly interspaced short palindromic repeats (CRISPR) has heralded the start of a new age in genome engineering (Pennisi 2013; Lander 2016). Initially repurposed for targeted gene editing, CRISPR has been adapted for seemingly countless applications (H. Kim & J.-S. Kim 2014; Mali, Esvelt, et al. 2013; Hsu et al. 2014; Reardon 2016). Unprecedented in its ease to use, low cost, specificity, and efficiency (H. Kim & J.-S. Kim 2014; Cox et al. 2015), CRISPR has revolutionised and democratised the genome engineering field (Ledford 2015; Travis 2015; Barrangou & Doudna 2016; Barrangou & Horvath 2017).

1.1.3.1 Mechanism of clustered regularly interspaced short palindromic repeat (CRISPR)-associated nuclease 9 (Cas9)

CRISPR (Ishino et al. 1987; Mojica et al. 1995; Jansen et al. 2002) constitutes an adaptive bacteriophage defence system in prokaryotes (Barrangou et al. 2007; Barrangou 2015; Horvath & Barrangou 2010; Bhaya et al. 2011; Labrie et al. 2010). Based on the nature of CRISPR-associated nucleases (Cas), there are 2 classes of the CRISPR-Cas system (Makarova et al. 2015). Class 1 deploys Cas3 exonuclease together with the multi-protein complex Cascade (CRISPR-associated complex for antiviral defence), whereas class 2 contains only one effector nuclease. Class 2 is further subdivided into type II and type V CRISPR systems, which employ Cas9 (also known as Csn1) and Cpf1 (Cas12a) as endonuclease, respectively (Makarova et al. 2015).

Both classes, however, typically encode Cas1 and Cas2 proteins in their *cas* operon which are essential for the acquisition (first step) of virally or plasmid DNA-derived spacers sequences for the integration into the CRISPR array (Nuñez, Lee, et al. 2015; Nuñez, Harrington, et al. 2015; Jiuyu Wang et al. 2015). The CRISPR array is composed of many short palindromic repeats which are interspaced by newly acquired spacers (~30 nucleotides in type II) (Jansen et al. 2002; Barrangou et al. 2007; Heler et al. 2015; Levy et al. 2015). In a second step (expression), the CRISPR array is expressed as a long pre-CRISPR RNA (pre crRNA) (Barrangou et al. 2007; Haurwitz et al. 2010; Deltcheva et al. 2011). In the type II CRISPR system of *Streptococcus pyogenes*, this pre-crRNA is processed by an endogenous RNase III into single crRNAs after the recruitment of trans-activating crRNA (tracrRNA) and Cas9, forming a crRNA:tracrRNA Cas9 complex (Deltcheva et al. 2011; Jinek et al. 2012; Karvelis et al. 2013; Y. Lim et al. 2016). The crRNA contains a variable target-specific spacer sequence at the 5' end as well as the non-variable repeat sequence at the 3' end which is complementary to the target-independent tracrRNA (Deltcheva et al. 2011; Jinek et al. 2012; Karvelis et al. 2013). Subsequently, the 5' spacer sequence of crRNA is trimmed down to 20 nucleotides (Deltcheva et al. 2011; Jinek et al. 2012). In a third step (interference), crRNA:tracrRNA Cas9 complex scans the genomic/plasmid DNA for a complementary target site (Deltcheva et al. 2011; Gasiunas et al. 2012; Knight et al. 2015; Singh et al. 2016). A potential target site is first engaged upon the detection of a protospacer adjacent motif (PAM) (Figure 1.1) (Mojica et al. 2009; Garneau et al. 2010). *Streptococcus pyogenes* Cas9 (SpCas9) requires a 5'-NGG-3' PAM sequence for recognition and upon perfect or near-perfect spacer-protospacer sequence complementarity, DNA cleavage occurs (Jinek et al. 2012; Cong et al. 2013; W. Jiang et al. 2013; Sternberg et al. 2014; Singh et al. 2016). Cas9 possesses two nuclease domains, RuvC and HNH, each creating a nick in the DNA, yielding a double-strand

break 3 nucleotides downstream of the PAM sequence (Garneau et al. 2010; Gasiunas et al. 2012; Jinek et al. 2012).

1.1.3.2 Harnessing CRISPR/dCas9 for inducible control of gene expression

In 2012/2013, the RNA-guided DNA cleavage system from type II CRISPR/Cas9 *Streptococcus pyogenes* was repurposed for targeted genome editing in eukaryotic cells (Jinek et al. 2012; Cong et al. 2013; Mali, Yang, et al. 2013; Cho et al. 2013). Instead of targeting DNA sequences originating from viral DNA in bacteria, any piece of DNA in the human genome could be cleaved by equipping Cas9 with corresponding crRNA:tracrRNA guides. Jennifer Doudna's laboratory simplified the dual RNA guide system by fusing crRNA and tracrRNA at their complementary regions together, creating an artificial chimeric single guide RNA (sgRNA) (Jinek et al. 2012) (Figure 1.1).

By introducing mutations in SpCas9 catalytic domains, either D10A/H840A (Jinek et al. 2012) or D10A/D839A/H840A/N863A (Mali, Yang, et al. 2013), SpCas9 was rendered catalytically deficient (dCas9). Similarly, Cas9 homologs and orthologs were subsequently catalytically inactivated, such as *Streptococcus aureus* dSaCas9 (D10A/N580A) (Ran et al. 2015) and dCpf1 homologs (Tak et al. 2017; Y. Liu et al. 2017). This engineering step established Cas9 as versatile RNA-guided DBD.

dCas9 has been extensively used in conjunction with transcriptional regulators that are either directly fused to and/or associated with dCas9, allowing the regulation of mammalian gene expression *in situ* (Figure 1.2) (Chavez et al. 2016; Dominguez et al. 2016; Jusiak et al. 2016; Mitsunobu et al. 2017; Nihongaki et al. 2018).

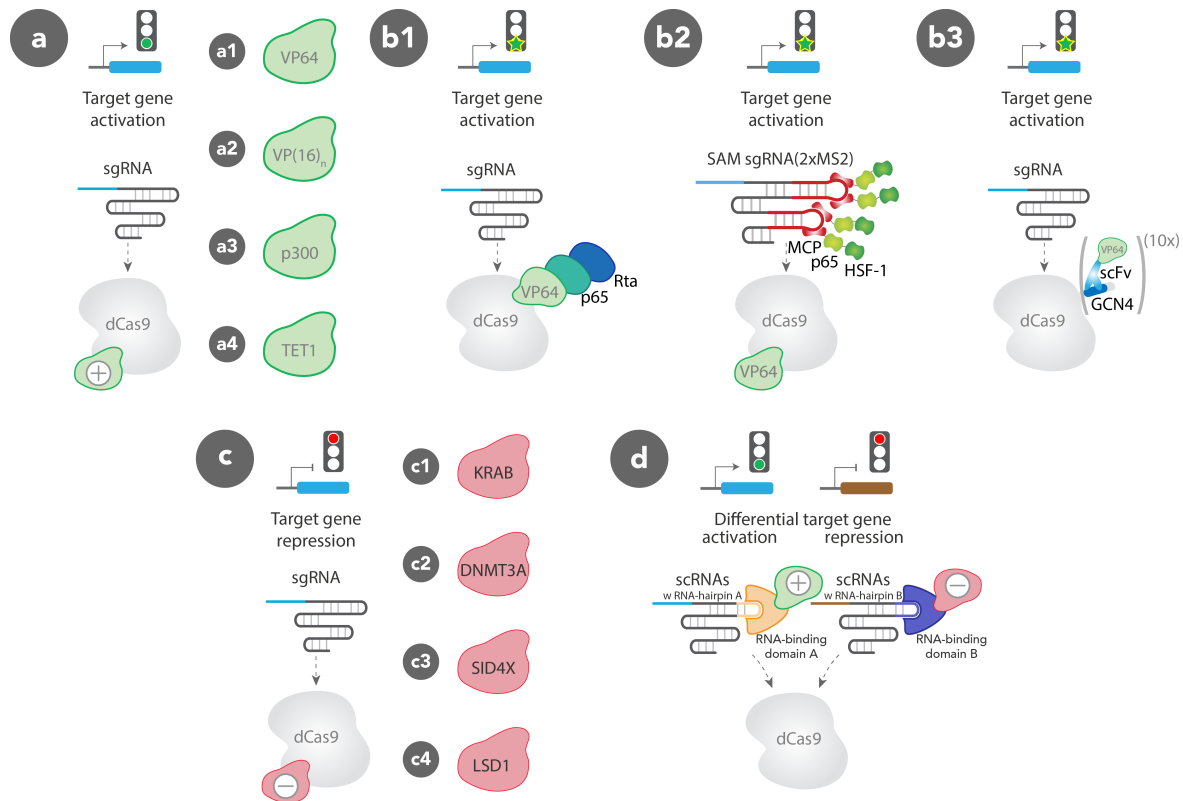


Figure 1.2 | dCas9-TF for transcriptional activation and repression. (a) Transcriptional activation has been achieved by directly fusing dCas9 to various transcriptional activators, such as (a1) VP64; (a2) VP160 or VP192; (a3) p300 core; (a4) TET1CD (catalytic domain of ten-eleven 1 hydroxylase). **(b)** Second-generation artificial transcriptional activators. (b1) VPR, an artificial fusion protein composed of VP64, the human NF- κ B p65 activation domain, and Epstein-Barr virus-derived R transactivator (Rta); (b2) SAM (synergistic activation mediator) system, employing a synthetic TF fusion protein comprised of MCP (MS2 bacteriophage coat protein), p65, and HSF-1 (activation domains of human heat-shock factor 1); (b3) SunTag, GCN4 epitope recognising scFv-VP64 fusion protein. **(c)** Transcriptional repression has been achieved by directly fusing dCas9 to (c1) KRAB; (c2) DNMT3A (DNA methyl-transferase domain); (c3) SID4X (synthetic repressor was constructed by concatenating 4 SID domains); (c4) LSD1 (lysine-specific histone demethylase 1A). **(d)** sgRNA scaffolds (scRNAs) containing RNA hairpins to recruit corresponding RNA hairpin binding proteins fused to various effectors.

1.1.3.3 CRISPR/dCas9 for transcriptional activation

To upregulate endogenous gene expression, dCas9-based transcription factors (dCas9-TF) have been created by coupling dCas9 directly to various transcriptional activators, such as VP64 (Maeder et al. 2013; Mali, Aach, et al. 2013; Perez-Pinera et al. 2013; Gilbert et al. 2013; Farzadfard et al. 2013; Black et al. 2016), VP160 (Cheng et al. 2013; X. Gao et al. 2014), VP192 (Balboa et al. 2015), the core domain of histone acetyltransferase p300 (Hilton et al. 2015; Shrimp et al. 2018; P. Liu et al. 2018), and catalytic domain of ten-eleven 1 hydroxylase (TET1CD) (X. S. Liu et al. 2016; X. Xu et al. 2016; Morita et al. 2016) (Figure 1.2a).

For more potent transcriptional gene upregulation, three strategies of second-generation transcriptional activators have been conceived (Chavez et al. 2016) (Figure 1.2b). (i) dCas9 can be directly fused to the artificial VPR transactivation chimera, consisting of VP64, the human NF-KB p65 activation domain and Epstein-Barr virus-derived R trans-activator (Rta) (Chavez et al. 2015) (Figure 1.2b1). (ii) The synergistic activation mediator (SAM) system deploys dCas9-VP64 in combination with a custom-made SAM sgRNA containing two MS2 aptamers. These aptamers are bound by the chimeric transactivator fusion protein MCP-p65-HSF-1 (MCP, MS2 bacteriophage coat protein recognising MS2 aptamers; HSF-1, activation domains of human heat-shock factor 1) (Konermann et al. 2015) (Figure 1.2b2). (iii) The SunTag system is composed of a dCas9 fused to a peptide tail containing 10 GCN4 epitopes. These epitopes are recognised by a soluble fusion protein composed of anti-GCN4 svFc and VP64 (Tanenbaum et al. 2014) (Figure 1.2b3).

Efficient activation of endogenous genes is facilitated when the RNA-guided dCas9-TF is targeted in proximity to the transcriptional start site (TSS), approximately -400/-200 to -50/+1 bp (Gilbert et al. 2014; Konermann et al. 2015; Polstein et al. 2015).

CRISPR activation (CRISPRa) genome-wide screens have been used to elucidate genes involved in cholera-diphtheria toxin sensitivity (Gilbert et al. 2014), to confer resistance to a particular drug treatment (Konermann et al. 2015), to decipher the molecular mechanism of an anti-cancer drug (Jost et al. 2017), to identify long non-coding RNAs (lncRNAs) involved in regulating neighbouring genes (Joung et al. 2017), and to reveal enhancer regions (Simeonov et al. 2017).

1.1.3.4 CRISPR/dCas9 for transcriptional repression

To downregulate endogenous gene expression, dCas9 was coupled to transcriptional repressor domains, such as KRAB (Gilbert et al. 2013; X. Gao et al. 2014; Thakore et al. 2015; Amabile et al. 2016; Mandegar et al. 2016), DNMT3A (DNA methyltransferase domain) (Amabile et al. 2016; X. S. Liu et al. 2016; Vojta et al. 2016; Stepper et al. 2017; Zheng et al. 2018), SID4X (four concatenated SID domains) (Konermann et al. 2013), and LSD1 (lysine-specific histone demethylase 1A) (Kearns et al. 2015) (Figure 1.2b).

The development of RNA hairpin-containing sgRNA scaffolds (scRNAs) enabled the differential or simultaneous recruitment of activation and/or repression domains to multiple gene targets, thereby multiplexing the transcriptional output (Zalatan et al. 2015). RNA hairpin examples include MS2, PP7, or com RNA which bind to their respective cognate RNA-binding proteins MCP, PCP and com, respectively (Zalatan et al. 2015) (Figure 1.2d). Adaptations of this system implemented various alternative RNA motifs to actively recruit effector domains thereby increasing combinatorial complexity (Shechner et al. 2015; Cheng et al. 2016).

In contrast to CRISPRa, CRISPR interference (CRISPRi) is most effective when dCas9-KRAB is directed to a window of -50 to +300 bp relative to TSS (Gilbert et al. 2014). Genome-wide CRISPRi screens have been employed to identify genes implicated in

conferring sensitivity to a bacterial toxin (Gilbert et al. 2014), to characterise genes implicated in an unfolded protein response (Adamson et al. 2016), to identified functional long noncoding RNA loci in human cells (S. J. Liu et al. 2017), to elucidate the molecular mechanism of an anti-cancer drug (Jost et al. 2017), to characterise genetic interactions (Du et al. 2017), and to identify essential genes (Rosenbluh et al. 2017).

To reliably and confidently deploy RNA-guided dCas9-TF as a versatile transcriptional regulator, studies investigated dCas9 binding specificity by chromatin immunoprecipitation followed by deep DNA sequencing (ChIP-seq) experiments. These experiments revealed that RNA-guided dSpCas9 preferably binds to open chromatin and only requires sequence complementarity to the 5-8 PAM-proximal nucleotides and a matching PAM. These binding properties resulted in numerous genome-wide off-target binding events (Kuscu et al. 2014; X. Wu et al. 2014; O'Geen et al. 2015). In agreement with these results were the binding profiles of dCas9 with or without effector domains (VP64 or KRAB), indicating that dCas9-TF binding is unaffected by covalently coupled effector domain(s) (Thakore et al. 2015; O'Geen et al. 2015; Polstein et al. 2015; Zheng et al. 2018).

Owing to the fact that dCas9-TFs exert their transcriptional potency only in a narrow window surrounding the TSS (Gilbert et al. 2014; Konermann et al. 2015; Polstein et al. 2015), it was more reasonable to examine off-target gene regulation rather than off-target binding. Transcriptome analysis via RNAseq revealed highly target-specific transcriptional regulation for dCas9-VP64 (Perez-Pinera et al. 2013; Polstein et al. 2015) and dCas9-KRAB (Gilbert et al. 2013; Thakore et al. 2015). In conclusion, the transient transcriptional up- or downregulation facilitated by dCas9-TFs is highly specific.

1.1.3.5 Inducible dCas9-TF systems

The simplicity with which dCas9-TFs can be directed to any genomic locus of choice has led to the development of a broad repertoire of inducible systems. Exogenously delivered input stimuli have enabled the fine-tuned and spatiotemporally tight regulation of dCas9-TFs and enabled more sophisticated implementations of differential target gene activation and/or repression. Examples include:

- Figure 1.3a. sgRNA-based inducibility concepts. (a1) Inducible spacer-blocking hairpin (iSBH). A spacer-blocking hairpin on the sgRNA prevents sgRNA-guided dCas9-TF recruitment to the desired genomic locus. In the presence of inducer input, either genetically encoded Csy4 endoribonuclease or exogenously delivered antisense oligonucleotides, the blocking sequence on the sgRNA is removed, reconstituting functional sgRNA (Ferry et al. 2017). (a2) Small molecule responsive, self-cleaving ribozyme (aptazyme)-embedded sgRNAs. Only in the presence of a small molecule input, the obstructing ribozyme is removed, liberating functional sgRNA (Y. Liu et al. 2016; Tang et al. 2017).
- Figure 1.3b. Orthogonal, multiplexed, or differential gene expression can be achieved by the dimerisation domain-mediated recruitment of transcriptional regulators to dCas9. Examples of inducer/dimerisation domain pairs include: plant hormone abscisic acid (ABA)-inducible ABI-PYL1 (Y. Gao et al. 2016; T. Chen et al. 2017); plant hormone gibberellin (GA)-inducible GID1-GAI (Y. Gao et al. 2016; Bao et al. 2017); blue light-inducible CRY2PHR-CIBN (Nihongaki, Yamamoto, et al. 2015; Polstein & Gersbach 2015; Y. Gao et al. 2016), FKF1-GI (Y. Gao et al. 2016); red light-controlled PHYB-PIF (Y. Gao et al. 2016); small molecule ABT-737-inducible BCL-xL-antibody (AbCIDs, human antibody-based chemically induced dimerisers (Hill et al. 2018)).

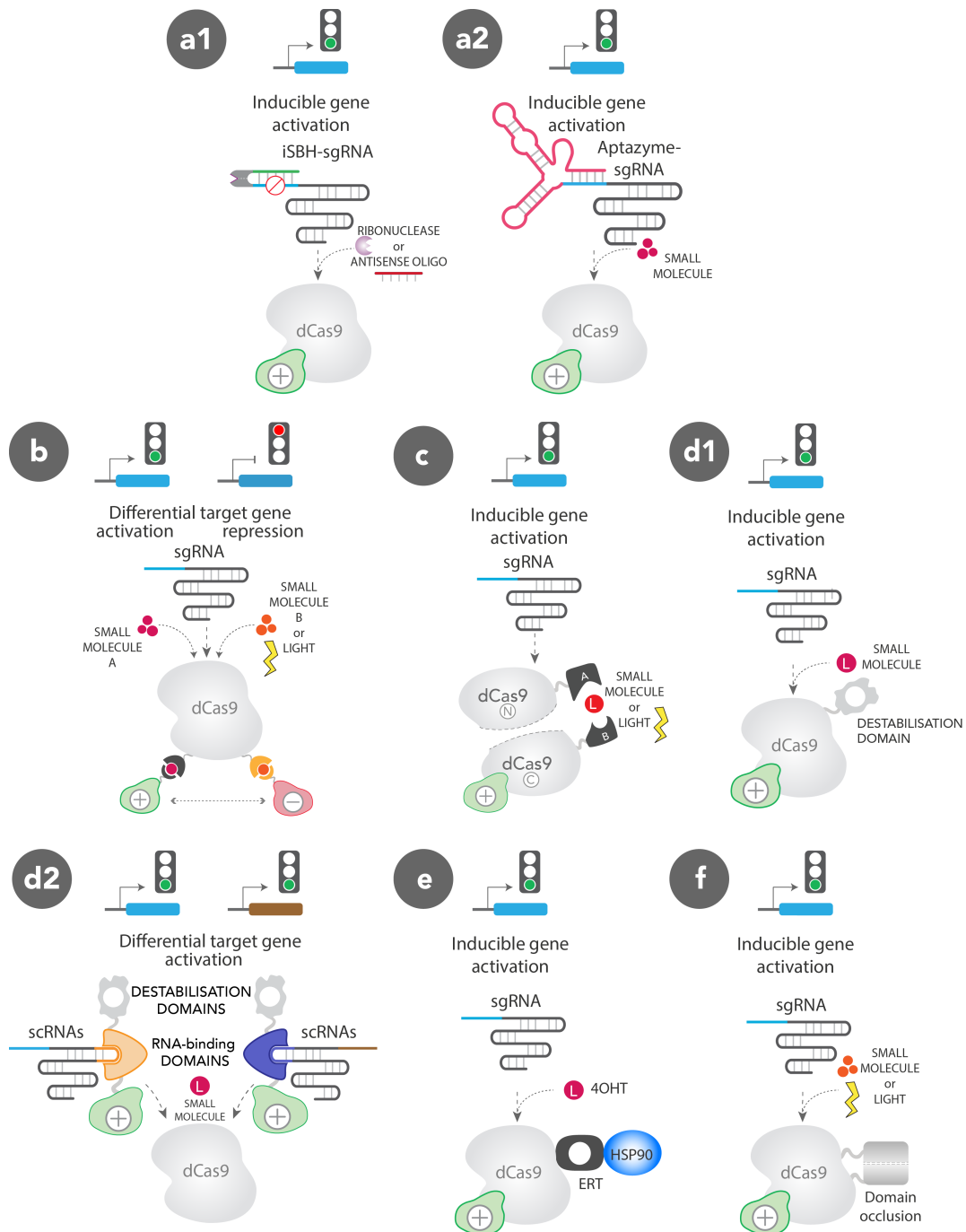


Figure 1.3 | Inducible dCas9-TF for controlled endogenous gene regulation. (a) sgRNA-based inducibility. (a1) iSBH (inducible spacer-blocking hairpin); (a2) Small molecule responsive, self-cleaving ribozyme (aptazyme)-embedded sgRNAs. **(b)** Orthogonal, multiplexed, or differential gene expression achieved via dimerisation domain-mediated recruitment of transcriptional regulators to dCas9. **(c)** Dimerisation-induced reassembly of split dSpCas9. **(d)** Conditional destabilisation of dCas9-TF. (d1) Destabilisation domains (DD) directly fused to dCas9; (d2) DD fused to effector domains. **(e)** Cytoplasmic sequestration of dCas9-TF. HSP90, heat shock protein 90; ERT, oestrogen receptor. **(f)** Steric hindrance-mediated sequestration of dCas9-TF.

- Figure 1.3c. Reconstitution of functional dCas9-TF via dimerisation-induced reassembly of split dSpCas9 by small molecule rapamycin-mediated rapamycin domain dimerisation (Zetsche, Volz, et al. 2015) or by blue light-mediated pMag and nMag dimerisation (Nihongaki, Kawano, et al. 2015).
- Figure 1.3d. Conditional destabilisation of dCas9-TF. (d1) dCas9 can be directly fused to destabilisation domains, such as oestrogen receptor (ER50) destabilisation domain and dihydrofolate reductase (DHFR), promoting its rapid degradation unless supplying trimethoprim (TMP) and 4-hydroxytamoxifen (4OHT) for reversible stabilisation of dCas9 fusion proteins, respectively (Balboa et al. 2015; Maji et al. 2017; Kleinjan et al. 2017). Alternatively, fusion to the destabilising domain of mutant FK506-binding protein 12 (FKBP12) allows stabilisation with a synthetic ligand (Shield-1) (Senturk et al. 2017). (d2) Fusing destabilisation domains to effector domains enable orthogonal control over transcriptional potency (Maji et al. 2017; Kleinjan et al. 2017).
- Figure 1.3e. Cytoplasmic sequestration of dCas9-TF via heat shock protein 90 (HSP90). dCas9-TF is fused to the hormone-binding domain of the oestrogen receptor (ERT2), which recruits HSP90 and localises the fusion protein to the cytoplasm. Upon 4-hydroxytamoxifen (4OHT) treatment, HSP90 dissociates which then allows dCas9-TF to translocate into the nucleus (Nguyen et al. 2016). Implementations include fusing full length dCas9-TF to ERT domains (K. I. Liu et al. 2016; J. Lu et al. 2017), split dCas9-TF to ERT domains (Nguyen et al. 2016), or inserting ER-LBD into a designated region within Cas9 (Oakes et al. 2016).
- Figure 1.3f. Steric hindrance can conditionally sequester dCas9-TF transcriptional potency. Implementations include the engraftment of 4OHT-responsive intein into Cas9 that renders Cas9 inactive unless 4OHT is co-

administered, which then induces protein splicing and thereby restores Cas9 activity (K. M. Davis et al. 2015); or the coupling of dCas9-TF to dimeric fluorescent protein domains that dissociate upon light exposure (Richter et al. 2016; Zhou et al. 2017).

Various inducible systems have also been utilised in combination to design higher order gene regulatory networks (D. Ma et al. 2016; Nihongaki et al. 2017; J. Lu et al. 2017).

Apart from the presented concepts, Cas9/dCas9-TF activity can additionally be governed by utilising (i) small molecule-inducible promoters (Tim Wang et al. 2014; Mandegar et al. 2016; Ishida et al. 2018), (ii) transcriptional cascades (Kiani et al. 2014; Nissim et al. 2014; Kiani et al. 2015), and (iii) miRNA target site-bearing Cas9 mRNA (D. Ma et al. 2016; Hirose et al. 2017).

1.2 Signalling networks and receptor engineering

Naturally, mammalian cells are decorated with a plethora of cell-surface receptors that provide a link from the extracellular to the intracellular space (Figure 1.4). They recognise a broad spectrum of ligands, ranging from ions and small molecules to sugars, lipids, and proteins. Upon ligand sensing, receptors transmit the signal across the cell membrane into the intracellular part of the cell. From there, complex signalling cascades are triggered that involve choreographed interactions of tightly regulated proteins, mediators, and second messengers. Eventually, a specific cellular response is created such as upregulation of transcription, initiation of secretion, changes in the cytoskeleton or promotion of cell growth (Figure 1.4) (W. A. Lim 2010; W. Lim et al. 2014). This information processing guarantees that cells in higher order structures such as tissues and organs act and react appropriately to an ever-changing environment (Chapter 1.2.1).

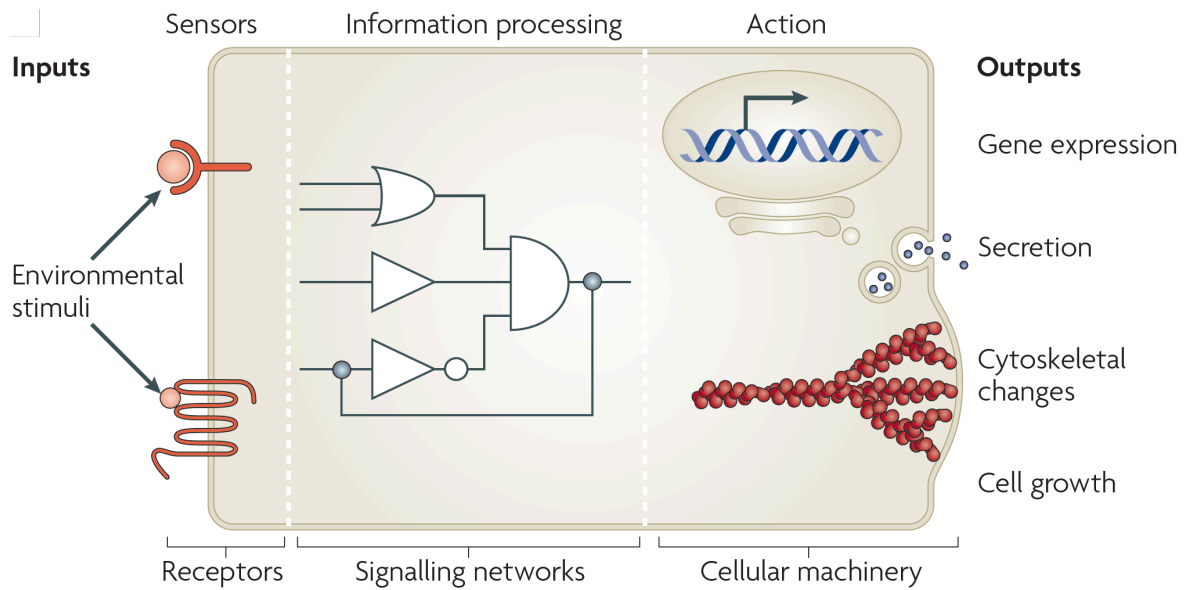


Figure 1.4 | Endogenous cell signalling. Cell-surface receptors enable the integration of a broad variety of environmental stimuli into the cell via complex signalling cascades. Eventually, the extracellular information is converted into a specific cellular response, e.g., transcriptional output. Reprinted by permission from Macmillan Publishers Ltd: [Nat Rev Mol Cell Biol.] (Lim, W.A., 2010. Designing customized cell signalling circuits. Nature Reviews Molecular Cell Biology, 11(6), pp.393–403.), copyright (2014). Licence #4243840331722.

If the intra- and intercellular communication is distorted, however, a large spectrum of diseases ensues, such as cancer, metabolic disorders, inflammatory diseases, and autoimmune ailments. To potentially correct these defects, a number of novel cell-based therapies have been conceptualised in addition to more traditional drug-based strategies. This emerging line of therapy revolves around cellular engineering. By exploiting the modular structure of receptors and signalling components and through substantial rational protein engineering, new artificial receptors have been created and embedded into cells (W. A. Lim 2010; Lienert et al. 2014; Arber et al. 2017; Brenner et al. 2018). These designer cells can now respond to pre-defined disease-associated molecules by initiating user-defined native or artificial signalling pathways and gene circuits, leading to the expression of therapeutically-relevant transcriptional programmes (Chapter 1.2.2 and 1.2.3) (S. Ausländer & Fussenegger 2013; Fesnak et al. 2016; W. A. Lim & June 2017; Teixeira & Fussenegger 2017; Kitada et al. 2018).

1.2.1 Endogenous mammalian cell-surface receptors

There are three major native cell-surface receptor families: G-protein-coupled receptors (GPCRs), receptor tyrosine kinases (RTKs) and ligand-gated and voltage-gated ion channels (Simons et al. 2016; Dorsam & Gutkind 2007). In the following sections, I detail characteristic properties of the former two families because they have been the subject of extensive engineering purposes (Chapter 1.2.3).

1.2.1.1 G-protein-coupled receptors (GPCRs)

G-protein-coupled receptors (GPCRs) are seven transmembrane-spanning receptors comprising the most diverse and largest class of cell-surface receptors in mammals (Pierce et al. 2002; Kroeze et al. 2003; Allen & Roth 2011). They account for roughly

2% of the human genome and are evolutionarily conserved across animals, plants, and fungi (Venter et al. 2001; Takeda et al. 2002; King et al. 2003; Fredriksson & Schiöth 2005; Schiöth et al. 2007; Allen & Roth 2011). GPCRs are involved in nearly every cellular and physiological process (Kroeze et al. 2003; Allen & Roth 2011). Their importance is highlighted by the fact that malfunctioning GPCRs are implicated in many human diseases. As a result they represent an attractive target for numerous therapeutic drugs (Hopkins & Groom 2002; M. D. Thompson et al. 2005; Armbruster & Roth 2005; Dorsam & Gutkind 2007; Lagerström & Schiöth 2008).

GPCR families

Human GPCRs are categorised into 5 families based on sequence homology and pharmacological similarities: A (rhodopsin), B (secretin), C (glutamate), adhesion, and frizzled/taste2 (Fredriksson et al. 2003; Lagerström & Schiöth 2008). Class A GPCRs constitute the largest and most diverse family in terms of structure and ligand preference, which include odorants, peptides, lipids, nucleotides, and photons (Figure 1.5). This class is composed of ~670 full-length receptors (Gloriam et al. 2007) and is subdivided into groups α , β , γ and δ (Fredriksson et al. 2003). Examples include arginine vasopressin receptor 2 (AVPR2), bradykinin receptor B2 (BDKRB2), lysophosphatidic receptors (LPARs), and C-X-C chemokine receptor type 4 (CXCR4). Class B GPCRs are activated by hormones and peptides (Lagerström & Schiöth 2008). Class C GPCRs (e.g., sweet and umami taste receptors, T1R2 and T1R3) bind to hormones and peptides, amino acids, ions, and taste-provoking chemical molecules (tastants) (Lagerström & Schiöth 2008; Laffitte et al. 2014). Adhesion GPCRs are predicted to interact with extracellular matrix or membrane-bound proteins (Lagerström & Schiöth 2008). Frizzled/taste2 GPCRs are activated by Wnt proteins and tastants (Lagerström & Schiöth 2008).

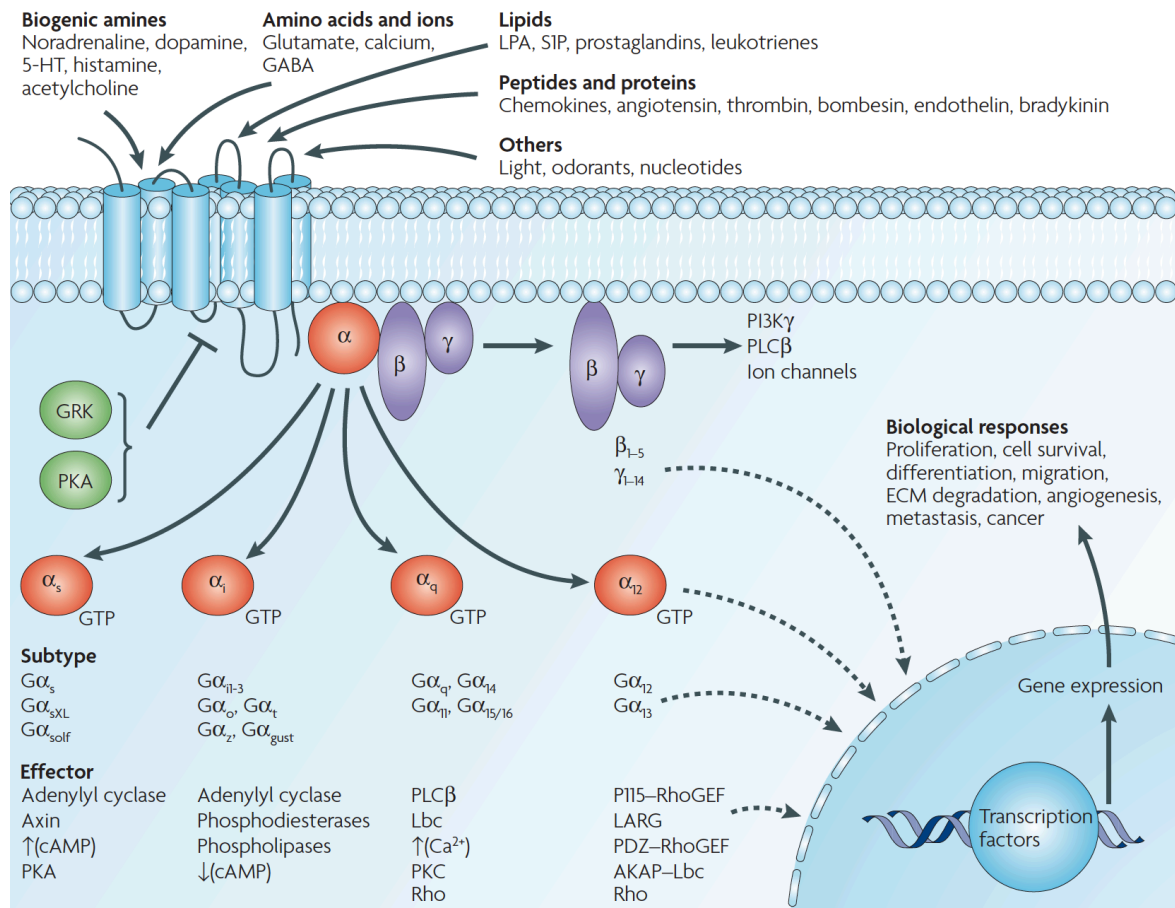


Figure 1.5 | G-protein-coupled receptor signalling. G-protein-coupled receptors (GPCRs) recognise a broad spectrum of ligands and subsequently initiate multifactorial downstream signalling cascades. Upon agonist-mediated conformational change within the receptor and subsequent GDP to GTP exchange in the $G\alpha$ subunit of the heterotrimeric G protein, $G\alpha$ dissociates from $G\beta\gamma$ subunits. Different GPCRs can interact with one or several different $G\alpha$ subunit subfamilies: $G\alpha_s$, $G\alpha_i$, $G\alpha_q$ and $G\alpha_{12}$. $G\alpha$ subtypes stimulate or inhibit downstream effectors, such as adenylyl cyclase or phospholipase C (PLC). $G\beta\gamma$ subunits can stimulate phospholipases, kinases and ion channels. Some $G\alpha$ subtypes may activate other signal transduction molecules, such as small GTP-binding proteins of the Ras and Rho families. Eventually, signalling leads to differential gene expression, regulating various cellular functions. Distorted signalling is attributed to favouring cancer development and metastasis. 5-HT, 5-hydroxytryptamine; ECM, extracellular matrix; GABA, gamma-aminobutyric acid; GEF, guanine nucleotide exchange factor; GRK, G-protein receptor kinase; LPA, lysophosphatidic acid; PI3K, phosphatidyl inositol 3-kinase; PKA and PKC, protein kinase A and C; S1P, sphingosine-1-phosphate. Reprinted by permission from Macmillan Publishers Ltd: [Nat Rev Cancer] (Dorsam, R.T. & Gutkind, J.S., 2007. G-protein-coupled receptors and cancer. Nature Reviews Cancer, 7(2), pp.79–94.), copyright (2007). Licence #4243851009975.

GPCR structure

Commonly, each of the seven α -helical segments spanning the transmembrane is separated by alternating intracellular and extracellular loop regions, whereas the third intracellular loop and intracellular C-terminal part of GPCRs (C-terminal tail) represent the primary point of contact for GPCR-associated proteins (adapter proteins). The N-terminal structures of GPCRs range from short and unstructured (e.g., rhodopsin-like) to large and globular (e.g., sweet and umami taste receptors) (Lagerström & Schiöth 2008; Assadi-Porter et al. 2010; Laffitte et al. 2014). The ligand binding domains (LBDs) for peptide hormones or proteins and small organic compounds (e.g., biogenic amines) are in the N-terminal domain and within the transmembrane segments, respectively (Armbruster & Roth 2005; Kobilka 2007). Ligand binding triggers conformational changes most pronounced in the transmembrane and intracellular domains (Katritch et al. 2012).

GPCR basic signalling principle

Generally, upon agonist binding, the GPCR receptor is activated and through its core transmembrane domain interacts with heterotrimeric G proteins ($G\alpha\beta\gamma$), favouring guanine nucleotide exchange (replacing GDP with GTP). This exchange leads to the dissociation of $G\alpha$ subunit from the $G\beta\gamma$ subunits and from the receptor (Figure 1.5) (Gilman 1987). Both subunits can then initiate signalling cascades that involve a broad variety of effectors (Gilman 1987; Pierce et al. 2002). Agonist-induced conformational changes lead to the phosphorylation of the intracellular GPCR tail by GPCR kinases (GRKs). These phosphorylation marks are subsequently recognised by arrestins (Reiter & Lefkowitz 2006). Individual GPCRs possess functional plasticity allowing them to signal through different G-protein subtypes and G-protein independent pathways (Kobilka & Deupi 2007).

GPCR signalling: focus on G proteins

The heterotrimeric G proteins represent the canonical mode of signalling. There are at least 18 different human $G\alpha$ proteins, 5 types of $G\beta$ subunits, and 11 different $G\gamma$ subunits (Figure 1.5) (Hermans 2003). $G\alpha$ proteins, such as $G\alpha_s$, $G\alpha_i$, and $G\alpha_q$, feed into a variety of pathways either by directly modulating downstream effector proteins or via second messengers (Figure 1.5) (Galandrin et al. 2007; Ritter & Hall 2009). Second messengers subsequently activate additional effectors further downstream, such as protein kinase A (PKA) and protein kinase C (PKC), or trigger intracellular calcium release. Dissociated $G\beta\gamma$ subunits can open ion channels or feed into the PLC β pathway (Figure 1.5) (Ritter & Hall 2009).

GPCR signalling: focus on GPCR kinases (GRKs)

GRKs in combination with arrestins display an antagonistic behaviour compared to the canonical G-protein-mediated response (Figure 1.6) (Reiter & Lefkowitz 2006; Gurevich et al. 2012). Agonist-occupied GPCRs are phosphorylated at serine and threonine residues in the C-terminal cytoplasmic tail and/or third cytoplasmic loop by GRKs, which subsequently promote the recruitment of cytosolic β -arrestins (Figure 1.6) (Shenoy & Lefkowitz 2003; Reiter & Lefkowitz 2006; Gurevich et al. 2012). There are 7 different GRKs (GRK1–GRK7) which are all differentially expressed in diverse tissues and organs and display preferential binding abilities to specific GPCRs (Reiter & Lefkowitz 2006; C. A. C. Moore et al. 2007; Premont & Gainetdinov 2007).

GPCR signalling: focus on arrestins

Arrestins constitute a family of four members in mammals: arrestin1 and arrestin4 (α -arrestins, confined to retinal rods and cones in the eye); and arrestin2 (known as β -arrestin1) and arrestin3 (β -arrestin2) which are ubiquitously expressed and interact with most GPCRs (Shenoy & Lefkowitz 2003; Reiter & Lefkowitz 2006).

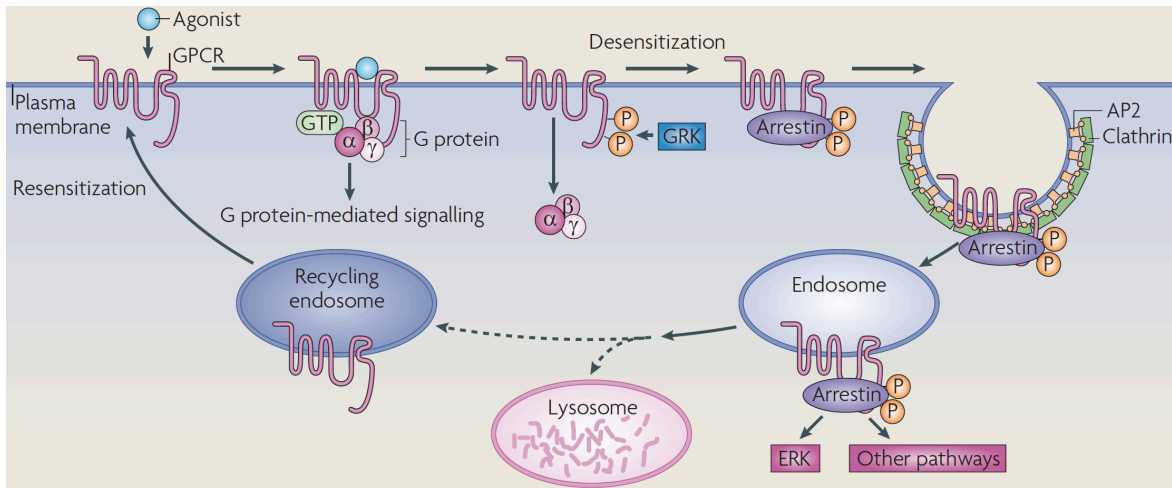


Figure 1.6 | Classical arrestin-mediated signalling by GPCRs. GPCR kinases (GRKs) catalyse the phosphorylation of agonist-activated GPCRs to facilitate termination of canonical signalling by initiating a degradation or recycling programme. The phosphorylated C-terminal tail of GPCRs attracts arrestin, which further recruits clathrin and the clathrin adaptor AP2 for the subsequent GPCR internalisation into endosomes (desensitisation). GPCR–arrestin interactions can trigger specific arrestin-mediated signalling, such as the activation of the mitogen-activated protein kinase (MAPK) extracellular signal-regulated kinase (ERK) pathway, involving small GTP-binding proteins of the Ras and Rho family. Internalised GPCR–arrestin complexes either feed into lysosome degradation processes or are recycled and shuttled back to the cell membrane (resensitization) to start another cycle of signalling. Reprinted by permission from Macmillan Publishers Ltd: [Nat Rev Mol Cell Biol.] (Ritter, S.L. & Hall, R.A., 2009. Fine-tuning of GPCR activity by receptor-interacting proteins. *Nature Reviews Molecular Cell Biology*, 10(12), pp.819–830.), copyright (2009). Licence #4243860799518.

For a long time, arrestins were primarily implicated in the silencing, internalisation and subsequent resensitisation/degradation of activated GPCRs (Figure 1.6). Actively recruited β -arrestin proteins to the phosphorylated C-terminal tail of activated GPCRs were believed to first, sterically impede continuous G-protein association to the receptor (binding competition model) and second, initiate an endocytosis cascade that would shut off G-protein-coupled signalling completely (Ferguson et al. 1996; Ferguson 2001; Shenoy & Lefkowitz 2003; Shenoy & Lefkowitz 2005; Kang et al. 2015; Shukla et al. 2014; Szczepek et al. 2014). GPCR internalisation is mediated through clathrin-coated pits, caveolae, or uncoated vesicles (Figure 1.6) (Claing et al. 2002; Marchese et al. 2003). The success of internalisation is, however, a function of the binding strength of β -arrestin to the GPCR. The presence or absence of enriched serine and threonine residues in the C-terminal tail of GPCRs determines the strength with which GPCRs interact with β -arrestins (Oakley et al. 2000; Oakley et al. 2001; Vilardaga et al. 2002). β -arrestins only transiently recruit to Class A GPCRs which do not possess serine and threonine clusters and dissociate from the GPCR after initial contact with clathrin-coated pits. Hence, β -arrestin offers a means for modulating GPCR signalling responsiveness to a continuously occurring stimulus (Figure 1.6).

Recently, emerging data for candidate agonist-stimulated GPCRs suggest sustained G-protein-dependent signalling capabilities concomitant with GPCR internalisation (Calebiro et al. 2009; Ferrandon et al. 2009; Feinstein et al. 2013; Irannejad et al. 2013; Werthmann et al. 2012; Calebiro et al. 2010; Kotowski et al. 2011; Mullershausen et al. 2009).

GPCR signalling: other modalities

An additional layer of GPCR signal modulation involves the formation of homo-/heterodimers/oligomers (Milligan 2004; Franco et al. 2007; Bulenger et al. 2005; Ferré et al. 2014; Franco et al. 2016). Heterodimerisation creates new functional signalling

units characteristically distinct to their monomeric versions enabling more complex signalling responses (e.g., sweet and umami taste receptors) (G. Q. Zhao et al. 2003; Franco et al. 2007; Bulenger et al. 2005).

1.2.1.2 Receptor tyrosine kinases (RTKs)

Receptor tyrosine kinases (RTKs) are single transmembrane-spanning receptors. Humans express 58 known RTKs originating from 20 subfamilies (Robinson et al. 2000; Blume-Jensen & Hunter 2001). They play crucial roles in the regulation of a broad range of cellular functions, such as proliferation, differentiation, cell migration, survival, and metabolism (Ullrich & Schlessinger 1990; Blume-Jensen & Hunter 2001; Lemmon & Schlessinger 2010). Derailed RTK signalling has implications in several diseases, such as cancer, inflammation, diabetes, and angiogenesis (Lemmon & Schlessinger 2010).

All RTKs have a similar modular architecture, with an extracellular domain for cognate ligand recognition, a single transmembrane domain (TMD), and an intracellular cytoplasmic domain which is comprised of a juxtamembrane regulatory region, the protein tyrosine kinase domain, and a C-terminal domain (Figure 1.7) (Lemmon & Schlessinger 2010).

RTKs share a common signalling mechanism upon cognate extracellular growth factor ligand recognition: homo-, hetero-, and/or oligomerization (Ullrich & Schlessinger 1990; Lemmon & Schlessinger 2010). Some subfamilies, however, rely to a greater extent on agonist-mediated dimerisation than others (Lemmon & Schlessinger 2010; Maruyama 2014; Maruyama 2015). For example, vascular epidermal growth factor receptor 2 (VEGFR-2) depends on bivalent agonist-facilitated dimerisation for gaining signalling competency (Figure 1.7) (Wiesmann et al. 1997; Ruch et al. 2007; Dosch & Ballmer-Hofer 2010). Other RTKs exclusively dependent on dimerisation include the cell factor receptor KIT (H. Liu et al. 2007), the nerve growth factor (NGF)/neurotrophin

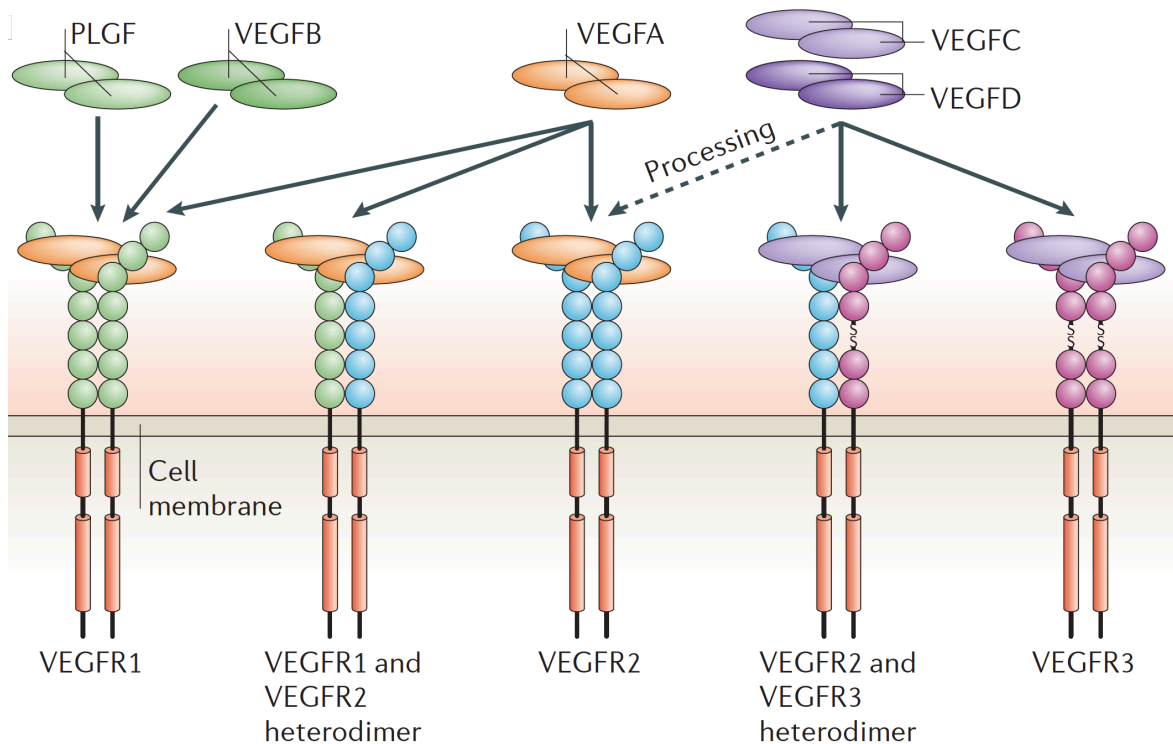


Figure 1.7 | A representative of RTKs: vascular endothelial growth factor receptors (VEGFRs). There are three VEGF receptor tyrosine kinases that bind to mammalian vascular endothelial growth factors (VEGFs), such as dimeric VEGF isoforms (A, B, C, D) and placenta growth factor (PLGF). Cognate agonist binding leads to homo- and/or heterodimer formation. These single transmembrane-spanning receptors share a modular structure, comprised of an immunoglobulin homology repeat-constituted extracellular domain for cognate ligand recognition (green, blue and purple ball structures in the cartoon), a single transmembrane domain (TMD), and an intracellular cytoplasmic domain which is comprised of 2 tyrosine kinase domains (orange rods) separated by a kinase insert and an additional C-terminal regulatory region (black line). VEGFR3 contains a disulfide bridge in the ectodomain (S-S). The dotted line indicates proteolytic processing of VEGFC and D prior to receptor binding. Reprinted by permission from Macmillan Publishers Ltd: [Nat Rev Mol Cell Biol.] (Olsson, A.-K. et al. 2006. VEGF receptor signalling - in control of vascular function. Nature Reviews Molecular Cell Biology, 7(5), pp.359–371.), copyright (2006). Licence #4246520172536.

receptor TrkA (Wiesmann et al. 1999), Axl (Sasaki et al. 2006), Tie2 (Barton et al. 2006), and ephrin receptors (Lackmann et al. 1998). In contrast, the insulin receptor, a disulfide-bonded dimer, undergoes conformational changes upon insulin binding (Ward et al. 2007). The epidermal growth factor receptor (EGFR, also known as ErbB1 or HER1) is proposed to be preassembled as oligomers before epidermal growth factor engagement (Gadella & Jovin 1995; Garrett et al. 2002; Ogiso et al. 2002; Clayton et al. 2005; Ward et al. 2007).

Irrespective of the receptor configuration prior to agonist binding, ligand engagement overcomes various modes of *cis*-autoinhibition in the tyrosine kinase domains, prompting intracellular receptor *trans*-phosphorylation (Nolen et al. 2004; Hubbard 2004; Lemmon & Schlessinger 2010). Subsequently, newly created phosphorylated residues serve as binding sites for the recruitment of Src homology-2 (SH2) and phosphotyrosine-binding (PTB) domain-containing cytoplasmic signalling proteins (Schlessinger & Lemmon 2003; Pawson 2004). One such example is the growth factor receptor-bound protein 2 (Grb2) that functions post-phosphorylation as docking site for successive protein recruitment (e.g., Grb2-associated binder (Gab1)) to eventually form a multiprotein complex. These complexes then trigger complex signalling pathways (Lemmon & Schlessinger 2010).

Vascular endothelial growth factor receptors (VEGFRs) comprise one of the 20 subfamilies of RTKs. Under normal physiological conditions, they are important regulators of angiogenesis in development and adulthood, and are primarily expressed on endothelial cells (Olsson et al. 2006; Simons et al. 2016). However, VEGFRs and their ligands (VEGFs) also play a crucial role in the tumour microenvironment, favouring tumour growth and aiding the spread of metastasis. Consequently, this signalling pathway has become an important target for therapeutic intervention in various forms of cancer, including metastatic colorectal cancer, non-small-cell lung cancer, and

ovarian cancer (Masoumi Moghaddam et al. 2012; Monk et al. 2016; Graybill et al. 2015; Ferrara & Adamis 2016).

VEGFRs comprise three closely related receptors that contain 7 immunoglobulin (Ig)-homology domain repeats in their ectodomain. In VEGFR3, one of the Ig domains is substituted for a disulfide bridge (Figure 1.7) (Olsson et al. 2006; Simons et al. 2016). The Ig 2 and 3 domains facilitate binding to dimeric VEGF isoforms (A, B, C, D) and placenta growth factor (PLGF) (Muller et al. 1997; Fuh et al. 1998; Christinger et al. 2004; Olsson et al. 2006). On the intracellular side, following a juxtamembrane domain, VEGFRs contain two tyrosine kinase domains which are separated by a short kinase insert, and finally a C-terminal tail (Figure 1.7) (Olsson et al. 2006; Simons et al. 2016).

Agonist binding facilitates receptor homo- or heterodimerisation, in particular between VEGFR1-VEGFR2 and VEGFR2-VEGFR3 (Dixelius et al. 2003; Nilsson et al. 2010; Harris et al. 2013; Yao et al. 2014; Y. Deng et al. 2015). Agonist-mediated receptor dimerisation activates canonical RTK signalling (Dixelius et al. 2003; Simons et al. 2016). Dimerised receptors auto-phosphorylate each other thereby creating binding sites for signal transducers and adapter proteins that trigger several signalling pathways, such as PLCγ–ERK1/2 pathway, the PI3K–AKT–mTOR pathway, or the SRC and small GTPases (Olsson et al. 2006; Simons et al. 2016). VEGFRs and VEGFs can also bind to co-receptors, such as heparan sulphate proteoglycans and neuropilins, providing an additional layer of canonical signal modulation (Grünewald et al. 2010; Simons et al. 2016).

Non-canonical activation of VEGFR2 can be achieved by non-VEGF ligands or non-VEGF-dependent factors, such as mechanical force (Z.-G. Jin et al. 2003; Tzima et al. 2005; Simons et al. 2016).

Similar to GPCR desensitisation and resensitisation, VEGFR2 is internalised by clathrin-coated pits and transported to early endosomes. From there, receptors are

either marked for degradation in lysosomes or recycled and transported back to the cell membrane (Simons et al. 2016).

1.2.2 Prosthetic gene networks

Synthetic biology advances in mammalian cells have generated complex synthetic networks that address pressing medical challenges in the treatment of metabolic and autoimmune diseases. Circuit designs rely on either engineered TFs that specifically detect and respond to endogenous, disease-associated diffusible biomarkers (Rössger, Charpin-El Hamri & Fussenegger 2013a; Kemmer et al. 2010) or native transmembrane receptors that propagate endogenous and/or exogenous stimuli through an engineered signalling cascade eventually leading to a therapeutically-relevant transcriptional response (Ye, Charpin-El Hamri, et al. 2013; Stanley et al. 2012; Stanley et al. 2014; Ye et al. 2011; Rössger, Charpin-El Hamri & Fussenegger 2013b; D. Ausländer et al. 2014; Bai et al. 2016; M. Xie et al. 2016; Ye et al. 2017; Haellman & Fussenegger 2017).

Genetically engineered cells that contain complex synthetic gene circuits can be deployed to alleviate metabolic disorders *in vivo* (Folcher & Fussenegger 2012; Bacchus et al. 2013; Heng et al. 2015). These transgenic cells can be encapsulated in biocompatible and semi-permeable packaging material (e.g., alginate-poly-L-lysine) that allows diffusion of small molecules and proteins, and at the same time hampers adverse host immune responses and potential leakage of genetic information (Drury & Mooney 2003). By fulfilling therapeutically-relevant tasks that are otherwise compromised in the host (i.e. a prosthetic gene network), designer cells respond to particular disease-relevant metabolites by secreting diagnostic, preventive, or therapeutic effector proteins (Folcher & Fussenegger 2012; S. Ausländer et al. 2012; Bacchus et al. 2013; Heng et al. 2014; Heng et al. 2015).

A powerful example is the use of prosthetic gene networks designed to regulate glucose homeostasis in mice. These include: (i) engineered HEK293 cells containing a synthetic, light-inducible transgene circuit for releasing glucagon-like peptide 1 (GLP1, attenuates hyperglycaemia) upon transdermal light stimulation (Ye et al. 2011; Shao et al. 2017); (ii) HEK293T cells engineered to express the temperature-sensitive calcium channel TRPV1 and elicit insulin transgene expression when exposed to low frequency radio waves (Stanley et al. 2012; Stanley et al. 2014); (iii) cells equipped with a synthetic signalling cascade that, upon being challenged with the anti-hypertension drug guanabenz, trigger the secretion of metabolically active peptides GLP1 and leptin (attenuates obesity) (Ye, Charpin-El Hamri, et al. 2013).

A more advanced class of prosthetic gene networks entail self-sufficient and autonomous close-looped systems, whereby an endogenous disease biomarker (input) automatically triggers the secretion of a therapeutic protein (output) (Heng et al. 2015). Martin Fussenegger's laboratory published the first implementation of a closed-loop prosthetic gene network that comprises engineered cells that sense high levels of uric acid and secrete urate oxidase, thereby controlling hyperuricemia (Kemmer et al. 2010). In another study using a diet-induced obesity mouse model, human fibrosarcoma cells were equipped with a chimeric nuclear lipid-sensing receptor (Rössger, Charpin-El Hamri & Fussenegger 2013a). This chimeric receptor initiates the expression of the appetite-suppressing peptide hormone pramlintide in presence of elevated blood fat levels, but only in absence of the dietary stimulus phloretin (Rössger, Charpin-El Hamri & Fussenegger 2013a).

Type 1 diabetes is characterised by a lack of insulin-producing pancreatic β -cells which eventually leads to greatly increased acidic ketone bodies in the liver that impair the CO_2 -bicarbonate pH-buffering in the bloodstream (Gosmanov et al. 2014). The increased amount of protons can be detected by engineered HEK293T cells that

express the pH-sensing GPCR TDAG8 (T-cell death-associated gene 8) and respond by secreting a transgenic insulin (furin-cleavable proinsulin I) thereby restoring glucose homeostasis (D. Ausländer et al. 2014).

Reports suggest that obesity-induced insulin resistance can be ameliorated by administering adiponectin which has insulin-sensitising properties. (H. Gao et al. 2013; Ziemke & Mantzoros 2010). Based on this observation, the Fussenegger laboratory configured a circuit in which the ectopically overexpressed human insulin receptor detects pathogenically high levels of insulin (Ye et al. 2017). This signal is transduced through the endogenous mitogen activated protein kinase (MAPK) signalling cascade culminating in the activation of a chimeric transcription factor which leads to the expression of transgenic adiponectin. Delivery of this genetically installed prosthetic gene network in microencapsulated HEK293 cells successfully attenuated insulin resistance (hyperinsulinaemia) in a mouse model (Ye et al. 2017).

Diabetic patients suffer from chronic hyperglycemia (elevated blood glucose levels) caused from either complete loss of insulin-secreting β -cells (type 1 diabetes, T1D) or progressive loss of β -cell numbers and/or their secretory capacity (type 2 diabetes, T2D). In the latest implementation of a closed-loop prosthetic gene network, the Fussenegger laboratory reported the first autonomously-operating therapeutic programme that directly senses glucose levels in the bloodstream (M. Xie et al. 2016). At the core of this synthetic programme is the ectopically expressed, voltage-gated Ca^{2+} channel. Glucose metabolism generates a large amount of ATP which triggers the closure of ATP-sensitive K^+ channels and concomitantly depolarises the membrane allowing Ca^{2+} channels to open (MacDonald et al. 2005). Ca^{2+} influx results in the phosphorylation of the nuclear factor of activated T cells (NFAT). Phosphorylated NFAT translocates into the nucleus and drives the expression of transgenes carrying a NFAT-responsive promoter (M. Xie et al. 2016). In T1D mouse models, microencapsulated

designer cells inducibly expressed an engineered insulin as a transgene, restoring glucose homeostasis. Importantly, mice did not experience too high blood glucose (hyperglycaemic) as well as too low blood glucose (hypoglycaemic) levels (M. Xie et al. 2016). These mimetic β -cells performed in a head-to-head comparison better than implanted immortalised β -cell lines and human islets (M. Xie et al. 2016).

In principle, prosthetic gene networks embedded into designer cells could offer great opportunities compared to traditional oral or intravenous administration of therapeutic drugs. Living cells detect, process and secrete therapeutically-relevant factors in a contained and compact format compared to implantable biomedical devices that require sensors, electronics and mechanical microfluidic system components (Heng et al. 2015). Furthermore, therapeutically engineered cells produce a limitless supply of therapeutic molecules unlike biomedical devices that require refilling and proper, temperature-wise storing of the drugs (Heng et al. 2015). Finally, cells with prosthetic gene networks can exploit the host organism metabolism as power source in contrast to biomedical devices which operate on rechargeable batteries (Heng et al. 2015).

1.2.3 Synthetic receptors

In 2017, the American Food and Drug Agency (FDA) approved the first chimeric antigen receptor (CAR) T cell therapy called Kymriah (Novartis) for the treatment of B cell acute lymphoblastic leukaemia (Dolgin 2017). One course of treatment costs US\$475,000 (Yip & R. M. Webster 2018). This decision marks a milestone in the development of personalised cell-based therapies as an alternative to standard small molecule-based drugs and biologics (e.g., antibodies) (Fischbach et al. 2013; Roybal & W. A. Lim 2017). Such “living drugs” are superior to the latter because they can autonomously register inputs, process and convert them into complex response programmes (Mason et al. 2011; Fischbach et al. 2013; Ye, Aubel, et al. 2013; Kitada et al. 2018) (Chapter 1.2.2).

In the last part of the introduction, I discuss: (i) the development of CAR and synNotch immunological synthetic receptors, which recognise cell-surface bound antigens; and (ii) chimeric receptors that enable the integration of various soluble ligands, such as RASSL/DREADD, GPCR TANGO, and MESA.

1.2.3.1 Chimeric antigen receptors (CARs)

CARs are modular artificial receptors that exploit antibody-based antigen recognition to elicit a full T cell response (Kershaw et al. 2013; Gill & June 2015; C.-Y. Wu, Rupp, et al. 2015; Fesnak et al. 2016; W. A. Lim & June 2017).

A CAR is a single polypeptide composed of three main parts: an extracellular target binding module, a transmembrane domain (TMD), and various intracellular signalling modules (Figure 1.8). The engineered extracellular antigen-recognition domain is composed of a single chain variable fragment (scFv). scFvs are derived from monoclonal antibodies and confer selectivity and specificity of CARs to a particular, pre-defined antigen. The scFv is linked via a flexible hinge region to a TMD that facilitates cell membrane integration. The CAR TMD is based primarily on the TMDs of CD8 or CD28 co-receptors. In the most basic CAR configuration, the intracellular domain consists of the T cell receptor (TCR)-originating CD3 zeta chain (CD3 ζ) for T cell activation (Figure 1.8) (Kershaw et al. 2013; Srivastava & Riddell 2015; Z. L. Chang & Y. Y. Chen 2017; Sadelain et al. 2017).

At the end of the 1980s, the first CARs were developed in an attempt to genetically redirect T cells to any cancer-associated antigen in a non-major histocompatibility complex (MHC)-restricted manner (Kuwana et al. 1987; Gross et al. 1989; Roberts et al. 1994; Altenschmidt et al. 1996). Generally, native T cells recognise antigens on antigen presenting cells (APC, comprising all nucleated cells in the body)

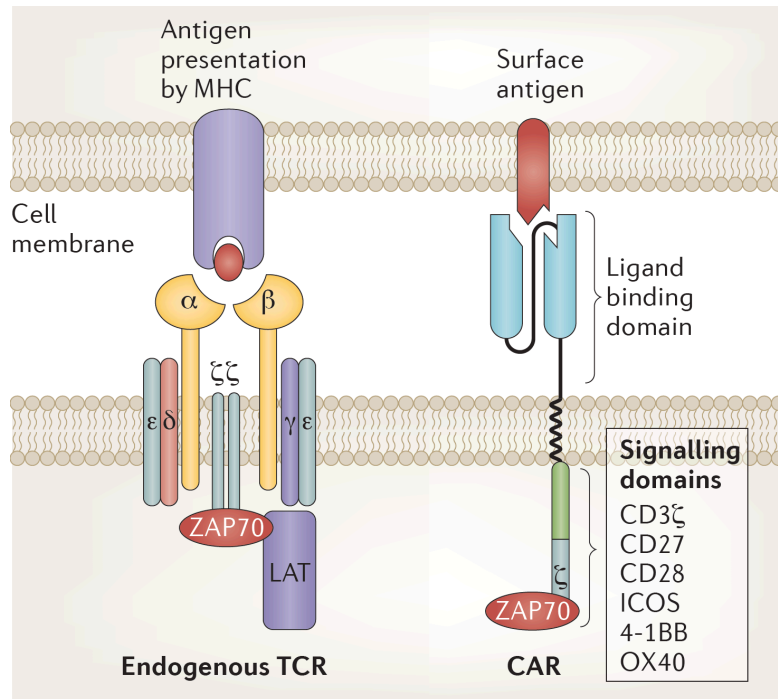


Figure 1.8 | Chimeric antigen receptor (CAR) design inspired by native T cell receptor (TCR) architecture. The endogenous, native TCR consists of a paired α and β chain that engage with an antigen presented by the major histocompatibility complex (MHC). The TCR is associated with additional chains comprising δ , ϵ , γ , and the signalling chain CD3 ζ (left). The activated TCR contributes signal 1 to T cell activation. Signal 2, the co-stimulatory signal, is provided *in trans* by activating co-stimulatory receptor domains, such as CD27, CD28, ICOS, 4-1BB, OX40, or by inhibitory co-stimulatory receptor domains, such as programmed cell death protein 1 (PD1) (not shown). Only signal 1 and activating signal 2 lead to potent T cell effector function. Inspired by the functional extracellular domains and intracellular signalling domain requirements, chimeric antigen receptors (CARs) were developed. CARs possess for antigen engagement on the extracellular side single-chain variable fragments (scFvs) derived from immunoglobulin (Ig) heavy chain variable and light chain variable domains (right). Downstream of the scFV is a hinge domain, transmembrane domain (TMD) and intracellular signalling domains. Second generation CARs contain two signalling domains, typically co-stimulatory domain 4-1BB and CD3 ζ *in cis*. Third generation CARs contain additional co-stimulatory domains (see box). Antigen-bound single CAR polypeptide elicits full T cell effector function. ICOS, inducible T cell co-stimulator; LAT, linker for activation of T cell family member 1; ZAP70, ζ -associated protein of 70 kDa. Reprinted by permission from Macmillan Publishers Ltd: [Nat Rev Cancer] and adopted from Fesnak, A.D., June, C.H. & Levine, B.L., 2016. Engineered T cells: the promise and challenges of cancer immunotherapy. Nature Reviews Cancer, 16(9), pp.566–581., copyright (2016). Licence #4245341241867.

via their endogenous TCR. In particular, the TCR binds tightly to an antigen that is presented in complex with MHC class I or II on the cell surface of a target cell (Figure 1.8). Ligand engagement by the ectodomain of the TCR leads to the phosphorylation of intracellular ITAM motifs (immunoreceptor tyrosine-based activation) of the CD3 ζ domain. These phosphorylations are specifically recognised by SH2 domain effector proteins (e.g., ZAP70 kinase) (Smith-Garvin et al. 2009; Chakraborty & Weiss 2014), providing initiating signal 1. To fully unfold T cell activation and expansion of tumour antigen-specific T cells, a second signal is required, which is obtained by the engagement of co-stimulatory T cell receptors (Fesnak et al. 2016). Inherently, cancer and tumour cells are recognised by T cells as “self” and therefore are “invisible” to the body’s immune system. This process is exacerbated by tumour cells actively down-regulating antigen presentation via MHC classes (Restifo et al. 1996).

Inspired by the TCR architecture, iterative engineering efforts led to three generations of CARs that eventually displayed desirable therapeutic properties in transgenic T cells in the fight of cancer and autoimmune diseases (Figure 1.8) (Kochenderfer & Rosenberg 2013; Kalos & June 2013; Sadelain et al. 2013; Kershaw et al. 2013; Gill & June 2015; Fesnak et al. 2016; W. A. Lim & June 2017; Z. L. Chang & Y. Y. Chen 2017).

In the process referred to as adoptive T cell therapy, T cells are isolated from blood, then purified, stimulated with antigens, and genetically modified with CARs *ex vivo*. After expansion *in vitro* and quality control, engineered T cells are re-infused into patients (Rosenberg & Restifo 2015; Fesnak et al. 2016; W. A. Lim & June 2017). T cells can either originate for the same patient (autologous) or from a donor (allogeneic).

Most notably, anti-CD19 CAR T cells have been shown to be especially effective in eliminating chemotherapy-resistant B cell leukaemia and lymphomas (response rates of 70%-90%) (Porter et al. 2011; Kalos et al. 2011; Kochenderfer et al. 2012; Davila et

al. 2014; Brentjens et al. 2013; Maude et al. 2014; Turtle et al. 2016; Kochenderfer & Rosenberg 2013; Maus et al. 2014; Gill & June 2015; Schuster et al. 2017; Neelapu et al. 2017). This is due to the fact that CD19, which drives T cell activation in CD19-CAR T cells, is exclusively and ubiquitously found in high abundance on most B cell malignancies and in the tissue of the B cell lineage but absent in other tissues (Y. S. Li et al. 1996). Since depleting all B cells is compatible with survival, CD19-CAR T cells can be effectively used to eradicate B cell malignancies (Gross & Eshhar 2016).

Immunological consideration and pitfalls

Although CARs hold great promise for next-generation therapeutic devices, there are still limitations associated with artificial receptors. A general problem in receptor engineering is the incorporation of heterologous polypeptide sequences. The scFv domains of CARs are derived from mice (Kershaw et al. 2006; Lamers et al. 2011), which pose a substantial risk of triggering adverse humoral and cellular immune responses in patients. A solution has been provided by the use of humanized (Westwood et al. 2005; Y. Zhao et al. 2009; Sun et al. 2014; Johnson et al. 2015) or even fully human scFvs (Morgan et al. 2012; Lanitis et al. 2012; Song et al. 2015; Sommermeyer et al. 2017).

Other complications related to CAR T cells include: (i) on-target, off-tumour responses (detecting antigens on bystander tissue, non-tumour cells); (ii) cytokine storm/cytokine release syndrome (CRS, release of proinflammatory cytokines on a systemic level, a function of large quantities of therapeutic T cells); (iii) tumour lysis syndrome (TLS, immune response as a consequence of rapidly eliminating large quantities of tumour cells) (Kalos & June 2013; W. A. Lim & June 2017).

Next-generation CAR designs: OFF switches and Boolean logic gates

Safety concerns associated with CAR T cells have been addressed by incorporating safety mechanisms, that either act as OFF switches (reactive approach,

endogenously/exogenously inactivated) or as ON switches (preventive approach, exogenously activated). In general, most of these concepts have been or can be applied to other synthetic receptors, highlighting their importance in synthetic receptor engineering (Kitada et al. 2018).

An OFF switch for a CAR T cell can be installed by integrating a suicide gene that converts an exogenously-administered prodrug into a cytotoxic agent thereby sacrificing transgenic T cells (Lupo-Stanghellini et al. 2010). Transducing T cells with the herpes simplex virus thymidine kinase (HSV-TK) gene renders them susceptible to ganciclovir administration (Bordignon et al. 1995; Tiberghien et al. 2001; Markt et al. 2003; Ciceri et al. 2009). Similarly, the engineered variant of human thymidylate kinase (tmpk) provides a selection opportunity in the presence of the small molecule drug 3-azido-3-deoxythymidine (AZT) (Sato et al. 2007).

Another approach is provided by the iCasp9 system, whereby T cells co-express a caspase 9-rapamycin-based dimerisation domain fusion protein. Upon administration of a chemical inducer, homodimerisation of iCasp9 is induced which triggers an endogenous apoptosis programme (Tey et al. 2007; Di Stasi et al. 2011).

Another layer of control can be achieved by labelling the engineered T cells with pre-defined cell-surface markers, which can be targeted by therapeutic antibodies leading to specific depletion of engineered T cell *in vivo*. Examples include: (i) co-expression of the CD20 B-cell antigen (Serafini et al. 2004; Griffioen et al. 2009; Vogler et al. 2010), (ii) a chimera of CD20 and CD34 antigens (RQR8) (Philip et al. 2014); (iii) a truncated version of human EGFR polypeptide (huEGFRt) (Xiuli Wang et al. 2011); (iv) engrafting a short c-myc tag into the TCR sequence (Kieback et al. 2008).

Some of the most powerful and robust ON switch systems described so far in engineered T cells are based on Boolean logic AND gates. Logic gates are part of Boolean logical computation in which only the correct combination of inputs (e.g.,

antigens, externally delivered small molecules) yield a predictable response (output). Specifically, an AND gate requires the presence of both antigen A and antigen B to be functional, performing an “A AND B” signal integration. When only antigen A or only antigen B is present, the requirements for the Boolean AND gate are not fulfilled, and the designer T cells remain inactivate. Such an AND gate configuration offers the opportunity to gain reversible temporal and spatial control over the designer cells by locally or systemically administering pharmacologically appropriate small molecules that exert dose-specific responses (Roybal & W. A. Lim 2017). Several successfully implemented strategies have been reported for creating AND gates in transgenic CAR T cells (Wilkie et al. 2012; Lanitis et al. 2013; Kloss et al. 2013; Roybal, Rupp, et al. 2016; Han et al. 2017).

For example, Wendell Lim’s laboratory conceived an alternative CAR design that requires a second input, an exogenously-delivered small molecule, to allow signal integration (C.-Y. Wu, Roybal, et al. 2015). This ON switch CAR was created by splitting and rearranging the original CAR into two fragments. One fragment contains the antigen-detecting scFv recognition domain that is linked via a TMD to a costimulatory domain and one half of a small molecule-dependent heterodimerisation domain (antigen-binding component). The other part of the receptor is membrane-tethered and contains the respective other half of the heterodimerisation domain as well as the T cell receptor CD3 ζ domain. In this instance, the receptor is only active –in a dose-dependent manner– in the presence of the small molecule that facilitates conditional association of the heterodimerisation domains to reconstitute a full length CAR that upon antigen engagement signals through its CD3 ζ domain.

As discussed so far, T cells have been subjected to extensive engineering efforts. Great success has been achieved in fighting B cell malignancies primarily due to the ease of access to B cells in the bloodstream and their exclusive expression of CD19. Expanding

the therapeutic potential of re-directed T cells to solid tumours poses a number of challenges owing to the complexity and pathologically diverse hallmarks of cancer (Hanahan & Weinberg 2000; Hanahan & Weinberg 2011). Currently, CAR T cells have been shown poor efficiency against solid tumours. This is due to multiple reasons, including (i) the immunosuppressive tumour microenvironment that limits T cell performance by lowering their efficacy and persistence; (ii) lower susceptibility of solid tumours to the effects of T cell-mediated cytotoxicity; and (iii) an inherent scarcity of a tumour/tissue-specific antigen profile (Motz & Coukos 2011; Sadelain et al. 2013; Joyce & Fearon 2015; W. A. Lim & June 2017; E. Zhang et al. 2018). To overcome these limitations, novel classes of CARs or alternative synthetic receptor scaffolds will have to be conceived, which take into consideration particular characteristics of different cancer types and integrate a broader portfolio of disease-specific biomarkers (Teng et al. 2015; Yuedi Wang et al. 2017).

1.2.3.2 Synthetic Notch receptor (synNotch)

Endogenous Notch is a single-pass transmembrane receptor consisting of 29 to 36 extracellular EGF-like repeats that mediate Delta ligand binding, a TMD, and an intracellular domain that functions as a TF (NICD) (Figure 1.9a) (Artavanis-Tsakonas et al. 1999; Kopan & Ilagan 2009). Notch signalling underlies a direct and straightforward signal integration pathway without additional mediators or second messengers for signal amplification (Kopan & Ilagan 2009). The endogenous Notch receptor recognises its cognate ligand Delta on the surface of other cells. This engagement generates mechanical force which triggers intramembrane proteolysis by metalloproteases and γ -secretases thereby releasing NICD which translocates into the nucleus to exert its transcriptional activation on target genes (Kopan & Ilagan 2009).

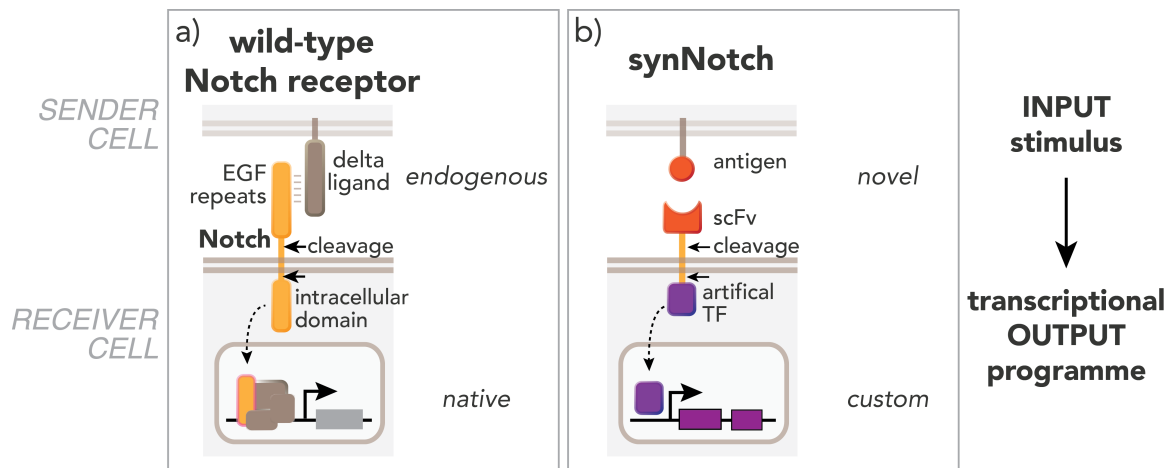


Figure 1.9 | Synthetic Notch receptor (synNotch) design inspired by native Notch receptor architecture. (a) The extracellular EGF repeats of endogenous Notch receptor recognise the ligand Delta on an opposing cell (sender cell). This ligand–receptor engagement triggers the intramembrane cleavage of Notch thereby releasing the Notch intracellular signalling domain (NICD). Liberated NICD translocates into the nucleus to induce endogenous target gene expression. **(b)** The synthetic Notch receptor (synNotch) describes a modular receptor architecture in which the extra- and intracellular domains can be exchanged for alternative domains. scFvs as novel extracellular recognition domains can recognise a particular antigen on a sender cell. Upon ligand engagement, the native minimal Notch transmembrane core undergoes proteolytic cleavage, thereby liberating the intracellularly coupled artificial transcription factor (TF), such as tTA or Gal4-VP64. The TF induces user-defined pre-integrated transgene expression in the nucleus. Adapted from “Engineering Customized Cell Sensing and Response Behaviors Using Synthetic Notch Receptors”. *Cell*, 164(4), pp.780–791, Morsut, L. et al. Copyright 2016, with permission from Elsevier. Licence #4245390566951.

The modularity, compact design and straight-forward signalling displayed by Notch receptors fostered implementations of novel synthetic receptor architectures. In the first iteration, the NICD was exchanged by Gal4-VP16. Studies in *Drosophila* (Struhl & Adachi 1998) and in mammalian cells (Sprinzak et al. 2010) revealed that upon Delta ligand engagement, Gal4-VP16 in this chimeric Notch receptor could translocate into the nucleus and drive transgene expression. To take full advantage of the modularity presented by Notch, both the input and output domains were subsequently modified. One such example is the Notch-based synthetic receptor (synNQ).

synNQ consists of an anti-GFP-binding single domain antibody (nanobody) for ligand-binding and an artificial activating TF (QF, originating from *Neurospora crassa qa*) which replaces the NICD. In this case, the cognate ligand for the nanobody is a membrane-tethered artificial GFP ligand provided *in trans*. This receptor was utilised to visualise cell-cell interactions in cultured *Drosophila* cells (Gordon et al. 2015) and between various tissues in *Drosophila in vivo* (He et al. 2017). Similarly, another group reported the construction of a chimeric Notch receptor (SNTGV) composed of an anti-mouse CD19 single-chain antibody domain (SCAD), Notch TMD, and Gal4-VP16 as TF. This receptor was used to visualise cell-cell interactions between neurons and glia cells in transgenic *Drosophila in vivo* (Huang et al. 2016).

In three seminal publications in 2016, the Wendell Lim laboratory established synthetic Notch receptors (synNotch) as a novel means for adoptive immunotherapy (Figure 1.9b) (Morsut et al. 2016; Roybal, Rupp, et al. 2016; Roybal, J. Z. Williams, et al. 2016). The synNotch receptor encompasses a modular structure of different scFvs or nanobodies as extracellular domains and various artificial transcription factors as endodomains (e.g., Gal4-VP64, tTA, Gal4-KRAB). The only constant part in all modular combinations is the native minimal Notch transmembrane core, which is indispensable for proteolysis. scFvs bind a variety of membrane-tethered endogenous and synthetic

ligands as input stimuli while the artificial TFs enable the transcription of custom-integrated transgenes as engineered output responses (Morsut et al. 2016) (Figure 1.9b).

Morsut et al. (2016) demonstrated (i) that synNotch is functional in many different immortalised mammalian cell lines, neurons, and primary T cells; (ii) that several synNotch signalling pathways can function independently to one another; and (iii) combinatorial input/output signalling (providing orthogonality) exemplified by a 2-input synNotch AND gate logic.

Roybal, Williams et al. (2016) showed that synNotch could be employed to construct T cells with novel therapeutically relevant properties. In this case, activated synNotch was used to drive customised cytokine programmes, express therapeutic agents (e.g., antibodies), and induce master regulator transcription factors to change T cell fates (Roybal, J. Z. Williams, et al. 2016).

The third paper by Roybal, Rupp et al. (2016) focused on serially connected synNotch/CAR AND gate circuits. The resulting designer T cells require synNotch receptor activation via antigen A to induce expression of a CAR that can then bind to antigen B, when presented. These synNotch/CAR T cells were active only in the presence of both antigens, which guaranteed confident discrimination between single antigen presenting bystander tumour cells and dual antigen presenting target tumour cells in a mouse model *in vivo* (Roybal, Rupp, et al. 2016).

However, the synNotch T cell input factors are inherently limited to cell-surface markers (CD19, GFP ligand) on target cells and the outputs generated by these engineered T cells originate from previously integrated transgenes or, in case of second input CARs, the endogenous T cell response programme. To overcome cell-surface restricted antigen recognition and further refine target specificity (bystander vs target cells), a new class of receptors is needed that can integrate a number of other tumour-specific

microenvironment markers such as expression profiles of chemokines, cytokines, and tumour-specific signalling molecules.

1.2.3.3 Engineered G-protein-coupled receptors

Since GPCRs can integrate the broadest variety of endogenous and exogenous ligands (Chapter 1.2.1.1), they embody an ideal receptor family for engineering complex synthetic gene networks (Heng et al. 2014). Much research has been dedicated to exploiting the input versatility of GPCRs in designing transcriptional programmes initiated by their agonists.

In the following subchapters, I first discuss input-modified GPCRs. These include receptor activated solely by a synthetic ligand (RASSL) and designer receptor exclusively activated by designer drugs (DREADDs), which are engineered GPCRs that lost binding specificity to their cognate ligands and instead are activated by synthetic compounds (Pei et al. 2008; Conklin et al. 2008; Jürgen Wess et al. 2013; D. J. Urban & Roth 2015). Second, I focus on output-modified GPCRs that integrate receptor stimulation via an artificial non-programmable TF (Barnea et al. 2008; Kroeze et al. 2015) (Figure 1.10).

1.2.3.3.1 Receptors Activated Solely by a Synthetic Ligand (RASSLs) and Designer Receptor Exclusively Activated by Designer Drugs (DREADDs)

RASSLs are synthetic GPCRs developed to possess novel ligand specificities. Instead of binding to their native cognate ligand, RASSLs display responsiveness to inert, drug-like small molecules while retaining specific G-protein-mediated signalling (Figure 1.10a,b) (Conklin et al. 2008).

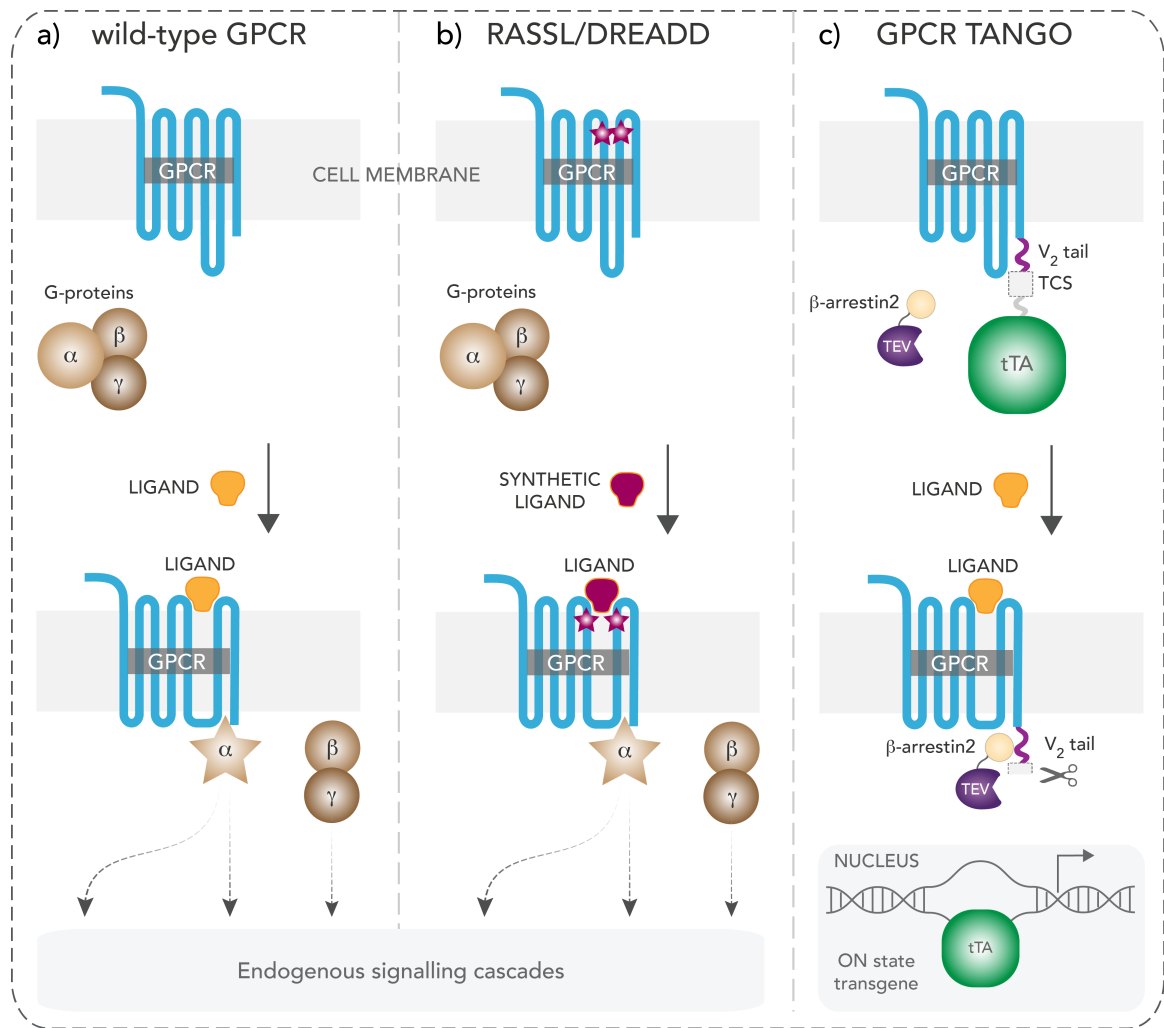


Figure 1.10 | Engineering GPCR signalling capabilities. (a) Wild-type GPCRs recognise their cognate ligand and trigger an endogenously pre-defined G-protein and G-protein-independent signalling cascade. $G\alpha$ (star shape) and $G\beta\gamma$ separated from each other symbolise G-protein signalling pathway activation. (b) Receptor activated solely by a synthetic ligand (RASSL) and Designer receptor exclusively activated by designer drugs (DREADDs) GPCR contain mutations in the GPCR scaffold (red stars) conferring new ligand binding selectivity while retaining endogenous signalling pathway activation. (c) Implementation of an artificial G-protein independent signalling pathway by utilising a synthetic TF. GPCR TANGO architecture describes a GPCR that is linked to a V_2 tail (originating from the vasopressin receptor), a TEV cleavage site (TCS), and tTA. Upon ligand engagement, the artificial fusion protein β -arrestin2-TEV is actively recruited to the phosphorylated V_2 tail thereby prompting TEV protease-mediated TCS cleavage. Liberated tTA then translocates into the nucleus to induce transgene expression.

RASSLs were created by site-directed mutagenesis and/or by replacing existing loop domains in the GPCR with domains from other GPCRs (Coward et al. 1998). One of the earliest examples of RASSLs is based on the human $G\alpha_i$ -coupled κ opioid receptor (human RASSL opioid $G\alpha_i$ -coupled receptor, hRO-i, old name Ro1). hRO-i and derivatives thereof exhibit an up to three orders of magnitude reduction in activation capability by their cognate ligand dynorphin (Coward et al. 1998). Instead, they are robustly activated by the small molecule drug spiradoline (Coward et al. 1998). hRO-i was functionally validated in a mouse model achieving inducible $G\alpha_i$ signalling in cardiomyocytes (Redfern et al. 1999). A second example is the hRS-s, a human RASSL based on the serotonin 4 receptor that signals through $G\alpha_s$. Unlike the wildtype serotonin 4 receptor which responds to serotonin (5-HT), the hRS-s is exclusively activated by small synthetic compounds due to a single amino acid mutation in the transmembrane domain (Claeysen et al. 2003; W. C. Chang et al. 2007; Hsiao et al. 2008).

Generally, first-generation RASSLs displayed compromised functionality *in vivo* because of: (i) low binding affinity to synthetic ligands; (ii) remaining residual binding affinity to endogenous ligands; (iii) displaying constitutive basal activation (leakiness); (iv) synthetic ligands activating other endogenous receptor signalling (off-target binding) (Conklin et al. 2008).

These limitations were overcome by utilising directed molecular evolution to design a second generation of RASSLs, which were called designer receptors exclusively activated by designer drugs (DREADDs) (Armbruster et al. 2007; S. Dong et al. 2010) (Figure 1.10b).

Prototype DREADDs are the clozapine-N-oxide (CNO)-sensitive designer GPCRs based on the human M3 muscarinic receptor family (class A GPCRs) (Jürgen Wess 2016). The unaltered, endogenous muscarinic receptors bind to their agonist

acetylcholine whereas the engineered hM3 DREADDs (hM3D) are activated exclusively by the inert drug CNO without compromising canonical signal transduction pathways (Armbruster et al. 2007). The initial hM3Dq retained the native signalling properties of the endogenous receptor, signalling through $G\alpha_q$ -type G proteins and β -arrestin (Armbruster et al. 2007; Nakajima & Juergen Wess 2012). Other variants of hM3Dq can engage in $G\alpha_i$ -mediated signalling (Armbruster et al. 2007), $G\alpha_s$ -coupled signalling (Guettier et al. 2009), signal exclusively via $G\alpha_q$ (J. Hu et al. 2016), or initiate signalling primarily through β -arrestin (Nakajima & Juergen Wess 2012).

hM3Dq and hM4Di were also employed to remotely excite and inactivate neurons, respectively, in a transient manner *in vivo* (Alexander et al. 2009; Scofield et al. 2015; Mahler et al. 2014; Stachniak et al. 2014). Specific hM3Dq (hM3D-Gq) overexpression in mouse pancreatic β -cells in combination with sustained CNO treatment improved β -cell proliferation and insulin production counteracting drug-induced diabetes in mice (Guettier et al. 2009; Jain et al. 2013). These studies not only showcased the potential for engineering synthetic signal pathways but also illustrated potential therapeutic applications for DREADDs.

Interestingly, CNO was later found to be metabolised in humans into the pharmacologically active compound clozapine which triggers a plethora of undesirable responses thereby compromising the specificity of DREADDs (Löffler et al. 2012; Roth et al. 2004). However, an alternative truly inert compound, termed “Compound 21”, was later identified (Xin Chen et al. 2015).

With the development of a second DREADD class, not only a different ligand could be used to trigger receptor activation but also multiplexing two DREADDs via 2-input logic gates was achievable. The KOR DREADD (KORD) is based on the $G\alpha_i$ -coupled κ -opioid receptor (KOR) and exclusively activated by salvinorin B (Vardy et al. 2015). This enabled remote and simultaneous or sequential modulation of neuronal

behaviour *in vivo* by multiplexing inhibitory KORD with excitatory hM3Dq receptors (Vardy et al. 2015; Aldrin-Kirk et al. 2016).

1.2.3.3.2 Non-programmable TF-based GPCRs (GPCR TANGO)

Barnea et al. (2008) conceived a synthetic GPCR-based scaffold that implements an orthologous signalling pathway by exploiting endogenous transient GPCR–arrestin interactions (GPCR TANGO, Figure 1.10c).

This GPCR TANGO architecture consists of two custom components. First, the human arginine vasopressin receptor 2 (AVPR2)-originating tail (V_2 tail), a TEV protease cleavage sequence (low-affinity TCS, amino acid sequence ENLYFQL) and the artificial TF, tTA, are fused to the GPCR C-terminal cytoplasmic tail. The addition of the well characterised V_2 tail increases the affinity for β -arrestin2 recruitment to the GPCR of interest, which form a more stable complex while simultaneously maintaining the pharmacological properties of a given GPCR (Hanyaloglu et al. 2001; Hamdan et al. 2005; K.-M. Kim & Caron 2008; Kroeze et al. 2015; Thomsen et al. 2016; Oakley et al. 2000; Vrecl et al. 2009). Moreover, the V_2 tail extension offers a generalisable scaffold for any GPCR regardless of their intrinsic β -arrestin2 affinity (Barnea et al. 2008). The low-affinity TCS (Parks et al. 1994; Kapust et al. 2002) reduces the likelihood of unspecific (e.g., agonist-independent) TEV protease-mediated cleavage events that lead to the liberation of tTA.

Second, β -arrestin2 is fused to the wild-type TEV protease (β -arrestin2-TEV, Figure 1.10c) (Parks et al. 1995; Kapust et al. 2001; Barnea et al. 2008). In this instance, the β -arrestin2-TEV fusion protein is stably integrated and constitutively expressed in HEK293 cells (HTLA cells) in contrast to the GPCR component, which can be supplied via transient transfections (Barnea et al. 2008). Ligand engagement leads to conformational changes and phosphorylations at serine and threonine residues of the

V₂ tail of the engineered GPCR (Oakley et al. 1999), which is specifically recognised by β -arrestin2-TEV. The proximity of β -arrestin2-TEV to the TCS in the agonist-activated GPCR facilitates proteolytic cleavage thereby releasing tTA. Then, tTA translocates into the nucleus where it can activate the expression of target transgenes. This process of transcriptional activation following arrestin translocation is depended on transient interactions between the two engineered protein components and referred to as GPCR TANGO (Barnea et al. 2008; Kroeze et al. 2015).

In 2015, Kroeze et al. (2015) expanded the repertoire of GPCR TANGO receptors and successfully adapted and validated a refined scaffold design for a subgroup of 135 human non-olfactory GPCRs. Barnea et al. (2008) and Kroeze et al. (2015) primarily used the TANGO assay to assist in identifying novel, endogenous ligands for orphan receptors and to more generally study receptor activation and protein-protein interactions.

Nonetheless, the GPCR TANGO system has subsequently been implemented to: (i) map the release of endogenous neurotransmitters in neurons in *Drosophila* with TANGO-map (Inagaki et al. 2012); (ii) study synaptic connections in the *Drosophila* brain with TANGO-Trace (Jagadish et al. 2014); (iii) map neural circuits in the *Drosophila* brain with *trans*-Tango (Talay et al. 2017); (iv) redirect the information flow in mammalian neurons *in vivo* via an alternative configuration, termed iTANGO2 (D. Lee et al. 2017).

iTANGO2 operates a 2-input AND gate logic in which light and cognate ligands induce gene expression in neurons to gain control over cellular behaviour. In iTANGO2, the original GPCR TANGO scaffold was extended by incorporating between the V₂ tail and the tTA one half of a split TEV protease (TEV(N)) and an engineered light-oxygen-voltage (LOV) domain that sterically covers the TCS. β -arrestin2 is fused to the second half of the split TEV(C), yielding a β -arrestin2-TEV(C) fusion protein. In the ligand-

induced state, β -arrestin2-TEV(C) is actively recruited to the V₂ tail of the activated GPCR. The proximity of split TEV protease fragments promote reconstitution of the catalytically functional full length TEV protease. However, only in the presence of a blue light stimulus, the LOV domain undergoes a conformational change, rendering the TCS accessible for cleavage by newly reconstituted TEV protease, leading to TF liberation and transgene expression (D. Lee et al. 2017).

The GPCR TANGO system provides a powerful tool to trigger a user-defined transcriptional response mediated by β -arrestin recruitment, alongside the canonical G-protein signalling pathway (Figure 1.10c) (Barnea et al. 2008; Hanson et al. 2009; Kroeze et al. 2015). This most likely creates a degree of crosstalk which may or may not elicit adverse effects, depending on each particular stimulus and cell type used.

1.2.3.4 Engineered receptor tyrosine kinases

Endogenous RTKs can integrate a broad variety of growth factors by activating complex signalling cascades (Chapter 1.2.1.2). RTK-mediated signalling has been the subject of engineering efforts at different levels: (i) ligand input specificity; (ii) identity of ectodomains; (iii) scaffold proteins for signal propagation and integration.

(i) RTKs with engineered ligand input specificity. Two research groups from the University of Tokyo pioneered the development of non-protein ligands for the hepatocyte growth factor (HGF) receptor (also known as MET receptor). These non-protein ligands, called HGF mimetics, facilitate MET dimerisation and trigger endogenous MET activation and signalling. The first reported example is artificial MET-activating dimeric macrocycles that consist of two MET-binding peptides connected through a chemical linker (Ito et al. 2015). Shinsuke Sando's laboratory further advanced this concept by creating oligonucleotide-based HGF mimetics, that are aptamers of 100-mer single-stranded DNA with nanomolar potency (Ueki et al. 2016).

In a follow-up study, they designed bispecific DNA aptamers enabling reprogramming of MET receptor ligand specificity (Ueki et al. 2017). These bispecific DNA aptamers consist of a Met receptor aptamer and a second aptamer that binds specifically to an user-defined external protein, such as platelet-derived growth factor (PDGF), which normally cannot bind to the MET receptor (Ueki et al. 2017). The Sando laboratory has showcased that genetically unaltered RTK receptors can still be externally manipulated by non-cognate ligands, gaining control over endogenous signal activation.

(ii) RTKs with engineered ectodomains. The extracellular recognition domains of RTKs can readily be replaced with LOV-sensing and dimerisation domains, to create optically controllable RTKs (Opto-RTKs) which retain native intracellular signalling capabilities (Grusch et al. 2014). This enables light-induced cellular signalling with spatial and temporal precision. Examples of Opto-RTKs include those based on the fibroblast growth factor receptor 1 (FGFR1), epidermal growth factor receptor (EGFR), and the rearranged during transfection (RET) receptor (Grusch et al. 2014).

(iii) RTKs with engineered scaffold proteins. The intracellular signalling of cell-surface receptors is crucially dependent on adaptor proteins that connect activated receptors to intracellular targets for efficient signal propagation. The combinatorial modular domain structure of signalling cascade components such as mitogen-activated protein kinase kinase kinase MAPKKK (Pawson & Nash 2003; Bhattacharyya et al. 2006) offer a great opportunity for network engineering. As an example of modularity, SH2 can be fused to other domains or proteins creating novel chimeric scaffold proteins (Schlessinger & Lemmon 2003). In one study, SH2 was fused to FADD (Fas-associated protein with death domain), which is a scaffold protein involved in the apoptosis signalling cascade. The resulting chimeric adaptor fusion redirects the information flow of stimulated RTKs towards cell death instead of growth and survival (Howard et al. 2003).

Barnea et al. (2008) designed a new orthogonal signal transduction pathway by fusing a SH2 domain or phosphotyrosine-binding (PTB) domain (Schlessinger & Lemmon 2003) to the TEV protease and covalently linking tTA to the cytoplasmic tail of an RTK via a canonical TEV cleavage site (TCS). Ligand binding induces the *trans*-phosphorylation of tyrosine residues in the intracellular domains of RTKs, which are subsequently recognised by the chimeric fusion protein SH2-TEV. Upon recruitment, TEV protease cleaves the TCS, liberating the artificial TF which in turn can induce the expression of a defined transgene. This receptor framework was termed RTK TANGO due to transient interactions between the two engineered protein components. This elegant design approach adds a novel orthogonal signalling pathway to the existing endogenous network triggered by RTK activation. Nonetheless, RTK activation still leads to native receptor signalling in addition to the engineered novel pathway, resulting in a certain degree of crosstalk.

1.2.3.5 Modular extracellular sensor architecture (MESA)

Joshua Leonard's laboratory conceived a novel receptor design that allows the integration of soluble factors and activation of desired transgenes via artificial TFs (Daringer et al. 2014; Schwarz et al. 2017; Hartfield et al. 2017).

Their receptor framework, termed modular extracellular sensor architecture (MESA), is composed of two polypeptide chains, a target chain (TC) and a protease chain (PC) (Daringer et al. 2014) (Figure 1.11). The PC contains a scFv as recognition domain (Schwarz et al. 2017), a standard TMD, and a TEV protease (Parks et al. 1994). The TC only differs in its intracellular part from the PC, in which an artificial TF (tTA) is fused to the TMD (derived from CD28) via a TCS. The scFvs of both chains recognise and bind to a soluble ligand which brings both chains into proximity, allowing PC's TEV

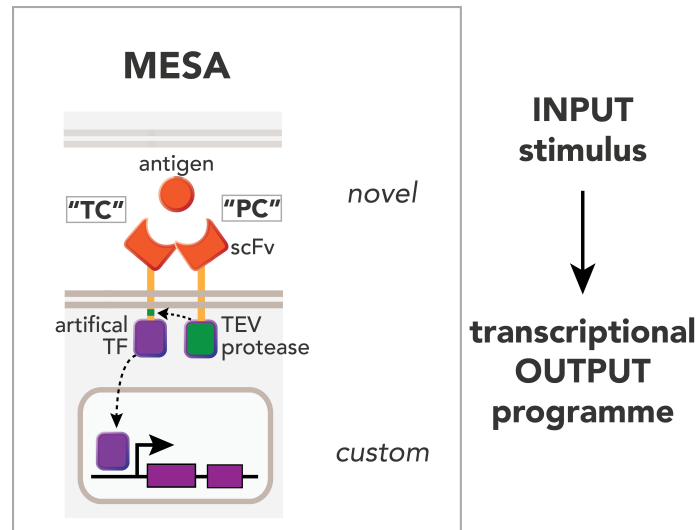


Figure 1.11 | Modular extracellular sensor architecture (MESA). MESA is composed of two polypeptide chains, target chain (TC) and protease chain (PC). The TC contains a scFv for ligand engagement, a transmembrane domain for cell membrane anchoring, a TEV cleavage site (TCS) and an artificial TF. The PC only differs from the TC in its intracellular configuration, where it possesses a TEV protease. Upon soluble ligand recognition, both chains come into proximity, prompting TEV protease-facilitated cleavage of the TCS thereby releasing the TF from the membrane. Liberated TF then induces the expression of a transgene, or in the case of dCas9-TF, any user-defined endogenous gene. Cartoon elements adapted from “Engineering Customized Cell Sensing and Response Behaviors Using Synthetic Notch Receptors”. *Cell*, 164(4), pp.780–791, Morsut, L. et al. Copyright 2016, with permission from Elsevier. Licence #4245390566951.

to cleave *in trans* the TCS on TC, thereby releasing the TF from the membrane and allowing its translocation into the nucleus (Daringer et al. 2014; Schwarz et al. 2017) (Figure 1.11).

Initially, the MESA system was characterised by employing homo-/heterodimerisation domains with different affinities as ectodomains (Daringer et al. 2014). Subsequently, Schwarz et al. (2016) took advantage of previously described anti-VEGF scFvs to re-direct the MESA system towards recognising and responding to soluble VEGF in culture media *in vitro*. However, even after extensive system optimisation only minimal ligand-induced reporter gene activation could be achieved. Eventually, Schwarz et al. (2016) exchanged tTA with a programmable, full length dCas9-TF (MESA-dCas9-TF) to induce transcription and promote protein secretion of endogenous human interleukin 2 (IL-2). The MESA-dCas9-TF also attained only modest induction rates in the presence of VEGF inducer, hardly distinguishable from the uninduced condition. Nonetheless, the authors propose that this system could be used for designing artificial signalling pathways, for example by converting an immunosuppressive stimulus into an immune-stimulating response (Schwarz et al. 2017).

One explanation for the minimal fold changes obtained in their study was that MESA receptor components interact while trafficking through the secretory pathway, thereby creating high local concentrations of both chains which favour ligand-independent activation (Daringer et al. 2014; Schwarz et al. 2017). Indeed, it was observed that by optimising the concentration of plasmids in transient transfections, unspecific dimerisation-induced system activation could be reduced to some degree (Daringer et al. 2014; Schwarz et al. 2017).

In a more recent study, the same group explored the idea of combining two sets of MESA receptors in an AND gate logic fashion (Hartfield et al. 2017). In this instance,

the AND gate configuration is built on a hybrid promoter for which two different MESA receptor-liberated TFs (tTA and Gal4-VP16) are required to fully drive reporter gene activation (Hartfield et al. 2017).

Conceptually, there are two main features of MESA that are distinct from native receptors (GPCR, RTK) and their associated signalling cascades. First, MESA is a self-contained sensing and signalling system since it does not necessitate other components to function once activated by a cognate antigen. Second, and similar to synNotch, MESA is completely orthogonal to endogenous signalling pathways because it does not require native second messengers or mediators.

1.3 Thesis Aims - Engineering a CRISPR/dCas9-based synthetic receptor scaffold

The overarching goal of my DPhil was the design, development and implementation of a framework for engineering synthetic dCas9-based chimeric receptors (dCas9-synRs) (Figure 1.12). This novel, modular, and generalisable receptor architecture brings together the capability of sensing a broad scope of soluble ligands by endogenous receptor ectodomains and the unprecedented potential of eliciting user-defined, multi-gene output response programmes by a highly programmable dCas9-based signal transduction module.

I demonstrate

- in Chapter 3 the necessity for the rational design of a minimal core split dCas9 architecture;
- in Chapter 4 and Chapter 5 the applicability and portability of this core framework to various GPCRs (dCas9-synGPCRs) and prototype RTK (dCas9-synRTK), eliciting a user-defined, non-native transcriptional response programme upon native ligand recognition in a dose-dependent manner;
- in Chapter 6 the therapeutic potential of dCas9-synRs in redirecting the cellular information flow to counteract the effects of specific disease biomarkers with customised therapeutically-relevant transcriptional multi-gene programmes in engineered designer cells.

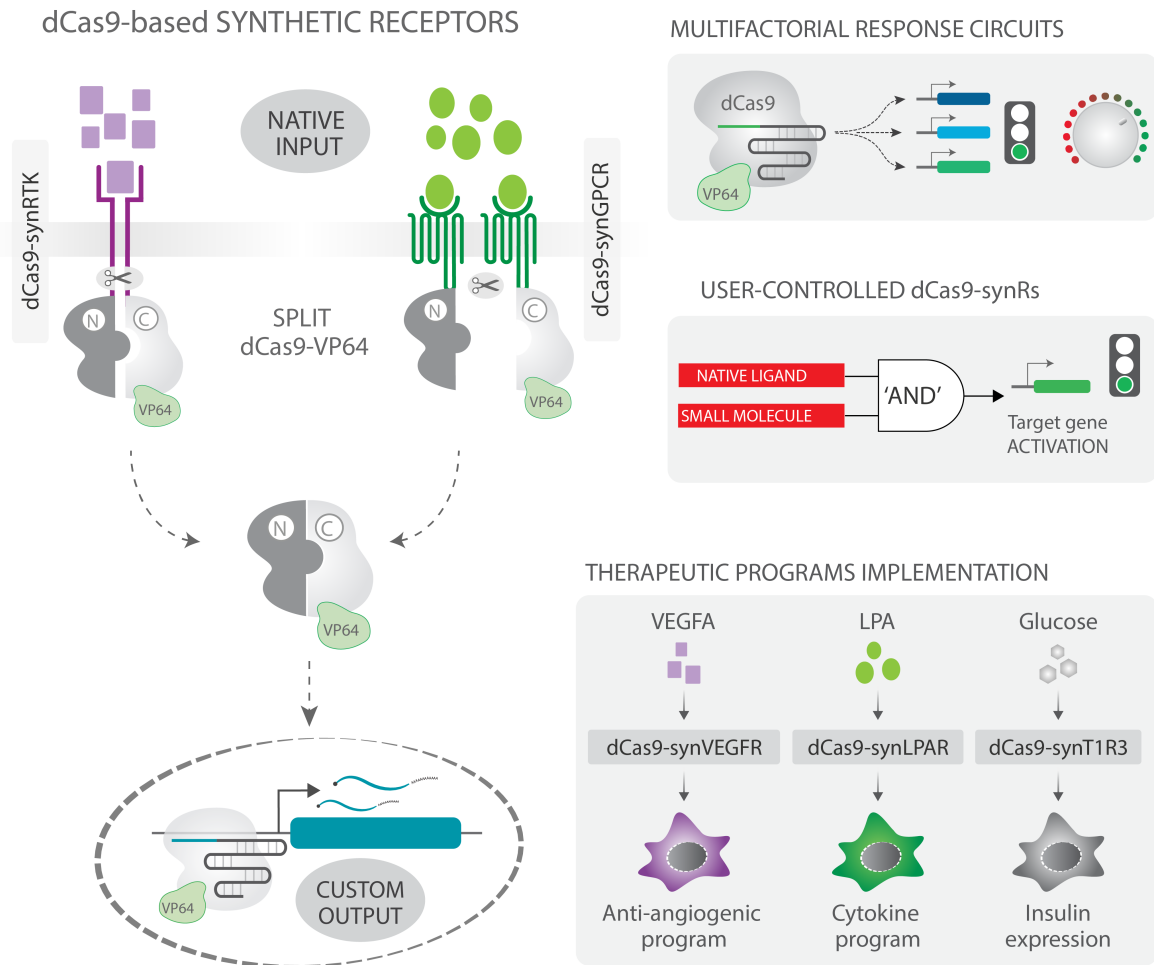


Figure 1.12 | Conceptual overview of dCas9-based synthetic receptors (dCas9-synRs).

dCas9-based synthetic receptors, dCas9-synRs, are a novel class of artificial receptors combining endogenous receptor recognition domains with the highly programmable signal transduction module dCas9-TF (left). This enables a multi-factorial response programme upon native ligand recognition in a dose-dependent manner (top right). Owing to their modular architecture, safety switch mechanisms can easily be installed, rendering receptor activation dependent on at least 2-input stimuli (middle right). The unprecedented versatility of dCas9-synRs in redirecting the cellular information flow makes them ideally suited to activate a variety of customised therapeutically-relevant transcriptional programmes in engineered designer cells (bottom right). Figure from Baeumler et al. (2017).

Chapter 2 : Materials and Methods

This chapter is an extended version of the corresponding experimental procedures and supplemental information including supplemental experimental procedures from: Baeumler, T.A., Ahmed, A.A. & Fulga, T.A., 2017. Engineering Synthetic Signaling Pathways with Programmable dCas9-Based Chimeric Receptors. Cell Reports, 20(11), pp.2639–2653.

2.1 Reagents and kits

Reagents	Cat. #	Supplier/notes
Antarctic phosphatase	M0289	New England Biolabs (NEB)
Bradykinin acetate salt powder	B3259	Sigma
DNA oligonucleotides and PCR/qPCR primer	-	Integrated DNA Technologies (IDT)
Coca-Cola® Classic	330 ml can, UK 2017	Degassed by extensive shaking, pH adjusted to 7.0 with 1 M NaOH. Contains 10.6 g sugar per 100 ml.
Doxycycline hyclate	D9891	Sigma
D-(-)-fructose	F3510	Sigma
D-glucose	G/0500/53	Fisher Scientific
MinElute PCR Purification Kit	28006	Qiagen
1-oleoyl lysophosphatidic acid (LPA)	10010093	Cayman Chemical
Phusion High-Fidelity PCR Master Mix with GC Buffer	M0532	NEB
Polyethylenimine, branched	408727	Sigma, was diluted in MilliQ water to 1 mg/ml, pH adjusted to 7, sterile filtered and kept in aliquots at -20°C as previously described (Aricescu et al. 2006)

QIAfilter Plasmid Midi Kit	12243	Qiagen
QIAprep Spin Miniprep Kit	27106	Qiagen
QIAquick Gel Extraction Kit	28706	Qiagen
QIAquick PCR Purification Kit	28106	Qiagen
Rapamycin	SM83	Cambridge Bioscience
Restriction enzymes	-	NEB or ThermoFisher Scientific
SDF-1 α (recombinant human stromal derived factor 1; CXCL12)	581202	Biolegend
Sucrose	S9378	Sigma
Trimethoprim (TMP)	CAY-16473	Cambridge Bioscience
T4 DNA Ligase	M0202	NEB
T4 Polynucleotide Kinase	M0201	NEB
Vasopressin ([deamino-Cys1, D-Arg8]-Vasopressin acetate salt hydrate)	V1005	Sigma

Table 1 | List of reagents, including chemicals, enzymes, and standard molecular biology kits. All enzymes were used according to the manufacturer protocols, unless otherwise stated. PCR reactions were performed in a C1000 Thermal Cycler (Bio-Rad). Standard molecular biology techniques and kits were used for all cloning experiments according to manufacturer protocols.

2.2 Plasmid constructs

Amino acid sequence for all constructs described below are provided in the Appendix.

A selection of plasmids is available on addgene (http://www.addgene.org/Tudor_Fulga/). DNA constructs were validated by diagnostic restriction digest and/or Sanger sequencing (performed by Source BioScience and Eurofins genomics).

2.2.1 EYFP reporter assay constructs

Control sgRNA (*sgSCR*): the sgRNA cassette (*U6 promoter/sgRNA scaffold/U6 terminator*) from pX330 vector (gift from Feng Zhang, Addgene plasmid #42230), f1 origin + SV40 promoter from pcDNA3.1 and mCherry gene (gift from Dr Fabien Pinaud, University of Southern California) were PCR amplified with primers containing *MluI(fwd)/KpnI(rev)*, *KpnI(fwd)/NheI(rev)* and *NheI(fwd)/EagI(rev)* respectively, and cloned into the *MluI* and *EagI* sites in pcDNA3.1 to generate plasmid *pU6-sgSCR_mCherry*.

EYFP-targeting sgRNA (sgEYFP): the EYFP targeting spacer (5'-GAGTCGCGTGTAGC GAAGCA-3') was synthesised (IDT) and cloned between *BbsI* sites in the *U6-sgSCR_mCherry* vector as previously described (Ran et al. 2013) to generate *pU6-sgEYFP_mCherry*.

EYFP reporter: the *P1-EYFP-pA* plasmid containing a synthetic operator (8 x target sequences 5'-AGTCGCGTGTAGCGAAGCA-3') recognized by the *sgEYFP* spacer placed upstream of the EYFP reporter gene (gift from Timothy K. Lu, Addgene plasmid #54781).

2.2.2 dCas9^{VP64}, Cas9, and associated constructs

dCas9^{VP64} (Mali et al. 2013): unaltered Cas9m4-VP64 plasmid was used in Figure 3.1 and 3.6 as a no HA/no myc-control dCas9-VP64 and in Figure 5.5 as a benchmark for fully active full length dCas9-VP64. It was a gift from George Church (Addgene plasmid #47319).

NLS(+)-dCas9^{VP64}: The pX330 vector (gift from Feng Zhang, Addgene plasmid #42230) was modified as follows: the *U6 promoter/sgRNA scaffold/U6 terminator* cassette was removed; the FLAG-tag NLS-Cas9 cassette was replaced with dCas9m4-VP64 (gift

from George Church, Addgene plasmid #47319) containing the original N-terminal SV40 NLS from pX330 and a new HA epitope tag, to generate plasmid *pNLS(+)-HA-dCas9m4-VP64*.

NLS-dCas9^{VP64}: based on *pNLS(+)-HA-dCas9m4-VP64*, however the N-terminal SV40 NLS constitutes now a shorter version to generate plasmid *pNLS-HA-dCas9m4-VP64*.

dCas9^{VP64}: based on *pNLS-HA-dCas9m4-VP64*, however excluding the N-terminal SV40 NLS and without the connecting NLS between dCas9m4 and VP64 that was replaced with a (GGGS)₂ linker to generate plasmid *pHA-dCas9m4-VP64*.

NES-dCas9^{VP64}: based on *pHA-dCas9m4-VP64*, however the HA-dCas9m4-VP64 was flanked by an NES sequence (from pX855 (gift from Feng Zhang, Addgene plasmid #62887)) on either of the N-terminal and C-terminal end to generate *pNES-HA-dCas9m4-VP64*.

SAM sgRNAs: the spacer sequences for all sgRNA targeting endogenous genes were synthesized (IDT) and cloned between BbsI sites in the sgRNA(MS2) cloning backbone (gift from Feng Zhang, Addgene plasmid #61424) as previously described (Ran et al. 2013). All sgRNA spacer sequences used in this thesis are provided in the table below.

SAM MCP: Plasmid MS2-P65-HSF1_GFP was a gift from Feng Zhang, Addgene plasmid #61423.

2.2.3 TMt-dCas9^{VP64}, TMt-Cas9, and associated constructs

TMt-NLS-dCas9^{VP64}: a transmembrane tether (TMt; modified from the pDisplay Vector (Invitrogen)) containing the Igκ signal sequence, (GGGS)₂ linker, myc epitope tag, PDGF receptor transmembrane domain and the XTEN linker (Schellenberger et al. 2009), was synthesized as a gBlock (IDT). This transmembrane tether was then fused

to the N-terminus of *dCas9-VP64* via a TEV cleavage site (ENLYFQG) to generate plasmid *pTMt_TCS(Q'G)_NLS-HA-dCas9m4-VP64*.

TMt^{ΔTCS}-NLS-dCas9^{VP64}: the ENLYFQG TEV cleavage site in *pTMt_TCS(Q'G)_NLS-HA-dCas9m4-VP64* was replaced by one GGGS linker to generate *pTMt_NLS-HA-dCas9m4-VP64*.

TMt-dCas9^{VP64}: the gBlock synthesized TMt from *pTMt_TCS(Q'G)_NLS-HA-dCas9m4-VP64* was fused to the N-terminus of *pHA-dCas9m4-VP64* to generate *pTMt_TCS(Q'G)_HA-dCas9m4-VP64*.

TMt-NES-dCas9^{VP64}: the gBlock synthesized TMt from *pTMt_TCS(Q'G)_NLS-HA-dCas9m4-VP64* was fused to the N-terminus of *pNES-HA-dCas9m4-VP64* to generate *pTMt_TCS(Q'G)_NES-HA-dCas9m4-VP64*.

TMt^{NES}-dCas9^{VP64}: the NES sequence from pX855 (gift from Feng Zhang, Addgene plasmid #62887) was cloned between the TMt and the TEV cleavage site in *pTMt_TCS(Q'G)_NLS-HA-dCas9m4-VP64*. In addition, the N-terminal NLS of *dCas9m4-VP64* was removed while the C-terminal NLS was replaced by a (GGGS)₂ linker, to generate plasmid *pTMt_NES_TCS(Q'G)_HA-dCas9m4-VP64*.

TMt^{NESΔTCS}-dCas9^{VP64}: the ENLYFQG TEV cleavage site in *pTMt_NES_TCS(Q'G)_HA-dCas9m4-VP64* was replaced by one GGGS linker to generate *pTMt_NES_HA-dCas9m4-VP64*.

TMt-NLS-dCas9^{VP64[Dox]}: the *TMt-NLS-dCas9^{VP64}* from *pTMt-TCS(Q'G)-NLS-HA-dCas9m4-VP64* was PCR amplified and cloned between XbaI and FseI in pCW-Cas9 (gift from Eric Lander and David Sabatini, Addgene plasmid #50661) to generate pTREtight-TMt_TCS(Q'G)_NLS-dCas9-VP64_PGK1-Puro-T2A-rtTA plasmid.

NLS-Cas9-mCherry: The pX330 vector (gift from Feng Zhang, Addgene plasmid #42230) was modified as follows: the original *sgRNA scaffold* cassette was replaced

with a gBlock containing the advanced sgRNA^{F+E} scaffold (B. Chen et al. 2013) and the C-terminal part of the FLAG-NLS-Cas9-NLS cassette was extended by the XTEN linker (Schellenberger et al. 2009) and mCherry (gift from Dr Fabien Pinaud, University of Southern California) with a C-terminal NLS sequence in frame, to generate plasmid *pFLAG-NLS-Cas9-XTEN-mCherry-NLS*. The spacer sequences for sgSCR (5'-GAGCACTGTCCTCCGAACGT-3', originally Gal4 targeting sequence, non-targeting in human genome (B. Chen et al. 2013)) and sgEGFP (5'-GGTGAACCGCATCGAGCTGAA-3', (Shalem et al. 2014)) were synthesized (IDT) and cloned between BspQI sites in the sgRNA^{F+E} cloning site as previously described (Ran et al. 2013).

Cas9-mCherry: The pX330 vector (gift from Feng Zhang, Addgene plasmid #42230) was modified as follows: the original *sgRNA scaffold* cassette was replaced with a gBlock containing the advanced sgRNA^{F+E} scaffold (B. Chen et al. 2013), and the N-terminal FLAG-NLS as well as the N-terminal NLS sequence of Cas9 was removed via PCR and the XTEN linker (Schellenberger et al. 2009) and mCherry (gift from Dr Fabien Pinaud, University of Southern California) sequences were fused in frame to the C-terminal part of the Cas9, to generate *pCas9-XTEN-mCherry*. Guide cloning equivalent to as described for NLS-Cas9-mCherry.

TMt-NLS-Cas9-mCherry: based on *pFLAG-NLS-Cas9-XTEN-mCherry-NLS* in which the gBlock synthesized TMt was fused in frame to the N-terminal part of FLAG-NLS-Cas9, to generate *pTMt_TCS(Q'G)_FLAG-NLS-Cas9-XTEN-mCherry-NLS*. Guide cloning equivalent to as described for NLS-Cas9-mCherry.

TMt^{ΔTCS}-NLS-Cas9-mCherry: based on *pTMt_TCS(Q'G)_FLAG-NLS-Cas9-XTEN-mCherry-NLS* in which the TCS was deleted, to generate *pTMt_FLAG-NLS-Cas9-XTEN-mCherry-NLS*. Guide cloning equivalent to as described for NLS-Cas9-mCherry.

TMt-Cas9-mCherry: based on *pCas9-XTEN-mCherry* in which the gBlock synthesized TMt was fused in frame to the N-terminal part of Cas9, to generate *pTMt_TCS(Q'G)_Cas9-XTEN-mCherry*. Guide cloning equivalent to as described for NLS-Cas9-mCherry.

TEV protease: the TEV protease coding sequence was PCR amplified from plasmid DNA (gift from Dr. Jon Elkins, Nuffield Department of Medicine, University of Oxford) and cloned between BamHI and XhoI in pcDNA3.1 to generate *pcDNA3.1_TEV* plasmid.

TEV^{C151A} protease: based on *pcDNA3.1_TEV* in which the codon for alanine was introduced to replace the cysteine codon to generate *pcDNA3.1_TEV^{C151A}*.

dEGFP: based on the pCSGW plasmid (gift from Jan Rehwinkel, WIMM, University of Oxford) in which the mRNA stabilising element WPRE and the stop codon of EGFP were replaced by an in-frame protein destabilising sequence, called PEST domain (originating from ornithine decarboxylase (X. Li et al. 1998)) and a new stop codon, to generate *pCSGW_dEGFP_noWPRE*.

2.2.4 Split mono and dual TMt-dCas9^{VP64} and associated constructs

NLS(+)-dCas9(C)^{VP64}-FKBP: U6 promoter/sgRNA scaffold/U6 terminator cassette was removed from pX856 vector (gift from Feng Zhang, Addgene plasmid #62888), to generate *pNLS(+)-FKBP-dCas9(C)-VP64*.

TMt-NLS-dCas9(C)^{VP64}-FKBP: based on *pNLS(+)-FKBP-dCas9(C)-VP64* in which the N-terminal NLS(+) was removed and replaced with the gBlock synthesized TMt, a new

SV40 NLS and HA tag sequence to generate *pTmt_TCS(Q'G)_NLS-HA-FKBP-dCas9(C)-VP64*.

NES-dCas9(N)-FRB: *U6 promoter/sgRNA scaffold/U6 terminator* cassette was removed from pX855 vector (gift from Feng Zhang, Addgene plasmid #62887), to generate *pNES-dCas9(N)-FRB*.

dCas9(N)-FRB: based on *pNES-dCas9(N)-FRB* in which the NES sequences were removed via PCR to generate *pdCas9(N)-FRB*.

Tmt^{NES}-dCas9(N)-FRB: based on *pdCas9(N)-FRB* in which the N-terminus was extended by the gBlock synthesized Tmt, a reintroduced NES sequence, followed by an efficient TCS(Q'G) and an HA tag sequence, to generate *pTmt_NES_TCS(Q'G)_HA-dCas9(N)-FRB*.

Tmt-NLS-dCas9(C)^{VP64}: the pX856 vector (gift from Feng Zhang, Addgene plasmid #62888) was modified as follows: the *U6 promoter/sgRNA scaffold/U6 terminator* cassette was removed; the *dCas9(C)^{VP64}* N-terminal NLS+FKBP were also removed; the Tmt-TCS(Q'G)-NLS-HA cassette from *pTmt_TCS(Q'G)_NLS-HA-dCas9m4-VP64* was fused to the N-terminus of *dCas9(C)^{VP64}*; the MCP-P65-HSF1 from plasmid MS2-P65-HSF1_GFP (gift from Feng Zhang, Addgene plasmid #61423) was fused to the C-terminus of *dCas9(C)^{VP64}* via at T2A site to generate plasmid *pTmt_TCS(Q'G)_NLS-HA-dCas9(C)-VP64_T2A_MCP-P65-HSF1*. Note: for confocal microscopy experiments the HA tag was removed from this construct.

NLS-dCas9(C)^{VP64}: based on *pTmt_TCS(Q'G)_NLS-HA-dCas9(C)-VP64_T2A_MCP-P65-HSF1* in which the Tmt(Q'G) was removed to generate *pNLS-HA-dCas9(C)-VP64_T2A_MCP-P65-HSF1*. Note: for confocal microscopy experiments the HA tag was removed from this construct.

NLS(+)-dCas9(C)^{VP64}: based on *pNLS-HA-dCas9(C)-VP64_T2A_MCP-P65-HSF1* in which the N-terminal SV40 NLS sequence was replaced with the longer NLS originating from pX330 vector (gift from Feng Zhang, Addgene plasmid #42230), to generate *pNLS(+)-HA-dCas9(C)-VP64_T2A_MCP-P65-HSF1*.

TMt^{NES}-dCas9(N): the pX855 vector (gift from Feng Zhang, Addgene plasmid #62887) was modified as follows: the *U6 promoter/sgRNA scaffold/U6 terminator* cassette was removed; the dCas9(N) N-terminal NES and the C-terminal FRB+NES were also removed; the TMt-NES-TCS(Q'G)-HA cassette from *pTMt_NES_TCS(Q'G)_HA-dCas9m4-VP64* was fused to the N-terminus of dCas9(N); the puromycin resistance gene and the WPRE stabilising element from pCW-Cas9 (gift from Eric Lander and David Sabatini, Addgene plasmid #50661) were inserted downstream of dCas9(N) to generate plasmid *pTMt_NES_TCS(Q'G)_HA-dCas9(N)_P2A-Puro-WPRE*.

TMt^{NES}-FKBP-TEV(N)-dCas9(N): based on *pTMt_NES_TCS(Q'G)_HA-dCas9(N)_P2A-Puro-WPRE* in which the cassette FKBP-TEV(N) from *pcDNA3.1_FKBP-TEV(N)* was inserted after the XTEN linker and before the NES sequence to generate *pTMt_FKBP-TEV(N)_NES_TCS(Q'G)_HA-dCas9(N)_P2A-Puro-WPRE*. Design inspired by (Wehr et al. 2006).

TMt-FKBP-TEV(N)-NLS-dCas9(C)^{VP64}: based on *pTMt_TCS(Q'G)_NLS-HA-dCas9(C)-VP64_T2A_MCP-P65-HSF1* in which the cassette FKBP-TEV(N) from *pcDNA3.1_FKBP-TEV(N)* was inserted after the XTEN linker and before the TCS sequence to generate *pTMt_FKBP_TEV(N)_TCS(Q'G)_NLS-HA-dCas9(C)-VP64_T2A_MCP-P65-HSF1*. Design inspired by (Wehr et al. 2006).

FKBP-TEV(N): the rapamycin inducible hetero-dimerisation FK506 binding protein 12 (FKBP) from pX856 vector (gift from Feng Zhang, Addgene plasmid #62888) was PCR amplified and fused in frame via a GGGS linker to the N-terminal TEV fragment

(amplified from full length TEV protease as previously described (Wehr et al. 2006)) and inserted together into the MCS of pcDNA3.1 to generate *pcDNA3.1_FKBP-TEV(N)*.

FRB-TEV(C): the FKBP rapamycin binding (FRB) from pX855 vector (gift from Feng Zhang, Addgene plasmid #62887) was PCR amplified and fused in frame via a GGGS linker to the C-terminal TEV fragment (amplified from full length TEV protease as previously described (Wehr et al. 2006)) and inserted together into the MCS of pcDNA3.1 to generate *pcDNA3.1_FRB-TEV(C)*.

2.2.5 dCas9-synGPCR and associated constructs

BDKRB2-NLS-dCas9^{VP64}: a sequence containing the membrane localisation signal, FLAG tag, BDKBR2 coding sequence and the V_2 tail were PCR amplified from plasmid BDKBR2-Tango (gift from Bryan Roth, Addgene plasmid #66230) and used to replace the TMt in *pTMt_TCS(Q'G)_NLS-HA-dCas9(C)-VP64_T2A_MCP-P65-HSF1*. Subsequently, the cassette NLS-HA-dCas9(C)-VP64_T2A_MCP-P65-HSF1 was replaced with NLS-HA-dCas9m4^{VP64} from *pNLS-HA-dCas9m4-VP64*, to generate *pBDKRB2-TCS(Q'L)_NLS-HA-dCas9m4-VP64*.

BDKRB2-dCas9^{VP64}: based on *pBDKRB2-TCS(Q'L)_NLS-HA-dCas9m4-VP64* cloning strategy but instead of inserting NLS-HA-dCas9m4^{VP64}, HA-dCas9m4^{VP64} from *pHA-dCas9m4-VP64* was used to generate *pBDKRB2-TCS(Q'L)_HA-dCas9m4-VP64*.

dCas9(C)-synBDKBR2: a sequence containing the membrane localisation signal, FLAG tag, BDKBR2 coding sequence and the V_2 tail were PCR amplified from plasmid BDKBR2-Tango (gift from Bryan Roth, Addgene plasmid #66230) and used to replace the TMt in *pTMt_TCS(Q'G)_NLS-HA-dCas9(C)-VP64_T2A_MCP-P65-HSF1* to generate *pBDKBR2_TCS(Q'G)_NLS-HA-dCas9(C)-VP64_T2A_MCP-P65-HSF1*. For

TEV cleavage optimization, the TCS sequences were iteratively replaced by ENLYFQG, ENLYFQY and ENLYFQL.

dCas9(N)-synBDKBR2: the TMt, NES and TCS(Q'G) sequences from *pTMt_NES_TCS(Q'G)_HA-dCas9(N)_P2A-Puro-WPRE* were removed and replaced with the membrane localisation signal/FLAG tag/BDKBR2 coding sequence/*V*₂ tail from plasmid BDKBR2-Tango and the TEV cleavage site ENLYFQL, to generate *pBDKBR2_TCS(Q'L)_HA-dCas9(N)_P2A-Puro-WPRE*. For the linker optimization of *dCas9(N)-synBDKBR2*, the following amino acid sequences were inserted between the *V*₂ tail and the TCS: for b) GGS and for c) EFLDLASLILGKLG. For TEV cleavage optimization, the TCS in *dCas9(N)-synBDKBR2* were consecutively replaced by ENLYFQY (for construct in d) and ENLYFQG (for construct in e). For the extracellular NES optimisation, in one case, *dCas9(N)-synBDKBR2* received a NES sequence from pX855 vector (gift from Feng Zhang, Addgene plasmid #62887) that was inserted downstream of the membrane localisation signal and FLAG tag and upstream of the BDKBR2 coding sequence, to generate *pNES-BDKBR2_TCS(Q'L)_HA-dCas9(N)_P2A-Puro-WPRE*. In the other case, the NES sequence was flanked by GGS short linker sequences was inserted in the same location, to generate *pNES(GGS)-BDKBR2_TCS(Q'L)_HA-dCas9(N)_P2A-Puro-WPRE*.

dCas9(N)-synBDKBR2^{DHFR}: the P2A-Puro-WPRE fragment from *pBDKBR2_TCS(Q'L)_HA-dCas9(N)_P2A-Puro-WPRE* was replaced with DHFR from plasmid *DHFR-PP7-VP64_T2A_GFP* (gift from Amit Choudhary, Addgene plasmid #86167) to generate *pBDKBR2_TCS(Q'L)_HA-dCas9(N)_DHFR*.

dCas9(N)-synBDKBR2^{RI}: the rapamycin inducible hetero-dimerisation FK506 binding protein 12 (FKBP) from pX856 vector (gift from Feng Zhang, Addgene plasmid #62888) was fused to the N-terminus of HA-dCas9(N) in *pBDKBR2_TCS(Q'L)_HA-*

dCas9(N)_P2A-Puro-WPRE, (design inspired by (Y. Gao et al. 2016; Bao et al. 2017)), to generate *pBDKBR2_TCS(Q'L)_FKBP-HA-dCas9(N)_P2A_Puro-WPRE*.

dCas9(C)-synBDKRB2^{(-)VP64}: the VP64_T2A_MCP-P65-HSF1 cassette from *pBDKBR2_TCS(Q'G)_NLS-HA-dCas9(C)-VP64_T2A_MCP-P65-HSF1* was removed, and DHFR-PCP-VP64 was placed downstream of dCas9(C) sequence separated by a T2A site. The DHFR-PCP-VP64 sequence was obtained from plasmid DHFR-PP7-VP64_T2A_GFP (gift from Amit Choudhary, Addgene plasmid #86167) to generate *pBDKBR2_TCS(Q'G)_NLS-HA-dCas9(C)_T2A_DHFR-PP7-VP64*.

dCas9(C)-synAVPR2: same strategy as *dCas9(C)-synBDKBR2* but instead of BDKBR2, the AVPR2 coding sequence was cloned from plasmid AVPR2 Tango (gift from Bryan Roth (Addgene plasmid #66227)) to generate *pAVPR2_TCS(Q'G)_NLS-HA-dCas9(C)-VP64_T2A_MCP-P65-HSF1*.

dCas9(N)-synAVPR2: same strategy as *dCas9(N)-synBDKBR2* but instead of BDKBR2, the AVPR2 coding sequence was cloned from plasmid AVPR2 Tango (gift from Bryan Roth, Addgene plasmid #66227) to generate *pAVPR2_TCS(Q'L)_HA-dCas9(N)_P2A-Puro-WPRE*. For the extracellular NES optimisation, dCas9(N)-synAVPR2 received a NES sequence from pX855 vector (gift from Feng Zhang, Addgene plasmid #62887) flanked by GGS short linker sequences, which was inserted downstream of the membrane localisation signal and FLAG tag and upstream of the AVPR2 coding sequence, to generate *pNES(GGS)-AVPR2_TCS(Q'L)_HA-dCas9(N)_P2A-Puro-WPRE*.

dCas9(C)-synCXCR4: same strategy as *dCas9(C)-synBDKBR2* but instead of BDKBR2, the CXCR4 coding sequence was cloned from plasmid CXCR4-Tango (gift from Bryan Roth, Addgene plasmid #66262) to generate *pCXCR4_TCS(Q'G)_NLS-HA-dCas9(C)-VP64_T2A_MCP-P65-HSF1*.

dCas9(C)-synCXCR4^{(-)VP64}: the VP64_T2A_MCP-P65-HSF1 cassette from *pCXCR4_TCS(Q'G)_NLS-HA-dCas9(C)-VP64_T2A_MCP-P65-HSF1* was removed, to generate *pCXCR4_TCS(Q'G)_NLS-HA-dCas9(C)*.

dCas9(N)-synCXCR4: same strategy as *dCas9(N)-synBDKBR2* but instead of BDKBR2, the CXCR4 coding sequence was cloned from plasmid CXCR4-Tango (gift from Bryan Roth, Addgene plasmid #66262) to generate *pCXCR4_TCS(Q'L)_HA-dCas9(N)_P2A-Puro-WPRE*.

dCas9(C)-synLAPR1: same strategy as *dCas9(C)-synBDKBR2* but instead of BDKBR2, the LAPR1 coding sequence was cloned from plasmid LAPR1-Tango (gift from Bryan Roth, Addgene plasmid #66418) to generate *pLAPR1_TCS(Q'G)_NLS-HA-dCas9(C)-VP64_T2A_MCP-P65-HSF1*.

dCas9(C)-synLAPR1^{(-)VP64}: the VP64_T2A_MCP-P65-HSF1 cassette from *pLAPR1_TCS(Q'G)_NLS-HA-dCas9(C)-VP64_T2A_MCP-P65-HSF1* was removed, to generate *pLAPR1_TCS(Q'G)_NLS-HA-dCas9(C)*.

dCas9(C)-synLAPR2: same strategy as *dCas9(C)-synBDKBR2* but instead of BDKBR2, the LAPR2 coding sequence was cloned from plasmid LAPR2-Tango (gift from Bryan Roth, Addgene plasmid #66419) to generate *pLAPR2_TCS(Q'G)_NLS-HA-dCas9(C)-VP64_T2A_MCP-P65-HSF1*.

dCas9(N)-synLAPR1: same strategy as *dCas9(N)-synBDKBR2* but instead of BDKBR2, the LAPR1 coding sequence was cloned from plasmid LAPR1-Tango (gift from Bryan Roth, Addgene plasmid #66418) to generate *pLAPR1_TCS(Q'L)_HA-dCas9(N)_P2A-Puro-WPRE*.

dCas9(N)-synLAPR2: same strategy as *dCas9(N)-synBDKBR2* but instead of BDKBR2, the LAPR2 coding sequence was cloned from plasmid LAPR1-Tango (gift

from Bryan Roth, Addgene plasmid #66419) to generate *pLPAR2_TCS(Q'L)_HA-dCas9(N)_P2A-Puro-WPRE*.

dCas9(C)-synT1R2: same strategy as *dCas9(C)-synBDKBR2* but instead of BDKBR2, the hT1R2 coding sequence was subcloned from cDNA (gift from Robert Margolskee, Monell Chemical Senses Center, under MTA agreement) to generate *pT1R2_TCS(Q'G)_NLS-HA-dCas9(C)-VP64_T2A_MCP-P65-HSF1*.

dCas9(C)-synT1R3: same strategy as *dCas9(C)-synBDKBR2* but instead of BDKBR2, the hT1R3 coding sequence was subcloned from cDNA (gift from Robert Margolskee, Monell Chemical Senses Center, under MTA agreement) to generate *pT1R3_TCS(Q'G)_NLS-HA-dCas9(C)-VP64_T2A_MCP-P65-HSF1*.

dCas9(N)-synT1R2: same strategy as *dCas9(N)-synBDKBR2* but instead of BDKBR2, the hT1R2 coding sequence was subcloned from cDNA (gift from Robert Margolskee, Monell Chemical Senses Center, under MTA agreement) to generate *pT1R2_TCS(Q'L)_HA-dCas9(N)_P2A-Puro-WPRE*.

dCas9(N)-synT1R3: same strategy as *dCas9(N)-synBDKBR2* but instead of BDKBR2, the hT1R3 coding sequence was subcloned from cDNA (gift from Robert Margolskee, Monell Chemical Senses Center, under MTA agreement) to generate *pT1R3_TCS(Q'L)_HA-dCas9(N)_P2A-Puro-WPRE*.

scSCR^{PP7}: the scaffold RNA containing the PP7 aptamer (gift from Amit Choudhary, Addgene plasmid #86168) was used to replace the sgSCR in *U6-sgSCR_mCherry* vector to create *pU6-scSCR^{PP7}_mCherry*.

EYFP-targeting scRNA^{PP7} (scEYFP^{PP7}): the EYFP targeting spacer (5'-GAGTCGCGTGTAGC GAAGCA-3') was synthesised (IDT) and cloned between BbsI sites in the *pU6-scSCR^{PP7}_mCherry* vector as previously described{Ran, 2013 #122} to generate *pU6-scEYFP^{PP7}_mCherry*.

2.2.6 dCas9-synRTK, MESA, and associated constructs

dCas9(C)-synVEGFR-1: a sequence containing the VEGFR1 (FLT1) leader peptide, extracellular domain and transmembrane domain were PCR amplified from plasmid pDONR223-FLT1 (gift from William Hahn & David Root, Addgene plasmid #23912) and used to replace the TMT in *pTMT_TCS(Q'G)_NLS-HA-dCas9(C)-VP64_T2A_MCP-P65-HSF1*. In addition, the C-terminal TEV fragment was amplified from full length TEV protease as previously described (Wehr et al. 2006) and cloned between the VEGFR1 transmembrane domain and the TEV cleavage site to generate *pVEGFR1_TEV(C)_TCS(Q'G)_NLS-HA-dCas9(C)-VP64_T2A_MCP-P65-HSF1*.

dCas9(N)-synVEGFR-2: a sequence containing the VEGFR2 (KDR) leader peptide, extracellular domain and transmembrane domain were PCR amplified from plasmid pDONR223-KDR (gift from William Hahn & David Root, Addgene plasmid #23925) and used to replace the TMT in *pTMT_NES_TCS(Q'G)_HA-dCas9(N)_P2A-Puro-WPRE*. The N-terminal TEV fragment was then amplified from full length TEV protease as previously described (Wehr et al. 2006) and fused to the C-terminus of the VEGFR2 transmembrane domain. In addition, a weak TEV cleavage site (ENLYFQL) was inserted instead of the TCS(Q'G) to generate plasmid *pVEGFR2_TEV(N)_NES_TCS(Q'L)_HA-dCas9(N)_P2A_Puro-WPRE*.

For the optimization of *dCas9-synVEGFR* dimer combinations, the VEGFR1 and VEGFR2 PCR products obtained above were interchangeably swapped to generate plasmids *pVEGFR2_TEV(C)_TCS(Q'G)_NLS-HA-dCas9(C)-VP64_T2A_MCP-P65-HSF1* and *pVEGFR1_TEV(N)_NES_TCS(Q'G)_HA-dCas9(N)_P2A_Puro-WPRE*. For TEV cleavage optimization the TCS in *dCas9(C)-synVEGFR-1* and *dCas9(N)-synVEGFR-2* were iteratively replaced by ENLYFQG, ENLYFQY and ENLYFQL.

dCas9(C)-synVEGFR1^{RI}: the sequences encoding VP64 and T2A-MCP-P65-HSF1 were removed from *pVEGFR1_TEV(C)_TCS(Q'G)_NLS-HA-dCas9(C)-VP64_T2A_MCP-P65-HSF1* to generate *pVEGFR1_TEV(C)_TCS(Q'G)_NLS-HA-dCas9(C)*.

dCas9(N)-synVEGFR2^{RI}: the rapamycin inducible hetero-dimerisation FK506 binding protein 12 (FKBP) from pX856 vector (gift from Feng Zhang, Addgene plasmid #62888) was fused to the N-terminus of HA-dCas9(N) in *pVEGFR2_TEV(N)_NES_TCS(Q'L)_HA-dCas9(N)_P2A_Puro-WPRE*, (design inspired by (Y. Gao et al. 2016; Bao et al. 2017)) to generate *pVEGFR2_TEV(N)_NES_TCS(Q'L)_FKBP-HA-dCas9(N)_P2A_Puro-WPRE*.

FRB-VP64: the FKBP rapamycin binding (FRB) from pX855 vector (gift from Feng Zhang, Addgene plasmid #62887) and VP64 were PCR amplified and cloned in between the HindIII and XhoI sites in pcDNA3.1 to generate *pcDNA3.1_FRB-VP64* plasmid.

VEGFR1-TEV(C): based on *pVEGFR1_TEV(C)_TCS(Q'G)_NLS-HA-dCas9(C)-VP64_T2A_MCP-P65-HSF1*, in which the cassette *NLS-HA-dCas9(C)-VP64_T2A_MCP-P65-HSF1* was removed, to generate *pVEGFR1_TEV(C)*.

V1-MESA-dCas9 and *V2-MESA-dCas9*: V1-MESA-45F-M-dCas9, V1-MESA-45F-Tev, V2-MESA-35F-M-dCas9, V2-MESA-35F-Tev were a gift from Joshua Leonard (Addgene plasmid #84504, #84501, #84506, #84503).

VEGFA121: the VEGFA121 coding sequence was PCR amplified from pQCXIP-VEGFA121 plasmid (gift from Michael Grusch, Addgene plasmid #73017) and cloned between the HindIII and XhoI sites in pcDNA3.1 to generate *pcDNA3.1_VEGFA121* plasmid.

2.3 sgRNA sequences

Target gene	Transcript ID	sgRNA spacer sequence (5'-3')	Reference
<i>ASCL1</i>	NM_004316	GCAGCCGCTCGCTGCAGCAG	Zetsche et al. 2015
		ATGGAGAGTTTGCAAGGAGC	
		GGCTGGGTGTCCCATGAAA	
		TGTTTATTCAGCCGGGAGTC	
<i>HBG1</i>	NM_000559	GGCTAGGGATGAAGAATAAA	Konermann et al. 2015
		CTTGACCAATAGCCTTGACA	
		AAAATTAGCAGTATCCTCTT	
		GTATCCTCTATGATGGGAGA	
<i>IL1B</i>	NM_000576	TTAGTATATGTGGGACAAAG	
		GAAAATCCAGTATTTTAATG	
		CTCTGGTTCATGGAAGGGCA	
		AGTATTGGTGGAAGCTTCTT	
<i>IL-2</i>	NM_000586	ACATCCATTGAGTCAGTCTT	
		ACCCCCAAAGACTGACTGAA	
		GTGGGCTAATGTAACAAAGA	
<i>IFNγ</i>	NM_000619	AACTAAGGTTTTGTGGCATT	
		AAGATGAGATGGTGACAGAT	
		TCTCATCGTCAAAGGACCCA	
<i>MIP-1α</i> (<i>CCL3</i>)	NM_002983	TAGCTCAAAGATGCTATTCT	
		TCAGGGTCCCTGGTGACCAC	
		TTGGATATCCTGAGCCCCTG	
<i>TNFα</i>	NM_000594	GAGAAACCCATGAGCTCATC	
		GGGCCCTGCACCTTCTGTCT	
		TTTCTTCTCCATCGCGGGGG	
<i>TSP-1</i> (<i>THBS1</i>)	NM_003246	AAAGTGAAGGGGGCGGGGGT	
		GCGGGAGGTGGGGGCCAGTC	
		TAGCTGGAAAGTTGCGCGCC	
<i>INS</i>	NM_001042376	GGGGCTGAGGCTGCAATTTT	Giménez et al. 2016
		CCAGCACCAGGGAAATGGTC	
		CTAATGACCCGCTGGTCCTG	
		AGGTCTGGCCACCGGGCCCC	

Table 2 | List of spacer sequences for targeting endogenous genes with dCas9-TF.

2.4 Cell lines and culture conditions

HEK293T cells were purchased from ATCC (ATCC-CRL-11268) and cultured in Dulbecco's modified Eagle's medium (DMEM, Cat. #41966052, Gibco) supplemented with 15% (v/v) FBS (Cat. #10500064, Gibco), 100 U/ml penicillin and 100 µg/ml streptomycin (Cat. #15140122, Gibco) (HEK293T full media). HTLA cells (HEK293 cell line stably expressing a tTA-dependent luciferase reporter and β -arrestin2-TEV fusion protein) were a gift from Bryan Roth. HTLA cells were maintained in DMEM supplemented with 10% (v/v) FBS, 100 U/ml penicillin, 100 µg/ml streptomycin, 2 µg/ml puromycin (Cat. #A1113803, Gibco), and 100 µg/ml hygromycin B (Cat. #10687010, Gibco) (HTLA full media). Both cell lines were cultured at 37°C and 5% CO₂, and passaged every 2 days at a 1:6 ratio for 2-3 months before being replaced with a new batch. Cells were infrequently tested for mycoplasma contamination using the Venor®GeM OneStep Mycoplasma Detection Kit (Cat. #11-8025, Minerva Biolabs).

2.5 Transient transfection protocols

HEK293T or HTLA cells were seeded in 24-well plates (reporter activation assay) or 12-well plates (endogenous gene activation assays and confocal microscopy) and transfected next day at 80-90% confluency (or approximately 70% for confocal microscopy). All transfections were performed with Polyethylenimine (PEI Sigma-Aldrich 1 mg/ml) as previously described (Aricescu et al. 2006), unless otherwise stated. Briefly, plasmids were mixed in either 50 or 100µl Opti-MEM (Cat. #31985047, Gibco) for 24-well and 12-well plate transfections, respectively, and PEI was added proportional to the total amount of DNA as follows: *i)* for experiments in 24-well plates, if the total plasmid concentration was $\leq 600\text{ng}$, $\leq 800\text{ng}$, $>800\text{ng}$, the transfection mixtures were supplemented with 1.5µl PEI, 2µl PEI, and 2.5µl PEI, respectively; *ii)* for

experiments in 12-well plates, a total plasmid amount of 1µg was used and supplemented with 3µl PEI. Within each experiment, the same amount of DNA was maintained across all conditions by supplementing the transfection mix with pcDNA3.1 where necessary. A detailed description of the DNA constructs and corresponding amounts used for each transfection reaction is provided in the Appendix.

Transfection mixtures were vortexed for 10 seconds and incubated at room temperature for 20-30 minutes. Full media was removed from cells and replaced with experiment-specific transfection media prior to adding the DNA:PEI transfection mix as follows.

For dCas9(N/C)-synBDKBR2 experiments, HTLA cells were transfected in DMEM + 2% (v/v) FBS supplemented with bradykinin at indicated concentrations. The bradykinin transfection media was replaced after 20 hours with HTLA full media also supplemented with bradykinin, and incubated for an additional 24 hours.

For dCas9(N/C)-synBDKRB2^{(-)VP64} and dCas9(N/C)-synBDKRB2^{DHFR} AND gate experiments, the bradykinin-containing transfection media and bradykinin-containing HTLA full media were also supplemented with trimethoprim (TMP) at indicated concentrations. All other 2-input Boolean AND gate configurations involving two dCas9-synGPCRs, the bradykinin-containing transfection media and bradykinin-containing HTLA full media were additionally supplemented with respective second agonist at indicated concentrations. 3-input Boolean AND gate configuration involving dCas9(N)-synBDKRB2^{RI} and dCas9(C)-synCXCR4^{(-)VP64}, the bradykinin-containing transfection media was supplemented with SDF-1α and bradykinin-containing HTLA full media was supplemented with SDF-1α and rapamycin at indicated concentrations.

For dCas9(N/C)-synAVPR2 and dCas9(C)-synCXCR4 experiments, the same transfection conditions were used, but the transfection media and HTLA full media were

supplemented with vasopressin or recombinant SDF-1 α at indicated concentrations, respectively.

For dCas9(N/C)-synLPA α experiments, HTLA cells were transfected in DMEM supplemented with 1% (w/v) fatty acid free BSA (Cat. #A8806, Sigma) containing LPA at indicated concentrations. The LPA transfection media was replaced after 20 hours with DMEM 1% (w/v) fatty acid free BSA containing LPA, and incubated for an additional 24 hours. 3-input Boolean AND gate configuration involving dCas9(N)-synBDKRB2^{RI} and dCas9(C)-synLPA α 1^{(-)VP64}, the LPA-containing LPA transfection media was supplemented with bradykinin. The LPA-containing DMEM supplemented with 1% (w/v) fatty acid free BSA was additionally supplemented with bradykinin and rapamycin at indicated concentrations.

For dCas9(N/C)-synT1R experiments involving D-glucose agonist only, HTLA cells were transfected in DMEM (no glucose, no glutamine, no phenol red, Cat. #A1443001, Gibco) supplemented with 5 mM L-glutamine (Cat. #25030, GIBCO) and 2% (v/v) FBS containing D-glucose at indicated concentrations. The D-glucose transfection media was replaced after 20 hours with DMEM (no glucose, no glutamine, no phenol red), 5 mM L-glutamine and 10% (v/v) FBS containing D-glucose, and incubated for an additional 24 hours. For the dCas9(N/C)-synT1R experiment involving dietary sugars, HTLA cells were transfected in DMEM (no glucose, no glutamine, no phenol red), supplemented with 5 mM L-glutamine and 2% (v/v) FBS. 6 hours post transfection, D-glucose-free transfection media was replaced with DMEM (no glucose, no glutamine, no phenol red), 5 mM L-glutamine and 10% (v/v) FBS and incubated for another 18 hours. Dietary sugar stock solutions in 1x PBS (final concentration of ~100 mM) were added to replenished DMEM (no glucose, no glutamine, no phenol red), 5 mM L-glutamine and 10% (v/v) FBS and incubated for an additional 48 hours.

For all soluble dCas9, Tmt-dCas9, and dCas9-synVEGFR experiments (including the MESA comparison), HEK293T cells were transfected in DMEM + 2% (v/v) FBS and this media was replaced after 20 or 24 hours with HEK293T full media for an additional 24 hours. For dCas9(N/C)-synVEGFR1/2^{RI} AND gate experiments, the HEK293T full media added after transfections also contained rapamycin at indicated concentration.

For Tmt-NLS-dCas9^{VP64[Dox]} experiments using the stable HEK293T cell line, transfections were performed in DMEM + 2% (v/v) FBS. Transfection media was changed after 24 hours to HEK293T full media supplemented with 2 µg/ml puromycin and doxycycline at indicated concentrations for 24 hours, and replaced again with the same media for another 24 hours.

For all experiments involving the HEK293T_dEGFP stable cell line, transfections were performed in DMEM + 15% (v/v) FBS with Lipofectamine 3000 (Cat. #L3000015, Thermo Fisher Scientific), following manufacturer's instructions. Media was replenished daily with HEK293T full media.

For all confocal microscopy experiments, HEK293T cells were directly processed for antibody staining 24 hours after addition of DNA:PEI transfection mixtures to cells in DMEM + 2% (v/v) FBS transfection media.

2.6 Flow cytometry experiments and data analysis

For all EYFP reporter experiments, media was removed 44, 48, or 72 hours post-transfection and cells were washed with PBS (1x phosphate buffer saline), trypsinized (0.05% trypsin-EDTA, Cat. #25300062, Gibco), and kept in 1x PBS on ice. Flow cytometry measurements were carried out within 30-60 min from harvest on a BD LSR Fortessa Analyzer (BD Biosciences). The laser configurations and filter sets were maintained across experiments. Forward scatter and side scatter were used to identify

the cell population and subsequently live single cells. 100,000 total events were recorded for each condition. Data was analysed and compensated using the FlowJo package (FLOWJO LLC). To calculate an EYFP activation score which integrates both reporter fluorescence intensity and % of activated cells the following formula was used as previously described (Z. Xie et al. 2011).

$$(\%EYFP^{+ve} \times EYFP^{mean}) / (\%mCherry^{+ve} \times mCherry^{mean})$$

The numerator $(\%EYFP^{+ve} \times EYFP^{mean})$ provides a weighted mean fluorescence accounting both for the strength of reporter activation ($EYFP^{mean}$) as well as population level activation ($\%EYFP^{+ve}$), which penalizes OFF-state leakage. Since both values are calculated from the parent population (viable single cells) without gating on $mCherry^{+ve}$ cells, the same formula is applied to mCherry for the denominator in order to control for variation in transfection efficiency ($\%mCherry^{+ve}$) and sgRNA levels ($mCherry^{mean}$) between conditions. The fluorescence compensation protocol and the gating strategy used for calculating the EYFP activation score are provided in Figure 3.2.

It should be noted that due to intrinsic experimental variations (e.g., timing, total amount of plasmids transfected, cell density range, cell passage number) absolute values should only be compared within the same experiment. The variations in fold change activation scores observed for certain constructs is imputable to extremely low (near zero) $\%EYFP^{+ve}$ cells in the OFF-state conditions. For the dose-response curves (bradykinin, LPA, SDF-1 α , and vasopressin), the lowest concentration plotted represents the no-agonist condition.

2.7 Direct flow cytometry

Transiently transfected HEK293T cells were harvested, washed in 1x PBS and fixed in 1.6% paraformaldehyde (Cat. #15710, Electron Microscopy Sciences) for 10 minutes

at RT. After a wash with FACS buffer (5% FBS in 1x PBS), cells were stained with 2 μ l of VEGFR antibodies (PE-conjugated anti-VEGF R1 (clone 49560) and VEGF R2 (clone 89106), Cat. #FABSP3P, R&D systems) in 50 μ l FACS buffer for 30 min on ice. Before flow cytometry analysis, cells were washed again twice in FACS buffer.

2.8 RT-qPCR analysis

For quantification of endogenous genes expression, transiently transfected cells were harvested 44 or 48 hours post-transfection, washed twice in 1x PBS and total RNA was extracted using either RNeasy Mini Kit (Cat. #74106, Qiagen) or EZNA Total RNA Kit I (Cat. #R6834, Omega) following manufacturer's instructions. Complementary DNA (cDNA) was prepared from 1 μ g of total RNA using the QuantiTect Reverse Transcription Kit (Cat. #205313, Qiagen). Quantitative PCR (qPCR) was carried out using the SsoAdvanced™ Universal SYBR® Green Supermix kit (Cat. #1725272, Bio-Rad) on a CFX384 real-time system (Bio-Rad). Each reaction was run in technical triplicates. In the absence of a relevant PCR product (based on melt curve analysis), values were set to a max. Ct of 40 cycles.

Data was analysed using the $\Delta\Delta$ Ct method as previously described (Ferry et al. 2017). Δ Ct was calculated using the house keeping gene GAPDH to control for number of cells ($\text{GOI transcript levels} = 2^{(\text{Ct}_{\text{GAPDH}} - \text{Ct}_{\text{GOI}})}$). Before calculating the $\Delta\Delta$ Ct, the GOI transcript levels were normalized to EGFP/dCas9(C) Δ Ct for the same condition to account for variations in transfection efficiency ($\text{GOI normalized transcript levels} = 2^{(\text{Ct}_{\text{GAPDH}} - \text{Ct}_{\text{GOI}})} / 2^{(\text{Ct}_{\text{GAPDH}} - \text{Ct}_{\text{EGFP/dCas9(C)}})}$). $\Delta\Delta$ Ct values for each condition were then calculated and normalized to $\Delta\Delta$ Ct in the control conditions (untreated scramble sgRNA) using the formula below (e = experiment (GOI) and c = control (untreated

scramble sgRNA, except for Figure 3F where the untreated GOI sgRNAs were used instead)).

EGFP as control for transfection efficiency was only employed for soluble dCas9^{VP64} experiments in Figure 3.3.

Fold change

$$= 2^{(Ct_{GAPDH}^e - Ct_{GOI}^e) - (Ct_{GAPDH}^e - Ct_{EGFP/dCas9(C)}^e)} / 2^{(Ct_{GAPDH}^c - Ct_{GOI}^c) - (Ct_{GAPDH}^c - Ct_{EGFP/dCas9(C)}^c)}$$

qPCR primer	Sequence (5'-3')	Reference
ASCL1 fwd	CCCCAACTACTCCAACGACT	This thesis
ASCL1 rev	GGTGAAGTCGAGAAGCTCCT	
EGFP fwd	CAACTACAAGACCCGCGC	
EGFP rev	CTTGTGCCCCAGGATGTTG	
GAPDH fwd	AACAGCGACACCCACTCCTC	Ferry et al. 2017
GAPDH rev	CATACCAGGAAATGAGCTTGACAA	
HBG1 fwd	GTTGTCTACCCATGGACCCA	
HBG1 rev	TCTCCCAAGGAAGTCAGCAC	
IL1B fwd	CGAATCTCCGACCACCACTA	
IL1B rev	AGGGAAAGAAGGTGCTCAGG	
dCas9(C) fwd	AACCTATGCCCACCTGTTTCG	
dCas9(C) rev	ATCCAGGATTGTCTTGCCGG	
IL-2 fwd	ACCAGGATGCTCACATTTAAGTTTT	Mocellin et al. 2003
IL-2 rev	GAGGTTTGAGTTCTTCTTAGACAC TG	
IFN γ fwd	CCAACGCAAAGCAATACATGA	Body-Malapel et al. 2008
IFN γ rev	CCTTTTTCGCTTCCCTGTTTTA	
MIP-1 α (CCL3) fwd	GGCTCTCTGCAACCAGTTCT	C. Lin et al. 2013
MIP-1 α (CCL3) rev	TGAAATTCTGTGGAATCTGCC	
INS fwd	ATCAGAAGAGGCCATCAAGCA	Giménez et al. 2016
INS rev	TAGAGAGCTTCCACCAGGTGTGA	
TNF α fwd	CCCAGGGACCTCTCTAATCA	Chiu & N.-S. Yang 2007
TNF α rev	AGCTGCCCTCAGCTTGAG	

TSP-1 (THBS1) fwd	ACATGCCACGGCCAACAAA	Ottow et al. 2014
TSP-1 (THBS1) rev	AGTGGCCCAGGTAGTTGCACTT	

Table 3 | List of forward and reverse primers used for qPCR.

2.9 T7 endonuclease I (T7E1) assay

Transiently transfected HEK293T_dEGFP cells after flow cytometry analysis were pelleted and washed once in 1x PBS. Genomic DNA was extracted by boiling cells for 10 min at 95°C in 100 µl of 50 mM NaOH, then neutralised with 11 µl of 1 M TRIS-HCl (pH 8.0) and centrifuged at 13,000 g for 3 min. 2 µl of the resulting supernatant was used for the targeted amplification step. PCR was performed in 50 µl total reaction volume with Phusion High-Fidelity PCR Master Mix with GC Buffer (Cat. #M0532, NEB) and with a final concentration of 0.4 µM forward primer 5'-ACCTGAAATGACCCTGCGCCTTATTTG-3' and 0.4 µM reverse primer 5'-GGTGTCTGCTGGTAGTGGTCGGC-3'. PCR programme was as follows: 95°C for 3 min, 35x (95°C for 30 sec, 65°C for 25 sec, 72°C for 55 sec), 72°C for 5 min, hold at 12°C. PCR reactions were purified with a MinElute PCR Purification Kit (Cat. #28006) according to manufacturer protocol. Subsequently, purified PCR products were reannealed by a programme consisting of the following steps: 95°C for 5 min, then gradual ramping down at 0.1°C/1 sec until reaching 25°C. A fraction was digested for 30 min at 37°C with T7 Endonuclease I (Cat. #M0302, NEB) according to manufacturer's instructions. After inactivation with a final concentration of 25 mM EDTA, aliquots were run on a 2% agarose gel (UltraPure Agarose, Cat. #16500500, ThermoFisher Scientific) supplemented with GelRed (Cat. #730-2958 (41003), Biotium) for 50 min at 150V. Gel was exposed in a GelDoc XR+ (BioRad) chamber and an image was acquired with Image Lab 5.2.1 software (BioRad).

2.10 Confocal microscopy

HEK293T cells were transiently transfected on round coverslips (Cat. #631-1577, VWR), washed twice in 1x PBS, fixed in 4% paraformaldehyde (Cat. #15710, Electron Microscopy Sciences) for 3 min at room temperature and incubated overnight in 100% EtOH at -20°C. EtOH was then removed, cells were briefly washed in washing buffer (1x TBS, 0.2% Triton X-100, 0.04% SDS) and incubated for 1 hour at room temperature in blocking buffer (1.5 % BSA in 1x TBS). Polyclonal rabbit HA (Cat. #A190-108A, Bethyl Laboratories Inc.) and monoclonal mouse c-myc (9E 10-c, Developmental Studies Hybridoma Bank) primary antibodies were added to cells in blocking buffer and incubated for 1 hour at room temperature. Cells were washed three times in washing buffer and incubated for 1 hour in blocking buffer containing secondary antibodies goat anti-rabbit A488 (Cat. #A-11008, Thermo Fisher Scientific), goat anti-mouse A568 (Cat. #A-11004, Thermo Fisher Scientific), DAPI (Cat. #D1306, Invitrogen) and, if applicable, Alexa Fluor 647 Phalloidin (Cat. no. A22287, Thermo Fisher Scientific). Coverslips were washed three times in washing buffer and mounted with SlowFade Diamond Antifade Mountant (Cat. #S36972, Life Technologies). Images were acquired on a Zeiss LSM 780 Inverted confocal microscope with an oil immersion objective (Plan-Apochromat 63x/1.4 Oil DIC M27, Zeiss) at non-saturating parameters and processed using the ImageJ package.

2.11 Viral particle production and cell line generation

HEK293T cells were transfected in DMEM + 15% FBS with pCMV-dR8.91 and pMD2.G (gifts from Thomas Milne), and lentiviral backbone constructs (TMT-NLS-dCas9^{VP64[Dox]} or dEGFP) at a ratio of 1:1:1.5 using Lipofectamine 2000 (Cat. #11668027, Thermo Fisher Scientific). After 24 hours, media was replaced with HEK293T full media. After

another 24 hours, the supernatant containing lentiviral particles was collected, passed through a 0.22 μm filter (Cat. #10268401, Millipore) and added to low passage HEK293T at low multiplicity of infection (MOI).

For the TMt-NLS-dCas9^{VP64[Dox]} HEK293T cell line generation, 2 days after infection HEK293T full media supplemented with 5 $\mu\text{g/ml}$ puromycin was added, transduced cells were passaged three times and then maintained in HEK293T full media supplemented with 2 $\mu\text{g/ml}$ puromycin. Cells were treated with 1 $\mu\text{g/ml}$ doxycycline, indirectly stained with c-myc primary antibody and goat anti-mouse A568 secondary antibody to identify TMt-NLS-dCas9^{VP64[Dox]} expressing cells, and sorted as single cells into Terasaki plates (Cat. #653180, Greiner Bio-One). One clone displaying the most stringent doxycycline-dependent expression was chosen for subsequent experiments. For HEK293T_dEGFP cell line generation, transduced cells were passaged three times and then maintained in HEK293T full media. dEGFP expressing cells were FACS sorted as single cells into Terasaki plates and subsequently expanded. One clone displaying the most homogenous dEGFP expression signature was chosen for subsequent experiments.

Chapter 3 : Establishing split transmembrane-tethered dCas9^{VP64} as signal transduction module

Most of the data presented in this chapter has been published in: Baeumler, T.A., Ahmed, A.A. & Fulga, T.A., 2017. Engineering Synthetic Signaling Pathways with Programmable dCas9-Based Chimeric Receptors. Cell Reports, 20(11), pp.2639–2653.

3.1 Chapter Aims

- Investigating the effects of different localisation tags on the potency of dCas9^{VP64} transcriptional activators.
- Establishing an artificial transmembrane-tethered (TMt) dCas9^{VP64} (TMt-dCas9^{VP64}) scaffold, which recapitulates an intracellular receptor-like bound transcription factor.
- Implementing a TEV protease-mediated release mechanism for TMt-dCas9^{VP64}.
- Demonstrating that full length TMt-dCas9^{VP64} and TMt-Cas9 scaffolds do not provide stringent OFF state characteristics.
- Determining and validating a split TMt-dCas9^{VP64} signal transduction module with robust OFF/ON state transition properties.

3.2 Results

3.2.1 Impact of localisation tags on the transcriptional potency of soluble dCas9^{VP64}

To successfully implement dCas9^{VP64} as potent signal transduction modality in a receptor-like setting, dCas9^{VP64} has to transition from a cytoplasmic, transcriptionally

inactive factor (OFF state) into a nuclear-localised, potent transcription factor (ON state). Because Cas9 originates from prokaryotes (Chapter 1.1.3.1) that do not possess a nucleus, I first investigated the effect of localisation tags on dCas9^{VP64} to modulate its transcriptional potency in eukaryotic cells.

To this end, I fused to the N- and C-terminus of dCas9 either nuclear localisation sequences (NLS) or nuclear export sequences (NES, amino acid sequence: LDLASLIL, (Ossovskaya et al. 2008)) (Figure 3.1a,d). I used either a strong NLS (designated as NLS+, amino acid sequence: PKKKRKVGIIHGVPAA, of unknown origin, reported in Mali et al. 2013) (Figure 3.1a) or a standard NLS (referred to as NLS, monopartite NLS from SV40 Large T antigen, amino acid sequence: PKKKRKV) (Kalderon et al. 1984; Makkerh et al. 1996; Ray et al. 2015) (Figure 3.1b). In addition, I created a dCas9 lacking any localisation tag (Figure 3.1c). In all instances, dCas9^{VP64} also contained an HA-epitope tag for subcellular visualisation.

To qualitatively assess the localisation pattern of designated dCas9^{VP64} constructs, I transiently transfected HEK293T cells and used an anti-*HA* antibody to stain for HA-dCas9 variants as well as the nucleus of cells (DAPI stain) and F-actin in the cytoplasm (Phalloidin). Anti-*HA* immunofluorescence revealed that NLS(+)-dCas9^{VP64} is exclusively found in the nucleus whereas NLS-dCas9^{VP64} is distributed both in the nucleus as well as in the cytoplasm (Figure 3.1a,b). In contrast, dCas9^{VP64} and NES-dCas9^{VP64} variants are solely found in the cytoplasm (Figure 3.1c,d). The anti-*HA* immunofluorescence staining was specific to the HA- epitope tag since the control dCas9^{VP64} (Mali et al. 2013) without an HA-tag did not produce any signal (Figure 3.1e).

For a robust assessment of the transcriptional activation potential of differently localised dCas9^{VP64} variants, I utilised a previously validated fluorescence reporter assay

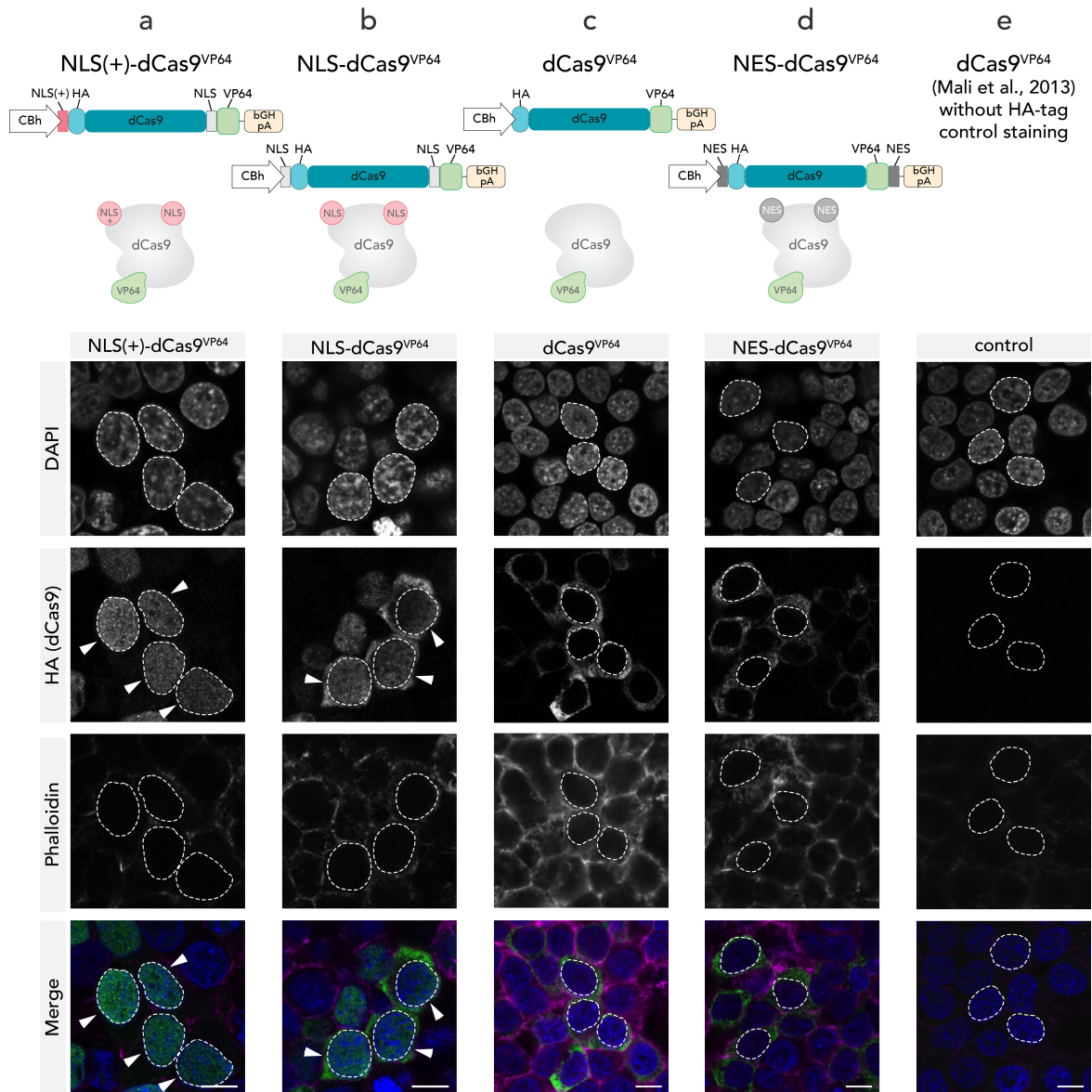


Figure 3.1 | Subcellular localisation of soluble dCas9^{VP64} variants. (a-d) Top: Molecular structures and cartoons of NLS(+)-dCas9^{VP64} (a), NLS-dCas9^{VP64} (b), dCas9^{VP64} (c), NES-dCas9^{VP64} (d), and control dCas9^{VP64} lacking an HA-tag (e). **Bottom:** Confocal images of HEK293T cells that were transiently transfected with dCas9^{VP64} variants and stained with anti-HA antibody. NLS(+)-dCas9^{VP64} shows exclusive nuclear localisation, NLS-dCas9^{VP64} cytosolic and nuclear localisation, and dCas9^{VP64} and NES-dCas9^{VP64} exclusive cytosolic localisation. Control dCas9^{VP64} lacking an HA-tag did not result in positive staining, confirming antibody specificity. DAPI = nuclear staining, indicating the nuclear compartment, Phalloidin = F-actin, indicating the cytoplasm. White dotted lines outline nuclei. White arrows indicate nuclear localisation of NLS-tagged dCas9^{VP64} variants. Scale bar = 10µm.

(Farzadfard et al. 2013; Nissim et al. 2014). This assay deploys a customised plasmid that encodes for a conditionally activatable EYFP fluorescence protein (Figure 3.2a). Upstream of the EYFP is a minimal CMV promoter that is regulated by a synthetic operator region containing a repetitive motif. This repeat sequence is complementary to the spacer sequence of a single guide RNA (sgEYFP). When co-transfected, dCas9^{VP64} forms a complex with sgEYFP and targets the synthetic operator region thereby eliciting EYFP transcription. As a transfection control, the sgRNA encoding plasmid also contains a constitutively expressed mCherry fluorescent protein (Figure 3.2a). 48 hours post transfection of all system components, HEK293T cells were subjected to flow cytometry (Figure 3.2b). First, the viable cell population was identified (FSC-A vs SSC-A) and single cells isolated (FSC-A vs FSC-H). Then, transfected cells were visualised and gated on using a flow cytometry scatter plot with the transfection control mCherry displayed on the x-axis and the experimental variable EYFP on the y-axis (Figure 3.2b). This gating strategy allowed for the subsequent calculation of a designated activation score that takes into account fluorescent protein expression (percent positive cells) and intensity (mean fluorescence intensity) of both EYFP and mCherry (Figure 3.2c). When all system components (EYFP reporter plasmid, dCas9^{VP64} plasmid, and the sgEYFP plasmid) were co-delivered to the same cells, strong EYFP fluorescence was detected (Figure 3.2c,d). However, if one of the components was missing or a non-targeting, scramble sgRNA (sgSCR) was supplied, no EYFP was detected and the activation score was near zero (Figure 3.2c,d). These results confirm the suitability of this assay for the assessment and quantification of dCas9^{VP64}-induced transcriptional activation.

Notably, two experimental practicalities had to be considered. First, to maintain the total amount of transfected plasmid DNA for each condition within the same

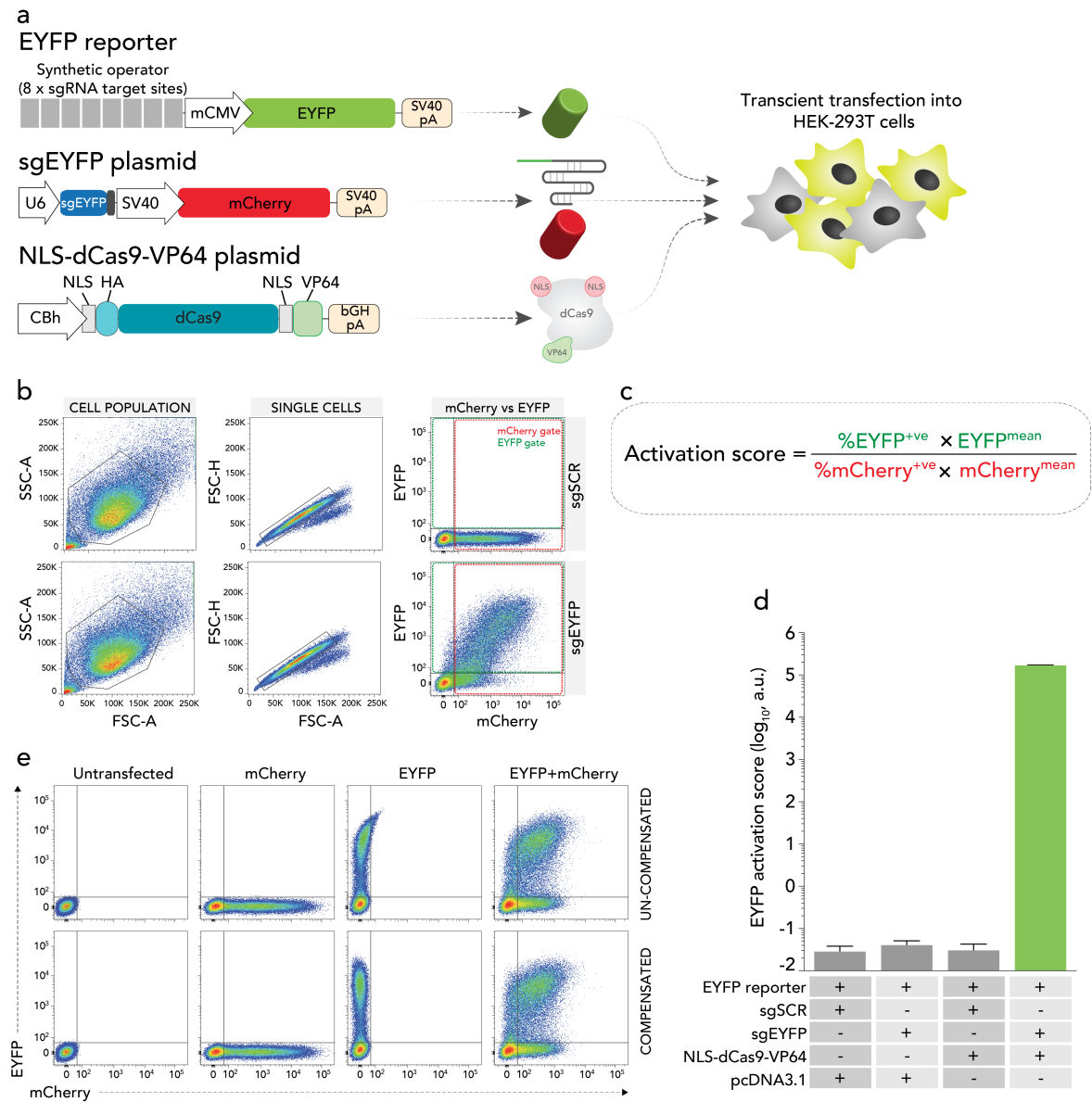


Figure 3.2 | EYFP reporter assay and quantification. (a) Schematic representation of the EYFP reporter assay. HEK293T cells were transiently co-transfected with the EYFP reporter plasmid, a plasmid expressing an EYFP-targeting sgRNA (sgEYFP) together with the transfection control mCherry, and a dCas9^{VP64} variant. sgEYFP directs dCas9^{VP64} to the minimal promoter of the reporter thereby inducing the transcription of EYFP. (b) Flow cytometry gating strategy. Depicted flow cytometry scatter plots are representative of the last two conditions in (d). Only in the presence of all system components (EYFP reporter, sgEYFP, and NLS-dCas9^{VP64}), EYFP reporter is detectable 48 hours post transfection. (c) EYFP activation score formula adapted from Xie et al. (2011). (d) Activation score quantification and reporter assay validation. Only in the presence of all system components, EYFP reporter fluorescence is detected. (e) Flow cytometry compensation strategy for dual fluorophore (mCherry/EYFP) analysis.

experiment, an empty cloning plasmid (pcDNA3.1) was added when necessary (Figure 3.2d). Second, EYFP fluorescence had to be compensated for to avoid false positive signal because very high EYFP emission values were detectable in the mCherry detection channel (bleed-through) (Figure 3.2e).

Having established the EYFP reporter assay, I investigated in a head-to-head comparison the transcriptional activation potential of differently localised dCas9^{VP64} variants (Figure 3.3a-d). As expected, no EYFP activation was detectable in the presence of control sgSCR (Figure 3.3a-e). Surprisingly, however, all dCas9^{VP64} variants in combination with sgEYFP induced very strong and robust EYFP expression (Figure 3.3a-e). This suggests that the subcellular localisation of transiently transfected dCas9^{VP64} has a limited impact on its transcriptional potency.

To substantiate these EYFP reporter assay studies, I then programmed dCas9^{VP64} variants to target an endogenous gene. For all testing and optimisation experiments carried out in this thesis, I focused on *ASCL1* as a preferred gene for several reasons: (i) *ASCL1* is expressed at very low levels in HEK293/HEK293T cells; (ii) sgRNAs capable of mediating strong transcriptional upregulation of *ASCL1* have been tested and optimised by several groups; and (iii) *ASCL1* is routinely used in the CRISPR field for testing novel dCas9-mediated transcriptional activation systems (Perez-Pinera et al. 2013; Chavez et al. 2015; Nihongaki, Yamamoto, et al. 2015; Konermann et al. 2015; Zetsche, Volz, et al. 2015; Chavez et al. 2016; Black et al. 2016; Kleinjan et al. 2017; H.-K. Liao et al. 2017). Notably, the dCas9-TF-facilitated overexpression of *ASCL1* and all other endogenous gene targets throughout this thesis were only assessed on the transcript level.

To target *ASCL1*, I utilised a second generation synthetic transcriptional potentiator, the synergistic activation mediator (SAM) (Konermann et al. 2015). This

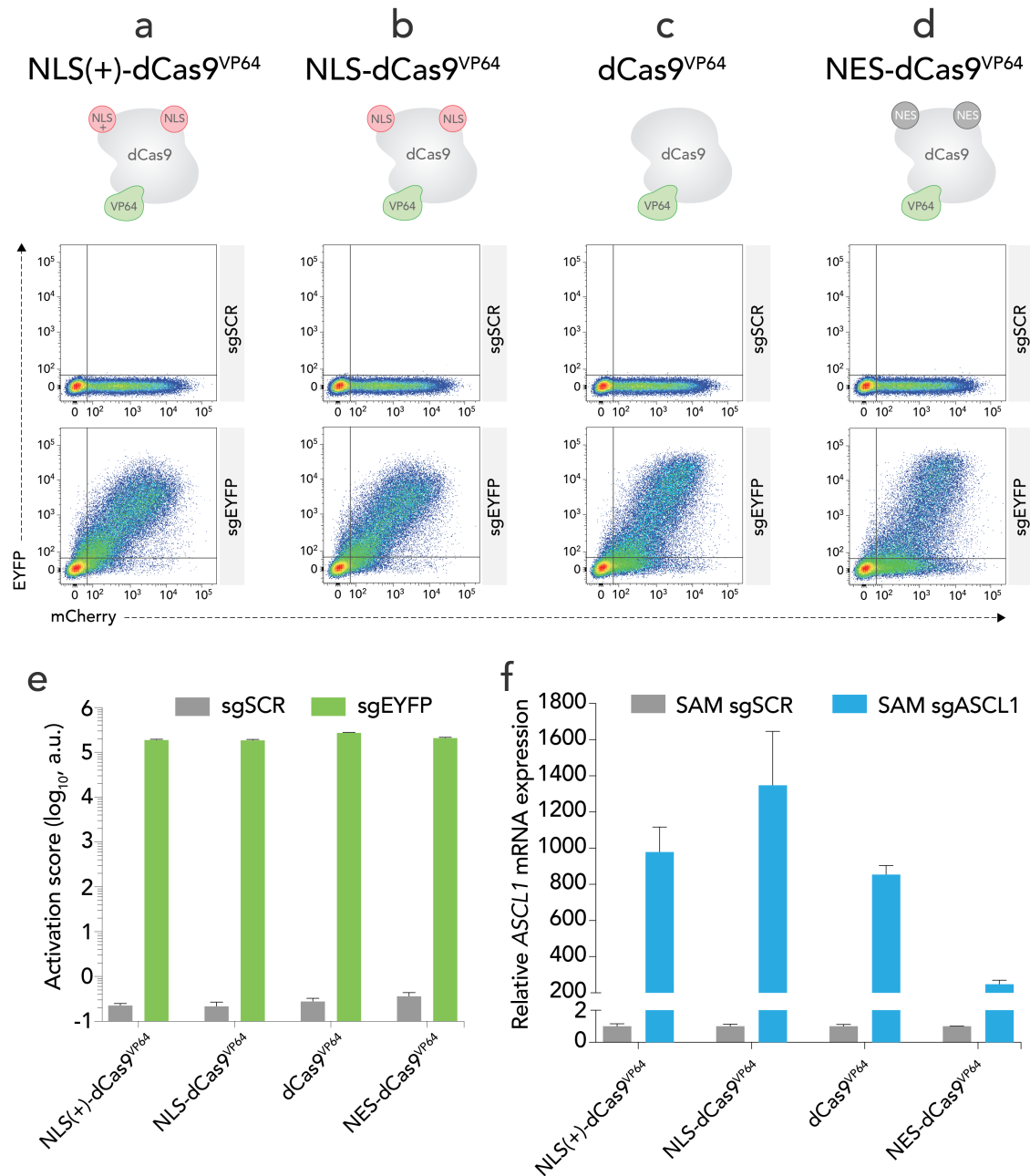


Figure 3.3 | Transcriptional activation potential of differently, subcellular-localised soluble dCas9^{VP64}. (a-e) EYFP activation reporter assay for differently tagged soluble dCas9^{VP64} variants (a-d, top). Plasmids were transiently transfected into HEK293T cells together with the EYFP reporter plasmid and a guide RNA-encoding plasmid, as indicated (sgSCR or sgEYFP). Representative flow cytometry scatter plots are shown (a-d, bottom) as well as quantification of EYFP reporter activation (e). N = 3 biological replicates from one experiment, mean +/- s.d., a.u. arbitrary units; sgSCR = scramble sgRNA control; sgEYFP = artificial EYFP operator region-targeting sgRNA. (f) RT-qPCR of endogenously targeted *ASCL1* in HEK293T cells with different soluble dCas9^{VP64} and SAM system components. Data displayed as fold change relative to each sgSCR condition. N = 3 biological replicates (x3 technical replicates), mean +/- s.d.

system offers two major advantages: (i) SAM generates higher fold change gene induction rates than any other reported system (Konermann et al. 2015; Chavez et al. 2016), a feature that will become important in downstream applications (Chapter 6); (ii) utilising an already established transcriptional activation system highlights the compatibility of my receptor platform and showcases its ease of adaptation and implementation (Chapter 4-6).

To harness the power of the SAM system, I replaced the sgRNA with a SAM sgRNA and co-transfected an additional plasmid, encoding the chimeric activator fusion protein MCP-p65-HSF1 (Figure 1.2b2) (Konermann et al. 2015). The SAM sgRNA contains two MS2 loops that each recruit two engineered MCP-p65-HSF1 activator fusion proteins thereby potentiating the transcriptional power of dCas9^{VP64}.

HEK293T cells transfected with all SAM system components and differently tagged dCas9^{VP64} variants were harvested and their RNA extracted and subjected to RT-qPCR analysis (Figure 3.3f). Both soluble NLS-tagged dCas9^{VP64} variants showed strong *ASCL1* gene induction (~1000-fold for NLS(+)-dCas9^{VP64} and >1300-fold for NLS-dCas9^{VP64}). dCas9^{VP64} upregulated *ASCL1* gene expression by ~850-fold while NES-dCas9^{VP64} displayed the lowest transcriptional activation of ~250-fold. Taken together, these results illustrate that all dCas9^{VP64} variants induce robust endogenous gene expression and that the localisation tags have only a minor impact on overall transcriptional potency.

Two striking conclusions can be drawn: (i) NLS sequences do not improve transcriptional potency compared to untagged dCas9^{VP64}, although NLS-dCas9^{VP64} is preferentially found in the nucleus; (ii) the lack of NLS or the addition of NES sequences are insufficient to efficiently retain dCas9^{VP64} in the cytoplasm and to silence its transcriptional activity. It was previously shown that catalytically active Cas9 nucleases, with or without flanking NLS sequences, display equivalent potency in disrupting

genomic DNA in HEK293T cells (Oakes et al. 2016). My data substantiates this surprising observation, as further detailed in the next subchapter. This phenomenon may be intrinsic to rapidly dividing cells (Oakes et al. 2016). Continuous cell division accompanied by nuclear membrane disintegration may expose genomic DNA to Cas9 cleavage and/or allows Cas9 and dCas9^{VP64} to be trapped in the nucleus during the reassembly of the nuclear envelope.

Importantly, the data on endogenous gene activation (Figure 3.3f) correlates nicely with the EYFP reporter system data (Figure 3.3e), in that all dCas9^{VP64} alternatives perform as strong transcriptional activators. Therefore, for subsequent characterisation, optimisation, and performance evaluations of dCas9^{VP64}-based systems, I utilised the EYFP reporter system.

3.2.2 Engineering a conditional transmembrane-tethered dCas9^{VP64} (TMt-dCas9^{VP64}) scaffold

The observation that Cas9/dCas9^{VP64} can exert its programmed function in the nucleus irrespective of the presence or absence of a localisation sequence may create a significant problem for implementing robust OFF/ON state transition in engineered dCas9-based chimeric receptors. To further investigate this behaviour and model dCas9^{VP64} in a receptor-relevant setting, I conditionally tethered dCas9^{VP64} to the membrane via an artificial transmembrane tether and incorporated a TEV-based release mechanism (Figure 3.4). The generic transmembrane tether (TMt) fused to dCas9^{VP64} facilitates the transport and the integration of dCas9^{VP64} into the cell membrane (OFF state). Only in the presence of the Nla tobacco etch virus (TEV) protease, dCas9^{VP64} is liberated from the membrane via the proteolytic activity exerted on a TMt-engrafted TEV cleavage site (TCS). Subsequently, dCas9^{VP64} can translocate into the nucleus to initiate gene expression (ON state, Figure 3.4).

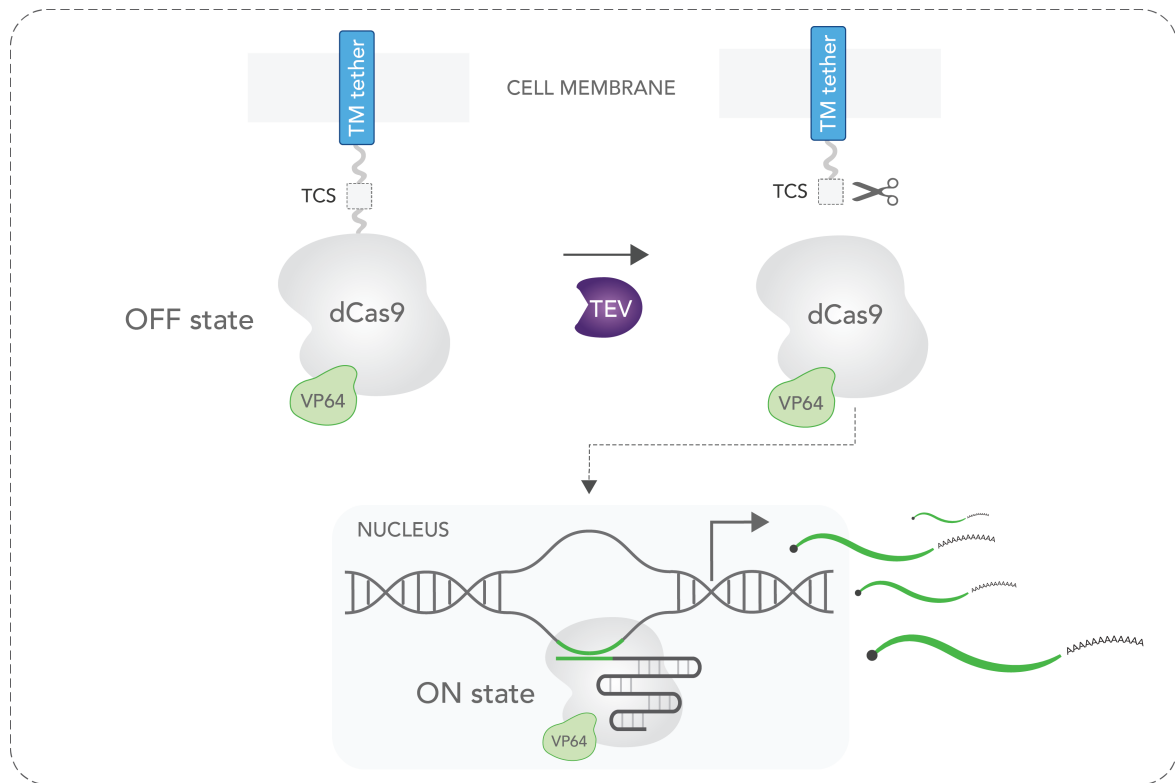


Figure 3.4 | Conceptual overview of transmembrane-tethered dCas9^{VP64} (TMt-dCas9^{VP64}).

A generic transmembrane tether (TMt) integrates dCas9^{VP64} into the membrane thereby rendering it transcriptionally silent. Upon the TEV protease-mediated release from the membrane, dCas9^{VP64} translocates into the nucleus to drive the expression of any customised endogenous gene, guided by its programmable sgRNA.

The endopeptidase TEV (Parks et al. 1995; Kapust et al. 2001; Phan et al. 2002) was chosen for several reasons: (i) it has a well-characterised cleavage site (Dougherty et al. 1988; Dougherty & Dawn Parks 1989; Phan et al. 2002; Kapust et al. 2002); (ii) it is unlikely to have endogenous mammalian target sites, minimising potential cytotoxicity (Wehr et al. 2006; Xi Chen et al. 2010); (iii) it has been successfully harnessed for numerous biotechnological applications (Cesaratto et al. 2016); (iv) it has been effectively utilised in mammalian cell lines and in mice *in vivo* in a receptor-related setting (Barnea et al. 2008; Kroeze et al. 2015; Wehr et al. 2006; Copeland et al. 2016; Wenjing Wang et al. 2017).

The TMt used in these experiments contains the murine IgK-originating leader signal sequence (directs proteins to the secretory pathway), a short flexible (GGGS)₂ linker sequence, a myc-epitope tag (for staining purposes) followed by a platelet-derived growth factor (PDGF) receptor transmembrane domain, and another linker sequence (XTEN) (Schellenberger et al. 2009). The TMt was then appended with the efficiently cleavable canonical TEV cleavage site ENLYFQG (TCS(Q'G)) (Kapust et al. 2001; Kapust et al. 2002), and directly fused to the N-terminus of dCas9^{VP64} variants, as described in Figure 3.3b-d to generate TMt-NLS-dCas9^{VP64}, TMt-dCas9^{VP64}, and TMt-NES-dCas9^{VP64} (Figure 3.5a-c). The standard NLS was chosen in this case because when fused to dCas9^{VP64} it performed slightly better than NLS+ in terms of activation (Figure 3.3e,g). An additional construct was generated which contained an NES sequence upstream of the TCS sequence instead of being fused directly to dCas9^{VP64} (TMt^{NES}-dCas9^{VP64}, Figure 3.5d). The rationale for this configuration is two-fold: (i) it could improve cytoplasmic sequestration since both tethering and the NES target dCas9^{VP64} to the cytoplasm; (ii) following TEV protease-mediated cleavage of the TCS, dCas9^{VP64} is separated from the NES thereby recapitulating soluble, untagged

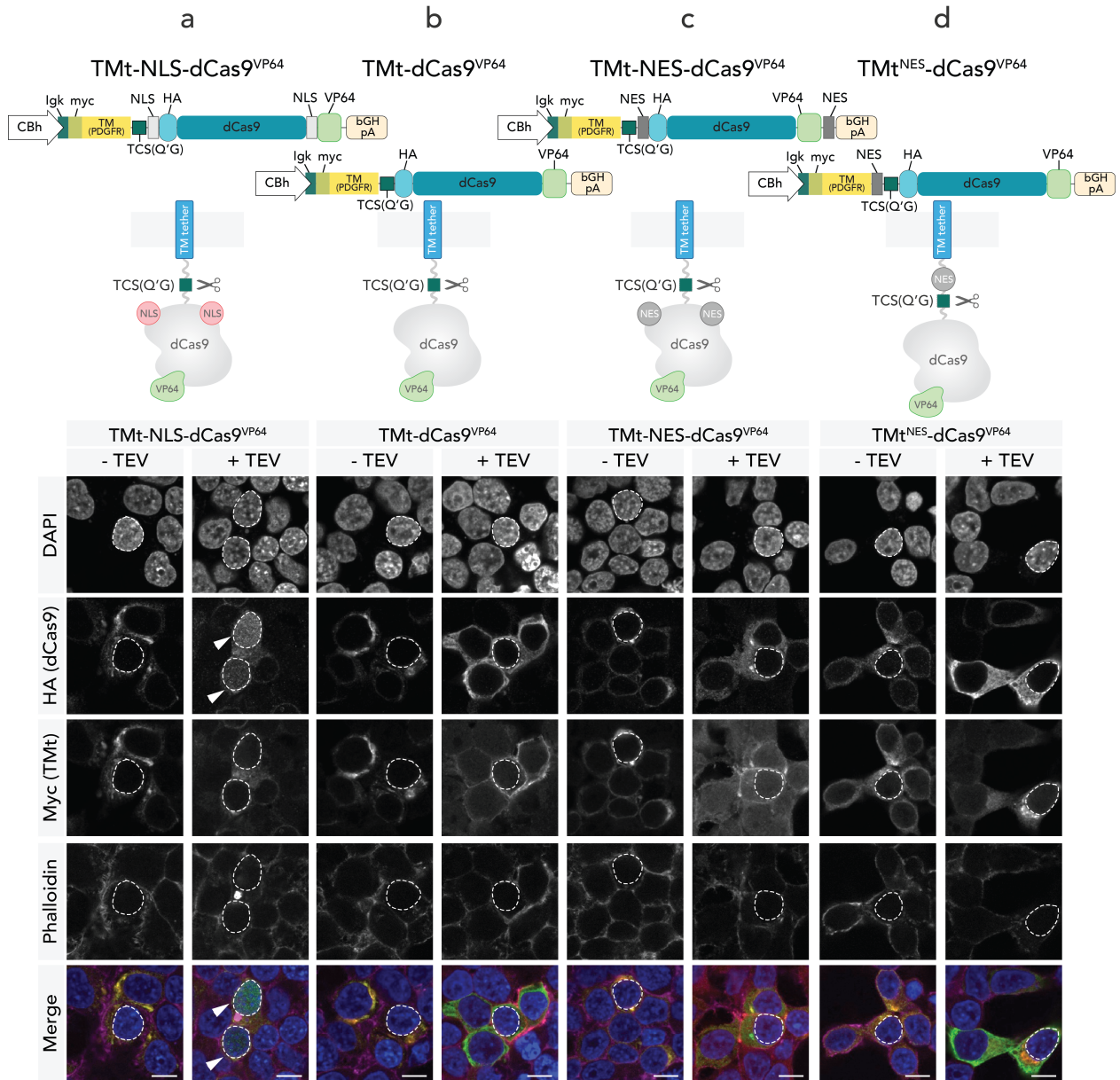


Figure 3.5 | Subcellular localisation of transmembrane-tethered dCas9^{VP64}. (a-d) Top: Molecular structures and cartoons of TMt-NLS-dCas9^{VP64} (a), TMt-dCas9^{VP64} (b), TMt-NES-dCas9^{VP64} (c), and TMt^{NES}-dCas9^{VP64} (d). **Bottom:** Anti-HA and anti-myc confocal microscopy of transiently transfected HEK293T cells with TMt-dCas9^{VP64} variants in the absence or presence of TEV protease. As expected, only NLS-HA-dCas9^{VP64} translocates into the nucleus upon the successful cleavage of the TCS in TMt-NLS-dCas9^{VP64} by TEV protease. Anti-HA staining indicates dCas9^{VP64} location in the cell whereas anti-myc staining visualises TMt-dCas9^{VP64} in the absence of TEV protease or only TMt in the presence of TEV protease. DAPI = nuclear staining, Phalloidin = F-actin. White dotted lines outline nuclei. White arrows indicate nuclear localisation of NLS-tagged dCas9^{VP64}. Scale bar = 10 μm.

dCas9VP64 that showed strong transcriptional potency (Figure 3.3e,g). I reasoned that these features may result in an improved OFF/ON state transition characteristic.

As expected, anti-*HA* and anti-*myc* confocal microscopy revealed that the myc-labelled TMt was only detected in the cytoplasm and/or transmembrane structures in all tested TMt-dCas9^{VP64} configurations, regardless of the presence or absence of TEV protease (Figure 3.5a-d). Similarly, the HA-tagged dCas9^{VP64} variants were all cytoplasmic in both conditions, except for TMt-NLS-dCas9^{VP64}. Only in the presence of TEV protease, the HA-tagged NLS-dCas9^{VP64} was separated from its TMt anchor, allowing NLS-dCas9^{VP64} to translocate into the nucleus, as revealed by strong anti-*HA* and non-detectable anti-*myc* signal in the nucleus (Figure 3.5a).

Confirming the specificity of the antibodies, no cross-reactivity was observed between anti-*HA* and anti-*myc* and no signal was detected in an antibody tag-free control condition (Figure 3.6).

The transcriptional activation potential of each TMt-dCas9^{VP64} configurations was then tested using the EYFP reporter system (Figure 3.7). Strikingly, TMt-NLS-dCas9^{VP64} displayed extremely high EYFP reporter activation, comparable to soluble dCas9^{VP64} variants (Figure 3.3g). Furthermore, TEV protease had apparently no impact on the overall activation potency, as evidenced by the lack of an OFF/ON state transition (Figure 3.7a,e). TMt-dCas9^{VP64} maintained similar absolute activation strength as TMt-NLS-dCas9^{VP64}. However, it generated a two-fold transactivation change in the presence of TEV (Figure 3.7b,e).

As expected (see Figure 3.3g), the overall transcriptional activity of TMt-NES-dCas9^{VP64} was slightly diminished, indicating that the flanking NES sequences on Cas9^{VP64} imposed moderate cytoplasmic sequestration. Nonetheless,

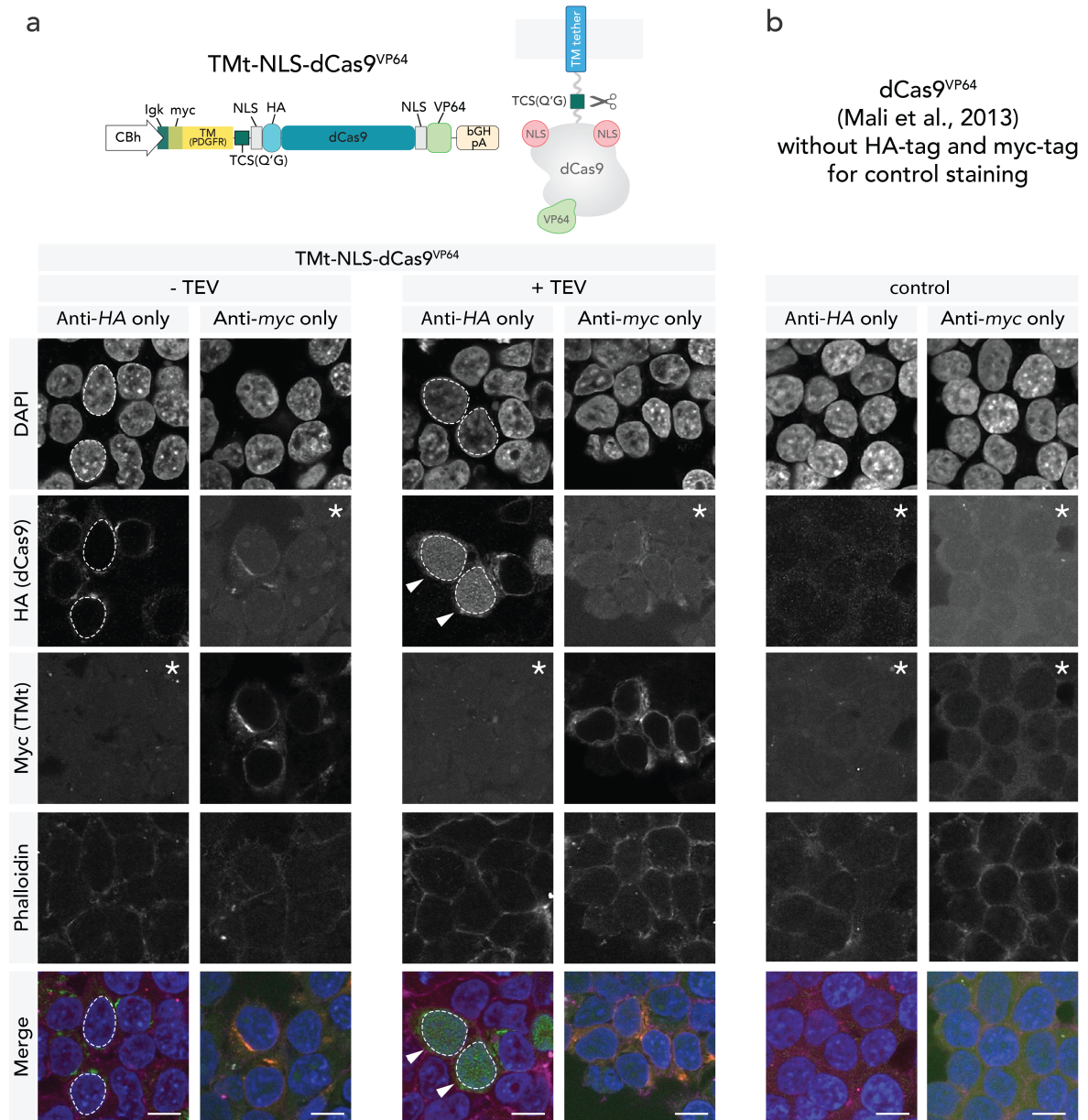


Figure 3.6 | Control antibody staining for anti-*HA* and anti-*myc*. (a) Transiently transfected HEK293T cells with TMt-NLS-dCas9^{VP64} and with or without TEV protease were subjected to either anti-*HA* antibody or anti-*myc* staining. (b) HEK293T cells transiently transfected with dCas9^{VP64} (no HA or myc tag) were also individually treated with either anti-*HA* antibody or anti-*myc* only. There was no crosstalk observed for any antibody combination or condition apart from unspecific background staining (denoted with white asterisks). DAPI = nuclear staining, Phalloidin = F-actin. White dotted lines outline nuclei. White arrows indicate nuclear localisation of NLS-tagged dCas9^{VP64}. Scale bar = 10µm.

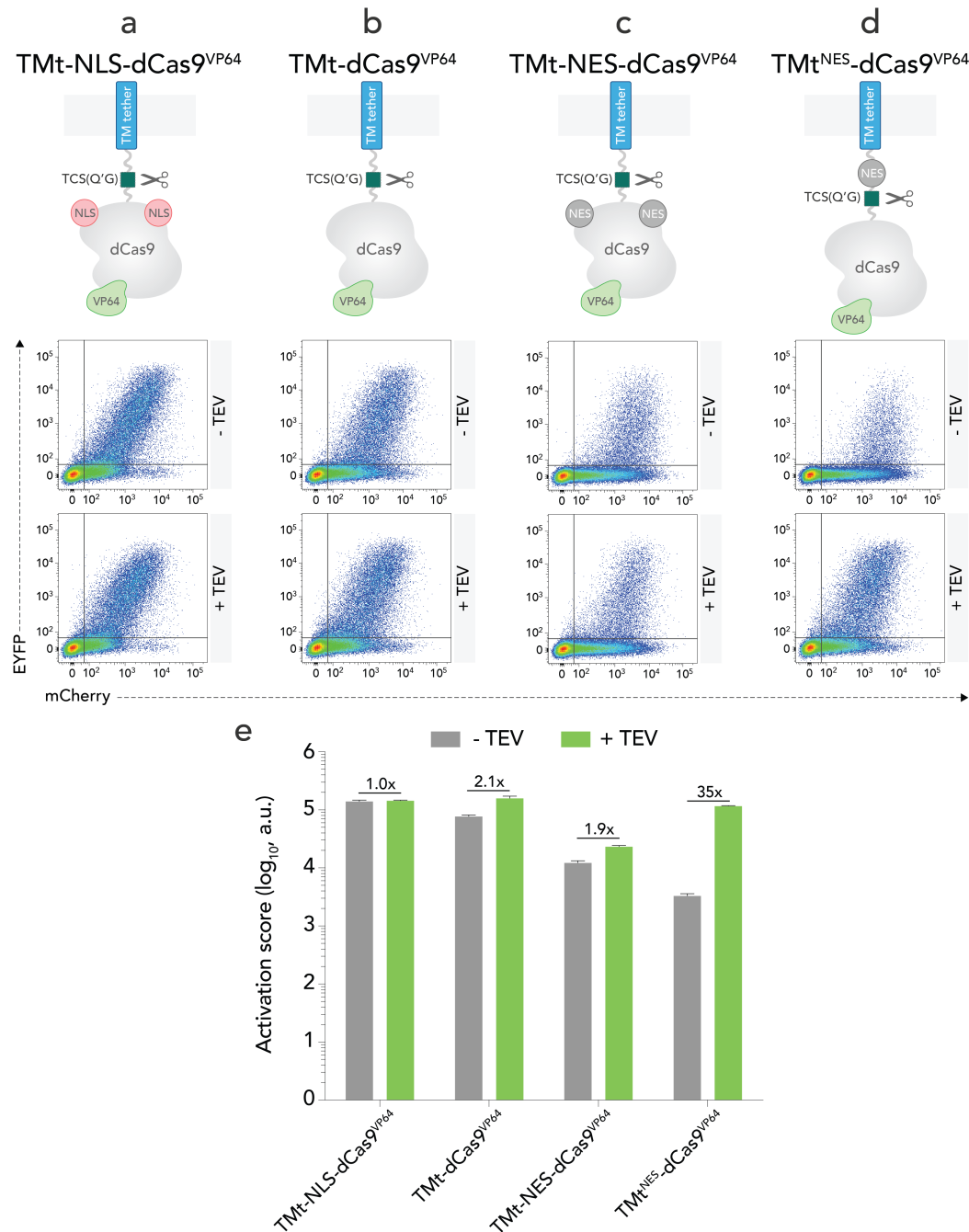


Figure 3.7 | Transcriptional activation potential of conditionally transmembrane-tethered dCas9^{VP64} (TMt-dCas9^{VP64}). (a-e) EYFP reporter assay for transmembrane-tethered (TMt), differently tagged dCas9^{VP64} variants (a-d, top). dCas9^{VP64} variants were transiently transfected into HEK293T cells together with the EYFP reporter plasmid, EYFP targeting guide plasmid (sgEYFP) in the presence or absence of TEV protease. Representative flow cytometry scatter plots are shown (a-d, bottom) as well as the quantification of EYFP reporter activation (e). Only the configuration in which the NES sequence is separable from dCas9^{VP64} (and conversely retained on the TMt) upon TEV protease cleavage results in a 35-fold OFF/ON state transition. In all instances, high unspecific background activation was observed. N = 3 biological replicates from one experiment, mean +/- s.d., a.u. arbitrary units.

TMt-NES-dCas9^{VP64} generated an approximately two-fold OFF/ON state transition change (Figure 3.7c,e). The effect of reduced baseline activation mediated by NES sequences was more pronounced for TMt^{NES}-dCas9^{VP64}. This construct achieved a decent 35-fold change of transcriptional activation in the presence of TEV protease (Figure 3.7d,e). TEV-mediated proteolytic cleavage liberated dCas9^{VP64} from both the TMt and the NES (TMt^{NES}). With regard to transcriptional potency, TEV-processed TMt^{NES}-dCas9^{VP64} recapitulated liberated dCas9^{VP64} (Figure 3.7e).

The apparent discrepancy between subcellular localisation (Figure 3.5) and the transcriptional activity of dCas9^{VP64} (Figure 3.7) prompted further investigation. In the following, I first investigate and validate the role of TEV-mediated TCS cleavage in the release and activation of dCas9^{VP64} and second, the transcript level of TMt-NLS-dCas9^{VP64} and its potential implications for poor OFF/ON state transition characteristics.

To validate TEV-facilitated release of transmembrane-tethered dCas9^{VP64} variants, I first used a catalytically deficient TEV protease (TEV^{C151A}, Phan et al. 2002, Figure 3.8a bottom). Confocal microscopy verified that catalytically inactive TEV^{C151A} did not liberate TMt-NLS-dCas9^{VP64} from its membrane-bound state (Figure 3.8b left). I then created a TMt-NLS-dCas9^{VP64} variant that lacks the TCS, TMt^{ΔTCS}-NLS-dCas9^{VP64} (Figure 3.8a). As expected, TMt^{ΔTCS}-NLS-dCas9^{VP64} could not be processed by TEV protease, impairing liberation of NLS-dCas9^{VP64} as revealed by the lack of anti-HA signal in the nucleus (Figure 3.8b right).

Reporter assays, performed with a combination of different TEV proteases and TMt-NLS-dCas9^{VP64} constructs, disclosed a similar level of EYFP activation throughout (Figure 3.9). This observation supports the previous finding that the potency of

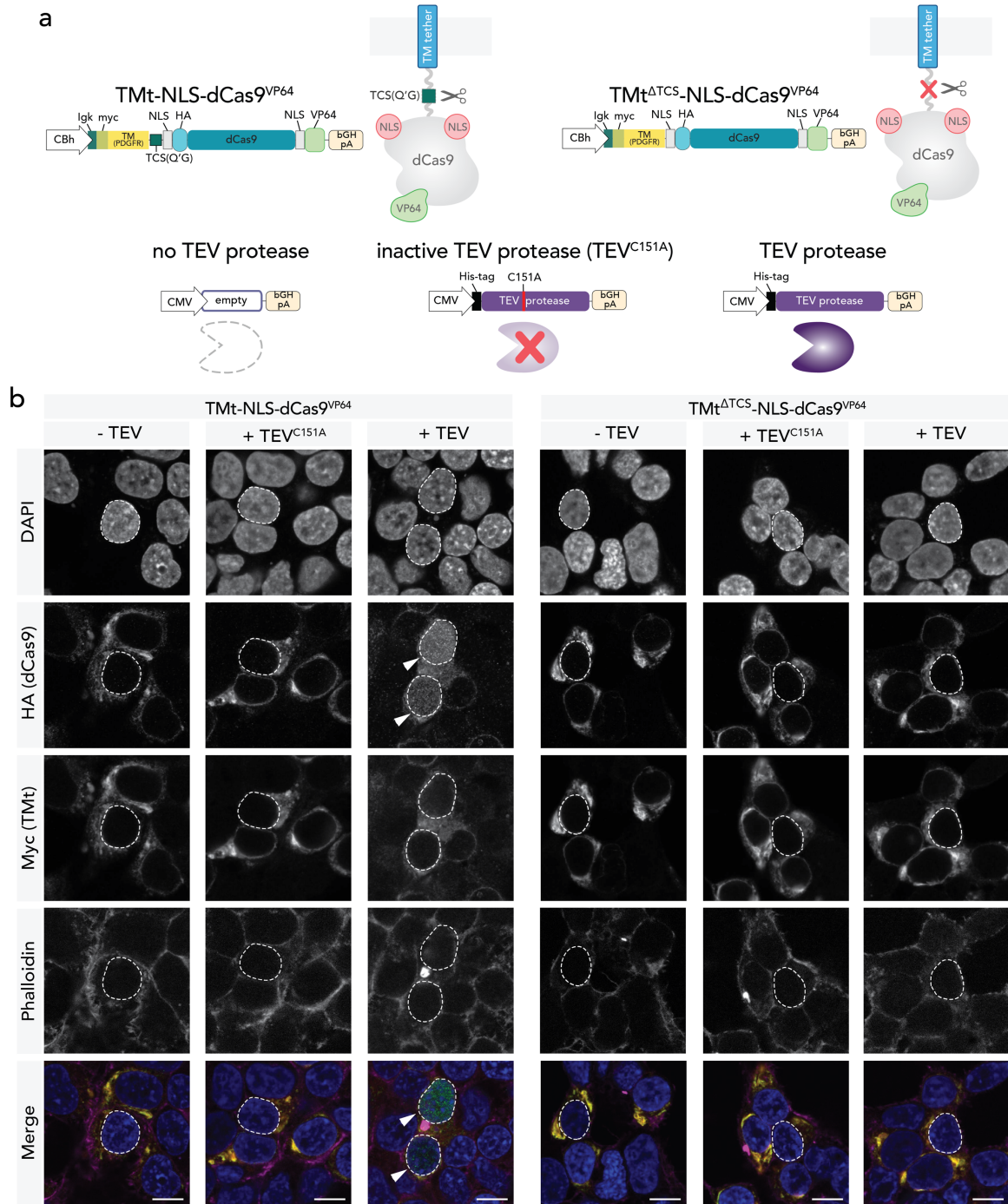


Figure 3.8 | TEV protease specifically releases transmembrane-bound NLS-dCas9^{VP64} from the membrane. (a) Construct designs and cartoons of TMt-NLS-dCas9^{VP64} (left); TMt^{ΔTCS}-NLS-dCas9^{VP64} (right); and TEV protease constructs, no TEV, catalytically inactive TEV (TEV^{C151A}), and catalytically active TEV (bottom). (b) Anti-HA and anti-myc confocal microscopy of transiently transfected HEK293T cells with either TMt-NLS-dCas9^{VP64} (left) or TMt^{ΔTCS}-NLS-dCas9^{VP64} in combination with TEV protease variants. Only TMt-NLS-dCas9^{VP64} co-transfected with catalytically active TEV protease led to nuclear NLS-dCas9^{VP64}. Arrows indicate nuclear localisation of NLS-tagged dCas9^{VP64}. DAPI = nuclear staining, Phalloidin = F-actin. White dotted lines outline nuclei. Scale bar = 10μm.

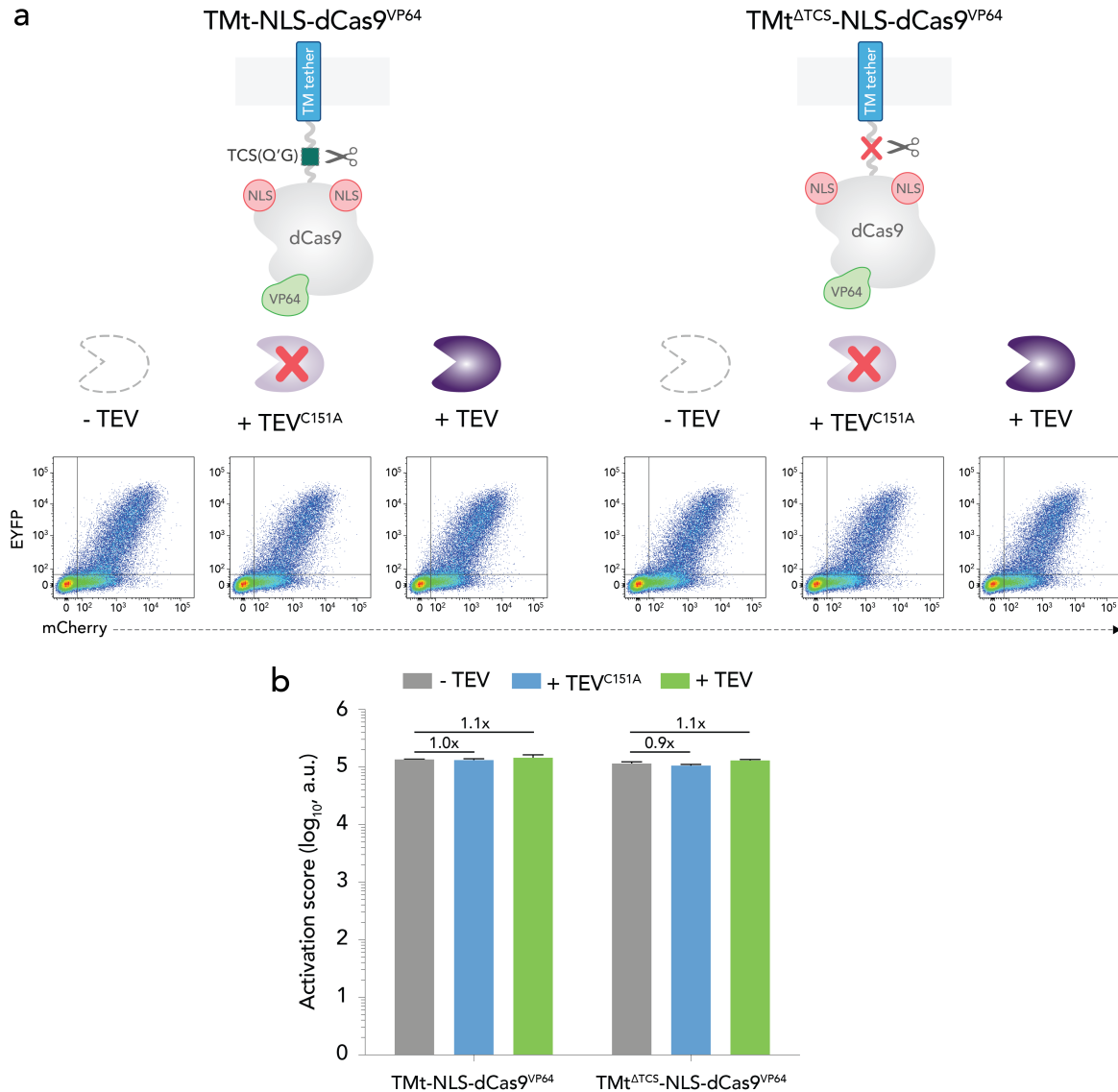


Figure 3.9 | TMt-NLS-dCas9^{VP64} transcriptional activation is independent of TEV protease. **(a)** EYFP reporter assay for TMt-NLS-dCas9^{VP64} and TMt^{ΔTCS}-NLS-dCas9^{VP64} in combination with either no TEV, catalytically inactive TEV (TEV^{C151A}), or catalytically active TEV protease. Representative flow cytometry scatter plots are shown (bottom). **(b)** Quantification of EYFP reporter activation. Independent of the presence of catalytically active TEV or an engrafted TCS in TMt-NLS-dCas9^{VP64}, all tested conditions display strong transcriptional reporter activation. N = 3 biological replicates from one experiment, mean $\pm$ s.d., a.u. arbitrary units.

TMt-NLS-dCas9^{VP64} is seemingly independent of TEV protease (Figure 3.7). Furthermore, TMt^{ΔTCS}-NLS-dCas9^{VP64} showed no evidence of being functionally affected by TEV protease co-delivery (Figure 3.9). This observation implies that: (i) there is no alternative endogenous TEV protease orthologue capable of cleaving the TCS site; (ii) there must be another mechanism that allows for TMt-NLS-dCas9^{VP64} to be transcriptionally active in the nucleus.

TMt-NLS-dCas9^{VP64} did not display an OFF/ON state transition behaviour, which may reflect a detection issue due to extremely high background leakiness. To address this possibility, I repeated the same experiment, however, using TMt^{NES}-dCas9^{VP64} and TMt^{NESΔTCS}-dCas9^{VP64} (Figure 3.10a). TMt^{NES}-dCas9^{VP64} co-transfected with TEV protease reproduced the observed OFF/ON state transition from Figure 3.7 (Figure 3.10b left). As expected, TMt^{NES}-dCas9^{VP64} in combination with TEV^{C151A} did not result in any increased transcriptional activity (Figure 3.10b left). Conversely, TMt^{NESΔTCS}-dCas9^{VP64} maintained its baseline transcriptional activity irrespectively of which TEV protease was used (Figure 3.10b right). This data unequivocally demonstrates that catalytically active TEV protease in combination with a cognate TCS is necessary and sufficient to release transmembrane-tethered dCas9^{VP64} (Figure 3.10b).

Next, I examined whether TMt-NLS-dCas9^{VP64} transcript abundance might have a negative impact on both the overall high background activity in the OFF state and the absence of inducibility in the presence of TEV protease (ON state). To this end, I created HEK293T cells carrying a stably integrated TMt-NLS-dCas9^{VP64} under the control of a doxycycline-inducible promoter (Figure 3.11a). This design allowed to determine whether low TMt-NLS-dCas9^{VP64} expression is a prerequisite for enabling potent OFF/ON state transition by fine-tuning its transcript levels. By comparison,

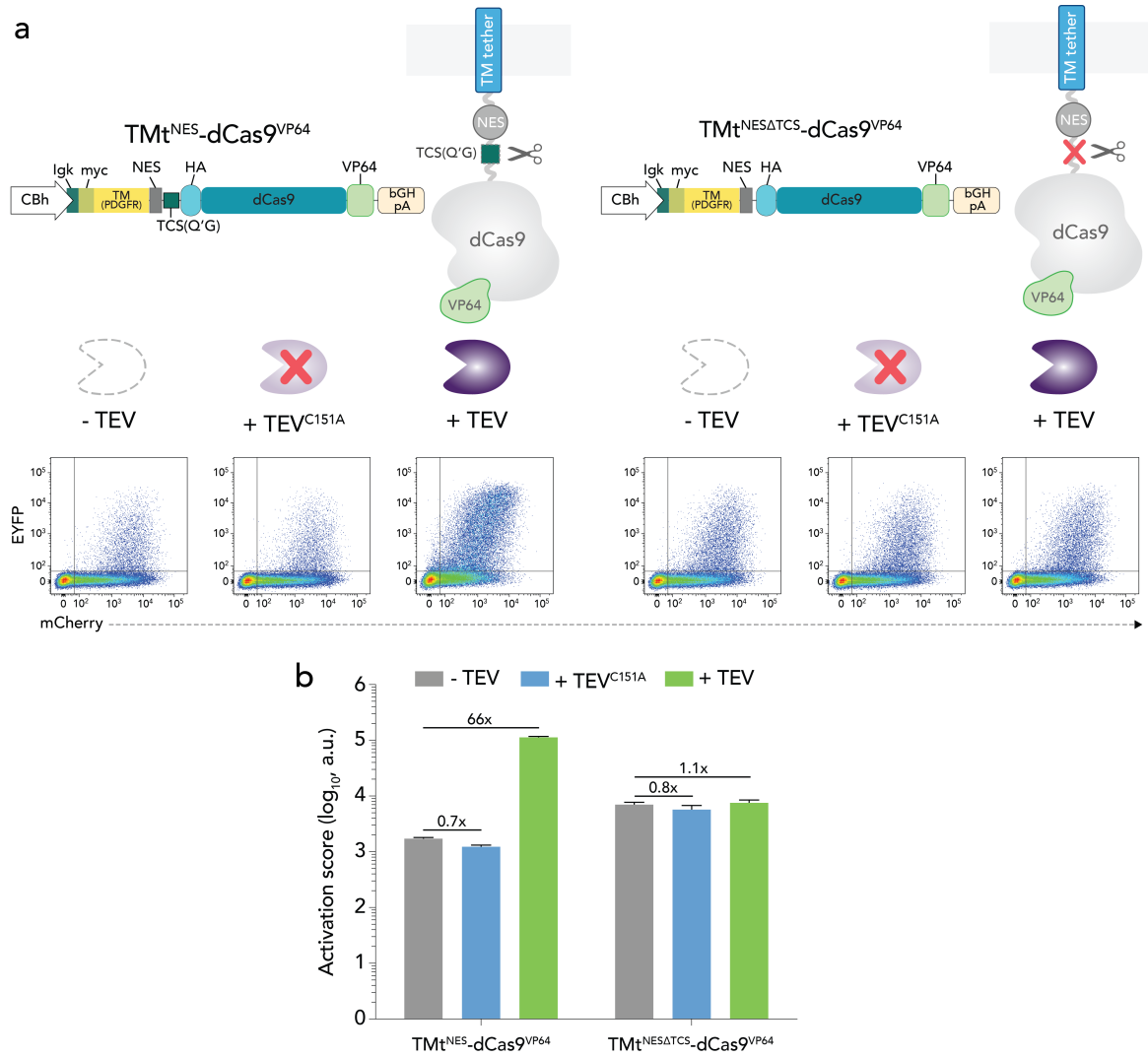


Figure 3.10 | TMt^{NES}-dCas9^{VP64} transcriptional OFF/ON transition is dependent on TEV protease. (a) EYFP reporter assay of transiently transfected TMt^{NES}-dCas9^{VP64} or TMt^{NESΔTCS}-dCas9^{VP64} HEK293T cells in combination with either no TEV, catalytically inactive TEV (TEV^{C151A}), or catalytically active TEV protease. Representative flow cytometry scatter plots are shown (bottom). (b) Quantification of EYFP reporter activation. Only in the presence of catalytically active TEV and an engrafted TCS in TMt^{NES}-dCas9^{VP64}, TMt^{NES}-dCas9^{VP64} displays a robust OFF/ON state transition in reporter output. N = 3 biological replicates from one experiment, mean +/- s.d., a.u. arbitrary units.

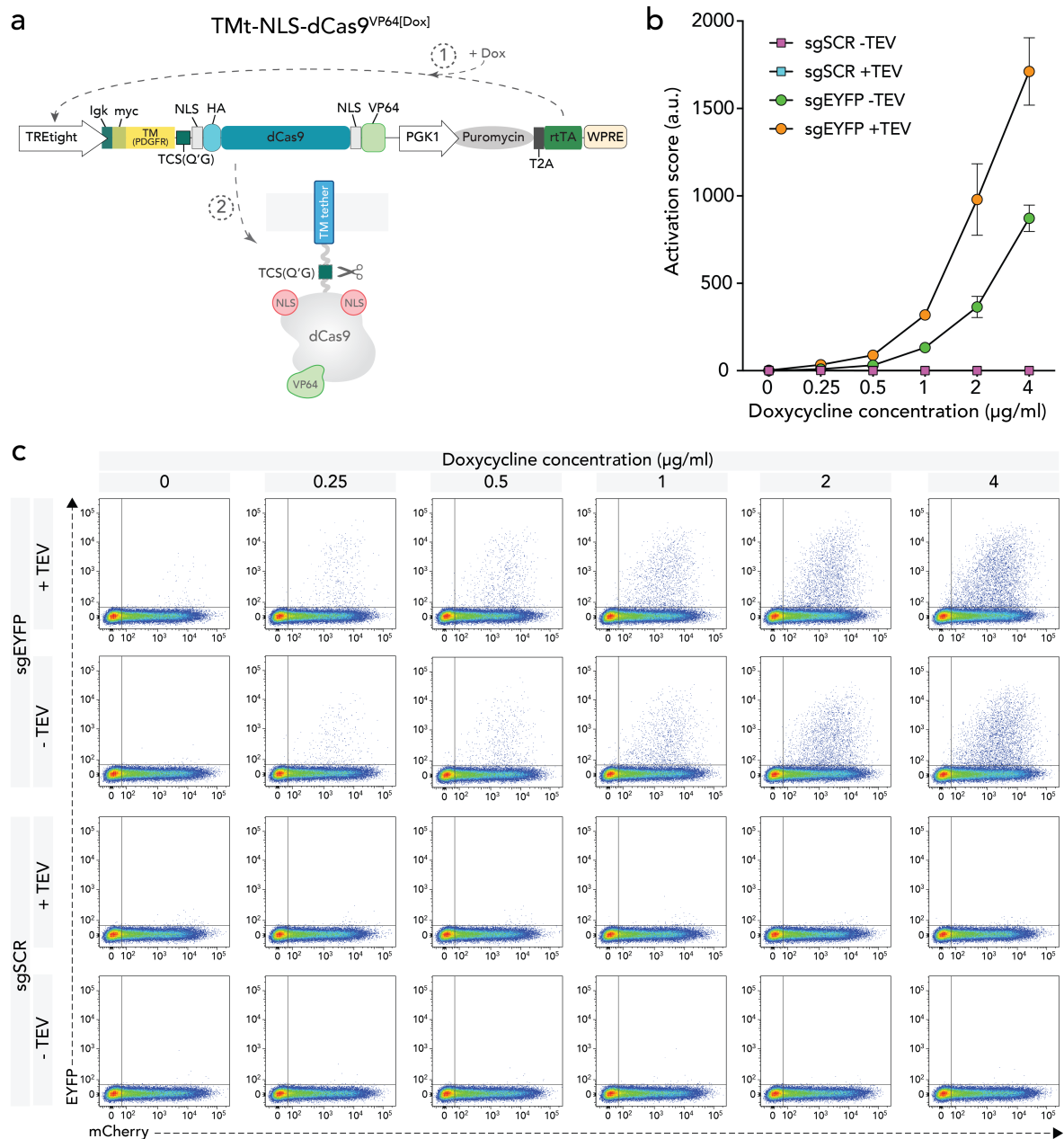


Figure 3.11 | TEV independent reporter activation of lowly expressed TMt-NLS-dCas9^{VP64}[DOX]. (a) Construct design and cartoon of stably integrated, doxycycline inducible TMt-NLS-dCas9^{VP64}[DOX]. Constitutively expressed rTA binds under doxycycline treatment (+ Dox, step 1) to TRE_{tight} promoter to induce the transcription of TMt-NLS-dCas9^{VP64} (step 2). (b) Quantification of EYFP reporter activation of transiently transfected, doxycycline inducible TMt-NLS- dCas9^{VP64}[DOX] HEK293T cells with the EYFP reporter, sgEYFP with mCherry, in the absence or presence of TEV protease. 24h post-transfection, doxycycline was supplemented to media at indicated concentrations and cells grown for an additional 48 hours. Data represent the mean $\pm$ s.d. (N = 3) from one experiment using biological replicates, a.u. arbitrary units. (c) Representative flow cytometry scatter plots from (b).

transient transfections with Tmt-NLS-dCas9^{VP64} expressing plasmid produce very high levels of protein as confirmed by microscopy (Figure 3.5a and 3.8b).

Only in the presence of doxycycline, constitutively expressed rtTA bound to doxycycline engages with the TRE_{tight} promoter, initiating Tmt-NLS-dCas9^{VP64} expression. In the absence of doxycycline, no transgene expression was observed (Figure 3.11b,c). Surprisingly, however, even at low doxycycline concentrations (low Tmt-NLS-dCas9^{VP64} expression level), background EYFP reporter expression was detected in the absence of TEV protease (Figure 3.11b,c). TEV presence led to approximately two-fold OFF/ON state transition at increasing concentrations of doxycycline. Taken together, this data suggests that Tmt-NLS-dCas9^{VP64} background transcriptional activity is mostly independent of expression levels.

To substantiate these observations, I then carried out a parallel set of experiments utilising catalytically active Cas9. To this end, I fused Cas9 and NLS-Cas9 to the transmembrane linker, an efficiently cleavable TCS sequence, and the fluorescent protein mCherry (Figure 3.12a-d). Confocal microscopy revealed that Tmt-NLS-Cas9-mCherry is directed to discrete structures of the secretory pathway in the cytoplasm in absence of TEV protease (Figure 3.12a left). In the presence of TEV protease, NLS-Cas9-mCherry is released from the membrane and can freely translocate into the nucleus (Figure 3.12a right), recapitulating the phenotype of soluble NLS-Cas9-mCherry (Figure 3.12c). Tmt-Cas9-mCherry also localises to discrete structures in the cytoplasm in the absence of TEV protease (Figure 3.12b, left) and once processed by TEV protease, it is evenly distributed in the cytoplasm (Figure 3.12b right). These results are in agreement with previous data on the localisation of Tmt-dCas9^{VP64} variants (Figure 3.5).

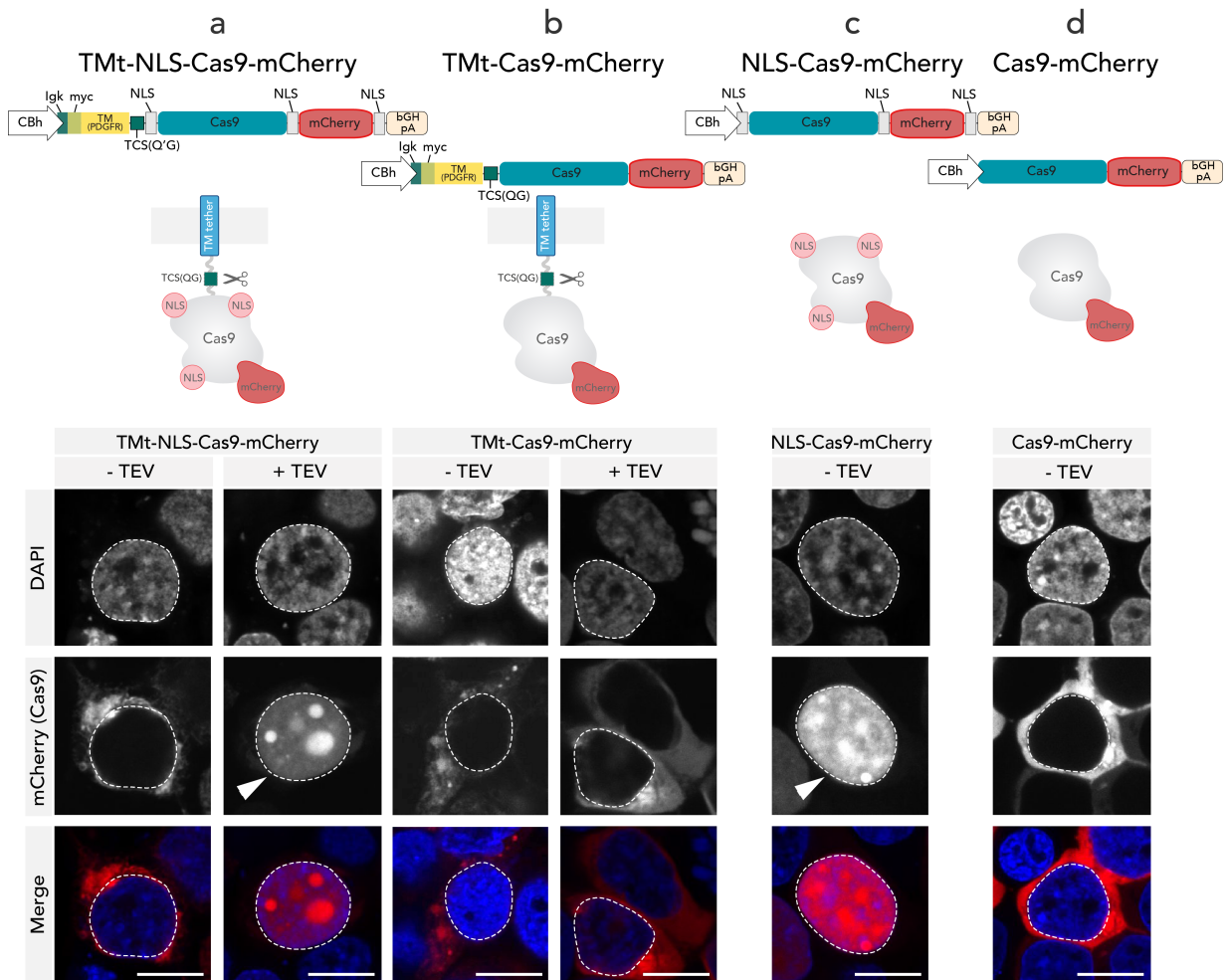


Figure 3.12 | Subcellular localisation of transmembrane-tethered Cas9-mCherry variants. **(a-d) Top:** Molecular structures and cartoons of TMt-NLS-Cas9-mCherry (a), TMt-Cas9-mCherry (b), NLS-Cas9-mCherry (c), and Cas9-mCherry (d). **Bottom:** Confocal microscopy of transiently transfected HEK293T cells with TMt-Cas9-mCherry variants in the absence or presence of TEV protease. As expected, only NLS-Cas9-mCherry translocated into the nucleus upon the successful cleavage of the TCS in TMt-NLS-Cas9-mCherry by TEV protease thereby recapitulating the characteristic nuclear distribution of NLS-Cas9-mCherry (c). Cas9-mCherry (d) served as control for the cytoplasmic localisation of TMt-Cas9-mCherry (b). White dotted lines outline nuclei. Arrows indicate nuclear localisation of NLS-tagged dCas9-mCherry variants. DAPI = nuclear staining. Scale bar = 10µm.

To assess the functional potency of catalytically active Cas9, I established a genomic EGFP disruption assay. In this assay, a destabilised EGFP (dEGFP) is integrated into the genome of HEK293T cells (HEK293T_dEGFP) and targeted for disruption by transiently transfecting EGFP specific sgRNA (sgEGFP)-programmed Cas9 (Tsai et al. 2014; Oakes et al. 2016) (Figure 3.13a). Insertions and/or deletions (indels) generated by catalytically active Cas9 distort the open reading frame of EGFP, leading to a loss of EGFP expression. The addition of a destabilisation tag on EGFP (PEST domain, dEGFP) guarantees a short half-life (~2 hours, Oakes et al. 2016), allowing the assessment of successful cleavage events by flow cytometry (Figure 3.13a,b). In this system, EGFP disruption is rigorously quantified in a binary manner: transfected Cas9-mCherry cells are either positive or negative for EGFP expression after Cas9 cleavage.

To generate a clonal HEK293T_dEGFP cell line I transduced the parental cell line HEK293T with a customised dEGFP lentivirus (Figure 3.13b) at low virus titer and sorted single cells by FACS. Following clonal expansion, a HEK293T_dEGFP clone that displayed homogenous dEGFP expression was isolated (Figure 3.13c).

The HEK293T_dEGFP cell line was then transiently transfected with previously described Cas9-mCherry variants (Figure 3.12a-d) and subjected to flow cytometry (Figure 3.14a,b). Only mCherry positive, Cas9-expressing cells were analysed for EGFP expression (Figure 3.14a). As expected, loss of EGFP expression was observed in cells that received soluble NLS-Cas9-mCherry or Cas9-mCherry programmed with sgEGFP (Figure 3.14a,b). The same result was obtained for TMt-NLS-Cas9-mCherry and TMt-Cas9-mCherry guided by sgEGFP irrespective of TEV protease co-delivery (Figure 3.14a,b). This observation is consistent with previous transcription activity data deploying dCas9^{VP64} variants (Figure 3.7e).

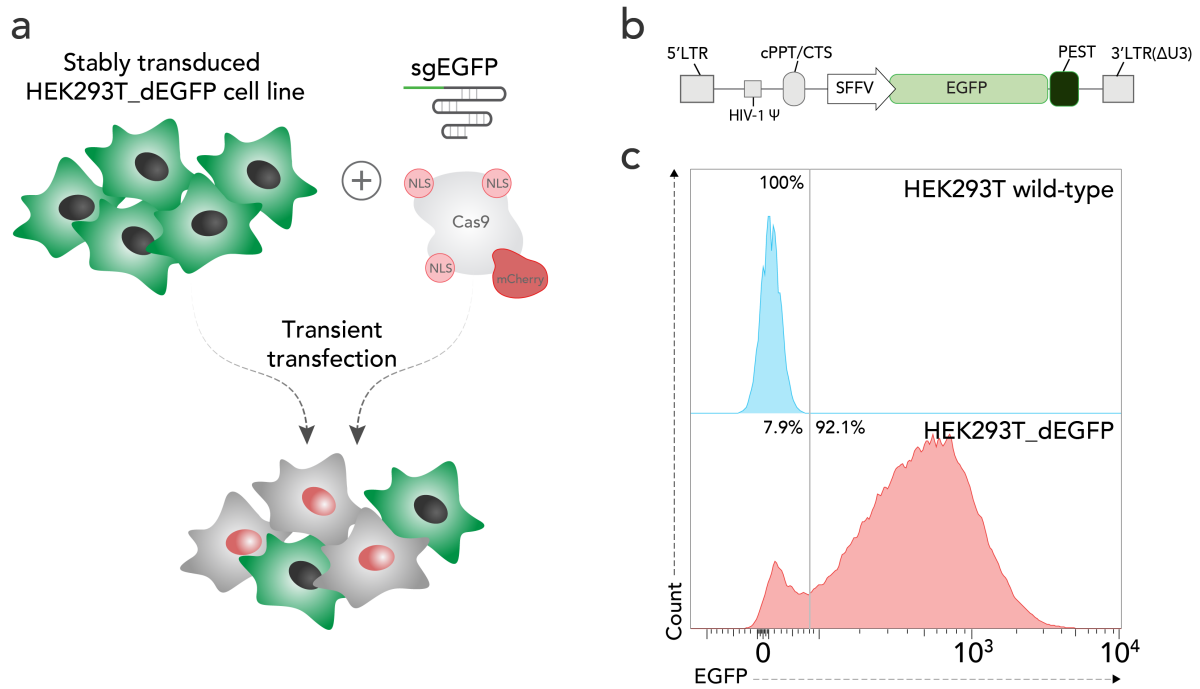


Figure 3.13 | Establishing HEK293T_dEGFP for a genomic dEGFP disruption assay. (a) Cartoon of the EGFP disruption assay. Stably transduced HEK293T cell line with destabilised EGFP (dEGFP) were transiently transfected with catalytically active Cas9-mCherry variants and a sgRNA targeting the ORF of dEGFP (sgEGFP). Successfully transfected HEK293T_dEGFP cells were mCherry positive and lost dEGFP expression. **(b)** Schematic of the lentivirally integrated dEGFP construct into HEK293T cells. LTR, long terminal repeat; 3'LTR(ΔU3), 3'LTR without U3 region; HIV-1 Ψ, human immunodeficiency virus 1 packaging signal; cPPT, central polypurine tract; CTS, central termination sequence; SFFV, spleen focus-forming virus. **(c)** Flow cytometry histogram of EGFP fluorescence for wild-type HEK293T cells (dEGFP negative) and monoclonal HEK293T_dEGFP cell line (dEGFP positive). The Y-axis of the histograms was set to modal in FlowJo software for improved visualisation/readability purposes.

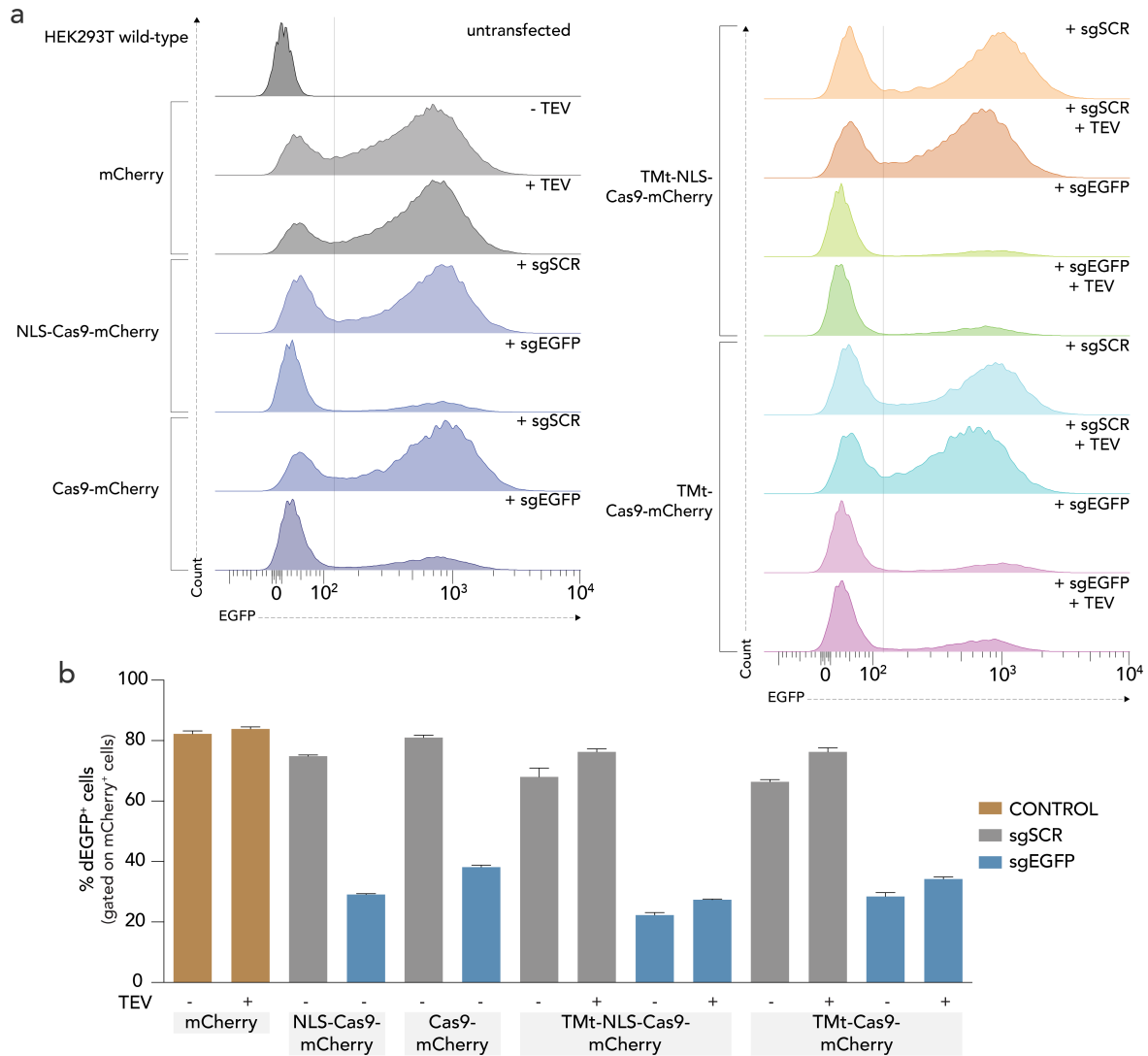


Figure 3.14 | TMt-Cas9-mCherry variants diminish dEGFP fluorescence in HEK293T_dEGFP cells. (a) Representative flow cytometry histograms of transiently transfected HEK293T_dEGFP cells with different Cas9-mCherry and TMt-Cas9-mCherry constructs as well as control plasmid (mCherry only) in the absence or presence of co-transfected TEV protease. Depicted is the dEGFP fluorescence of pre-gated, mCherry positive cells (receiver of Cas9 constructs or mCherry control plasmid). The Y-axis of the histograms was set to modal in FlowJo software for improved visualisation/readability purposes. (b) Quantification of the dEGFP disruption assay as described in (d). N = 3 biological replicates from one experiment, mean +/- s.d.

EGFP disruption on the DNA level, rather than on the protein level, was confirmed via a T7 endonuclease I (T7E1) assay. The T7E1 assay detects mismatches within the PCR amplicon that were introduced as a consequence of Cas9 cleavage events. If Cas9 cutting occurred, fragments of the size of 545 and 184 nucleotides are visible, in addition to the non-mutated, wildtype 729 nt amplicons (Figure 3.15a). In agreement with the flow cytometry data, loss of EGFP expression was accompanied by detectable EGFP DNA mutagenesis (Figure 3.15b). No distinct fragment pattern (545 and 184 nt) were observed in any of the control conditions. Furthermore, NLS-Cas9-mCherry and TMt-NLS-Cas9-mCherry appeared to have a more robust effect than Cas9-mCherry or TMt-Cas9-mCherry, both with regard to loss of EGFP expression (Figure 3.14b) and indel formation (band intensity, Figure 3.15b). This is indicative of an increased DNA cutting potency of NLS-tagged Cas9 compared to its untagged counterparts.

Since Cas9 is directly fused to mCherry, the fluorescence intensity of mCherry should correlate with Cas9 expression levels. Taking advantage of this feature, I sorted NLS-Cas9-mCherry and TMt-NLS-Cas9-mCherry expressing cells into several fluorescence bins (Figure 3.15c) and assessed dEGFP fluorescence in each bin. Even low levels of sgEGFP-programmed NLS-Cas9-mCherry or TMt-NLS-Cas9-mCherry (without co-delivered TEV protease) led to loss of EGFP fluorescence and this effect amplified with increasing mCherry fluorescence (Figure 3.15d). Interestingly, control sgSCR-guided NLS-Cas9-mCherry or TMt-NLS-Cas9-mCherry fractions also displayed reduced dEGFP fluorescence, indicating that transfection of HEK293T_dEGFP cells compromised their dEGFP intensity, albeit this effect was only observed in the high mCherry fluorescence bins (Figure 3.15d).

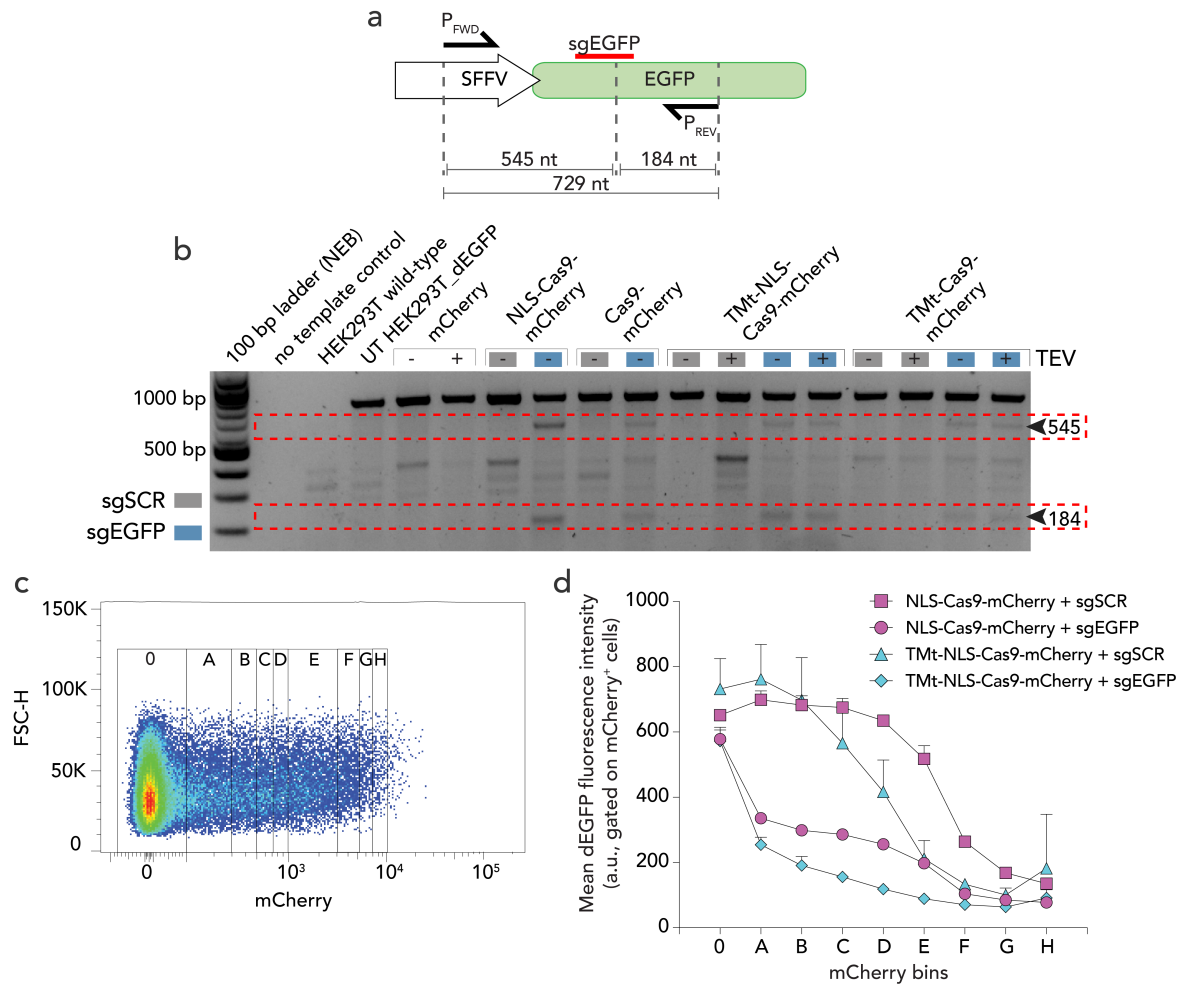


Figure 3.15 | TMT-Cas9-mCherry variants disrupt dEGFP in the genome of HEK293T_dEGFP cells independent of TEV protease-mediated liberation. (a, b) T7 endonuclease I (T7E1) assay for the detection of targeted Cas9 cleavage events. (a) Schematic illustration of the PCR strategy. Indicated are amplicon lengths of uncut genomic *EGFP* as well as expected Cas9 cleavage fragment sizes after the T7E1 assay. nt, nucleotides; P_{FWD} , forward primer; P_{REV} , reverse primer. (b) Agarose gel of T7E1 digested PCR amplicons of genomic DNA from Cas9-transfected HEK293T_dEGFP cells from Figure 3.14a,b. To note, this digest was performed on unsorted/non-enriched cell populations. Transfection efficiency in Figure 3.14a,b was approximately 30%. Only in the presence of *sgEGFP*-guided Cas9, cutting events were detected (see arrows and red dotted boxes). N = 1. UT, transfected cells. (c) Representative flow cytometry scatter plot of transiently transfected HEK293T_dEGFP cells with NLS-Cas9-mCherry and *sgSCR*. Bins were generated for different mCherry mean fluorescence intensities. Same experiment as in Figure 3.14a,b. (d) Quantification of dEGFP mean fluorescence of mCherry bins (g) for various Cas9-mCherry constructs. N = 3 biological replicates from one experiment, mean $\pm$ s.d., a.u. arbitrary units.

These results suggest that TMt-NLS-Cas9-mCherry is functional even at low expression levels, which is in accordance with data relating to doxycycline-inducible TMt-NLS-dCas9^{VP64} (Figure 3.11). TEV protease is indeed dispensable for TMt-NLS-Cas9-mCherry functionality, because TMt^{ΔTCS}-NLS-Cas9-mCherry is as potent as TMt-NLS-Cas9-mCherry in the absence of TEV protease (Figure 3.16a,b).

In all described EGFP disruption experiments, I observed that whenever sgEGFP is co-expressed with Cas9-mCherry, EGFP mutagenesis occurred. To conclusively verify that this is not due to the mere presence of the EGFP-targeting guide RNA, I transiently transfected HEK293T_dEGFP cells with either sgEGFP on a mCherry backbone alone or with a scramble control (sgSCR) (Figure 3.16a,b). This control experiment confirmed that sgEGFP without Cas9 does not lead to EGFP protein loss in HEK293T_dEGFP cells (Figure 3.16a,b).

In summary of subchapter 3.2.2, I demonstrated that a transmembrane anchor is not sufficient for functionally sequestering full length dCas9^{VP64} or Cas9 in HEK293T cells. Tethering, however, resulted in distinct subcellular localisation which can be conditionally regulated by an engineered TEV protease-mediated release mechanism.

3.2.3 A split TMt-dCas9^{VP64} scaffold design for stringent OFF/ON state transitions

In the previous section, I showed that a transmembrane tether, in combination with an NES sequence preceding the TCS facilitated limited conditional sequestration of full length dCas9^{VP64} and transcriptional upregulation in the presence of TEV protease (Figure 3.7d,e). However, the high background activity of TMt^{NES}-dCas9^{VP64} prohibited its implementation as a receptor with dCas9^{VP64} as intracellular signal transduction module.

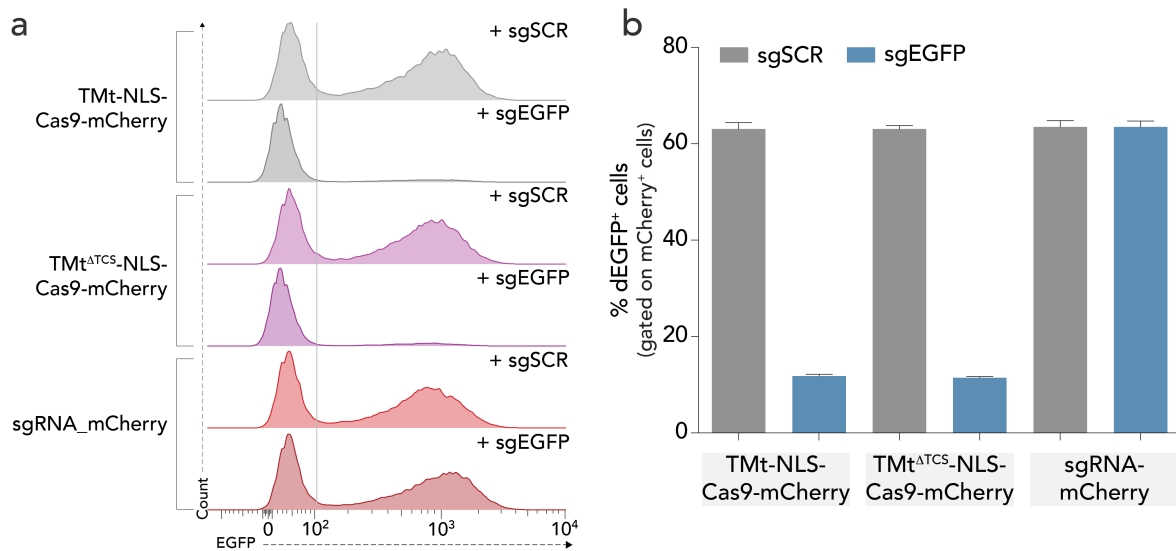


Figure 3.16 | Genomic dEGFP disruption is mediated by sgEGFP-guided TMt-Cas9-mCherry variants. (a) Representative flow cytometry histograms of transiently transfected HEK293T_dEGFP cells with TMt-NLS-Cas9-mCherry, TMt^{ΔTCS}-NLS-Cas9-mCherry, or mCherry in combination with either sgSCR or sgEGFP (in all cases without co-delivered TEV protease). Depicted is the dEGFP fluorescence intensity of pre-gated, mCherry positive cells (receiver of Cas9 constructs or mCherry control plasmid). The Y-axis of the histograms was set to modal in FlowJo software for improved visualisation/readability purposes. (b) Quantification of the dEGFP disruption assay as described in (a). N = 3 biological replicates from one experiment, mean +/- s.d.

I hypothesised that adapting a split dCas9^{VP64} design may provide a solution to reduce or completely eliminate this high OFF state leakiness. Feng Zhang's laboratory previously reported that conditional splitting of full length dCas9^{VP64} into an N-terminal part, dCas9(N), and C-terminal part, dCas9(C)^{VP64}, almost completely abrogated its transcriptional activity (Zetsche, Volz, et al. 2015). To efficiently reconstitute full length dCas9^{VP64} in an inducible manner, they fused the hetero-dimerisation FK506 binding protein 12 (FKBP) domain to dCas9(C)^{VP64} and the FKBP rapamycin binding (FRB) domain to dCas9(N) (Choi et al. 1996; Banaszynski et al. 2005; Zetsche, Volz, et al. 2015). In the presence of rapamycin, the dimerisation pair FKBP/FRB facilitates the reassembly of a full length, functional dCas9^{VP64} transcription factor. Notably, to separate both dCas9 moieties in the cell, NES sequences were appended to the dCas9(N) fragment and NLS sequences to dCas9(C)^{VP64} (Zetsche, Volz, et al. 2015).

To integrate the system developed by Zetsche, Volz, et al. (2015), I tested various split/tethered dCas9^{VP64} pair combinations. These included the original NLS(+)-dCas9(C)^{VP64}-FKBP construct as well as an appended version with a transmembrane tether (TMt-NLS-dCas9(C)^{VP64}-FKBP) (Figure 3.17a, top). In addition to the original NES-dCas9(N)-FRB fragment, I included a dCas9(N) variant without NES sequences (dCas9(N)-FRB) and a tethered version in which the NES precedes the TCS sequence (TMt^{NES}-dCas9(N)-FRB) (Figure 3.17a, bottom), by analogy with TMt^{NES}-dCas9^{VP64} (Figure 3.7d,e). The subsequent experiments were carried out in the absence of rapamycin since the trigger for transcriptional activation in my case ought to be TEV protease only. As expected, soluble NLS(+)-dCas9(C)^{VP64}-FKBP in combination with soluble NES-dCas9(N)-FRB or dCas9(N)-FRB did not display any OFF/ON state transition characteristic in the presence of TEV protease since they do not possess a TMt or the TCS (Figure 3.17b left). The NLS(+)-dCas9(C)^{VP64}-FKBP and

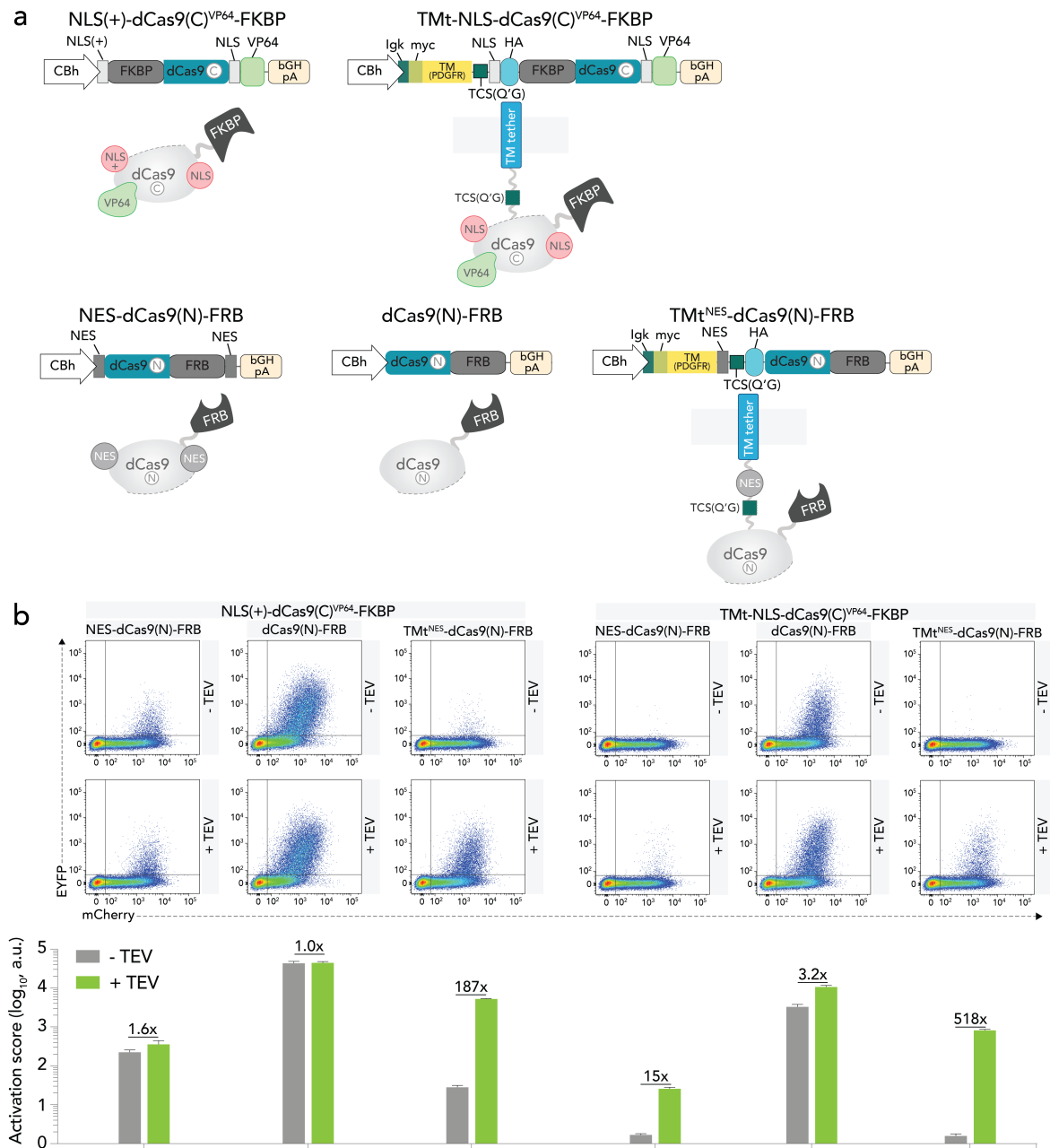


Figure 3.17 | Screen to identify split dCas9^{VP64} configurations with improved OFF/ON state transition characteristics. (a) Molecular structures and cartoons of NLS(+)-dCas9(C)^{VP64}-FKBP and TMt-NLS-dCas9(C)^{VP64}-FKBP (top); and NES-dCas9(N)-FRB, dCas9(N)-FRB and TMt^{NES}-dCas9(N)-FRB (bottom). (b) **Top:** Representative flow cytometry scatter plots of EYFP reporter assay (without rapamycin) with a combination of different dCas9(C)^{VP64}-FKBP and dCas9(N)-FRB constructs in presence or absence of TEV protease transiently transfected into HEK293T cells. **Bottom:** Quantification of EYFP reporter activation assay (without rapamycin) as in (b, top). N = 3 biological replicates from one experiment, mean +/- s.d., a.u. arbitrary units.

NES-dCas9(N)-FRB pair displayed reduced background activation, due to the NES-tagged dCas9(N)-FRB fragment, replicating previous results (Zetsche, Volz, et al. 2015). Strikingly, soluble NLS(+)-dCas9(C)^{VP64}-FKBP and TMt^{NES}-dCas9(N)-FRB displayed a very strong OFF/ON state transition with a 187-fold change in EYFP reporter activation in the presence of TEV protease (Figure 3.17b left). This result demonstrates that, as previously reported, dCas9 fragments can spontaneously assemble in the absence of rapamycin (Zetsche, Volz, et al. 2015). The affinity of dCas9(C)^{VP64} and dCas9(N) for each other is a fundamental property for creating a split dCas9-TF signal transduction module.

TMt-NLS-dCas9(C)^{VP64}-FKBP in combination with either NES-dCas9(N)-FRB or dCas9(N)-FRB yielded low OFF/ON state transitions (Figure 3.17b). Remarkably, the split dual-tethered dCas9 fragments, TMt-NLS-dCas9(C)^{VP64}-FKBP and TMt^{NES}-dCas9(N)-FRB, showed extremely low background activation in the absence of TEV protease, resulting in a 518-fold change EYFP activation (Figure 3.17b). This striking result demonstrates that tethering NLS-dCas9(C)^{VP64} provides effective sequestration and concomitantly reduces background leakiness which is in stark contrast to tethering full length NLS-dCas9^{VP64} (Figure 3.7a,e). In addition, it validates that dual tethering counteracts the inherent auto-assembly properties of split dCas9^{VP64} fragments. Finally, TEV protease appears to effectively release dual-tethered dCas9^{VP64} fragments and rapamycin is not required for system activation.

To determine if rapamycin could, however, confer an additional, beneficial layer to control the activity of split dCas9 pairs, I repeated the EYFP activation assay utilising TMt^{NES}-dCas9(N)-FRB in combination with either NLS(+)-dCas9(C)^{VP64}-FKBP or TMt-NLS-dCas9(C)^{VP64}-FKBP, in the presence of 200 nM rapamycin (Figure 3.18a). Surprisingly, rapamycin completely abrogated TEV-mediated system inducibility,

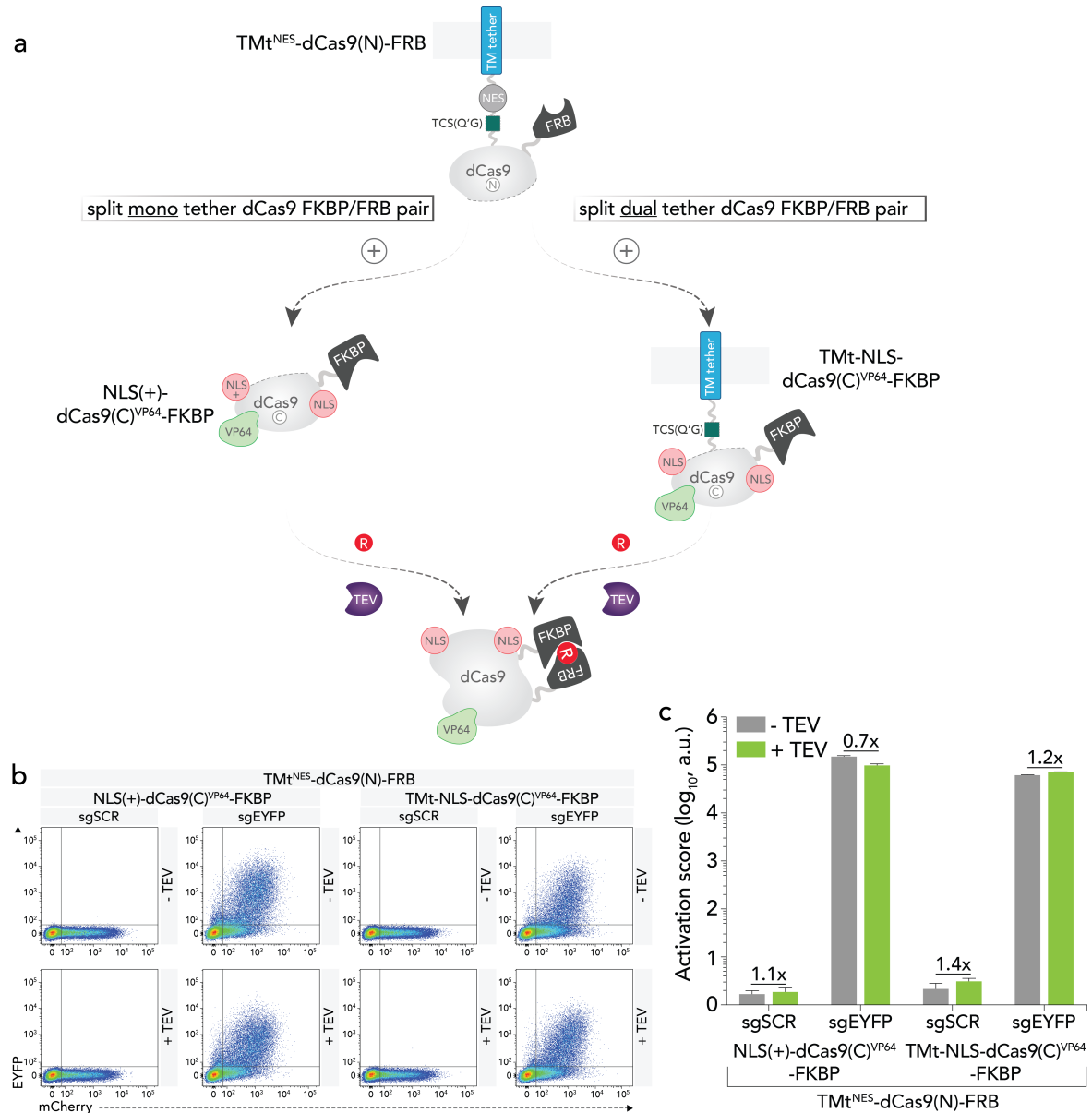


Figure 3.18 | Rapamycin abrogates OFF/ON state transition in split dCas9 FKBP/FRB pair configurations. **(a)** Schematic representation of the split mono and dual tether dCas9 FKBP/FRB pair. TMt^{NES}-dCas9(N)-FRB with either soluble dCas9(C)^{VP64}-FKBP or tethered TMt-NLS-dCas9(C)^{VP64}-FKBP reconstitute in the presence of TEV protease and rapamycin. Rapamycin facilitates the dimerisation of the FKBP/FRB pair. **(b)** Representative flow cytometry scatter plots of EYFP reporter assay with a combination of different dCas9(C)^{VP64}-FKBP and dCas9(N)-FRB constructs transiently transfected into HEK293T cells. Additionally, TEV protease was supplemented in some conditions (as indicated). Rapamycin was co-supplied in all instances. **(c)** Quantification of EYFP reporter activation assay in the presence of 200nM rapamycin for 24h in all conditions. N = 3 biological replicates from one experiment, mean +/- s.d., a.u. arbitrary units.

leading to full EYFP activation (Figure 3.18b,c). This might be explained by the rapamycin-induced formation of dimerised dCas9 fragments, resembling in essence full length TMt-NLS-dCas9(C)^{VP64} (Figure 3.5a).

3.2.3.1 Split mono and dual TMt-dCas9^{VP64}

Since fusing the FKBP/FRB dimerisation domains to the two best performing split dCas9 variants did not improve system performance, I removed these domains from subsequent TMt scaffolds. Additionally, I incorporated the SAM system into the NLS-dCas9^{VP64} plasmid variants which will allow the efficient induction of endogenous genes in downstream experiments. These changes in construct designs led to the creation of an optimised set of split transmembrane-tethered dCas9^{VP64} variants: **split mono TMt-dCas9^{VP64}**, comprised of NLS-dCas9(C)^{VP64} and TMt^{NES}-dCas9(N) and **split dual TMt-dCas9^{VP64}**, consisting of TMt-NLS-dCas9(C)^{VP64} and TMt^{NES}-dCas9(N) (Figure 3.19a,b).

Confocal microscopy on both split mono and dual TMt-dCas9^{VP64} systems using an HA-tagged dCas9(N) fragment (TMt^{NES}-HA-dCas9(N)) revealed that in the absence TEV protease, no HA-signal was detected in the nucleus in either configuration (Figure 3.20a). In the presence of TEV protease, however, both soluble NLS-dCas9(C)^{VP64} and previously tethered TMt-NLS-dCas9(C)^{VP64} redirected proteolytically processed TMt^{NES}-HA-dCas9(N), HA-dCas9(N), into the nucleus (Figure 3.20a). This indicates that split dCas9^{VP64} fragments auto-assemble after TEV-mediated cleavage.

Next, I tested the transcriptional potency of split mono and dual TMt-dCas9^{VP64} in transiently transfected HEK293T cells. Astonishingly, both tethered configurations displayed higher background leakiness and lower fold activation than their FKBP/FRB counterparts (Figure 3.20b,c compared to Figure 3.17b). Split mono TMt-dCas9^{VP64}

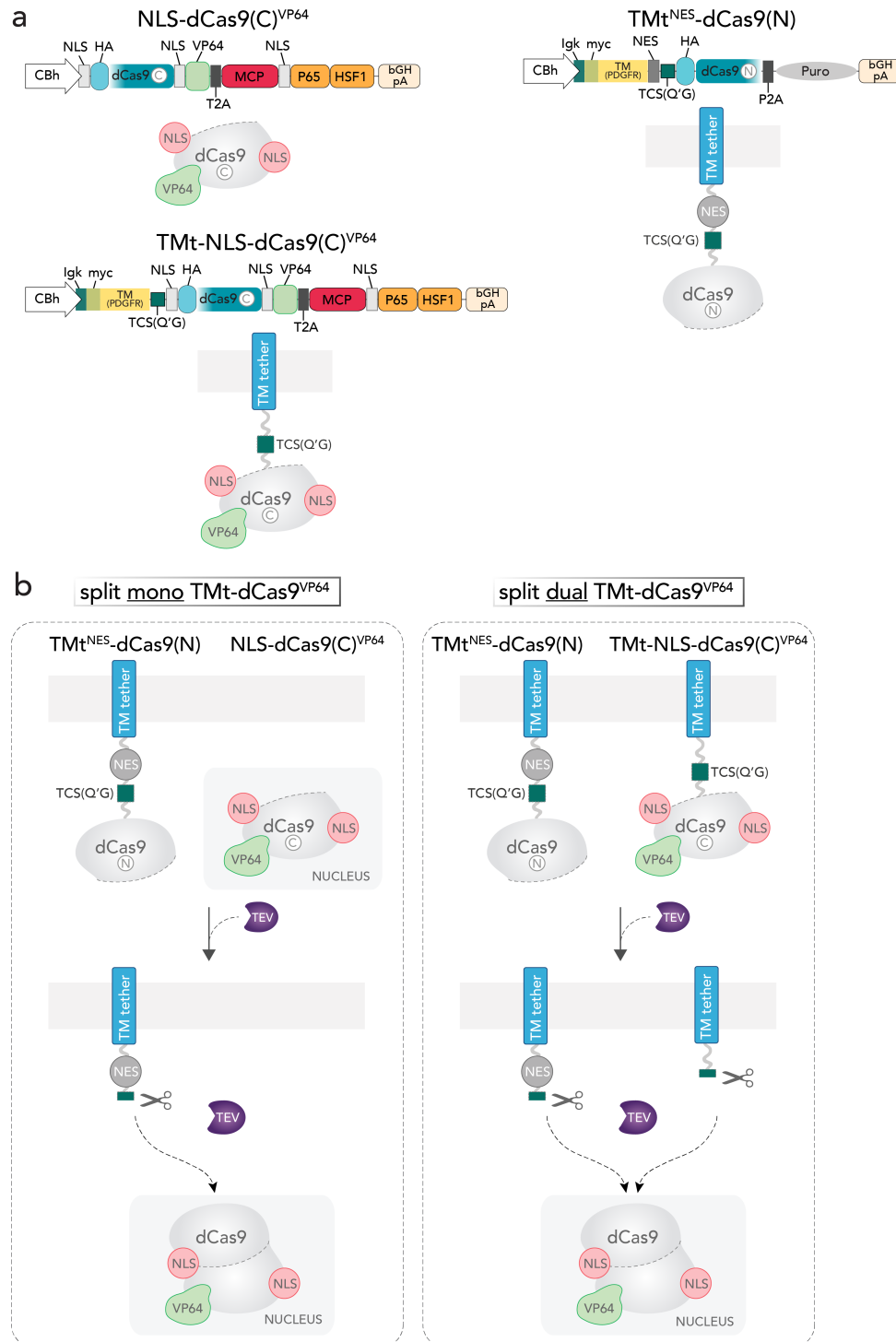


Figure 3.19 | Conceptual overview of split mono and dual TMt-dCas9^{VP64}. (a) Molecular structures and cartoons of NLS-dCas9(C)^{VP64}, TMt-NLS-dCas9(C)^{VP64}, and TMt^{NES}-dCas9(N). (b) Schematic representation of the split mono and dual tethered dCas9^{VP64} concept. A generic transmembrane tether (TMt) anchors dCas9(N) in the membrane as well as TMt-NLS-dCas9(C)^{VP64}, whereas NLS-dCas9(C)^{VP64} resides soluble in the nucleus. Upon TEV-mediated cleavage of the TCS in the tether, both dCas9 fragments reassemble to form a transcriptionally active TF in the nucleus.

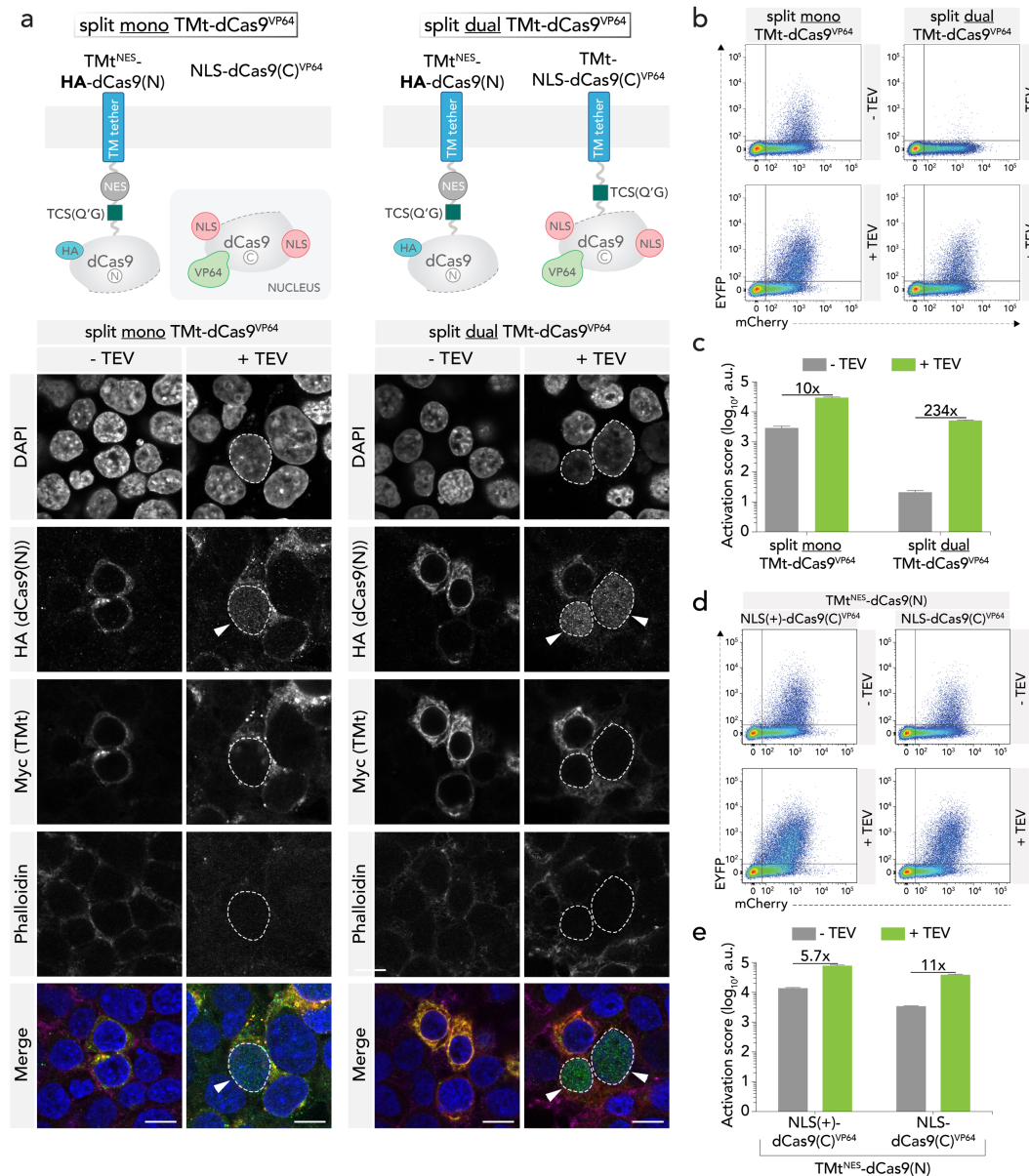


Figure 3.20 | dCas9(C)^{VP64} reassembles with liberated dCas9(N) fragment. (a) Anti-HA and anti-myc confocal microscopy of transiently transfected HEK293T cells with NLS-dCas9(C)^{VP64} variants (without an HA tag) in the absence or presence of TEV protease. Only liberated HA-tagged TMt^{NES}-dCas9(N) translocates into the nucleus by reassembling with NLS-dCas9(C)^{VP64}. DAPI = nuclear staining, Phalloidin = F-actin. White dotted lines outline nuclei. Arrows indicate nuclear localisation of dCas9(N). Scale bar = 10μm. (b) Representative flow cytometry scatter plots of EYFP reporter assay of split mono and dual Tmt-dCas9 in the presence or absence of TEV protease in transiently transfected HEK293T cells. (c) Quantification of EYFP reporter assay as described in (b). (d) Representative flow cytometry scatter plots of EYFP reporter assay with a combination of differently NLS-tagged dCas9(C)^{VP64} and TMt^{NES}-dCas9(N) in transiently transfected HEK293T cells. (e) Quantification of EYFP reporter activation assay. (c,e) N = 3 biological replicates from one experiment, mean +/- s.d., a.u. arbitrary units.

attained only 10-fold activation while split dual TMt-dCas9^{VP64} yielded 234-fold activation (Figure 3.20b,c). The diminished OFF/ON state transitions may be a consequence of removing the FKBP/FRB domains. FKBP/FRB domains were fused to the dCas9 fragments via short linkers thereby potentially disfavours the inherent split dCas9 auto-reconstitution. Furthermore, the FKBP/FRB domains may also contribute to reduced affinity between Cas9 fragments since they themselves do not possess any affinity for one another in the absence of rapamycin (Banaszynski et al. 2005).

In particular, the fold-change of split mono TMt-dCas9^{VP64} deteriorated compared to its FKBP/FRB equivalent, although a weak NLS sequence was used in the NLS-dCas9(C)^{VP64} variant. The original NLS-dCas9(C)^{VP64} fragment had a NLS(+)sequence on its N-terminus (NLS(+)-dCas9(C)^{VP64}-FKBP). However, the stronger NLS(+) led to an even poorer fold-change activation (5.7-fold) compared to the weaker NLS variant (11-fold) when tested in a head-to-head experiment (Figure 3.20d,e).

3.2.3.2 Inducible split mono and dual TMt-dCas9^{VP64}

As demonstrated in the previous section, stringent OFF/ON state transitions of split mono and dual TMt-dCas9^{VP64} scaffolds are dependent on co-delivered TEV protease. Engineering a chimeric receptor based on this conceptual framework may benefit from a strategy in which TEV protease activity is regulated by extracellular factors/exogenously delivered compounds. To render TEV protease inducible, I split constitutively active, full length TEV protease into functionally inactive TEV(N) and TEV(C) parts, as previously described (Wehr et al. 2006; D. J. Williams et al. 2009; Gray et al. 2010; D. Lee et al. 2017). Next, I fused FKBP to the C-terminus of TEV(N), generating FKBP-TEV(N), and FRB to TEV(C), creating FRB-TEV(C) (Figure 3.21a,b)

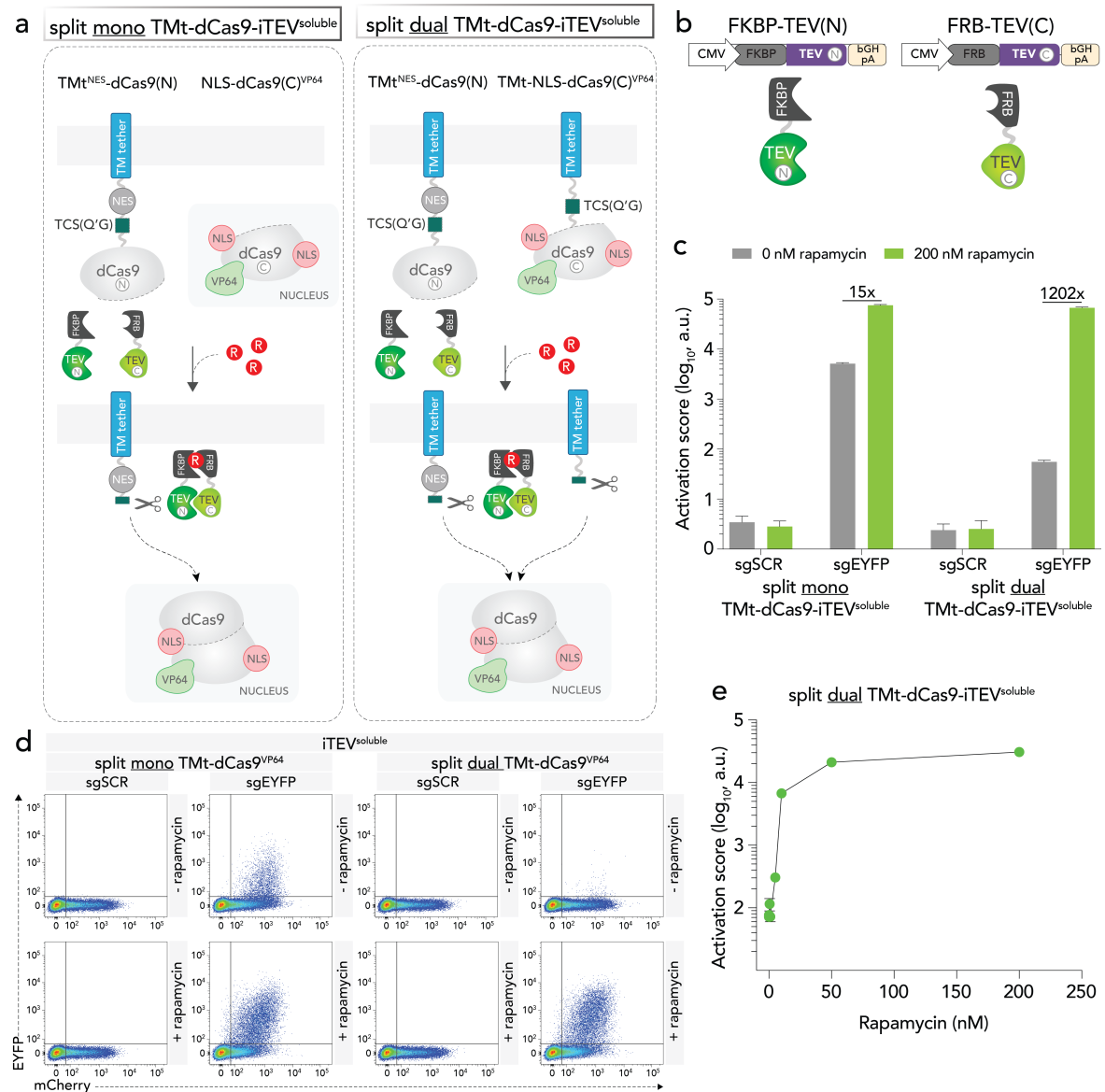


Figure 3.21 | Inducible transcriptional activation with soluble split TEV for split mono and dual TMT-dCas9^{VP64} (TMT-dCas9-iTEV^{soluble}). (a) Schematic representation of the split mono and dual TMT-dCas9-iTEV^{soluble} concept. Rapamycin triggers the reconstitution of catalytically active TEV protease that subsequently liberates TMt-NLS-dCas9(C)^{VP64} and/or TMt^{NES}-dCas9(N). (b) Molecular structures and cartoons of FKBP-TEV(N) and FRB-TEV(C). (c) Quantification of EYFP reporter assay and (d) representative flow cytometry scatter plots of split mono and dual TMT-dCas9-iTEV^{soluble} in the presence or absence of 200 nM rapamycin in transiently transfected HEK293T cells. N = 3 biological replicates from one experiment, mean $\pm$ s.d., a.u. arbitrary units. (e) Dose-dependent transcriptional reporter activation of split dual TMT-dCas9-iTEV^{soluble}. N = 3 biological replicates for each concentration data point, data from one experiment, mean $\pm$ s.d., a.u. arbitrary units.

(Wehr et al. 2006). I then delivered the newly created FKBP-TEV(N)/FRB-TEV(C) pair together with the split mono and dual Tmt-dCas9^{VP64} (Figure 3.21a). Upon rapamycin administration, the FKBP-TEV(N)/FRB-TEV(C) pair reconstitutes a catalytically active TEV protease, which is competent for liberating the tethered dCas9 fragments (Figure 3.21a). These rapamycin-dependent designs are referred to as **split mono Tmt-dCas9-iTEV^{soluble}** and **split dual Tmt-dCas9-iTEV^{soluble}**, respectively.

The transcriptional activity of inducible split mono and dual Tmt-dCas9-iTEV^{soluble} constructs was first assessed using the EYFP reporter assay. Split mono Tmt-dCas9-iTEV^{soluble} generated a modest 15-fold change EYFP upregulation upon induction with 200 nM rapamycin (Figure 3.21c,d). The inducible split dual Tmt-dCas9-iTEV^{soluble}, however, yielded a staggering 1202-fold activation, while maintaining a substantially lower background activity compared to the split mono configuration (Figure 3.21c,d). Notably, the split dual Tmt-dCas9-iTEV^{soluble} quickly reached its maximum transcriptional potential at low rapamycin concentrations (Figure 3.21e).

This characteristic suggests a narrow window of fine-tuning opportunity which I further tested on the endogenous gene *ASCL1*. Transient transfections of HEK293T cells with split dual Tmt-dCas9-iTEV^{soluble} as well as a pool of previously validated SAM sgASCL1 guide RNAs (Zetsche, Volz, et al. 2015) confirmed dose-dependent endogenous gene activation of up to 11-fold at 200 nM rapamycin compared to 0 nM (Figure 3.22). This performance was comparable to the full length TEV protease condition, which, as expected, did not display rapamycin-mediated inducibility. Importantly, rapamycin in combination with a control guide RNA (sgSCR) did not trigger *ASCL1* upregulation (Figure 3.22).

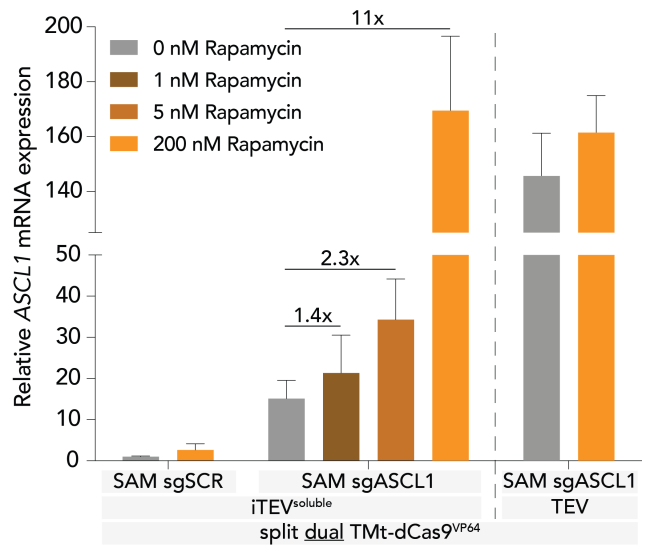


Figure 3.22 | Inducible transcriptional activation of an endogenous gene with split dual TMt-dCas9-iTEV^{soluble}. RT-qPCR of endogenously targeted *ASCL1* in HEK293T cells transiently transfected with either split dual TMt-dCas9-iTEV^{soluble} at different concentrations of inducer or with full length TEV protease. N = 3 biological replicates (x3 technical replicates), mean +/- s.d.

In an attempt to further enhance the transcriptional potency of inducible split mono and dual TMT-dCas9-iTEV^{soluble} constructs, I adapted an integrative design (Wehr et al. 2006). To this end, FKBP-TEV(N) was incorporated into the split tethered dCas9 fragments, creating **split mono TMT-dCas9-iTEV^{engrafted}** and **split dual TMT-dCas9-iTEV^{engrafted}** (Figure 3.23). I reasoned that the resulting proximity of reconstituted TEV protease to its TCS may favour enhanced overall performance, compared to a freely diffusible, soluble iTEV (Figure 3.21a). Indeed, these designs improved OFF/ON state transition properties of both split mono TMT-dCas9-iTEV^{engrafted} and split dual TMT-dCas9-iTEV^{engrafted} to 60-fold and 35900-fold, respectively (Figure 3.24). These enhanced fold-changes were accomplished by massively reducing background leakiness while maintaining high overall activation potency. Importantly, both iTEV^{engrafted} systems did not operate in control reactions in which there was no soluble FRB-TEV(C) fragments supplied or in the presence of only soluble FKBP-TEV(N) fragments (Figure 3.24). Strikingly, split dual TMT-dCas9-iTEV^{engrafted} displayed virtually no background leakiness (Figure 3.24). Subsequently, I benchmarked both iTEV^{engrafted} systems against the original rapamycin-inducible, dimerisation-dependent split dCas9^{VP64} constructs (Zetsche, Volz, et al. 2015) (Figure 3.24). Split dCas9^{VP64} outperformed mono TMT-dCas9-iTEV^{engrafted} but was inferior to split dual TMT-dCas9-iTEV^{engrafted} (Figure 3.24).

Finally, I tested split dual TMT-dCas9-iTEV^{engrafted} in the context of endogenous gene regulation. At 2 nM, 20 nM, and 200 nM rapamycin, split dual TMT-dCas9-iTEV^{engrafted} yielded robust 3.5-, 63-, and 107-fold *ASCL1* upregulation, respectively (Figure 3.25), vastly outperforming split dual TMT-dCas9-iTEV^{soluble} (Figure 3.22).

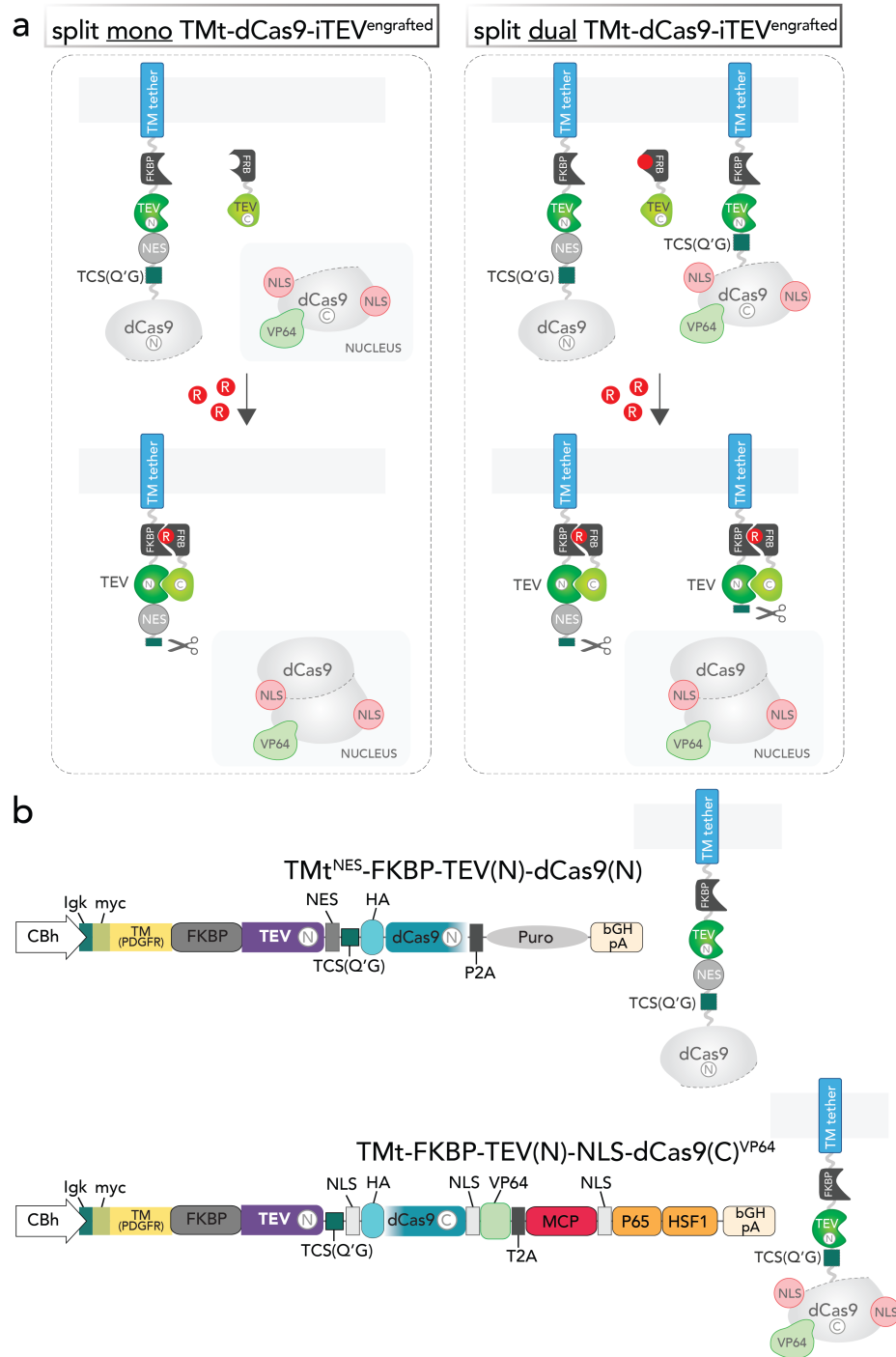


Figure 3.23 | Concept of an engrafted split TEV in split mono and dual TMt-dCas9^{VP64} (TMt-dCas9-iTEV^{engrafted}). (a) Schematic representation of the split mono and dual TMt-dCas9-iTEV^{engrafted} concept. Rapamycin induces the recruitment of TEV(C) to the tethered and engrafted TEV(N) fragment, thereby allowing the reassembly of catalytically active TEV protease. Reconstituted TEV protease cleaves the in proximity located TCS, which leads to the liberation of TMt-FKBP-TEV(N)-NLS-dCas9(C)^{VP64} and/or TMt^{NES}-FKBP-TEV(N)-dCas9(N). Subsequently, both dCas9 fragments reassemble to form a transcriptionally active TF. (b) Molecular structures and cartoons of TMt^{NES}-FKBP-TEV(N)-dCas9(N) (top) and TMt-FKBP-TEV(N)-NLS-dCas9(C)^{VP64} (bottom).

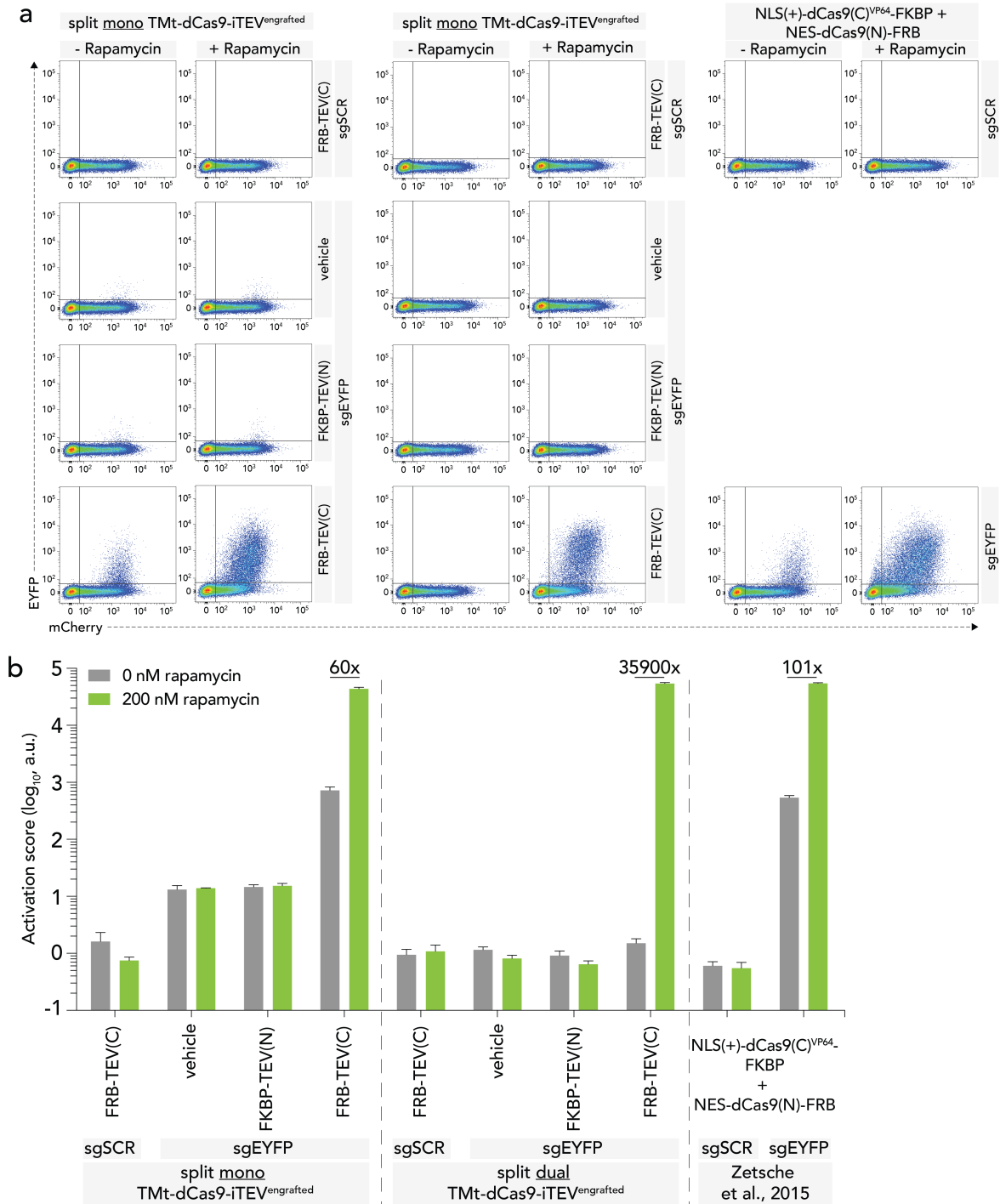


Figure 3.24 | Inducible transcriptional activation with split mono and dual Tmt-dCas9-iTEV^{engrafted}. (a) Representative flow cytometry scatter plots of EYFP reporter assay of split mono and dual Tmt-dCas9-iTEV^{engrafted} in the presence or absence of 200 nM rapamycin in transiently transfected HEK293T cells and (b) the quantification thereof. In addition, the originally reported split dCas9^{VP64} constructs from Zetsche, Volz, et al. (2015) were run in parallel. N = 3 biological replicates from one experiment, mean +/- s.d., a.u. arbitrary units.

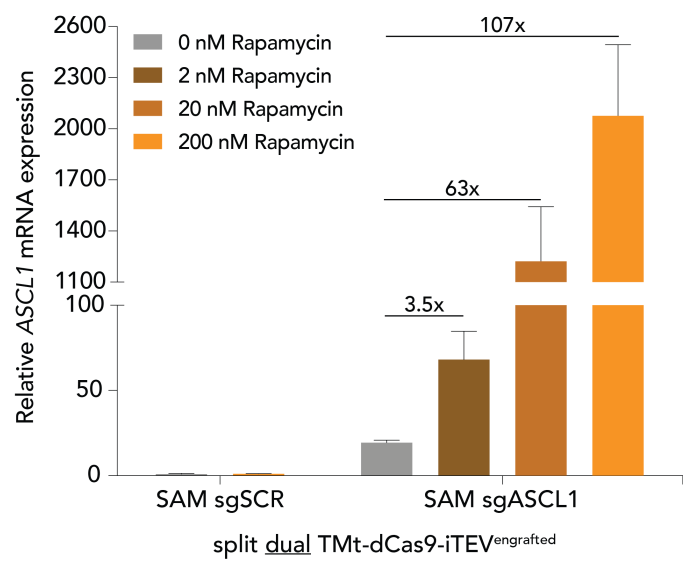


Figure 3.25 | Inducible transcriptional activation of an endogenous gene with split dual TMT-dCas9-iTEV^{engrafted}. RT-qPCR of endogenously targeted *ASCL1* in HEK293T cells with split dual TMT-dCas9-iTEV^{engrafted} at different concentrations of rapamycin. N = 3 biological replicates (x3 technical replicates), mean +/- s.d.

In summary of subchapter 3.2.3, I demonstrated that splitting dCas9^{VP64} into two fragments and simultaneously tethering them to the membrane is sufficient for functionally sequestering dCas9^{VP64} in HEK293T cells. Upon TEV-mediated proteolysis, split dual TMt-dCas9^{VP64} transitions into a potent signal transduction module, achieving stringent OFF/ON state transitions.

Split dual TMt-dCas9-iTEV^{engrafted} represents an effective approach to render split dual TMt-dCas9^{VP64} responsive to an exogenously-delivered compound, enabling temporally-controlled and user-defined endogenous gene expression modulation.

3.3 Discussion

In this chapter, I set out to design, implement, and optimise a programmable dCas9-TF-based signal transduction module. To create a functional module for synthetic receptor engineering, two critical conditions had to be met: (i) by analogy with native/endogenous receptors, it was indispensable to attain minimal background activation (leakiness) in the absence of an inducer (OFF state); (ii) in the presence of the inducer, the system should display robust activation (ON state).

To recapitulate the intracellular signal transduction module of a cell-surface receptor, I created a generic transmembrane-tethered dCas9-TF scaffold. Subsequently, I implemented a TEV protease-mediated mechanism to reliably release dCas9-TF from its membrane anchor. Surprisingly, full length tethered dCas9-TF was extremely potent in activating target gene transcription. I have demonstrated that this leakiness is not caused by extensive protein expression, but rather might be explained by the cell type of choice and cellular processes. Jennifer Doudna's group has shown that in rapidly dividing HEK293T cells, Cas9 does not require an NLS to efficiently disrupt genomic DNA (Oakes et al. 2016). Additionally, the conclusion that full length transmembrane-bound dCas9 is inadequate for generating a stringent OFF/ON state

transition has been confirmed by studies of alternative chimeric receptor scaffolds in HEK293T, HEK293FT, and Jurkat cells (Schwarz et al. 2017; Kipniss et al. 2017; Y. Liu et al. 2017).

One possible, although speculative, explanation is that this property might be a result of nuclear membrane disintegration, providing an access opportunity for highly effective dCas9-TF. This behaviour would be significantly exacerbated in rapidly dividing cells, such as HEK293T cells. However, this also suggests that testing and optimising a system that works robustly in HEK293T cells will presumably exceed its performance in other cell types that display moderate rates of cell division. This would be highly relevant for downstream applications of this new class of chimeric receptors, which would likely entail delivery to induced pluripotent stem cells (iPSC), human stem cells, and primary cells.

To control dCas9-TF activity in a membrane-tethered configuration, I have implemented several modifications to the basic architecture: (i) I removed NLS sequences, thereby reducing dCas9-TF translocation to the nucleus; (ii) I included an NES sequence upstream of the TCS thereby potentially increasing cytoplasmic sequestration in the transmembrane-bound state; and (iii) I segregated dCas9-TF into two fragments, completely rendering it functionally inactive in the absence of an inducer. These three steps of iterative dCas9-TF engineering resulted in a design that displayed stringent OFF/ON state transition properties.

Chapter 4 : Engineering programmable chimeric dCas9-synGPCR scaffolds

Most of the data presented in this chapter has been published in: Baeumler, T.A., Ahmed, A.A. & Fulga, T.A., 2017. Engineering Synthetic Signaling Pathways with Programmable dCas9-Based Chimeric Receptors. Cell Reports, 20(11), pp.2639–2653.

4.1 Chapter Aims

- Verifying that grafting full length dCas9-TF to the GPCR TANGO scaffold is inadequate for attaining low OFF state activity.
- Adapting the core split dCas9-based architecture to the GPCR TANGO scaffold to create chimeric dCas9-based GPCRs (dCas9-synGPCRs).
- Implementing and characterising a TCS-optimised, split dual dCas9-synGPCR as ideal configuration for attaining robust OFF/ON state transition properties in the presence of cognate ligands.
- Demonstrating the generalisability of this dCas9-based chimeric receptor architecture by designing a variety of dCas9-synGPCRs.
- Engineering Boolean AND logic gates for multi-factorial input processing.

4.2 Results

4.2.1 Chimeric, full length dCas9-based GPCRs are inadequate for attaining robust OFF/ON state transitions

After establishing that transmembrane-bound full length dCas9^{VP64} is insufficient in generating a suitable OFF/ON state transition (i.e. low background activation level,

Chapter 3.2.2), I first sought to ascertain this conclusion by appending dCas9^{VP64} onto a native, cell-surface receptor.

To this end, I exploited the previously reported “transcriptional activation following β -arrestin translocation” (GPCR TANGO) scaffold (Barnea et al. 2008; Kroeze et al. 2015) as a template for the development of chimeric full length dCas9-based GPCRs (GPCR-dCas9) (Figure 4.1). The basic GPCR-dCas9 design contains the GPCR coding DNA sequence (CDS) followed by a V₂ tail sequence, originating from the vasopressin receptor. Downstream of the V₂ tail, I have inserted an inefficiently cleavable TCS (amino acid sequence ENLYFQL, cleavage occurs between Q and L, abbreviated as TCS(Q’L)). The TCS(Q’L) is a TEV cleavage site variant that is proteolytically processed with reduced efficiency compared to the canonical TCS(Q’G) (Kapust et al. 2001; Kapust et al. 2002). Upon cognate ligand engagement, the receptor undergoes a conformational change that subsequently leads to V₂ tail phosphorylation. The artificial and constitutively overexpressed β -arrestin2-TEV fusion protein is then actively recruited to the phosphorylated V₂ tail, thereby bringing the TEV protease and its cognate TCS in proximity, prompting its proteolytic cleavage. Liberated dCas9^{VP64} translocates into the nucleus to induce transcriptional upregulation of any chosen gene (Figure 4.1). Under this premise, a prototype GPCR-dCas9 should be capable of rewiring specific input/output functions by linking a native ligand to the activity of a synthetic, programmable dCas9^{VP64} as novel signal transduction module (Figure 4.1).

I grafted NLS-dCas9^{VP64} and dCas9^{VP64} onto the original GPCR TANGO scaffold of the bradykinin B2 receptor (BDKRB2), generating BDKRB2-NLS-dCas9^{VP64} and BDKRB2-dCas9^{VP64}, respectively (Figure 4.2a).

For all optimisation experiments, I chose BDKRB2 TANGO because it displayed potent OFF/ON state transition characteristics in the original paper (Kroeze et al. 2015)

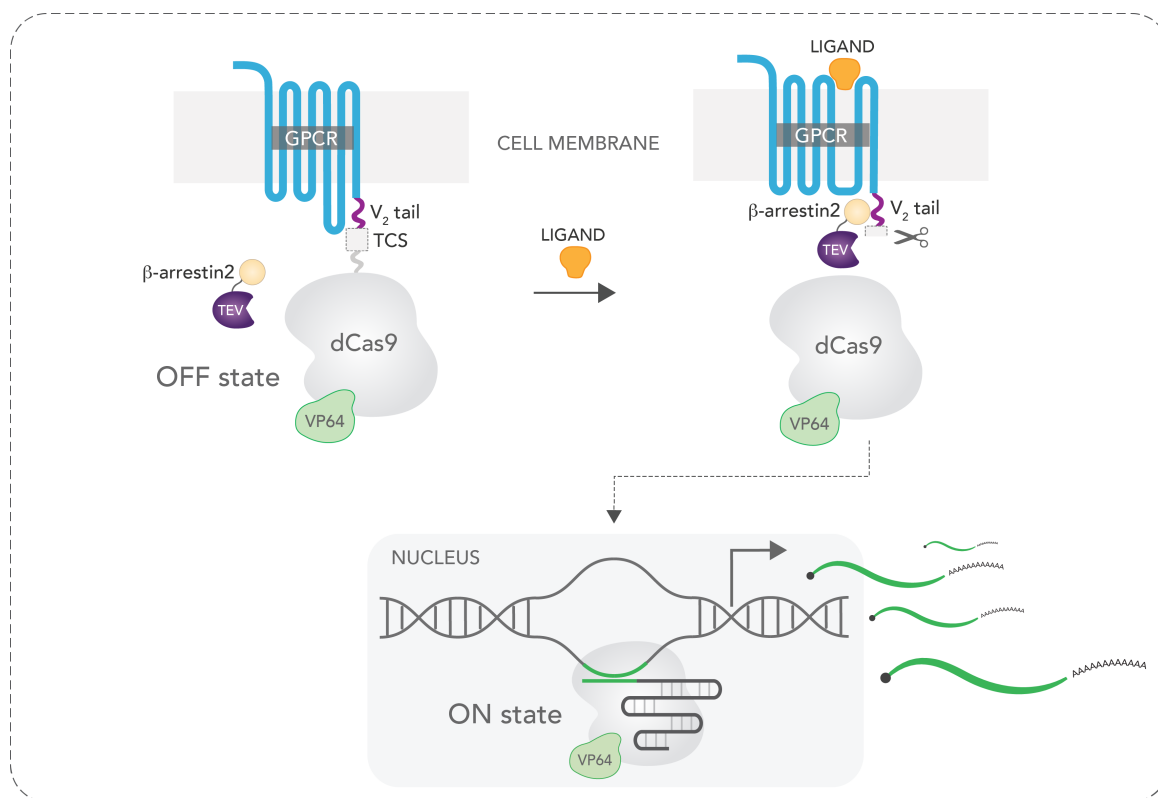


Figure 4.1 | Concept of GPCR-coupled dCas9^{VP64} as novel modality for the rewiring of signal pathways. dCas9^{VP64} is fused to the C-terminal part of an artificial GPCR scaffold, which was originally reported with a non-programmable TF (Barnea et al. 2008; Kroeze et al. 2015). This design contains the GPCR coding DNA sequence (CDS) followed by the V₂ tail of the vasopressin receptor and a TEV protease cleavage site (TCS). Upon cognate ligand engagement, the receptor undergoes a conformational change which leads to the phosphorylation of its C-terminal V₂ tail. The artificial and constitutively expressed β-arrestin2-TEV fusion protein is then recruited to the phosphorylated V₂ tail thereby bringing the TEV protease and its cognate TCS in proximity prompting its proteolytic cleavage. Liberated dCas9^{VP64} can then translocate into the nucleus to exert its transcriptional activity.

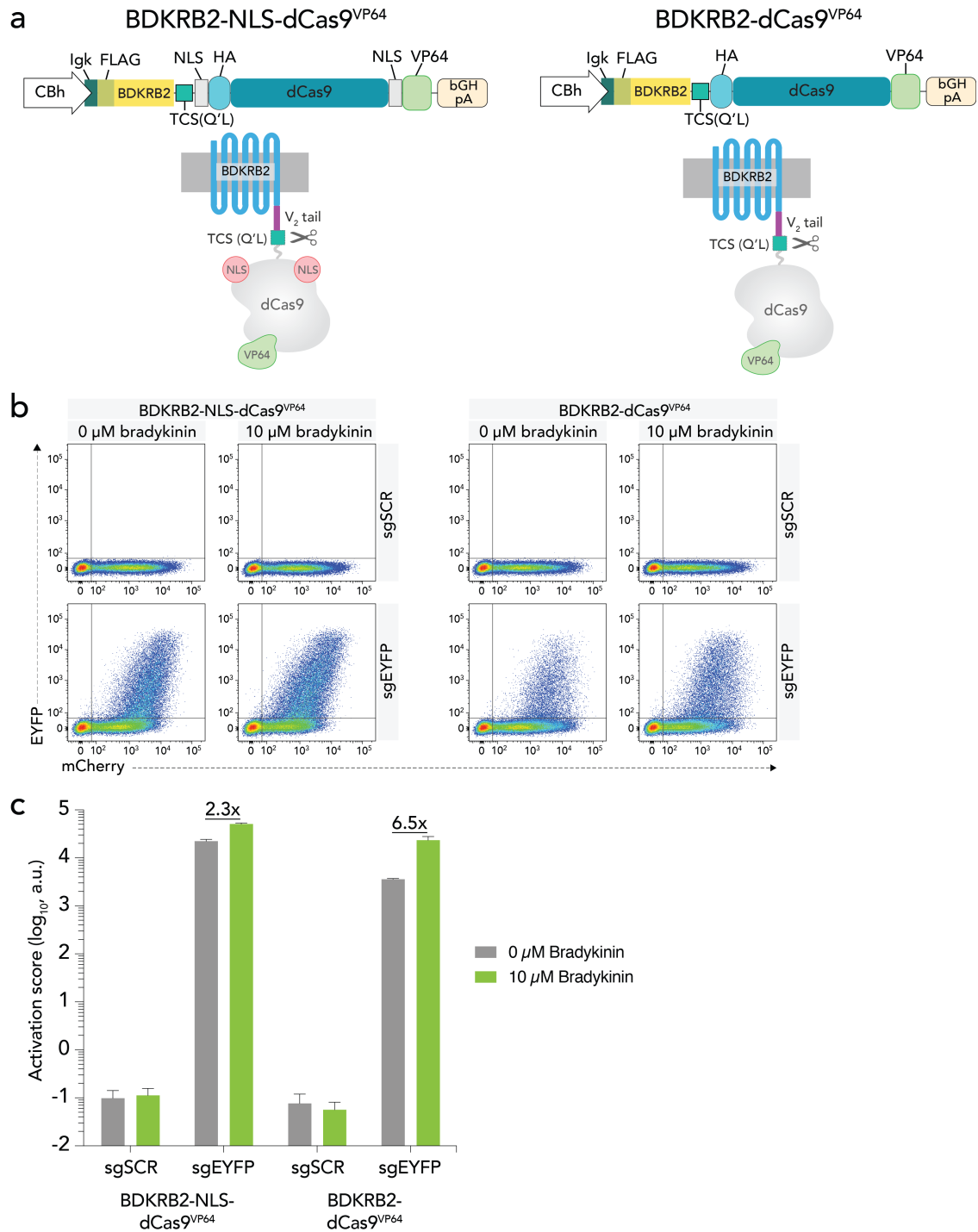


Figure 4.2 | BDKRB2-dCas9^{VP64} recapitulates poor OFF/ON state transition behaviour of TMT-dCas9^{VP64}. (a) Construct designs and cartoons of BDKRB2-NLS-dCas9^{VP64} and BDKRB2-dCas9^{VP64}. (b) Representative flow cytometry scatter plots of EYFP reporter assay of transiently transfected HTLA cells with EYFP reporter, respective sgRNA and variants of BDKRB2-dCas9^{VP64} in the absence or presence of 10 μ M bradykinin and (c) the quantification thereof. GPCR-tethered full length dCas9^{VP64} variants demonstrated extremely high background activities, compromising robust OFF/ON state transition characteristics. N = 3 biological replicates from one experiment, mean $\pm$ s.d., a.u. arbitrary units.

and its agonist, bradykinin, is a relatively inexpensive short peptide. Furthermore, bradykinin is produced by cancer cells to stimulate proliferation and thus may be relevant for therapeutic interventions (da Costa et al. 2014). For all transient transfections involving GPCRs, I utilised a stably transduced HEK293 clone that constitutively expresses the β -arrestin2-TEV fusion protein at low levels, referred to as HTLA cells (Kroeze et al. 2015).

Analysis of the reporter expression disclosed that the magnitude of activation by both BDKRB2-NLS-dCas9^{VP64} and BDKRB2-dCas9^{VP64} in the absence or presence of bradykinin is similar to TMT-NLS-dCas9^{VP64} and TMT-dCas9^{VP64} in the absence or presence of TEV protease, or their untethered derivatives (compare Figure 4.2b,c with Figure 3.7e and Figure 3.3e). As a consequence of extremely high background activity, these chimeric receptors displayed poor fold change activation (~2-fold for BDKRB2-NLS-dCas9^{VP64} and ~6-fold for BDKRB2- dCas9^{VP64}). This data confirms that transmembrane-tethered full length dCas9-TF is indeed not an adequate means for engineering dCas9-based chimeric receptors, as demonstrated in Chapter 3 and subsequently by others (Kipniss et al. 2017; Y. Liu et al. 2017).

4.2.2 Exploring design principles for synthetic split dCas9-based GPCRs (dCas9-synGPCRs)

Having conceived a split, conditionally membrane-bound and functionally sequestered dCas9^{VP64} architecture (Chapter 3.2.3), I sought to append this core design onto the original GPCR TANGO scaffold in order to mitigate the extremely high OFF state activation observed for full length dCas9^{VP64} (Figure 4.2).

To this end, I adapted the split **mono** TMT-dCas9^{VP64} design to the original GPCR TANGO architecture (Figure 4.3a). In the first instance, I fused the dCas9(N) fragment to the BDKRB2 TANGO receptor (dCas9(N)-synBDKRB2), and retained NLS-

dCas9(C)^{VP64} in a soluble form (Figure 4.3a). Having only one dCas9 fragment tethered at a time, enabled me to disentangle the effects of introducing modifications into the original scaffold in pursuit of achieving improved OFF/ON state transitions. In contrast, if I had adapted the split **dual** Tmt-dCas9^{VP64} configuration for the GPCR TANGO architecture, I would have had to investigate the effects of amendments in the scaffolds of both tethered variants in an alternating fashion.

First, I tested the original GPCR TANGO design in which the inefficient TCS(Q'L) follows the V₂ tail directly (Figure 4.3a). Second, I conceived a modified design in which I incorporated an NES preceding the TCS(Q'L), by analogy with Tmt^{NES}-dCas9^{VP64} (Figure 3.7d,e). Third, I generated a control design in which a short linker sequence (GGS) was introduced instead of the NES (Figure 4.3a). Surprisingly, by introducing additional amino acids/linker sequence between the V₂ tail and TCS(Q'L) in form of an NES or GGS, the overall activation potential diminished (Figure 4.3b,c). Simultaneously, I observed an improved fold change in the activation level (Figure 4.3b,c). This indicates that the linker length is a critical component impacting the overall magnitude of reporter activation and responsiveness upon ligand stimulation. In particular, the V₂-NES-TCS(Q'L) version of dCas9(N)-synBDKRB2, as well as the V₂-GSS-TCS(Q'L) control variant, showed a ~24-fold activation upon bradykinin treatment compared to ~11-fold activation in the original V₂-TCS(Q'L) design (Figure 4.3b,c).

To potentially compensate for the reduced overall activation potential, I explored the idea of swapping the TCS sequence for other variants that enhance TEV protease binding affinity (Barnea et al. 2008; Wenjing Wang et al. 2017) (ordered by increasing TEV protease binding affinity, TCS(Q'L), ENLYFQ'L; TCS(Q'Y), ENLYFQ'Y; TCS(Q'G), ENLYFQ'G) (Figure 4.3a). The overall activation potentials of the split mono dCas9(N)-synBDKRB2 designs with more efficiently cleavable TCS displayed increased

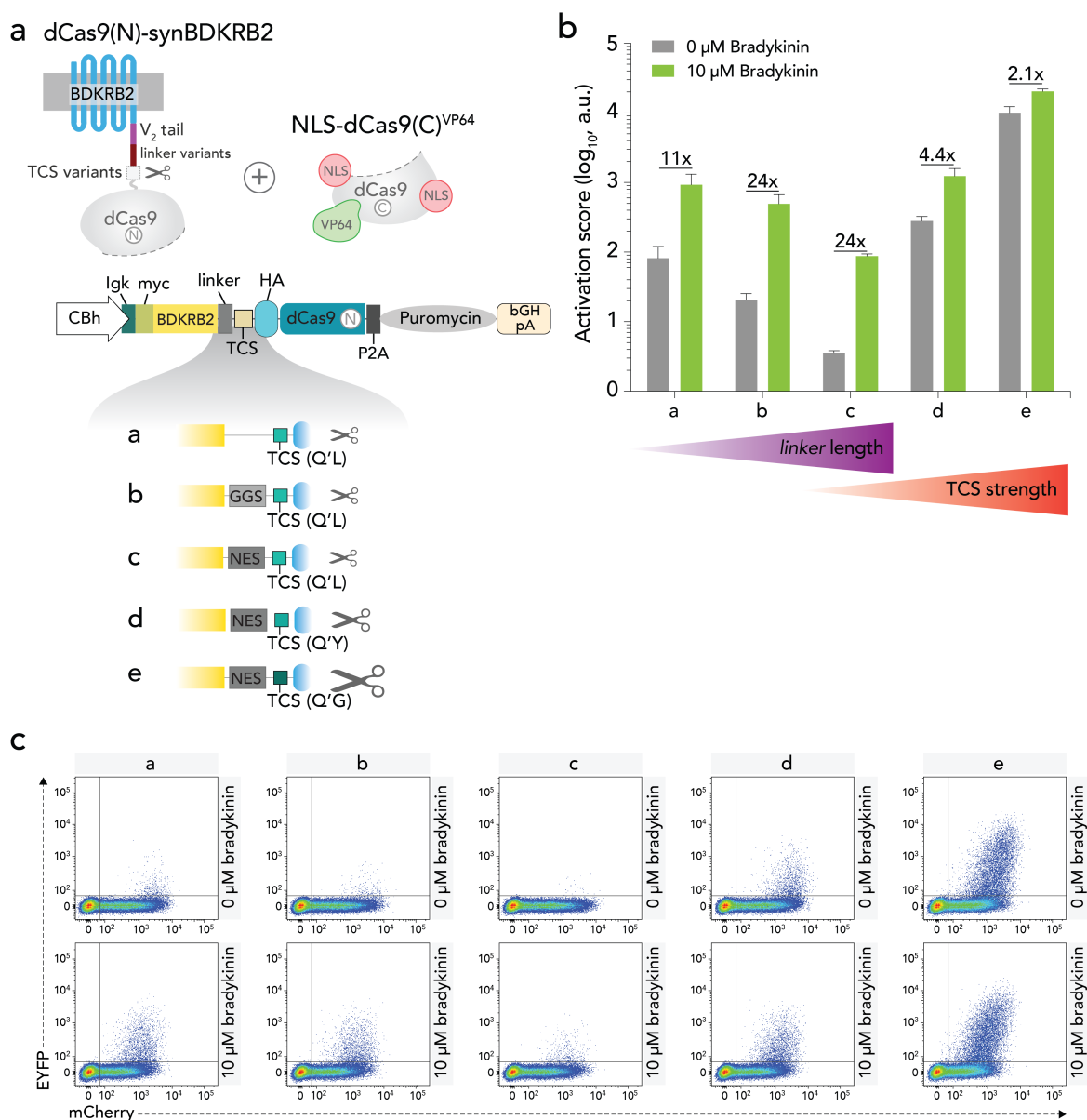


Figure 4.3 | Scaffold optimisation for split mono dCas9(N)-synBDKRB2. (a) Construct designs and cartoons of dCas9(N)-synBDKRB2 variants in combination with soluble NLS-dCas9(C)^{VP64}. dCas9(N)-synBDKRB2 constructs were tested with different TCS variants and linker lengths between the V₂-tail and the TCS (ordered by decreasing TEV protease binding affinity): TCS(Q'G), ENLYFQ'G; TCS(Q'Y), ENLYFQ'Y; TCS(Q'L), ENLYFQ'L. **(b)** Quantification of EYFP reporter activation of transiently transfected HTLA cells with the EYFP reporter system, NLS-dCas9(C)^{VP64}, and different split dCas9(N)-synBDKRB2 configurations in the absence or presence of 10 μM bradykinin. N = 3 biological replicates from one experiment, mean $\pm$ s.d., a.u. arbitrary units. **(c)** Representative flow cytometry scatter plots from (b).

transcriptional potencies compared to the TCS(Q'L) variant, albeit at the expense of OFF/ON state transitions (Figure 4.3b,c).

In summary, the original V₂-TCS(Q'L) GPCR TANGO scaffold displayed the best performance in the split mono dCas9(N)-synBDKRB2 configuration.

Next, I contemplated whether incorporating an NES sequence upstream of the BDKRB2 CDS could improve receptor performance. An NES may reduce background leakiness by enforcing cytoplasmic localisation of dCas9-based receptor fusion proteins. I designed two configurations in which the split mono dCas9(N)-synBDKRB2 is extended by only an NES or an NES flanked by GGS sequences, generating dCas9(N)-synNES-BDKRB2 and dCas9(N)-synNES(GGS)BDKRB2, respectively (Figure 4.4a). Benchmarking against the original split mono dCas9(N)-synBDKRB2, the alternative NES-bearing versions did not appear to improve the OFF/ON state transition (Figure 4.4b,c). One plausible explanation may be that the introduced sequences in the extracellular scaffold of the GPCR compromised its bradykinin binding ability and hence led to an overall reduction in receptor performance. Taken together, these results indicate that the split mono dCas9(N)-synBDKRB2 in the original GPCR TANGO scaffold displayed the best OFF/ON state transition.

The property of endogenous receptors to modulate transcriptional programmes in response to varying concentrations of input signals is of paramount importance. To this end, I investigated the ability of split mono dCas9(N)-synBDKRB2 to induce gene expression in a dose-dependent manner (Figure 4.4d). At increasing concentrations of bradykinin (0.01 nM – 10 μ M), an increase in reporter gene expression was observed, reaching a maximum level at 10 μ M bradykinin. The resulting dose-response curve revealed bradykinin-dependent EYFP activation at a half maximal effective ligand concentration of EC₅₀ = 223 nM (Figure 4.4d).

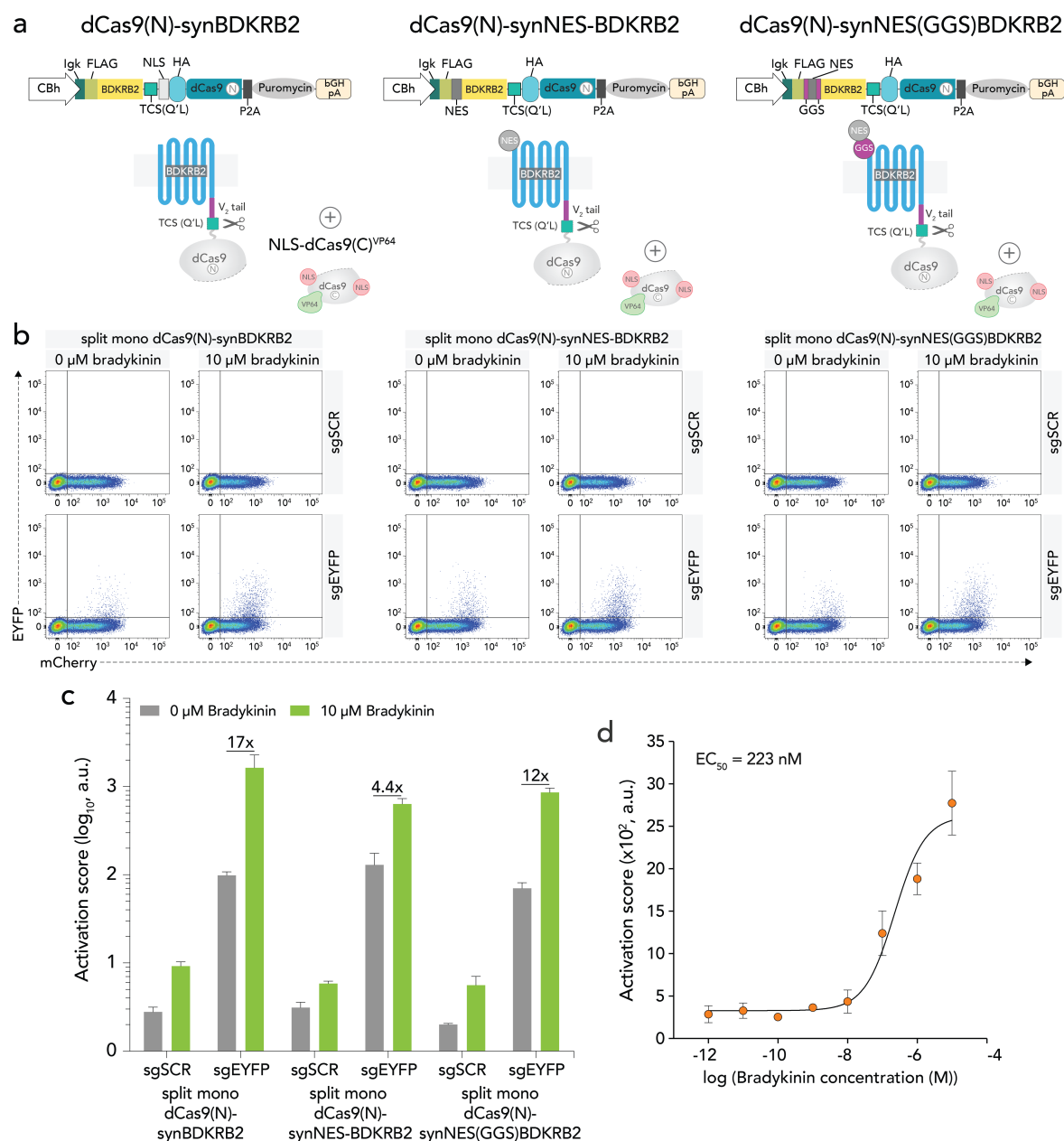


Figure 4.4 | Extracellular domain optimisation for split mono dCas9(N)-synBDKRB2. (a) Construct designs and cartoons of dCas9(N)-synBDKRB2, dCas9(N)-synNES-BDKRB2, and dCas9(N)-synNES(GGS)BDKRB2. **(b)** Representative flow cytometry scatter plots of transiently transfected HTLA cells with EYFP reporter, respective sgRNAs, and variants of dCas9(N)-synBDKRB2 in the absence or presence of 10 μM bradykinin and **(c)** the quantification thereof. $N = 3$ biological replicates from one experiment, mean $\pm$ s.d., a.u. arbitrary units. **(d)** Dose-response curve for dCas9(N)-synBDKRB2 with NLS-dCas9(C)^{VP64} in the EYFP reporter assay at increasing concentrations of bradykinin (EC_{50} = half-maximal effective concentration; each data point represents EYFP activation score from 3 biological replicates, mean $\pm$ s.d., a.u. arbitrary units; curve was fitted using a non-linear variable slope (three parameters) function in GraphPad Prism).

To assess whether this core split dCas9-based architecture is portable to other GPCRs, I adapted this particular configuration as well as the NES(GGS)-bearing version to the arginine vasopressin receptor 2 (AVPR2) (Figure 4.5a). dCas9(N)-synAVPR2 yielded a 2.4-fold activation while dCas9(N)-synNES(GGS)AVPR2 a 2.0-fold activation in the presence of vasopressin (Figure 4.5c). This data suggests that this scaffold architecture is indeed portable to other GPCR CDS and that the original GPCR TANGO design is more suitable for achieving robust OFF/ON state transitions (Figure 4.5b,c).

Subsequently, I explored the ability of split mono dCas9(N)-synGPCRs to activate the expression of several endogenous genes (*ASCL1*, *IL1B*, and *HBG1*) (Figure 4.6a-c). HTLA cells transiently transfected with split mono dCas9(N)-synBDKRB2 components and respective pools of SAM sgRNAs yielded low fold change upregulation of target genes, in the range from 2.6 to 4.7, at 10 μ M bradykinin (Figure 4.6a-c). Importantly, none of the genes were transcriptionally upregulated by bradykinin treatment alone when co-transfected with a control SAM sgRNA (SAM sgSCR) (Figure 4.6a-c). Split mono dCas9-synAVPR2 displayed a similar system performance when programmed to target the endogenous locus of *ASCL1*, yielding a 2.3-fold activation upon 1 μ M vasopressin stimulation (Figure 4.6d).

In summary, the original, unaltered GPCR TANGO core scaffold appears to be suitable for accommodating a split dCas9^{VP64} configuration. However, the performance of split mono dCas9-synGPCRs, in terms of OFF state leakiness and fold change activation, remained unsatisfactory for rewiring signal transduction pathways, irrespective of the GPCR or downstream target genes.

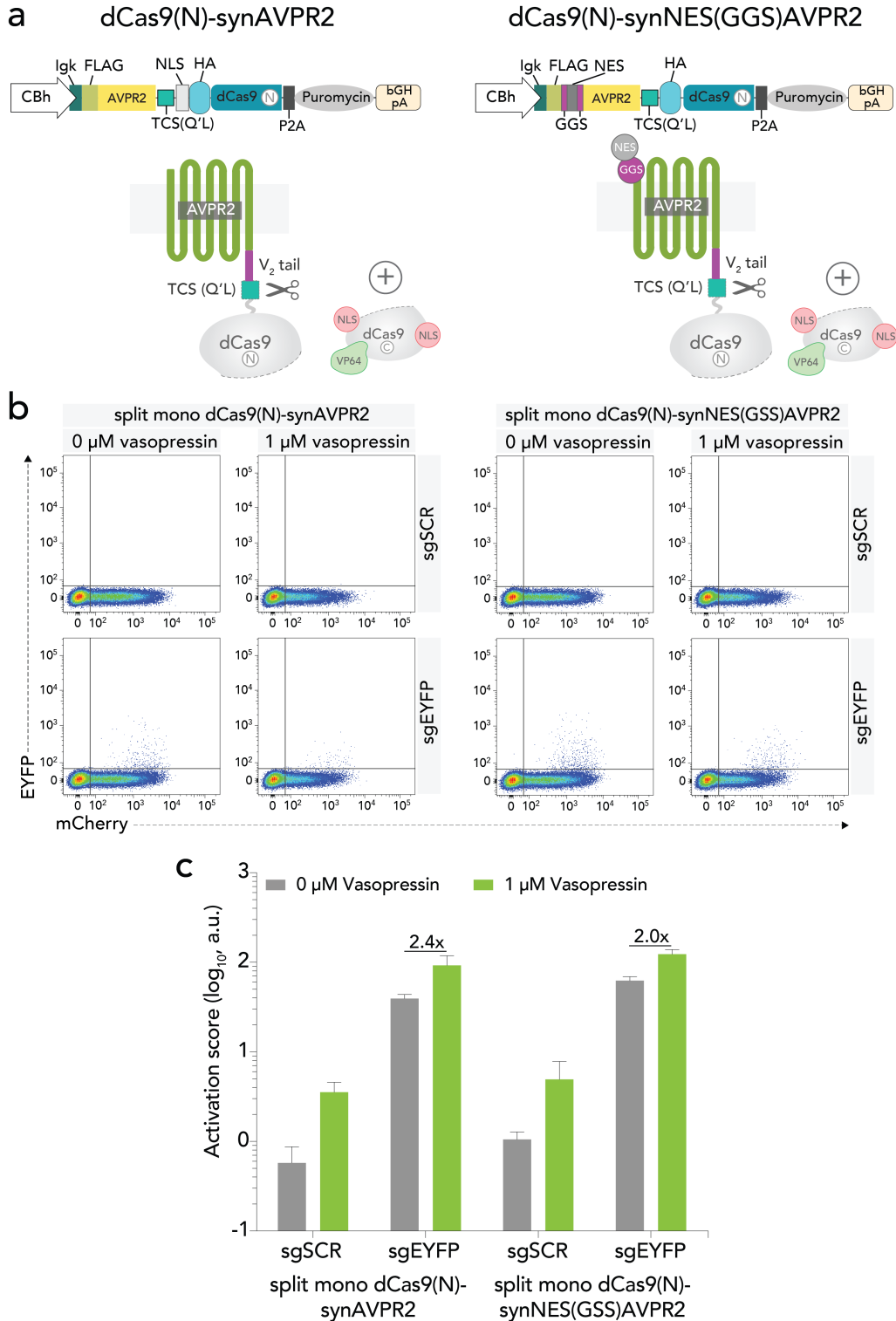


Figure 4.5 | Implementation of the split mono dCas9-synGPCR designs for AVPR2. (a) Construct designs and cartoons of dCas9(N)-synAVPR2 and dCas9(N)-synNES(GGS)AVPR2. (b) Representative flow cytometry scatter plots of transiently transfected HTLA cells with EYFP reporter, respective sgRNAs, and variants of dCas9(N)-synAVPR2 in the absence or presence of 1 μ M vasopressin. (c) Quantification of the EYFP reporter assay from (b). N = 3 biological replicates from one experiment, mean $\pm$ s.d., a.u. arbitrary units.

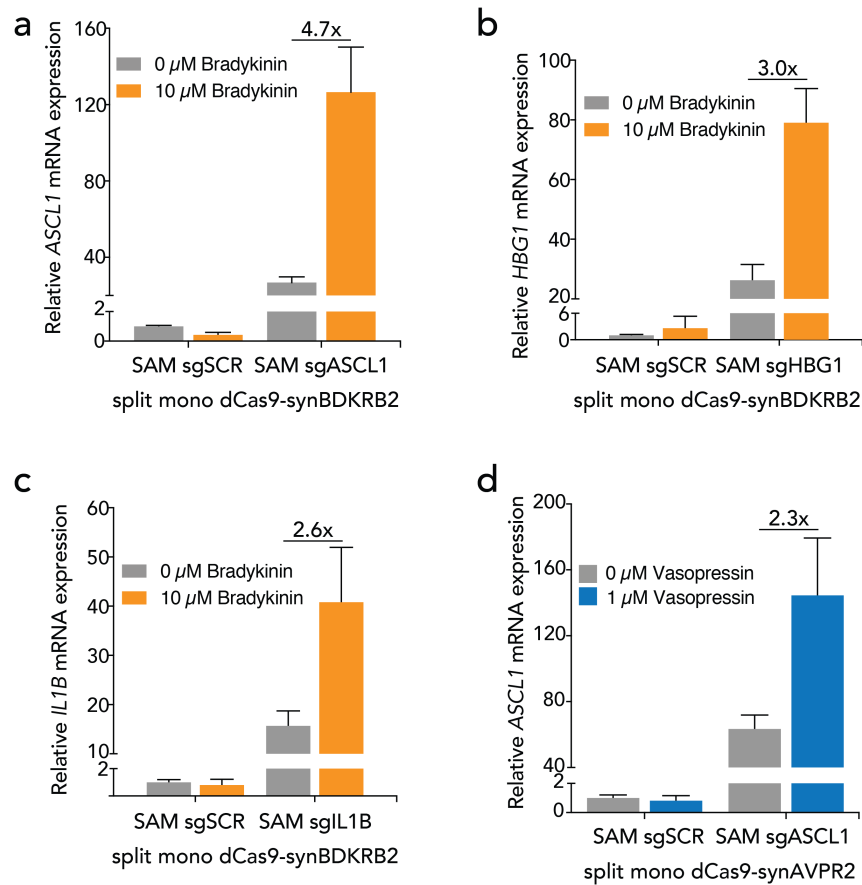


Figure 4.6 | Endogenous gene upregulation with split mono dCas9-synBDKRB2 and split mono dCas9-synAVPR2. (a-c) RT-qPCR of targeted activation of three individual endogenous genes: *ASCL1* (a), *IL1B* (b), and *HBG1* (c) in HTLA cells by transiently transfected split mono dCas9(N)-synBDKRB2 in the presence of 10 μ M bradykinin. **(d)** Employing split mono dCas9(N)-synAVPR2 to induce *ASCL1* transcription upregulation upon the administration of 1 μ M vasopressin. In all instances, data is represented as fold change relative to respective no-agonist condition. N = 3 biological replicates (x3 technical replicates), mean $\pm$ s.d.

4.2.3 Engineering an optimised split dual dCas9-synGPCR architecture with stringent OFF/ON state transition characteristics

4.2.3.1 Establishing and characterising a prototype split dual dCas9-synBDKRB2 receptor

Since the split **mono** dCas9-synGPCRs displayed relatively poor OFF/ON state transition characteristics, I sought to adapt a split **dual** dCas9-synGPCR architecture, based on the original GPCR TANGO scaffold. To this end, I fused the NLS-dCas9(C)^{VP64} downstream of the TCS(Q'L) sequence of BDKRB2 TANGO, creating dCas9(C)-synBDKRB2^{TCS(Q'L)} (Figure 4.7a). In this instance, both dCas9(N) and dCas9(C)^{VP64} are coupled to the BDKRB2 TANGO scaffold, collectively referred to as **split dual dCas9-synBDKRB2**. Surprisingly, HTLA cells transiently transfected with the EYFP reporter system and split dual dCas9-synBDKRB2 components yielded no EYFP activation under stimulation conditions (Figure 4.7b,c).

To assess whether the dCas9(C)-synBDKRB2^{TCS(Q'L)} fusion is competent for transcriptional activation *per se*, I combined it with TMT^{NES}-dCas9(N) (Figure 4.7a). Intriguingly, even in this combination, dCas9(C)-synBDKRB2^{TCS(Q'L)} did not lead to any reporter expression (Figure 4.7b,c).

I then transfected HEK293T cells with TMT^{NES}-dCas9(N) and dCas9(C)-synBDKRB2^{TCS(Q'L)} in the presence of co-delivered TEV protease (Figure 4.7d). I hypothesised that if the constructs are by itself functional, TEV protease-mediated liberation of dCas9 fragments should lead to EYFP expression. Indeed, under these experimental conditions, reporter activation was observed, indicating that the constructs are functional and TCS(Q'L) in dCas9(C)-synBDKRB2^{TCS(Q'L)} is accessible for proteolytic cleavage (Figure 4.7e,f). Relative to receptor behaviour, this result may suggest that the transient interaction of lowly expressed β -arrestin2-TEV in HTLA cells

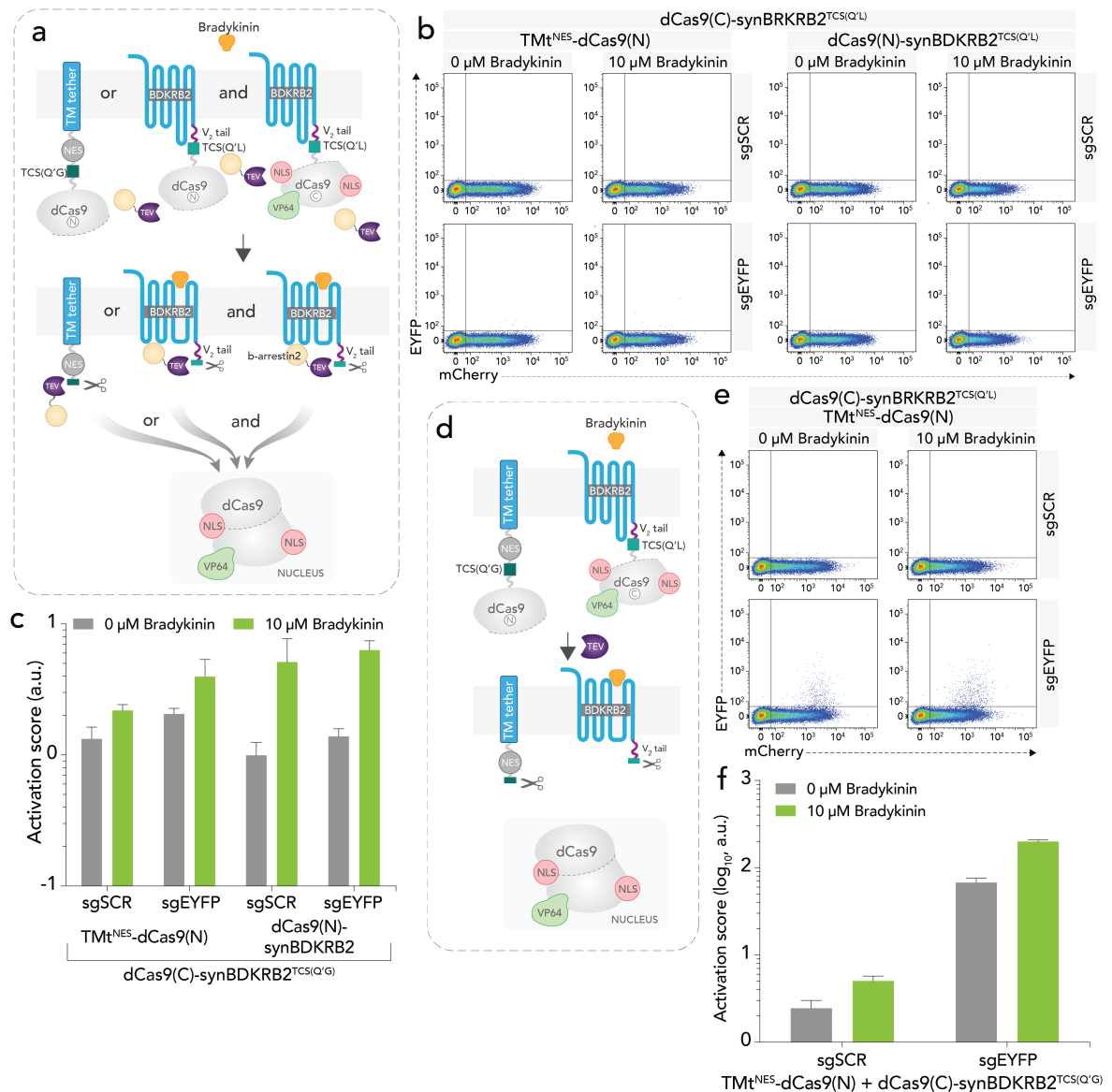


Figure 4.7 | Design and functional validation of a non-optimised split dual dCas9-synBDKRB2 architecture. (a) In HTLA cells, constitutively expressed β -arrestin2-TEV cleaves the TCS in TMt^{NES} -dCas9(N), in ligand-activated dCas9(N)-synBDKRB2^{TCS(Q'L)}, and in dCas9(C)-synBDKRB2^{TCS(Q'G)}. TMt^{NES} -dCas9(N) serves as a control for the supply of dCas9(N). (b) Representative flow cytometry scatter plots of transiently transfected HTLA cells with the EYFP reporter system in the absence or presence of 10 μ M bradykinin and (c) the quantification thereof. No reporter activation was observed. (d) In HEK293T cells, TMt^{NES} -dCas9(N) and dCas9(C)-synBDKRB2^{TCS(Q'G)} are co-expressed with TEV protease. (e) Representative flow cytometry scatter plots of transiently transfected HEK293T cells with the EYFP reporter system in the absence or presence of 10 μ M bradykinin and (f) the quantification thereof. dCas9(C)-synBDKRB2^{TCS(Q'G)} is functional and responsive to agonist treatment. (c,f) N = 3 biological replicates from one experiment, mean $\pm$ s.d., a.u. arbitrary units.

with an TCS(Q'L) in dCas9(C)-synBDKRB2^{TCS(Q'L)} is too short lived to facilitate cleavage. Furthermore, the β -arrestin2-TEV fusion protein in HTLA cells consists of the wild-type TEV protease that undergoes intramolecular auto-inactivation, thereby reducing its catalytic activity (Parks et al. 1995). The soluble TEV protease, however, is a stabilised mutant (S219V) with greater resistance to auto-inactivation and ~100-fold increased catalytic activity (Kapust et al. 2001).

To address this reasoning, I devised various TCS variants in dCas9(C)-synBDKRB2, similar to the experiments performed in Figure 4.3a (Figure 4.8a). One variant carrying the most efficient TCS, dCas9(C)-synBDKRB2^{TCS(Q'G)}, yielded an enormous 925-fold change upon bradykinin administration while maintaining minimal background activation in the uninduced state (Figure 4.8b,c). None of the other TCS variants generated such a satisfactory transcriptional response upon 10 μ M bradykinin stimulation (Figure 4.8b,c).

Taken together, these exploratory studies identified the combination of dCas9(N)-synBDKRB2^{TCS(Q'L)} and dCas9(C)-synBDKRB2^{TCS(Q'G)} as the most optimal receptor architectures in generating a robust OFF/ON state transition upon bradykinin administration. To the best of my knowledge, this configuration, collectively termed **dCas9(N/C)-synBDKRB2**, was the first robustly functioning synthetic receptor that deploys dCas9-TF as signal transduction module. dCas9(N/C)-synBDKRB2 was then used for all subsequent experiments and further characterised in detail.

To investigate the relationship between dCas9(N/C)-synBDKRB2 receptor abundance (plasmid copy number, input) and its performance (output), I first divided the reporter data into sequential mCherry fluorescence bins (transfection control) at increasing intensity (Figure 4.8d). Second, I calculated the EYFP mean fluorescence values and the percentage of activated cells within each mCherry bin (Figure 4.8e,f).

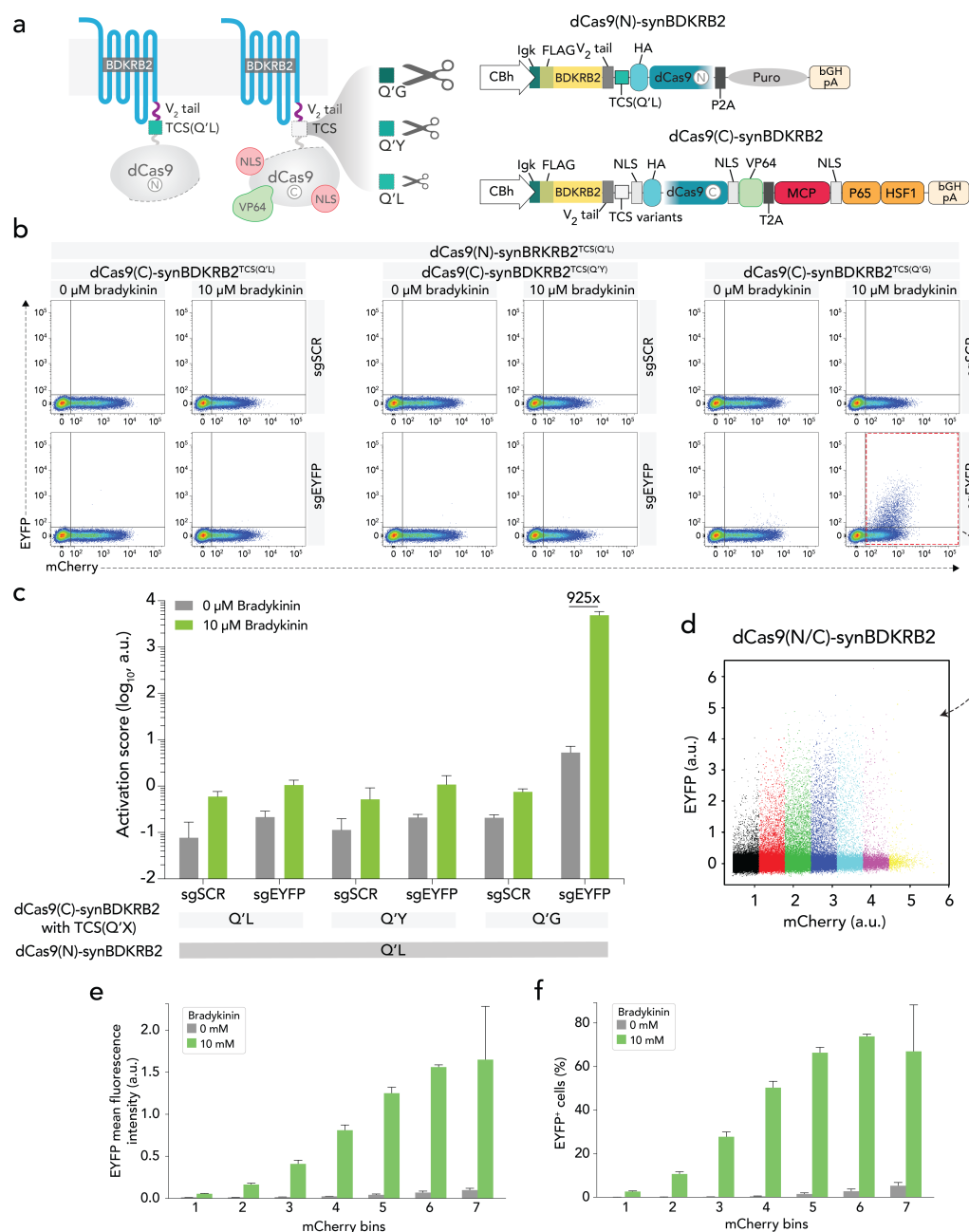


Figure 4.8 | Optimisation of split dual dCas9-synBDKRB2 designs. (a) Construct designs and cartoons of dCas9(N)-synBDKRB2 and dCas9(C)-synBDKRB2 TCS variants. **(b)** Representative flow cytometry scatter plots of transiently transfected HTLA cells with EYFP reporter system in absence or presence of 10 μ M bradykinin and **(c)** the quantification thereof. N = 3 biological replicates from one experiment, mean $\pm$ s.d., a.u. arbitrary units. **(d)** Data binning strategy based on mCherry fluorescence for best performing dCas9(N/C)-synBDKRB2 variant from (b). **(e)** Mean EYFP fluorescence intensity for each mCherry bin from (b), in the presence or absence of 10 μ M bradykinin (N = 3 biological replicates; mean $\pm$ s.d.; a.u. arbitrary units). **(f)** Percentage of EYFP positive cells per mCherry bin in the presence or absence of corresponding agonist (N = 3 biological replicates; mean $\pm$ s.d.). David Knapp (Fulga lab) performed the analysis in (d,e,f).

Since the mCherry construct is co-transfected with dCas9(N/C)-synBDKRB2, it provides a good approximation for the amount of dCas9(N/C)-synBDKRB2 in cells. This analysis revealed a positive correlation between the abundance of system components and the level of EYFP reporter activation. Across a broad range of expression levels, the activity of dCas9(N/C)-synBDKRB2 increased until reaching a maximum threshold, after which the system performance did not improve anymore (Figure 4.8e,f).

Next, I characterised the dose-response behaviour of dCas9(N/C)-synBDKRB2. At increasing concentrations of bradykinin agonist, this receptor displayed a robust ligand-mediated reporter gene activation with a half-maximal effective agonist concentration of 603 nM (Figure 4.9a).

Another important property of endogenous receptors is their temporal profile. In this regard, a question might be: How long does it take to generate a transcriptional output upon agonist stimulation? To this end, I investigated the temporal performance of dCas9(N/C)-synBDKRB2 by analysing transfected and agonist-stimulated HTLA cells at various time points (Figure 4.9b). Reporter expression was detected as early as 12 hours post stimulation and increased vigorously throughout the duration of the experiment (Figure 4.9b). This suggests that bradykinin is swiftly converted by activated and reconstituted dCas9 fragments into a transcriptional output response (Figure 4.9b).

Moving away from transgene activation, I tested the ability of this highly optimised dCas9(N/C)-synBDKRB2 receptor to control the induction of the endogenous gene *ASCL1*. In a dose-dependent manner, *ASCL1* was specifically upregulated by 5.2-, 11- and 13-fold in the presence of 0.4, 2 and 10 mM bradykinin, respectively (Figure 4.9c). These results indicate that dCas9(N/C)-synBDKRB2 surpassed the performance of split mono dCas9-synBDKRB2 (Figure 4.6a) by 277%.

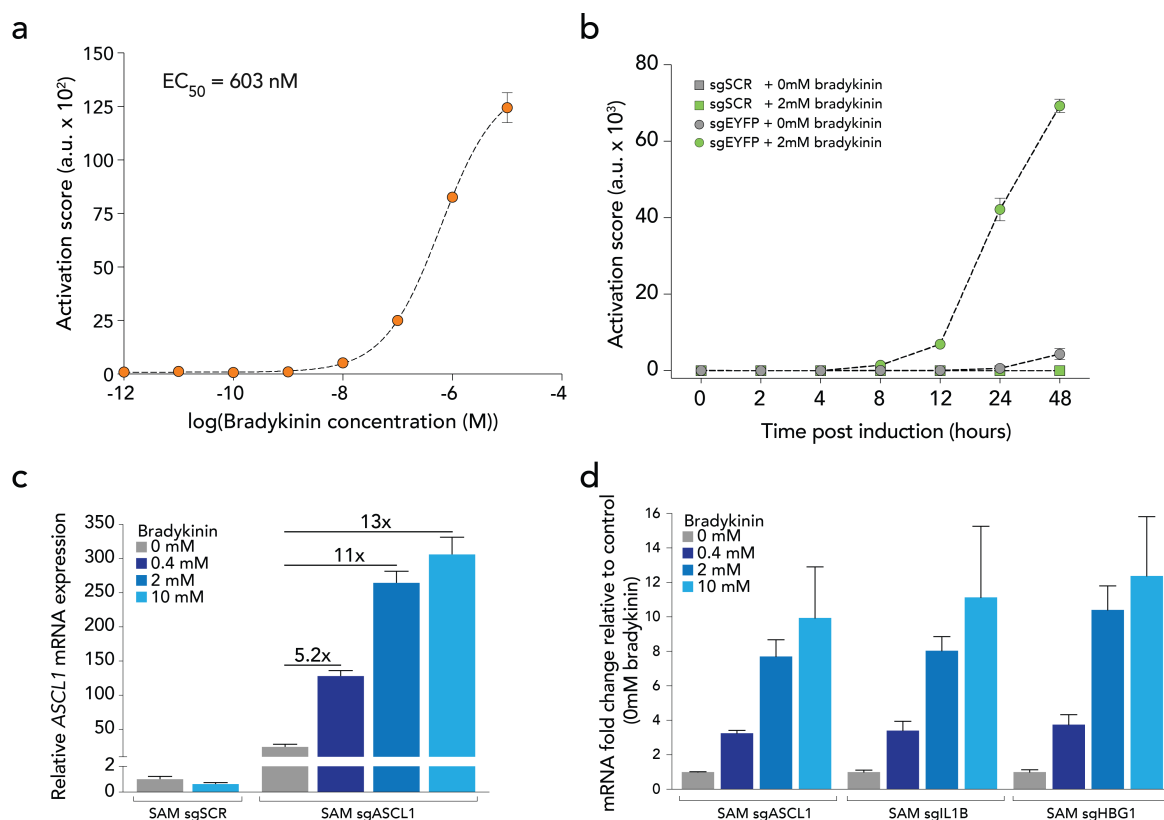


Figure 4.9 | Characterisation of dCas9(N/C)-synBDKBR2. (a) Dose-response curve for dCas9(N/C)-synBDKBR2 in the EYFP reporter assay at increasing concentrations of bradykinin (EC₅₀ = half-maximal effective concentration; each data point represents EYFP activation score from 3 biological replicates, mean $\pm$ s.d., a.u. arbitrary units; curve was fitted using a non-linear variable slope (four parameters) function in GraphPad Prism). (b) Time course analysis of dCas9(N/C)-synBDKBR2 agonist response in HTLA cells. After 20 hours, the media was changed to HTLA full media containing no agonist (0 μ M bradykinin) or supplemented with 2 μ M bradykinin. EYFP reporter expression was measured by flow cytometry at 0, 2, 4, 8, 12, 24, and 48 hours post induction. Data points for sgSCR with and without bradykinin overlap in the diagram. EYFP activation score was calculated from three biological replicates (n = 3, mean $\pm$ s.d.; a.u., arbitrary units). (c, d) Dose-dependent and multiplexed endogenous gene activation with split dual dCas9-synBDKBR2. (c) Endogenous *ASCL1* transcriptional gene activation in transiently transfected HTLA cells in a dose-responsive manner. Data is represented as fold change relative to sgSCR no-agonist condition. (d) Dose-dependent and multiplexed transcriptional activation of three endogenous genes (*ASCL1*, *IL1B*, *HBG1*) at increasing agonist concentrations (0.4, 2, 10 μ M bradykinin) at the same time. Displayed as fold change relative to respective no-agonist condition (0 μ M bradykinin). In all instances, data represents n = 3 biological replicates (x3 technical replicates), mean $\pm$ s.d.

A notable advantage of employing dCas9-TF is the opportunity to control virtually any user-defined endogenous gene or gene set by simply programming dCas9-TF with a combination of sgRNAs. To demonstrate this capacity in the context of dCas9(N/C)-synBDKRB2, I simultaneously induced the expression of a panel of endogenous genes (*ASCL1*, *IL1B*, and *HBG1*) using validated SAM sgRNAs (Konermann et al. 2015). This experiment revealed the robust upregulation of all three genes in a dose-dependent manner (Figure 4.9d).

Taken together, the scaffold design of dCas9(N/C)-synBDKRB2 represents the first successful installation of a highly potent, dCas9-based synthetic receptor configuration, characterised by minimal background activation in the uninduced state and stringent OFF/ON state transition in the presence of its ligand.

4.2.3.2 Expanding the repertoire of ligand-activated dCas9-synGPCRs

To harness the diversity of input-sensing GPCRs, I next investigated whether the core split dCas9 architecture of dCas9(N/C)-synBDKRB2 could be applied to other GPCRs.

First, I converted vasopressin-responsive AVPR2 into dCas9(N/C)-synAVPR2. Upon hormone stimulation at different concentrations, a potent dose-response curve was observed in the EYFP reporter assay (Figure 4.10a).

Second, I adapted the C-X-C chemokine receptor type 4 (CXCR4) into the dCas9-synGPCRs scaffold, generating dCas9(N/C)-synCXCR4. At high concentrations of the chemokine protein SDF-1 (stromal cell-derived factor 1, also known as C-X-C motif chemokine 12 (CXCL12)), a robust reporter gene expression was observed, indicating an agonist threshold level required for the transition from OFF to ON state (Figure 4.10b).

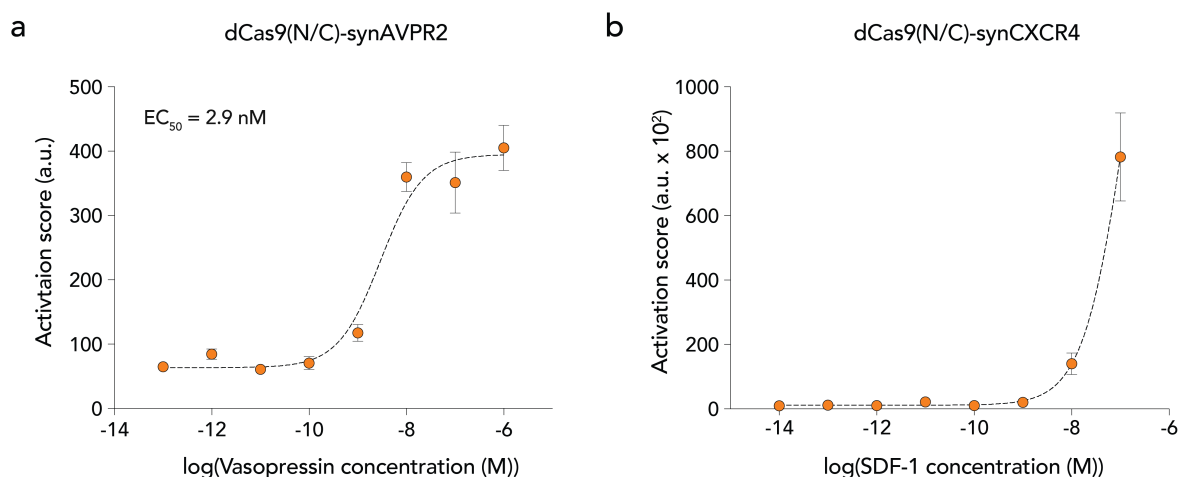


Figure 4.10 | Vasopressin and SDF-1 responsive dCas9-synGPCRs. (a) Dose-response curve for dCas9(N/C)-synAVPR2 at increasing concentrations of vasopressin. EC_{50} = half-maximal effective concentration. (b) Dose-response curve for dCas9(N/C)-synCXCR4 at increasing concentrations of SDF-1. In all instances, each data point represents EYFP activation score from 3 biological replicates, mean $\pm$ s.d., a.u. arbitrary units; curve was fitted using a non-linear variable slope (three parameters) function in GraphPad Prism).

Third, I turned the lysophosphatidic acid receptors 1 and 2 (LPAR1 and LPAR2) into dCas9(N)-synLPAR1, dCas9(N)-synLPAR2, dCas9(C)-synLPAR1, and dCas9(C)-synLPAR2, and tested their reporter activation potency in a combinatorial screen in the presence or absence of 20 μ M lysophosphatidic acid (Figure 4.11a,b). dCas9(N/C)-synLPAR1, in which each dCas9 fragment is fused to LPAR1, performed best with a 191-fold change in EYFP induction. The dose-response curve for dCas9(N/C)-synLPAR1 revealed robust transgene activation at higher agonist concentration, with an $EC_{50} = 1.82$ mM (Figure 4.11c).

Last, I adapted the sweet taste receptors T1R2 and T1R3 to the dCas9-synGPCR scaffold, creating dCas9(N)-synT1R2, dCas9(N)-synT1R3, dCas9(C)-synT1R2, and dCas9(C)-synT1R3. I then tested their activation potency in a combinatorial screen in the presence or absence of 100 mM D-glucose (Figure 4.12a,b). All combinations revealed weak reporter activation, however, dCas9(N/C)-synT1R3 displayed a promising 7.2-fold change induction (Figure 4.12a,b). Furthermore, analysis at increasing concentrations of D-glucose demonstrated a clear dose-dependent behaviour (Figure 4.12c).

To further characterise the dCas9(N/C)-synT1R3 receptor, I investigated its capacity to respond to other nutritional sugars (Figure 4.12d). In direct comparison with 100 mM D-glucose, 100 mM D-fructose and Coca-Cola Classic (equivalent sugar content) yielded 9.9- and 10.6-fold EYFP reporter activation, respectively. This confirmed that dietary sugars can bind to the T1R3 receptor, as previously reported (Nelson et al. 2001; G. Q. Zhao et al. 2003; Y. Zhang et al. 2003; P. Jiang et al. 2004; Nie et al. 2005; Chandrashekar et al. 2006; Assadi-Porter et al. 2008; Assadi-Porter et al. 2010).

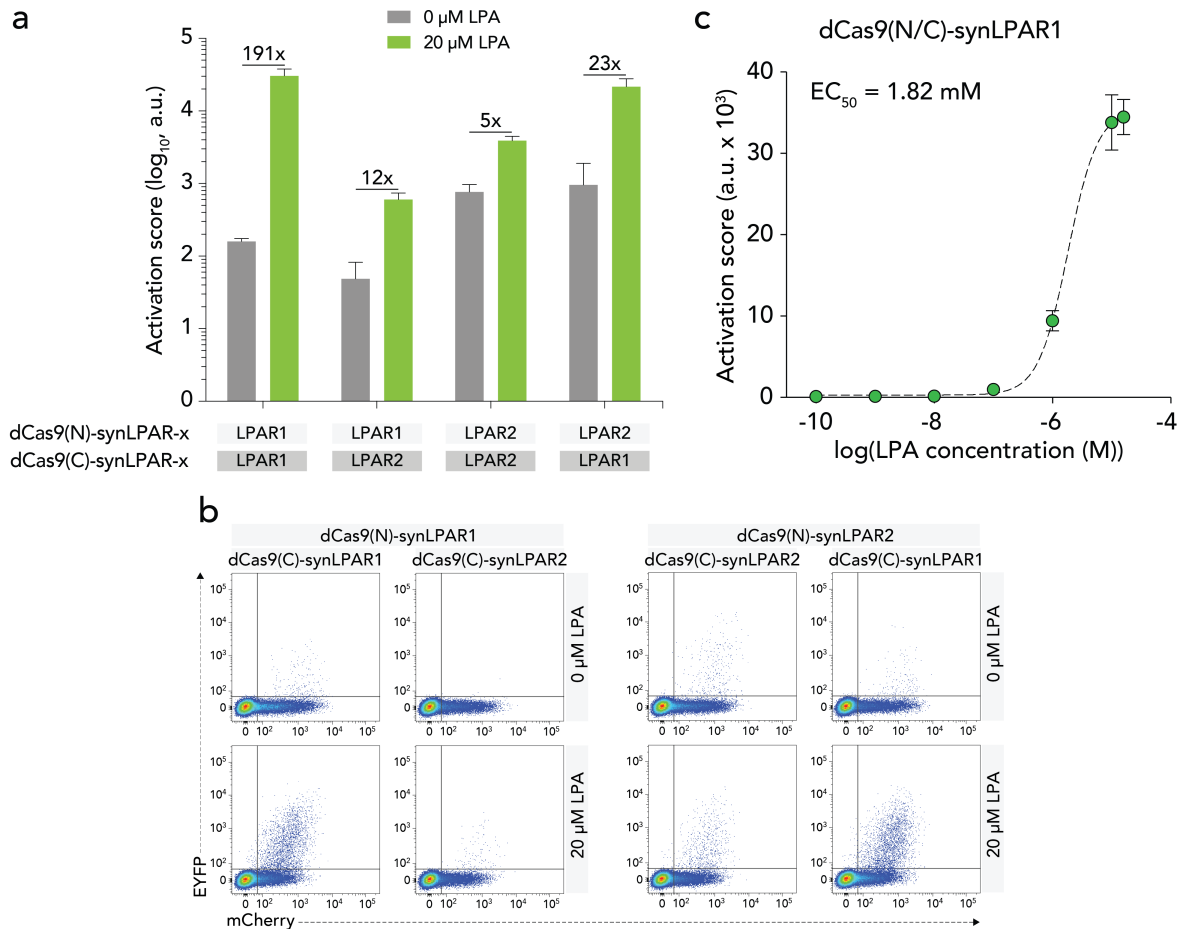


Figure 4.11 | Lipid responsive dCas9(N/C)-synLPARs. (a) Identification of the best performing dCas9(N/C)-synLPAR pair via the EYFP reporter assay in transiently transfected HTLA cells in the absence or presence of 20 μ M LPA. N = 3 biological replicates from one experiment, mean $\pm$ s.d., a.u. arbitrary units. (b) Representative flow cytometry scatter plots of (a). (c) Dose-response curve for dCas9(N/C)-synLPAR1 at increasing concentrations of LPA. (EC₅₀ = half-maximal effective concentration; each data point represents EYFP activation score from 3 biological replicates, mean $\pm$ s.d., a.u. arbitrary units; curve was fitted using a non-linear variable slope (four parameters) function in GraphPad Prism).

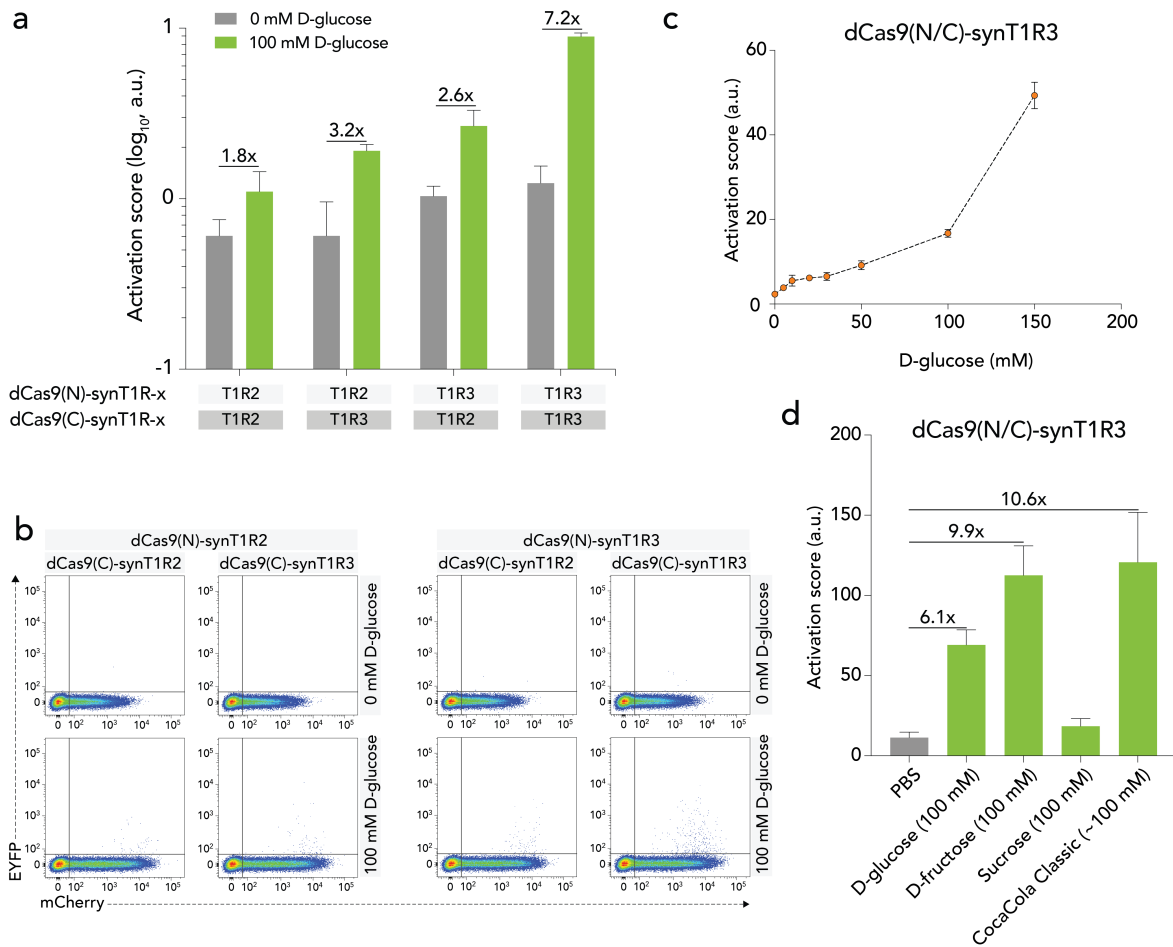


Figure 4.12 | D-glucose responsive dCas9(N/C)-synT1Rs. (a) Identification of the best performing dCas9(N/C)-synT1R pair via the EYFP reporter activation assay in transiently transfected HTLA cells in the absence or presence of 100 mM D-glucose. N = 3 biological replicates from one experiment, mean $\pm$ s.d., a.u. arbitrary units. (b) Representative flow cytometry scatter plots of (a). (c) Dose-response diagram for dCas9(N/C)-synT1R3 at increasing concentrations of D-glucose. Dotted line connecting data points was not fitted. (d) EYFP reporter assay for dCas9(N/C)-synT1R3-transfected HTLA cells with a variety of nutritional sugar compounds. N = 3 biological replicates from one experiment, mean $\pm$ s.d., a.u. arbitrary units.

4.2.4 Boolean AND logic gates for multi-factorial input processing

Building on the successful adaptation of a diverse group of GPCRs to the dCas9-synGPCR scaffold (section 4.2.3.2), I next asked whether it is possible to deploy two different GPCRs in combination. In doing so, I would create a Boolean AND logic gate operation (see Chapter 1.2.3.1). Theoretically, only in the presence of both respective ligands, a transcriptional output should be observed. Such AND gate configurations offer the opportunity to gain temporal and spatial control over receptor output programmes in designer cells.

In the first instance, I combined dCas9(N)-synAVPR2 with dCas9(C)-synBDKRB2 (Figure 4.13a,b). In the absence of both agonists or in the presence of only vasopressin, no reporter activation was observed (Figure 4.13b). In the presence of bradykinin only, however, the system was capable of eliciting an agonist dose-dependent EYFP activation (Figure 4.13b). Interestingly, increasing concentrations of vasopressin agonist appeared to compromise bradykinin-mediated activation and reduced overall reporter fluorescence (Figure 4.13b). This effect might be caused by the physiologically reciprocal properties of these two agonists. Vasopressin is a potent vasoconstrictor, whereas kinins (such as bradykinin) are generally vasodilatory and together they play major roles in regulating the ionic balance and water content in kidneys (Gonzalez & Figueroa 1999). When the receptors on the dCas9 fragments were interchanged (dCas9(N)-synBDKRB2 and dCas9(C)-synAVPR2), a similar reciprocal relationship was observed (Figure 4.13c,d). Surprisingly, this receptor AND gate combination resulted in high background activation across all agonist concentrations (Figure 4.13d). Because of these poor performances, I explored alternative receptor combination in pursuit of creating functional Boolean AND logic gates.

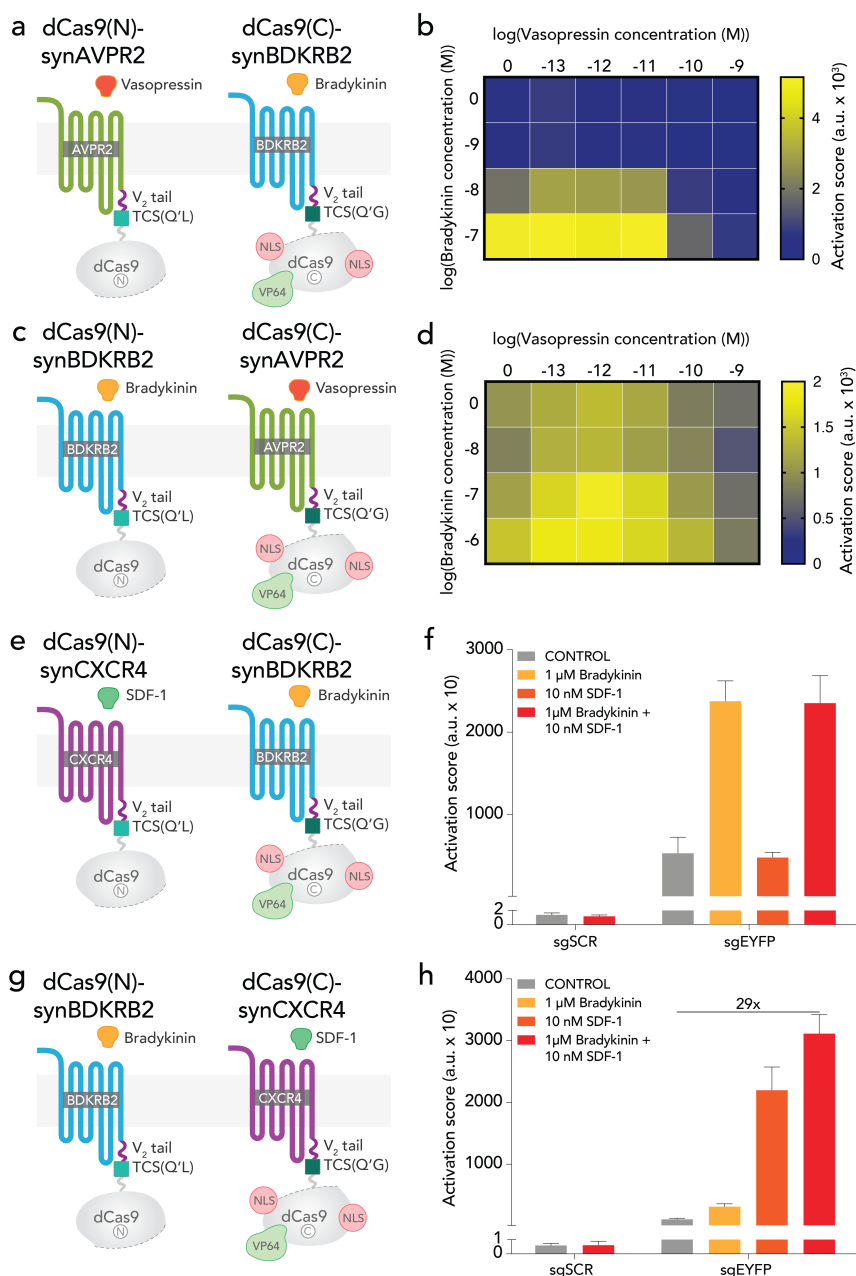


Figure 4.13 | 2-input Boolean AND logic gates with heterologous dCas9-synGPCRs. (a, c) Schematic representation of two AND logic gate configurations for the heterologous combination of dCas9(N/C)-synBDKRB2 and dCas9(N/C)-synAVPR2 at increasing concentrations of bradykinin and vasopressin, respectively. (b) EYFP reporter assay heatmap for dCas9(N)-synAVPR2 and dCas9(C)-synBDKRB2. (d) EYFP reporter assay heatmap for dCas9(N)-synBDKRB2 and dCas9(C)-synAVPR2. (b,d) N = 1 from one experiment, mean $\pm$ s.d.; a.u., arbitrary units. (e, g) Schematic representation of two AND logic gate configurations for the heterologous combination of dCas9(N/C)-synBDKRB2 and dCas9(N/C)-synCXCR4 at increasing concentrations of bradykinin and SDF-1, respectively. (f) EYFP reporter assay for dCas9(N)-synCXCR4 and dCas9(C)-synBDKRB2. (h) EYFP reporter assay for dCas9(N)-synBDKRB2 and dCas9(C)-synCXCR4. (f,g) N = 3 biological replicates from one experiment, mean $\pm$ s.d., a.u. arbitrary units.

To this end, I paired dCas9(N)-synCXCR4 with dCas9(C)-synBDKRB2 (Figure 4.13e) and observed strong EYFP activation only when the receptor linked to NLS-dCas9(C)^{VP64} was activated by 1 μ M bradykinin (Figure 4.13f). This level of activation was equivalent to dual agonist stimulation. Strikingly, equivalent results were obtained with the reciprocal pair of dCas9(N)-synBDKRB2 and dCas9(C)-synCXCR4 (Figure 4.13g,h). Stimulating dCas9(C)-synCXCR4 with 10 nM SDF-1 generated similar strong activation levels compared to dual agonist treatment. Together, these results suggest a pattern in which the liberation of NLS-dCas9(C)^{VP64} led to nearly full system activation. Therefore, the applicability of the multiplexed dCas9-synGPCRs using the current design would be rather limited and will require further optimisation.

With the aim to create a tightly regulated, dual input dependent dCas9-synGPCR configuration, I utilised an inert compound as second input stimulus that conditions the activity of a destabilisation domain. The well characterised, structurally unfolded domain from *Escherichia coli* dihydrofolate reductase (DHFR) has been successfully employed to induce accelerated proteosomal degradation of a number of fusion proteins (including dCas9-TF) in mammalian cells (Iwamoto et al. 2010; Maji et al. 2017; Kleinjan et al. 2017). Upon trimethoprim (TMP) administration, DHFR is reversibly stabilised, inhibiting rapid degradation of the fusion protein (Iwamoto et al. 2010). I fused DHFR to dCas9(N)-synBDKRB2, generating dCas9(N)-synBDKRB2^{DHFR}, and tested it in combination with dCas9(C)-synBDKRB2 (Figure 4.14a). Encouragingly, no transgene activation was observed in the absence of both agonists (bradykinin and TMP) or in the presence of single inducers, compared to control conditions (Figure 4.14b,c). However, concomitant administration yielded robust and gradual EYFP activation at increasing TMP concentrations (Figure 4.14b,c).

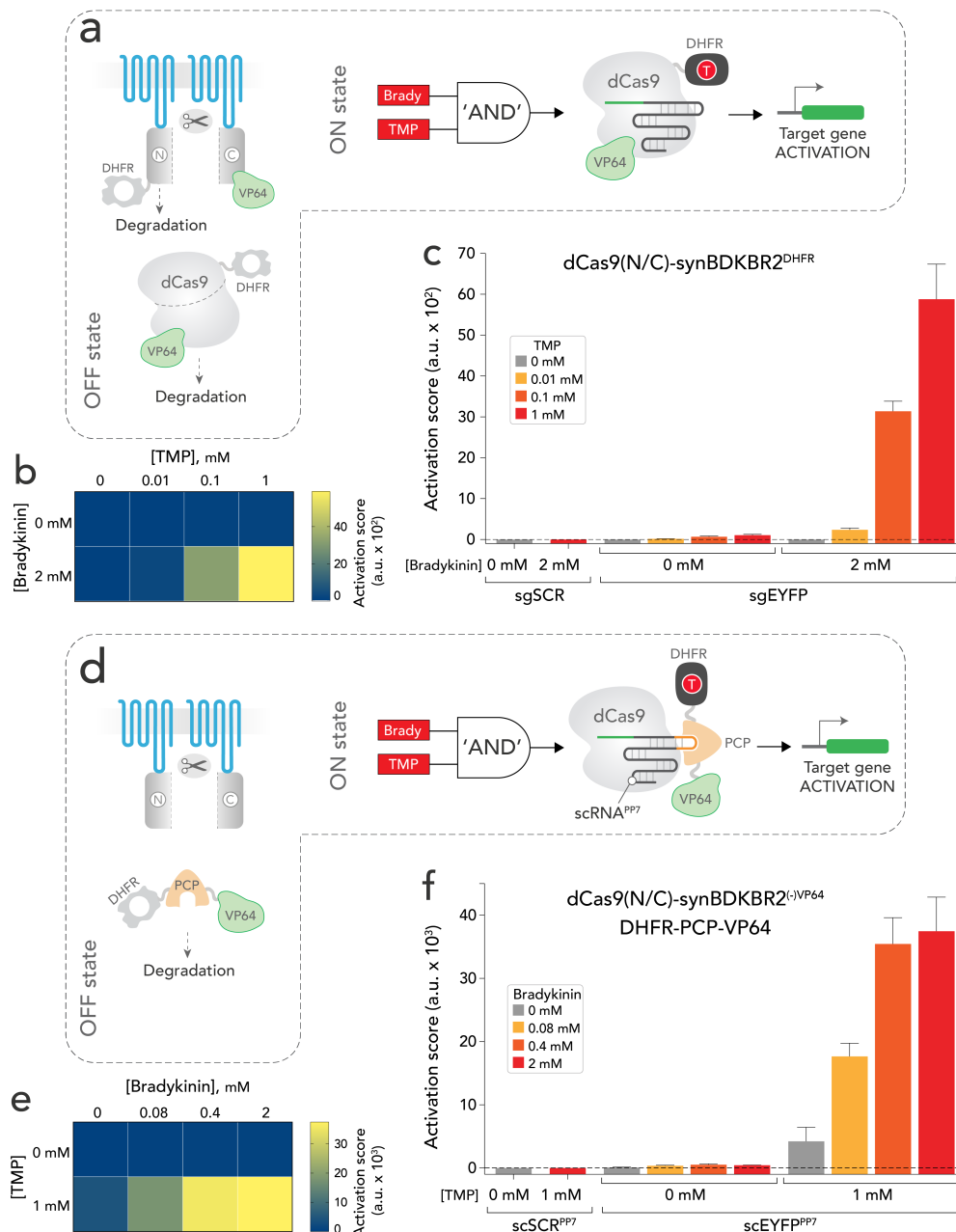


Figure 4.14 | 2-input Boolean AND logic gates with dCas9(N/C)-synBDKBR2 and an inert drug. (a, d) Schematic representation of two AND logic gate strategies for dCas9(N/C)-synGPCR activation. In the OFF state (absence of TMP) the DHFR domains target either dCas9(N) (or reconstituted dCas9-VP64) (a) or the VP64 effector fusion protein (d) for proteasomal degradation. ON state transition required simultaneous bradykinin-mediated receptor activation and TMP-dependent DHFR stabilisation. (b) Heatmap and (c) quantification of EYFP reporter assay in the absence or presence of bradykinin and increasing concentrations of TMP. (e) Heatmap and (f) quantification of reporter activation assay in the absence or presence of TMP and increasing concentrations of bradykinin. In all instances, N = 3 from one experiment, mean $\pm$ s.d.; a.u., arbitrary units; sgSCR = scramble sgRNA, scSCR^{PP7} = scramble PP7-aptamer scaffold guide RNA; scEYFP^{PP7} = EYFP-targeting PP7-aptamer scaffold guide RNA. Figure adapted from Baeumler et al. 2017.

In a slightly different iteration, I segregated the VP64 activation domain from dCas9(C)-synBDKRB2, creating dCas9(C)-synBDKRB2^{(-)VP64}, and instead utilised a soluble DHFR-PCP-VP64 effector fusion protein to conditionally control the activity of dCas9(N/C)-synBDKRB2^{(-)VP64} (Figure 4.14d) (Maji et al. 2017; Kleinjan et al. 2017). This is achieved by deploying a modified sgRNA that incorporates the PP7 aptamer (scaffold guide RNA (scRNA^{PP7})). scRNA^{PP7} is bound by the RNA-binding protein PCP, bringing the activation domain VP64 in proximity to dCas9 and thereby forming a potent transcriptional activator (Figure 4.14d) (Zalatan et al. 2015; Maji et al. 2017). In the absence of either inducer, or in the presence of 1 μ M bradykinin alone, no background leakiness was detected (Figure 4.14e,f). In contrast, a concentration-dependent reporter gene activation was observed at increasing doses of TMP (Figure 4.14e,f). It should be noted, however, that this AND logic gate configuration did not perform as tight as the dCas9-DHFR-based AND gate (Figure 4.14a-c).

Taken together, the successful implementation of these TMP-based AND gates demonstrate viable strategies for exogenously controlling the activity of dCas9-synGPCRs.

To expand the spectrum of applications, I next tried to converge the successful adaptation of an inert drug-responsive domain and heterologous dCas9-synGPCR pairs to design 3-input Boolean AND logic gates. Using a VP64-deficient dCas9(C)-synGPCR, dCas9(C)-synCXCR4^{(-)VP64}, I engrafted a FKRP domain downstream of the TCS(Q'L) in dCas9(N)-synBDKRB2, to generate dCas9(N)-synBDKRB2^{RI}. The VP64 domain was then fused to FRB (Figure 4.15a). This particular design addressed the observation in which liberated NLS-dCas9(C)^{VP64} leads to full system activation in the presence of its respective GPCR agonist alone (Figure 4.13). By segregating the VP64 from NLS-dCas9(C)^{VP64}, I aimed to reduced its background activity.

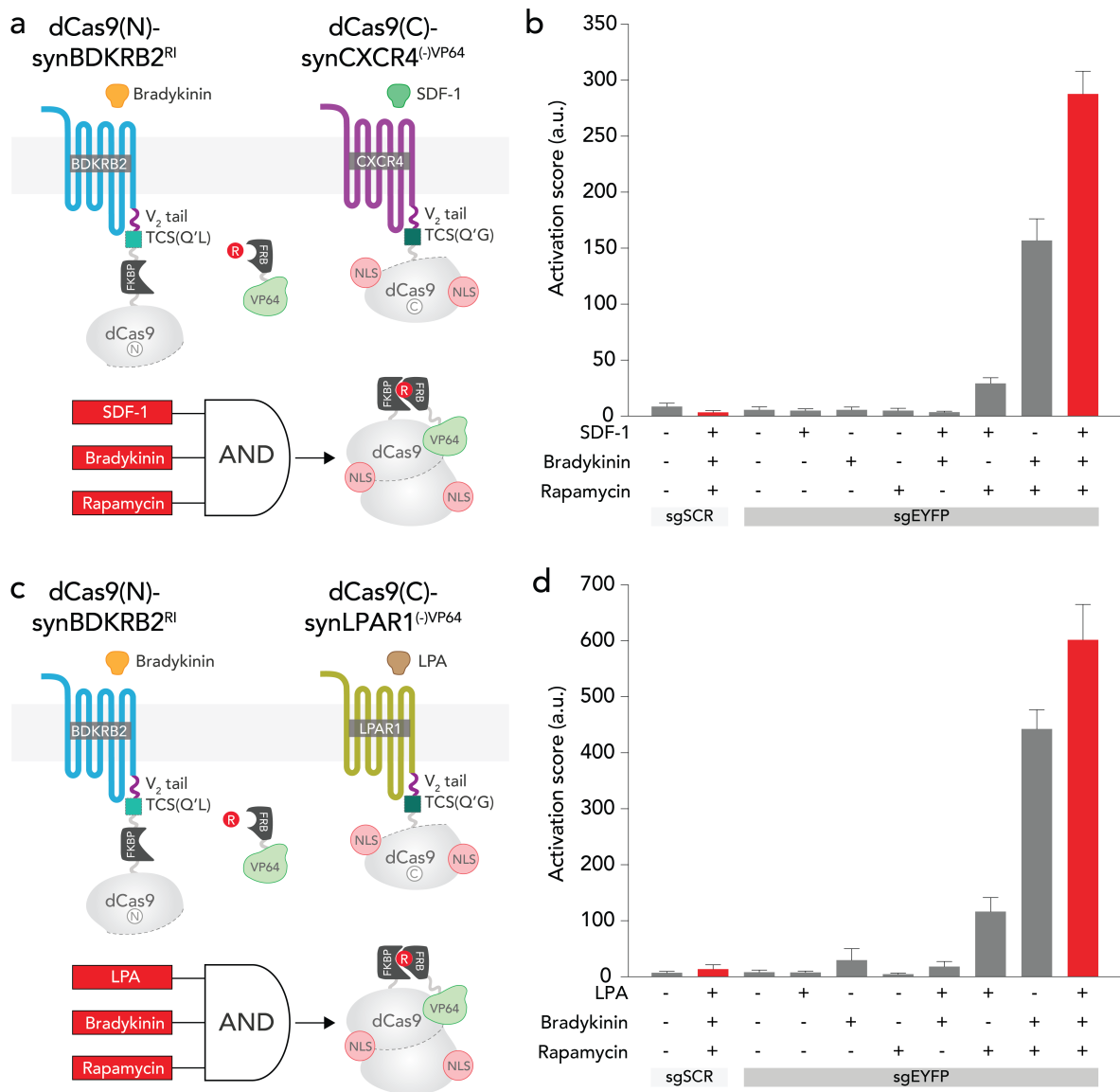


Figure 4.15 | 3-input Boolean AND logic gates with two dCas9-synGPCRs and one inert drug. (a, c) Schematic representation of two 3-input AND logic gate strategies for the heterologous combination of dCas9(N)-synBDKRB2^{RI} and either dCas9(C)-synCXCR4^{(-)VP64} or dCas9(C)-synLPAR1^{(-)VP64}. In both instances, the output response is dependent on a triple input: Bradykinin and SDF-1 or LPA mediate dCas9 fragments to be liberated from the membrane and rapamycin induces the recruitment of VP64 to the reconstituted dCas9, yielding a functional dCas9^{VP64} signal transduction module. (b) EYFP reporter assay for dCas9(N)-synBDKRB2^{RI} and dCas9(C)-synCXCR4^{(-)VP64} in the presence or absence of SDF-1, bradykinin, and rapamycin. N = 3 biological replicates from one experiment, mean +/- s.d., a.u. arbitrary units. (d) EYFP reporter assay for dCas9(N)-synBDKRB2^{RI} and dCas9(C)-synLPAR1^{(-)VP64} in the presence or absence of LPA, bradykinin, and rapamycin. N = 3 biological replicates from one experiment, mean +/- s.d., a.u. arbitrary units.

The output response in this 3-input Boolean AND gate is dependent on three inputs: bradykinin and SDF-1 mediate dCas9 fragment release from the membrane and its subsequent auto-assembly, while rapamycin induces the recruitment of VP64 to the reconstituted dCas9, yielding a functional dCas9-TF (Figure 4.15a). This configuration generated no reporter output above baseline in the absence of all inducers or when only one stimulus was present (Figure 4.15b). However, delivery of any two inducers resulted in system activation and EYFP expression, which reached maximum intensity in the presence of all three inputs (Figure 4.15b).

A similar behaviour was observed for a 3-input Boolean AND logic gate in which dCas9(C)-synCXCR4^{(-)VP64} had been replaced with the LPA-inducible dCas9(C)-synLPAR1^{(-)VP64} (Figure 4.15c,d).

Nonetheless, these 3-input Boolean AND logic gate configurations require further refinement since they show strong activation with only a subset of agonists compared to the full activation achieved by all ligands.

In summary, the collected data demonstrate that Boolean AND logic gates are powerful tools in regulating signalling circuits and their implementation will be imperative for stringent multi-input integration. However, my attempts at establishing such complex circuits uncovered certain limitations that will need to be addressed in future studies.

4.3 Discussion

In this chapter, I have successfully established a versatile dCas9-TF as intracellular signal transduction module for GPCRs. Based on the original GPCR TANGO scaffold (Kroeze et al. 2015), the optimised split dual dCas9-synGPCR represents the most advanced and potent programmable chimeric GPCR design to this date.

This chimeric receptor provides a novel link between the extracellular and intracellular environment, allowing the transduction of signals and input stimuli into user-defined, non-native, multi-gene output programmes. The core split dCas9 architecture was readily adaptable to various GPCRs demonstrating portability and generalisability. In total, four GPCRs were converted into dCas9-synGPCRs, all of which showed stringent OFF/ON state transitions. These GPCRs are representative of a pool of 135 class A and B receptors that have been recently incorporated into the GPCR TANGO scaffold (Kroeze et al. 2015). This suggests that the remaining GPCRs could also be converted into dCas9-synGPCR, expanding the repertoire of endogenous ectodomain/recognition domains for a plethora of future implementations.

Furthermore, I harnessed the glucose-responsive T1R3 receptor in the dCas9-synGPCR design, which is to my knowledge the first time a class C GPCR has been adapted to the GPCR TANGO scaffold. This is particularly noteworthy because there are 20 class C GPCRs (Bräuner-Osborne et al. 2007).

Overall, GPCRs adapted to dCas9-synGPCRs represent a great opportunity for integrating a variety of input signals, ranging from light, small molecules, nucleotides, sugars to lipids, hormones, and proteins (Pierce et al. 2002).

4.3.1 Improving the scaffold of dCas9-synGPCRs

The dCas9-synGPCR scaffold integrates the full length GPCR CDS with a programmable dCas9-TF. To facilitate dCas9-TF-mediated signalling, a V₂ tail is appended to all GPCR scaffolds by default. It has been shown that the additional C-terminal V₂ tail enhances recruitment of β -arrestin2-TEV to any given GPCR (Barnea et al. 2008; Kroeze et al. 2015). The transient interaction between β -arrestin2-TEV and the GPCR is crucial for the subsequent liberation of the transcription factor and its transcriptional activation (Barnea et al. 2008; Kroeze et al. 2015).

However, in some cases for class A and class B GPCRs, the native C-terminal tail sequence facilitated a more stable interaction with β -arrestin2 than appending the V₂ tail to the same GPCR (Hanson et al. 2009). In other cases, it was more beneficial to replace the endogenous C-terminal tail of the receptor with the V₂ tail or alternative C-terminal sequences from other GPCRs (Oakley et al. 1999; Oakley et al. 2000; Hamdan et al. 2005; Vrecl et al. 2009; Thomsen et al. 2016). Additionally, mutations introduced into wildtype β -arrestin2 can lead to enhanced affinity to the agonist-activated class A GPCRs (Oakley et al. 2001; Vrecl et al. 2004), potentially strengthening β -arrestin2-TEV-mediated transcription factor activation.

Taken together, to further improve current OFF/ON state transition properties and to facilitate the adaptation of additional GPCRs, the standardised dCas9-synGPCR scaffold design may benefit from customisation for each GPCR.

4.3.2 Considerations for the endogenous signalling pathways of GPCRs and their ligand specificities

There are at least two major areas of concern for exploiting dCas9-based GPCRs that require thorough consideration and further investigation: (i) canonical and non-canonical signalling pathways of GPCRs and (ii) the ligand specificities of GPCRs.

(i) Interference from endogenous downstream GPCR signalling. It is likely that agonist-activated dCas9-synGPCRs signal not only through dCas9-TF but also activate their canonical and non-canonical signalling pathways, mediated by G-proteins and G-protein independent transducers (Reiter & Lefkowitz 2006; Ritter & Hall 2009). It remains to be seen on a case by case basis, if such endogenous programmes interfere with the synthetic pathway or even generate adverse effects.

A promising approach to circumvent canonical and non-canonical signalling has been proposed by engineering the domains of GPCRs that interact with G-proteins and

β -arrestins (Arber et al. 2017). Introducing mutations or interchanging different GPCR domains (GPCR chimera) made it possible to favour signalling towards specific G-proteins or bias signalling predominantly towards β -arrestin (Oakley et al. 2000; Wei et al. 2003; Zhai et al. 2005; Shenoy et al. 2006; J. D. Urban et al. 2007; Violin & Lefkowitz 2007; Guettier et al. 2009; Kapp et al. 2012; Y. Xu et al. 2014; Siuda et al. 2015; Arber et al. 2017). For example, hM3Dq (a designer GPCR based on the human M3 muscarinic receptor) natively signals through G_{α_q} -type G-proteins and β -arrestin (Armbruster et al. 2007; Nakajima & Juergen Wess 2012). Various modifications were shown to bias its signalling towards G_{α_i} (Armbruster et al. 2007), G_{α_s} (Guettier et al. 2009), exclusively towards G_{α_q} (J. Hu et al. 2016), or primarily towards β -arrestin (Nakajima & Juergen Wess 2012).

In addition, it was proposed to design minimal orthogonal GPCR signalling systems in which all of the native signalling pathways are abolished (Arber et al. 2017). Translating this idea to dCas9-synGPCRs, potential cross-talk from native signalling pathways would be sacrificed in favour of the non-native, dCas9-TF-based pathway.

(ii) GPCR ligand specificities. Some agonists can trigger the activation of more than one receptor and *vice versa*, thereby contributing to unspecific receptor activation and cross reactivity (Lazennec & Richmond 2010; Viola et al. 2012; Scholten et al. 2012). One intriguing solution to counteract ambiguous receptor activation is to engineer the recognition domains of GPCRs. Very recently, the human dopamine receptors were genetically modified to exhibit skewed affinities for previously high- and low-affinity native ligands (X. Feng et al. 2017). By tuning the selectivity of the ligand binding domain, dCas9-synGPCRs could be customised to reliably detect only a particular endogenous agonist.

Engineering the affinity of a GPCR for a particular ligand also provides an opportunity to modulate the magnitude of receptor activation and as a consequence,

the strength of the downstream signalling response. This can be accomplished either by increasing the sensitivity towards a scarcely available agonist or by reducing the sensitivity in the case of an abundant ligand which would otherwise lead to maximum activation.

In some cases, GPCR signalling is susceptible to endogenous allosteric modulators (Wootten et al. 2013; Changeux & Christopoulos 2016; Foster & Conn 2017). Such allosteric regulators bind to topographically distinct sites within the GPCR and either positively or negatively potentiate agonist-mediated GPCR activation (Milligan & Smith 2007; May et al. 2007; Changeux & Christopoulos 2016; Foster & Conn 2017). Similarly, antagonists interfere with potent GPCR activation by binding to the agonist binding site (steric regulators) (Foster & Conn 2017). Future experiments involving allosteric modulators and antagonists might disclose novel means for the regulation and fine-tuning of dCas9-synGPCR response programmes.

In addition to tuning agonist sensitivity, ligand binding domains have also been engineered to abrogate endogenous agonist binding. Receptor activated solely by a synthetic ligand (RASSL) and designer receptor exclusively activated by designer drugs (DREADDs) lack specificity to their cognate ligands but exhibit binding affinity to synthetic compounds, circumventing the activation of bystander receptors (Pei et al. 2008; Conklin et al. 2008; Jürgen Wess et al. 2013; D. J. Urban & Roth 2015). This is potentially useful for controlling dCas9-synGPCR activity remotely and creating alternative logic gates in addition to those explored in this chapter.

4.3.3 Prospective Boolean AND logic gates with dCas9-synGPCRs

Generally, AND gates provide a great tool for a higher order of control over system activation and provide a safety switch for therapeutically-relevant downstream applications. In synthetic biology, these gates can be installed on the level of DNA,

RNA, and/or protein (Dominguez et al. 2016; Lo & L. Qi 2017; Richter et al. 2017). Dimerisation domains incorporated into dCas9-TF have been shown to conditionally generate highly robust and specific output responses (Y. Gao et al. 2016; Bao et al. 2017). Nonetheless, the frequently used FKBP/FRB dimerization domain pair requires induction by rapamycin, which has undesirable immunosuppressive effects (Liberles et al. 1997; Bayle et al. 2006). However, more favourable analogues, such as AP21967/rapalog, have been reported which appear to alleviate some of these concerns (Liberles et al. 1997; Bayle et al. 2006; C.-Y. Wu, Roybal, et al. 2015). Moreover, there are numerous alternative dimerisation domain pairs that offer potential orthogonal, multiplexed input processing (Liang et al. 2011; Miyamoto et al. 2012; Kennedy et al. 2010; Nihongaki, Kawano, et al. 2015; Nihongaki, Yamamoto, et al. 2015; Yazawa et al. 2009; Levskaya et al. 2009; Redchuk et al. 2017; Y. Gao et al. 2016; Bao et al. 2017).

4.3.4 Alternative CRISPR-TF-based synthetic GPCRs

Subsequent to my publication in September 2017, Lei Stanley Qi's laboratory reported an alternative approach for the implementation of dCas9-based GPCRs in December 2017 (Kipniss et al. 2017). This bipartite system was termed CRISPR ChaCha (Figure 4.16a). In this case, the GPCR receptor architecture is rudimentarily based on the original GPCR TANGO scaffold, in which a V_2 tail proceeds the GPCR CDS. Instead of tTA (in the case of GPCR TANGO) or one of the two split dCas9 fragments (in case of dCas9-synGPCRs), full length TEV protease was grafted downstream of the V_2 tail, to create GPCR- V_2 -TEVp (Kipniss et al. 2017). To complement this GPCR scaffold, the authors fused human β -arrestin2 via the low affinity TCS(Q'L) to the N-terminus of dCas9-VPR, generating ARRB2-TCS-dCas9-VPR (Figure 4.16a).

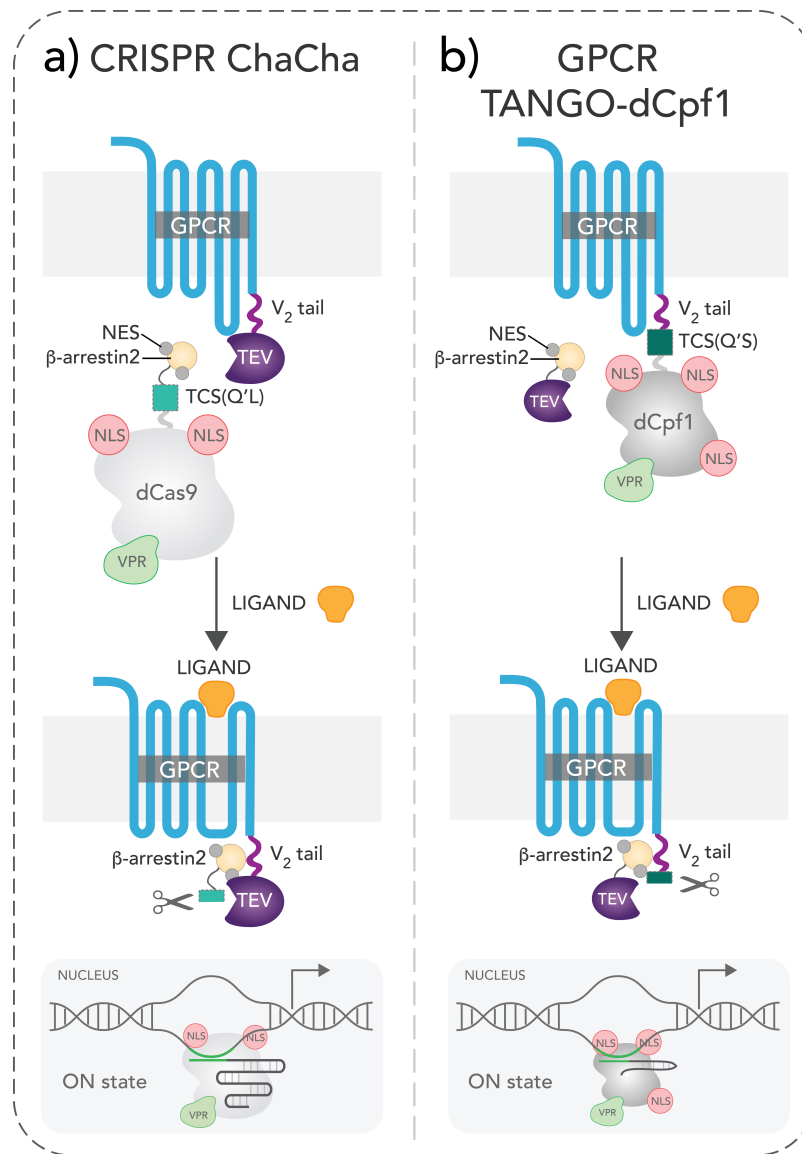


Figure 4.16 | Alternative CRISPR-TF-based GPCR architectures. (a) CRISPR ChaCha combines a GPCR fused to the V₂ tail and TEV protease, and as a second component, β-arrestin2 that is linked via a TCS to dCas9^{VPR}. Agonist stimulation leads to proteolytic cleavage of the TCS, facilitating the translocation of dCas9^{VPR} into the nucleus (Kipniss et al. 2017). (b) GPCR TANGO-dCpf1^{VPR} combines the original TANGO scaffold with an alternative CRISPR-TF, dCpf1^{VPR} (Y. Liu et al. 2017). The release mechanism of dCpf1^{VPR} is equivalent to tTA in the GPCR TANGO design (Barnea et al. 2008; Kroeze et al. 2015).

ARRB2-TCS-dCas9-VPR contains two different localisation sequences: (i) ARRB2 is flanked by NES sequences (amino acid sequence LALKLAGLDI) to facilitate cytoplasmic localisation of the chimeric fusion protein prior to TEV-mediated proteolytic cleavage; (ii) dCas9 fused to the second generation VPR transactivator (VP64-NLS-p65-RTA, (Chavez et al. 2015)) carries two SV40 NLS sequences, which direct dCas9^{VPR}, after TEV cleavage, into the nucleus.

The best performing CRISPR ChaCha receptor architecture resulted in a 9.4-fold EGFP induction (Kipniss et al. 2017). In a head-to-head comparison, the authors observed a similar level of absolute reporter gene activation for soluble full length dCas9^{VPR}. This suggests that, in relation to the small fold change of the CRISPR ChaCha receptor, there is very high unspecific background activation (leakiness) in the uninduced state (Kipniss et al. 2017). This is not surprising since tethered or non-tethered full length dCas9-TF is extremely potent when expressed in HEK293T cells (Chapter 3).

By analogy to my study, the authors also tested the original GPCR TANGO scaffold fused to full length dCas9^{VPR} (GPCR-TANGO-dCas9-VPR). Confirming my results, GPCR-TANGO-dCas9^{VPR} generated only a 1.2-fold change and displayed massive leakiness. This observation reinforces the finding that fusing full length dCas9-TF to the original GPCR TANGO scaffold is not an adequate means for creating chimeric GPCR receptors.

CRISPR ChaCha displays an improved performance of 7.3-fold (9.4x/1.2x) compared to GPCR TANGO. By analogy, dCas9-synBDKRB2 displayed ~400-fold activation (925x/2.3x) compared to BDKRB2-NLS-dCas9^{VP64} (Figure 4.8 and Figure 4.2, respectively).

In another study published in December 2017, a research group from Shenzhen University also observed that deploying full length dCas9-TF in the original GPCR

TANGO scaffold generated very high background activation (Y. Liu et al. 2017). In this instance, however, the authors replaced the dCas9-TF with the full length endonuclease-deficient Cpf1 from *Acidaminococcus* sp. BV3L6 (dAsCpf1) (Figure 4.16b). The GPCR-TANGO-dAsCpf1-TF construct contains a high affinity TCS, ENLYFQS, upstream of dAsCpf1-TF. The human β -arrestin2-TEV fusion protein is flanked by NES sequences to maintain cytoplasmic localisation, similar to Kipniss et al. (2017). This configuration achieved robust ligand-mediated transcriptional activation with GPCR-TANGO-dAsCpf1^{VPR} and transcriptional repression with GPCR-TANGO-dAsCpf1^{KRAB}. In all instances, they ascertained little unspecific background activation (Y. Liu et al. 2017).

Intriguingly, dAsCpf1 in the original GPCR TANGO scaffold displays robust OFF/ON state transition but not dCas9. Liu et al. (2017) argue that fusing a large protein downstream of the V₂ tail compromises the conformation of the GPCR. However, AsCpf1 (~144 kDa) is only 6kDa lighter than SpCas9 (~150 kDa). The same size argument was made by Kipniss et al. (2017) albeit comparing TEV protease (~28 kDa) to dCas9-VPR (~220 kDa). In support of this hypothesis, the original GPCR TANGO architecture utilises the small artificial transcription factor tTA (~37 kDa) (Barnea et al. 2008; Kroeze et al. 2015). However, definitively addressing this issue will require further studies.

Nonetheless, dCpf1 represents an interesting addition to CRISPR-TF-based GPCRs for several reasons. First, Cpf1 homologs contain an RNase domain that facilitates autonomous processing of multiple crRNAs from its respective CRISPR array, thus enabling simple multiplexed targeting of endogenous genes (Zetsche et al. 2017; Zhong et al. 2017; Fonfara et al. 2016; Y. Liu et al. 2017). Second, owing to different crRNAs scaffolds, dCpf1 and dCas9 can be used together without cross-talk, facilitating orthogonal gene regulation. Third, dAsCpf1 has a different PAM requirement

and can tolerate better adenosine thymine-rich protospacer sequences than dSpCas9, expanding the number of target sites in the genome (Jinek et al. 2012; Cong et al. 2013; Mali, Yang, et al. 2013; Zetsche, Gootenberg, et al. 2015).

Chapter 5 : Engineering programmable chimeric dCas9-synVEGFRs

Most of the data presented in this chapter has been published in: Baeumler, T.A., Ahmed, A.A. & Fulga, T.A., 2017. Engineering Synthetic Signaling Pathways with Programmable dCas9-Based Chimeric Receptors. *Cell Reports*, 20(11), pp.2639–2653.

5.1 Chapter Aims

- Conceiving and optimising a core split dCas9 minimal architecture for dimerisation-dependent chimeric dCas9-based receptor tyrosine kinases (dCas9-synRTKs).
- Characterising the performance of a prototype dCas9-based vascular endothelial growth factor receptor (dCas9-synVEGFR).
- Engineering Boolean AND logic gates for multi-factorial input processing.
- Head-to-head comparison of dCas9-synVEGFR to previously reported V-MESA-dCas9 receptor.

5.2 Results

5.2.1 Design and optimisation of dCas9-synVEGFR with stringent OFF/ON state transition characteristics

Having engineered a split, membrane-bound dCas9^{VP64} architecture (Chapter 3), I next sought to apply this system to create a chimeric dCas9-synRTK.

As a representative of the RTK family, the vascular endothelial growth factor receptors (VEGFRs) share a conserved, single transmembrane domain topology

(Lemmon & Schlessinger 2010). They are activated by specifically engaging extracellular bivalent growth factors that mediate receptor dimerisation (Olsson et al. 2006; Lemmon & Schlessinger 2010).

To engineer a dCas9-synVEGFR, I replaced the endogenous, intracellular signalling unit of the endogenous receptor with the split dCas9 fragments. Specifically, I fused the NES-TCS(Q'G)-dCas9(N) and the TCS(Q'G)-NLS-dCas9(C)^{VP64} fragments downstream of the VEGFR transmembrane domain (TMD). Removing the intracellular part abolishes the potential for native signalling via second messengers, mediators, and auxiliary proteins. As a means of a TEV protease-mediated release mechanism, I engrafted the split TEV(N) and TEV(C) fragments (Chapter 3.2.3.2) upstream of dCas9(N) and dCas9(C), respectively (Figure 5.1a,b).

Because endogenous VEGF receptor 1 (VEGFR1) and VEGF receptor 2 (VEGFR2) can form homodimers as well as heterodimers in the presence of the ligand VEGFA (Simons et al. 2016), I generated all possible receptor combinations and tested their performance using the EYFP reporter assay (Figure 5.1a,b). Notably, the parental cell line of HEK293T, HEK293, was shown to express extremely low transcript levels of VEGFR1 and no VEGFR2, based on RNA-sequencing data (Human Protein Atlas available from www.proteinatlas.org; Uhlén et al. 2015).

Transiently transfected HEK293T cells with all system components revealed that dCas9(N)-synVEGFR2^{TCS(Q'G)} paired with dCas9(C)-synVEGFR1^{TCS(Q'G)} yielded the highest fold change of 5.5 in the presence of 50 ng VEGFA121-expressing plasmid (Figure 5.1c,d). Although this pair displayed the lowest background activation in relation to the other constructs, the OFF state remained at a high level (Figure 5.1c,d). One hypothesis that may account for this observation is that the receptor components interact while being trafficked through the secretory pathway. This can lead to high local concentrations, (i) favouring the auto-assembly of split dCas9 fragments and/or (ii) TEV

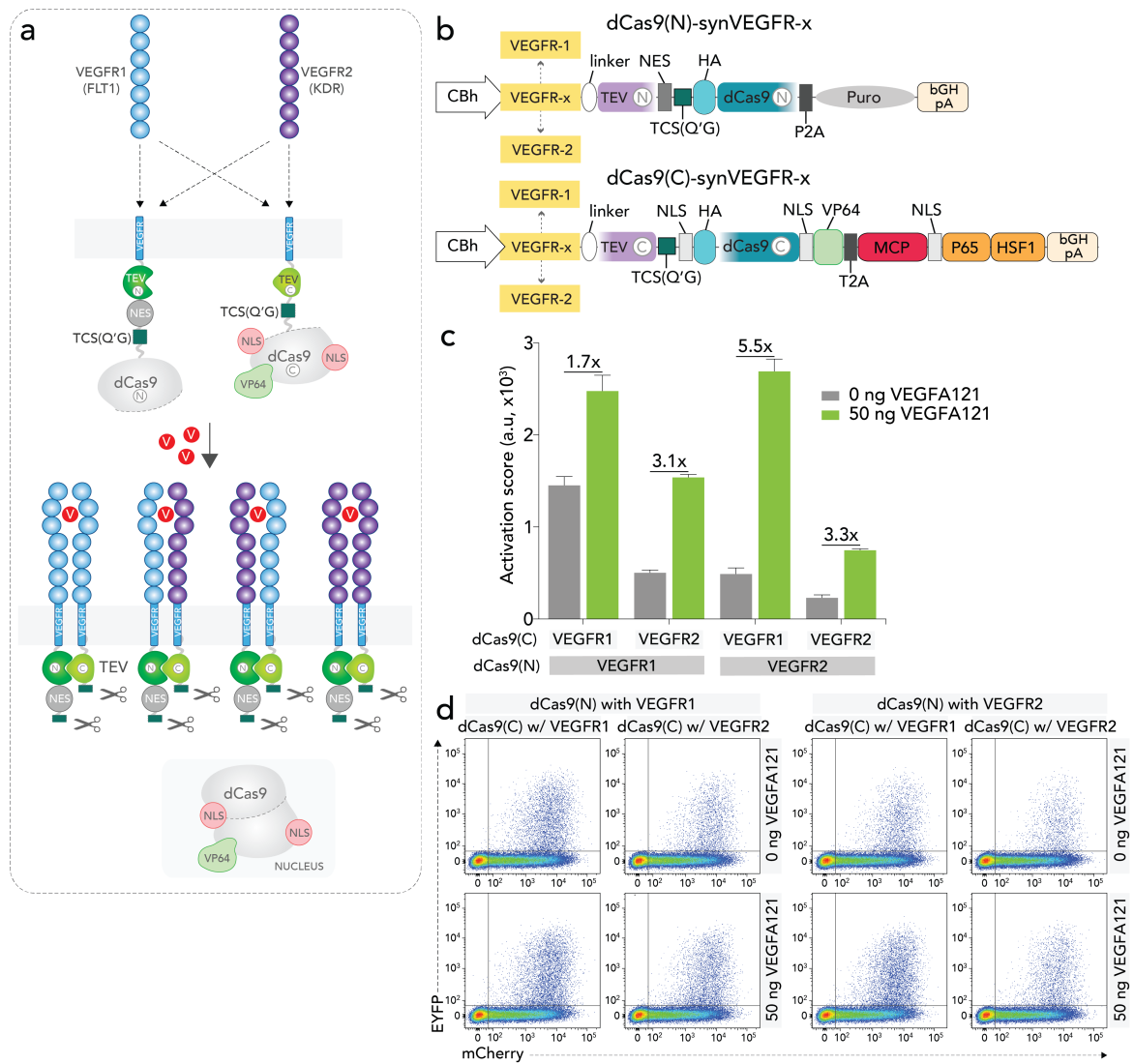


Figure 5.1 | Identification of optimal VEGFR combinations for dCas9-synVEGFR. (a) Conceptual overview: VEGF receptor 1 (VEGFR1) and VEGF receptor 2 (VEGFR2) are fused to either TEV(C)-TCS(Q'G)-dCas9(C)^{VP64} and TEV(N)-NES-TCS(Q'G)-dCas9(N). Ligand recognition via the extracellular domains triggers receptor dimerisation, which brings the split TEV fragments in proximity. This juxtaposition reconstitutes TEV protease and enables subsequent cleavage of TCS, releasing split dCas9 fragments. VEGFR1 and VEGFR2 describe the endogenous leader peptide, extracellular domain and transmembrane domain of VEGFR1 and VEGFR2, respectively. **(b)** Construct designs of dCas9(N)-synVEGFR-x and dCas9(C)-synVEGFR-x. **(c)** Representative flow cytometry scatter plots of transiently transfected HEK293T cells with the EYFP reporter system with or without co-delivered VEGFA121-expressing plasmid (50ng). **(d)** Quantification of EYFP reporter activation in (c). N = 3 biological replicates from one experiment, mean +/- s.d., a.u. arbitrary units.

protease reconstitution. Catalytically active TEV then cleaves the efficient TCS thereby liberating dCas9-TF in a ligand-independent manner, as previously proposed for another TEV-based synthetic receptor configuration (Daringer et al. 2014; Schwarz et al. 2017).

To address this problem, I performed an exhaustive screen to identify optimal TCS combinations for which dCas9(N)-synVEGFR2 and dCas9(C)-synVEGFR1 display enhanced OFF/ON state transition parameters (Figure 5.2a). Strikingly, only dCas9(N)-synVEGFR2 TCS variants in combination with dCas9(C)-synVEGFR1^{TCS(Q'G)} generated detectable outputs in the presence of 50 ng VEGFA121 (Figure 5.2b,c). From these, dCas9(N)-synVEGFR2^{TCS(Q'L)} displayed favourably low background leakiness while improving OFF/ON state transition characteristics. Henceforth, the combination of dCas9(N)-synVEGFR2^{TCS(Q'L)} and dCas9(C)-synVEGFR1^{TCS(Q'G)}, from here onwards referred to as **dCas9(N/C)-synVEGFR1/2**, was chosen for all subsequent experiments.

To demonstrate efficient cell surface expression of individually or combinatorically transfected dCas9(N)-synVEGFR2^{TCS(Q'L)} and dCas9(C)-synVEGFR1^{TCS(Q'G)}, I performed an antibody-based, direct flow cytometry experiment (Figure 5.3). This analysis illustrates that single chimeric receptors as well as dCas9(N/C)-synVEGFR1/2 were presented on the cell surface, in a range of ~6 to 11% of transiently transfected HEK293T cells (Figure 5.3).

Next, I characterised the performance of dCas9(N/C)-synVEGFR1/2 receptor in more detail. As evidenced in Figure 5.2b, dCas9(N/C)-synVEGFR1/2 reproducibly showed robust OFF/ON state transition while maintaining minimal background EYFP activation (Figure 5.4a,b,c). Importantly, delivery of VEGFA121 did not lead to any reporter gene expression in the control sgSCR conditions (Figure 5.4b,c).

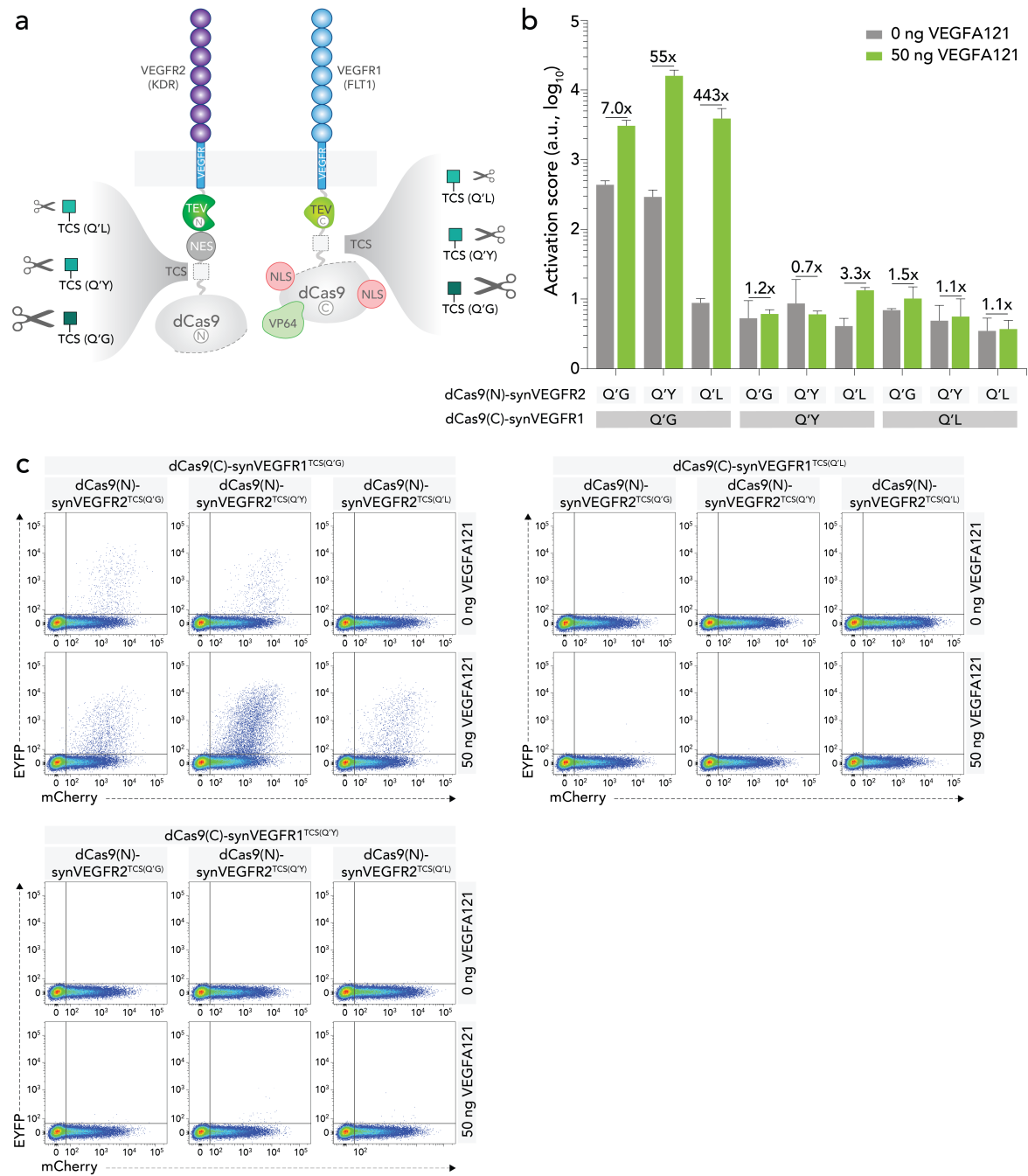


Figure 5.2 | Identification of an optimal TCS combination in dCas9(N/C)-VEGFR1/2. (a) Cartoon of dCas9(C)-synVEGFR1 and dCas9(N)-synVEGFR2 with three different TCS (ordered by decreasing TEV protease binding affinity): TCS(Q'G), ENLYFQ'G; TCS(Q'Y), ENLYFQ'Y; TCS(Q'L), ENLYFQ'L. **(b)** Quantification of EYFP reporter activation assay of HEK293T cells transiently transfected with the EYFP reporter system and an exhaustive combination of TCS permutations in dCas9(N/C)-synVEGFR1/2 in the presence or absence of co-transfected 50ng VEGFA121 plasmid. N = 3 biological replicates from one experiment, mean $\pm$ s.d., a.u. arbitrary units. **(c)** Representative flow cytometry scatter plots for the data described in (b).

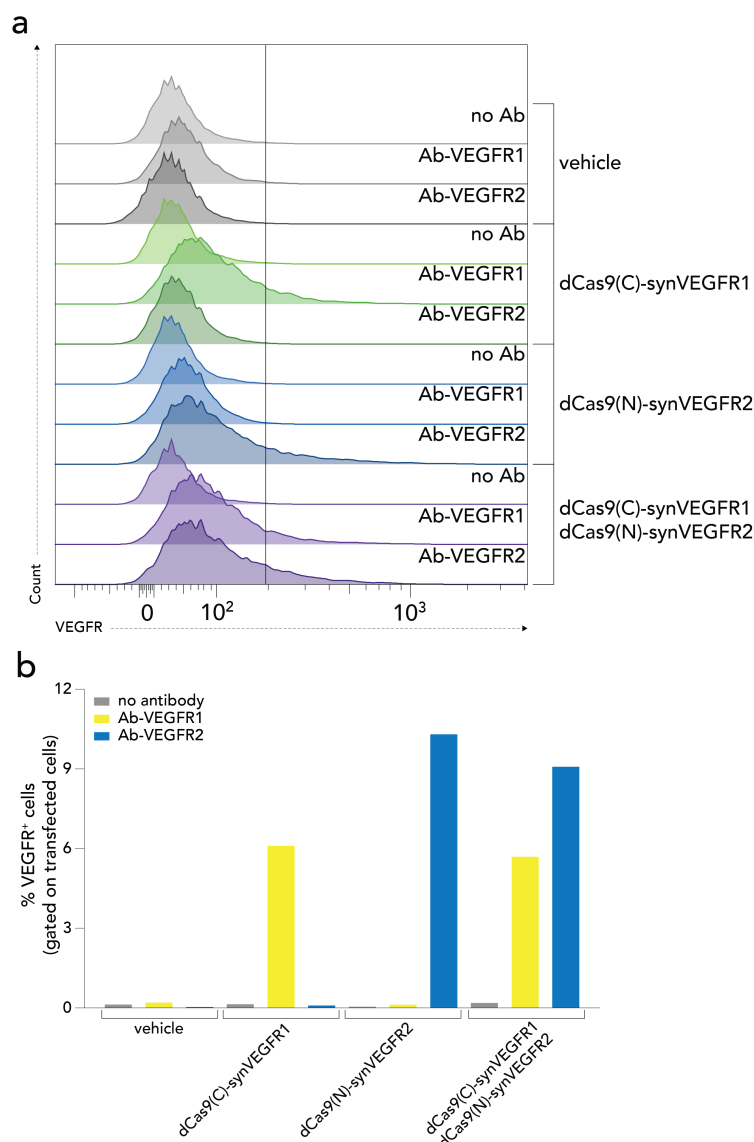


Figure 5.3 | dCas9(N/C)-VEGFR1/2 is expressed on the cell-surface. (a) Representative flow cytometry histograms of transiently transfected HEK293T cells with either dCas9(N)-synVEGFR2^{TCS(Q'L)}, dCas9(C)-synVEGFR1^{TCS(Q'G)}, or a combination thereof. Depicted are successfully transfected, VEGFR⁺ cells (receiver of chimeric receptor constructs and EGFP transfection control plasmid). Empty vector (vehicle) transfected cells served as negative control. The Y-axis of the histograms was set to modal in FlowJo software for improved visualisation/readability purposes. Ab, antibody. (b) Quantification of the antibody VEGFR⁺ stained cells as described in (a). N = 1 from one experiment.

To study the relationship between dCas9(N/C)-synVEGFR1/2 receptor abundance (plasmid copy number, input) and its performance (output), I first partitioned the reporter data into sequential mCherry fluorescence bins (transfection control) at increasing intensity (Figure 5.4d). Then, I calculated the EYFP mean fluorescence values and the percentage of activated cells within each mCherry bin (Figure 5.4e,f). Since mCherry is co-transfected with dCas9(N/C)-synVEGFR1/2, it provides a good approximation for the amount of dCas9(N/C)-synVEGFR1/2 in single cells. This analysis revealed a positive correlation between the abundance of system components and the level of EYFP reporter activation. Across a broad range of expression levels, the activity of dCas9(N/C)-synVEGFR1/2 increased until reaching a maximum threshold, after which the system performance did not improve anymore (Figure 5.4e,f).

Next, I characterised the agonist dose-response behaviour of dCas9(N/C)-synVEGFR1/2. At increasing amounts of co-transfected VEGFA121, the receptor displayed robust ligand-mediated reporter gene activation with a broad linear range between approximately 4 to 50 ng (Figure 5.4g).

Furthermore, dCas9(N/C)-synVEGFR1/2 in association with the SAM system robustly induced *ASCL1* gene expression in a VEGFA121-dependent manner. Relative to the uninduced condition, *ASCL1* was upregulated by 2.0-, 6.7-, and 48-fold in the presence of 1, 5, and 50 ng VEGFA121, respectively (Figure 5.4h). No *ASCL1* upregulation was observed in the presence of a control scramble SAM sgRNA (Figure 5.4h).

5.2.2 Boolean AND logic gates as inducible control switches

The integration of a second, exogenously-delivered trigger can provide a valuable safety switch for controlling the activity of dCas9(N/C)-synVEGFR1/2.

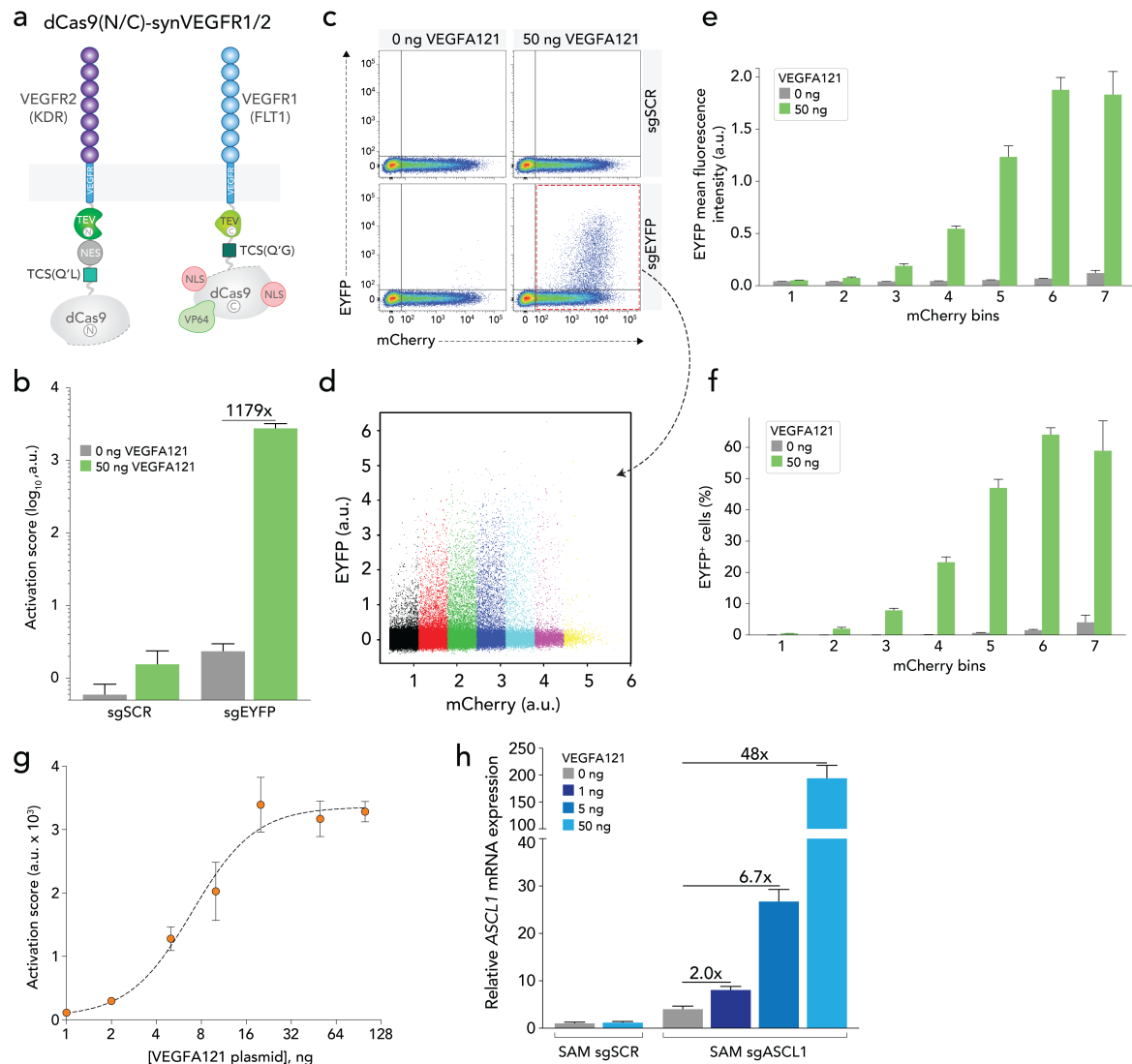


Figure 5.4 | Functional characterisation of dCas9(N/C)-VEGFR1/2. (a) Cartoon of dCas9(N)-synVEGFR2 with TCS(Q'L) and dCas9(C)-synVEGFR1 with TCS(Q'G). (b) Quantification of EYFP reporter assay in the presence or absence of co-transfected 50 ng VEGFA121 plasmid. (c) Representative flow cytometry scatter plots from (b). (d) Data binning strategy based on mCherry fluorescence for dCas9(N/C)-synVEGFR1/2. (e) Mean EYFP fluorescence intensity for each mCherry bin from (a), in the presence or absence of agonist. (f) Percentage of EYFP positive cells for each mCherry bin in the presence or absence of VEGFA121. (g) Dose-response curve for dCas9(N/C)-synVEGFR1/2. Each data point represents EYFP activation score from 3 biological replicates; curve was fitted using a non-linear variable slope (four parameters) function in GraphPad Prism. (h) Endogenous *ASCL1* transcriptional upregulation by dCas9(N/C)-synVEGFR1/2 in a dose-responsive manner. Data is represented as fold change relative to sgSCR no-agonist condition (N = 3 (x3 technical replicates)). (b-f) N=3 biological replicates; mean +/- s.d.; a.u. arbitrary units). David Knapp performed analysis and generated plots in (d,e,f).

For the first implementation, I designed an AND gate for each split dCas9 fragment. The first gate resides within the dCas9(N)-synVEGFR2^{TCS(Q'L)} and VEGFR1-TEV(C). VEGFR1-TEV(C) is an NLS-dCas9(C)^{VP64}-depleted variant of dCas9(C)-synVEGFR1^{TCS(Q'G)} (Figure 5.5a). VEGFA121 binding triggers dimerisation of VEGFRs, reconstituting a functional TEV protease. Catalytically active TEV protease then cleaves the proximal TCS thereby liberating dCas9(N) fragment. The second gate consists of TMt-FKBP-TEV(N)-NLS-dCas9(C)^{VP64} which had been previously described for split dual TMt-dCas9-iTEV^{engrafted} (section 3.2.3.2) and FRB-TEV(C). Rapamycin administration facilitates dimerisation of the FKBP-TEV(N)/FRB-TEV(C) pair, enabling the assembly of catalytically active TEV protease and as a consequence releasing the dCas9(C)^{VP64} fragment, thereby reconstituting a functional dCas9-TF. This AND logic gate is based on a rapamycin- and VEGFA121-mediated release of split dCas9 fragments, and is referred to as dCas9(N/C)-synVEGFR1/2^{RC} (Figure 5.5a).

In the presence of no agonist or 50 ng VEGFA121 alone, dCas9(N/C)-synVEGFR1/2^{RC} is completely inactive (Figure 5.5b,c). However, higher concentrations of rapamycin alone (20 nM and above) triggered unspecific reporter gene expression. Full system performance was achieved by the co-administration of 50 ng VEGFA121 and resulted in at least 2-fold higher reporter activation, compared to the rapamycin-specific background alone (Figure 5.5b,c).

To improve this AND gate configuration, I next implemented a different approach. By analogy with the previously validated split dual TMt-dCas9-iTEV^{engrafted} AND gate design (section 3.2.3.2), I conceptualised a variation in which the TMt domain was replaced by the VEGFRs, generating a rapamycin-inducible dCas9(N/C)-synVEGFR1/2^{RI} configuration (Figure 5.5d). VEGFA121 binding facilitates TEV

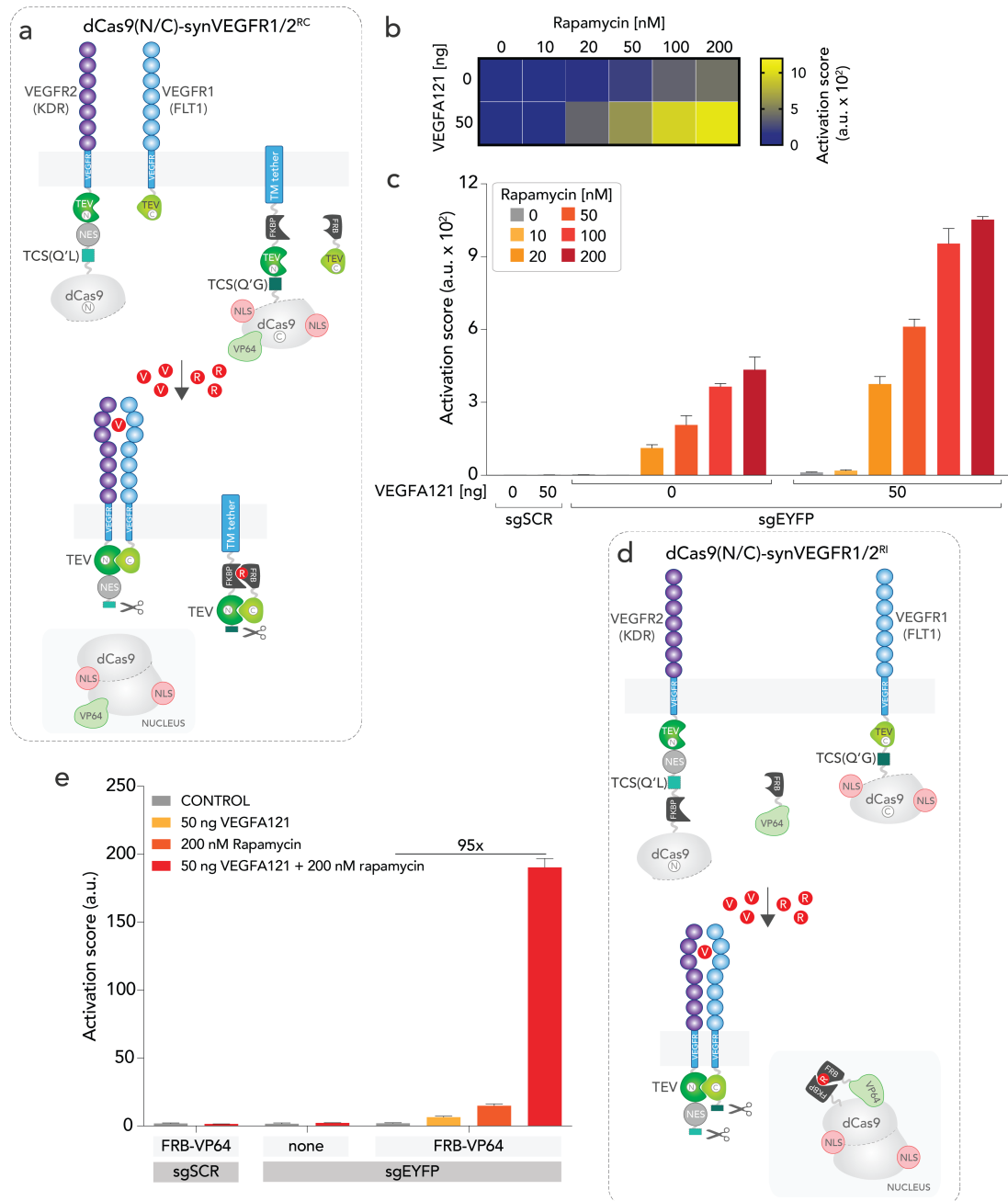


Figure 5.5 | 2-input Boolean AND logic gate strategies for dCas9(N/C)-synVEGFR1/2 with an inert drug. (a) Cartoon of dCas9(N/C)-synVEGFR1/2^{RC} AND gate configuration. VEGFA121 binding triggers dCas9(N) release and rapamycin facilitates the reconstitution of TEV protease, liberating dCas9(C) and enabling the reassembly of dCas9-TF. **(b)** EYFP reporter assay heatmap for dCas9(N)-synVEGFR1/2^{RC}. N = 3 from one experiment, no mean displayed; a.u., arbitrary units. **(c)** EYFP reporter assay quantification including mean $\pm$ s.d. from (b). **(d)** Cartoon of dCas9(N/C)-synVEGFR1/2^{RI} AND gate configuration. VEGFA121 binding triggers dCas9(N) and dCas9(N) release by reconstituted TEV protease. Rapamycin recruits VP64 to dCas9(N), reconstituting dCas9^{VP64}. **(e)** EYFP reporter assay quantification of 2-input dCas9(N/C)-synVEGFR1/2^{RI} AND gate. N = 3 biological replicates from one experiment, mean $\pm$ s.d., a.u. arbitrary units.

protease assembly and subsequent liberation of dCas9(N) and dCas9(C) fragments. The second input, rapamycin, recruits the effector domain VP64 to reassembled dCas9, reconstituting a dCas9-TF (Figure 5.5d). Using this configuration, a robust 95-fold EYFP reporter gene induction was observed only in the presence of all system components and both input stimuli (Figure 5.5e). This result highlights the power of a stringent AND logic gate configuration in containing dCas9-TF activity and rendering it susceptible to 2-input stimuli.

5.2.3 dCas9(N/C)-synVEGFR1/2 is superior to V-MESA-dCas9-TF

Recently, Joshua Leonard's laboratory reported a dCas9-TF-based receptor design that allows the integration of VEGFA and conversely the activation of endogenous interleukin 2 (*IL-2*) (Schwarz et al. 2017). In brief, this receptor system, termed modular extracellular sensor architecture (MESA, Chapter 1.2.3.5), is composed of two polypeptide chains, the target chain (TC) and the protease chain (PC) (Figure 5.6a) (Daringer et al. 2014). PCs (V1-MESA-45F-TEV or V2-MESA-35F-TEV) consist of anti-VEGF scFvs (V1 or V2) recognition domains, an extracellular linker (45 or 35 amino acids long), a TMD, a linker sequence and full length TEV protease (Schwarz et al. 2017). The TCs (V1-MESA-45F-M-dCas9 or V2-MESA-35F-M-dCas9) only differ in their intracellular part from PCs, in which full length NLS(+)-dCas9^{VP64} is fused to the TMD via an efficient TCS(Q'M). The VEGF-engaging scFvs on both chains recognise and bind to VEGFA, which brings both chains into proximity, allowing TEV to cleave in *trans* its cognate TCS on TC, thereby releasing dCas9-TF from the membrane and facilitating its subsequent translocation into the nucleus (Figure 5.6a) (Schwarz et al. 2017).

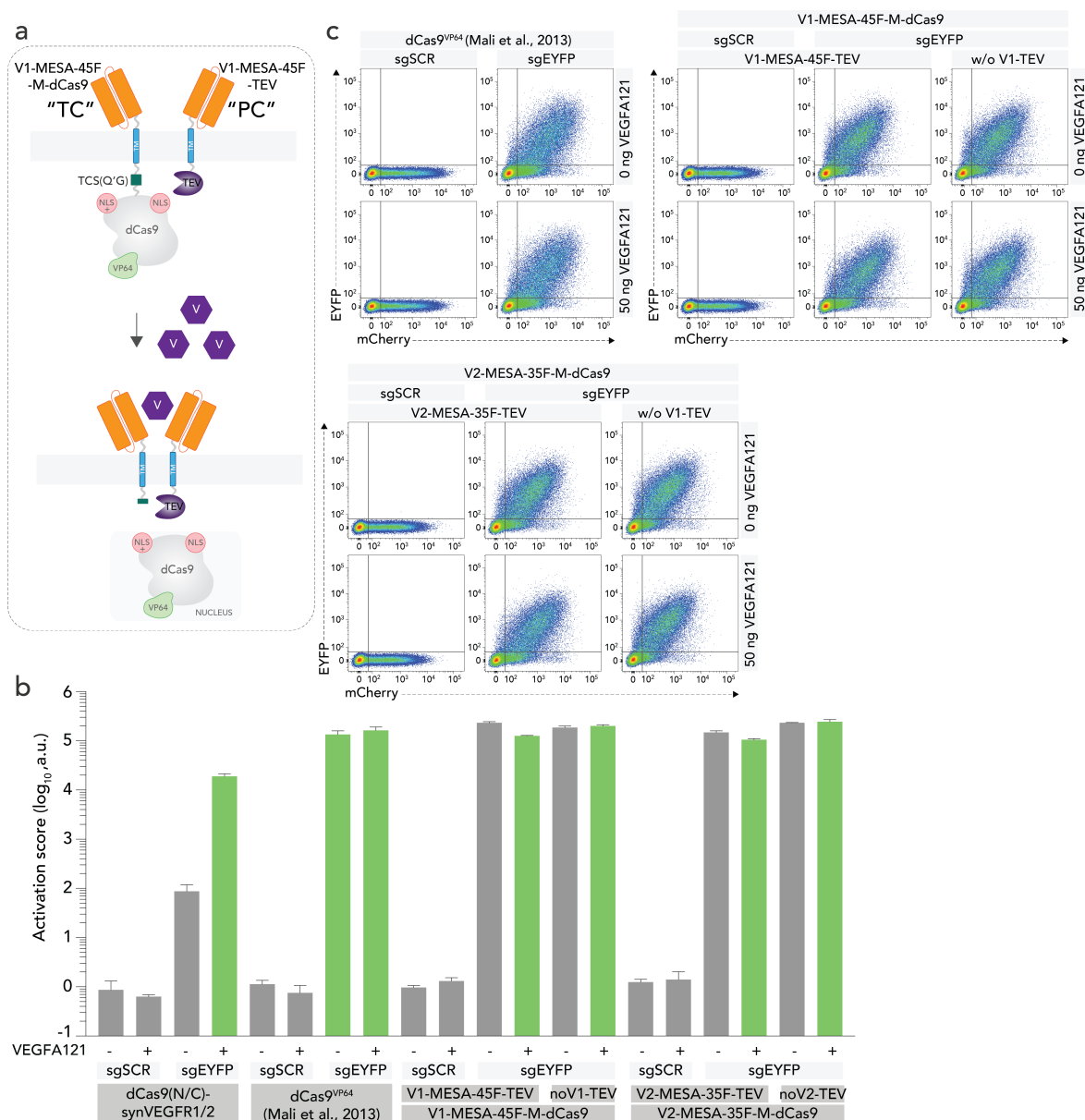


Figure 5.6 | Comparing dCas9(N/C)-synVEGFR1/2 with MESA-dCas9-TF receptors. (a) Cartoon and concept of the functional mechanism of dCas9-based MESA receptors for VEGFA. The VEGF-specific scFvs of MESA receptors bind to VEGFA, thereby bringing the TEV protease chain (PC, V1-MESA-45F-TEV) into proximity to its cognate TCS on the target chain, which contains dCas9^{VP64} (TC, V1-MESA-45F-M-dCas9). TEV protease cleaves the TCS thereby liberating dCas9-TF which subsequently translocates into the nucleus to exerts its transcriptional function. V2-MESA-35F-M-dCas9 and V2-MESA-35F-TEV are derivatives of V1-MESA (Schwarz et al. 2017). **(b)** Representative flow cytometry scatter plots of transiently transfected HEK293T cells with the EYFP reporter system with or without co-transfected VEGFA121 plasmid. **(d)** Quantification of EYFP reporter activation in (c). N = 3 biological replicates from one experiment, mean +/- s.d., a.u. arbitrary units.

In a head-to-head comparison, I benchmarked the established V-MESA-dCas9-TF with the engineered dCas9(N/C)-synVEGFR1/2 chimeric receptor. Of note, the reported optimal ratio of TC:PC (1:10) was maintained in the experimental setting, where applicable (Figure 5.6b,c) (Schwarz et al. 2017). Additionally, I included a previously published, unaltered full length dCas9-TF (Mali, Aach, et al. 2013) as control for maximum EYFP reporter activation (Figure 5.6b,c).

Analysis of reporter expression confirmed that, as shown in this chapter, dCas9(N/C)-synVEGFR1/2 displayed low OFF state activity in the absence of agonist and strong activation in the presence of VEGFA121 (Figure 5.6b,c). In my hands, both the V1-MESA-dCas9-TF and V2-MESA-dCas9-TF receptors revealed extremely high background leakiness in the uninduced condition, irrespective of the presence or absence of the PC components (Figure 5.6b,c). Strikingly, the level of baseline activity detected by V-MESA-dCas9-TF is equivalent to the maximum activation potential observed for soluble, full length dCas9-TF. Furthermore, this extremely high OFF state leakiness hampered the detection of agonist-mediated EYFP transcriptional upregulation (Figure 5.6b,c). Indeed, Schwarz et al. (2016) only obtained very modest induction rates in the presence of recombinant VEGFA inducer. Taken together, this data indicates that transmembrane-tethered full length dCas9-TF, in its current configuration, is incompatible with engineering a functional dCas9-based signal transduction module.

5.3 Discussion

In this chapter, I have successfully implemented a programmable dCas9-TF intracellular signal transduction module for a prototype chimeric RTK. The optimised core split dCas9 architecture (Chapter 3) was grafted onto the endogenous ectodomain of the VEGFR receptor family (dCas9(N/C)-synVEGFR1/2). The rationale for utilising

the ectodomains of native receptors was to alleviate the necessity of testing and optimising ligand binding specificities and affinities. Therefore, with regard to the sensing domains, dCas9(N/C)-synVEGFR1/2 should behave similarly to their endogenous counterparts. Concomitantly, this design takes advantage of the naturally occurring ligand-mediated dimerisation potential of VEGFRs. Heterodimerisation facilitates reassembly of the integrated TEV protease and subsequent liberation of downstream anchored split dCas9 fragments.

I propose that dCas9-synRTKs will be a valuable tool for converting endogenous, growth factor-related ligands into novel cellular response programmes.

5.3.1 Agonist repertoire for dCas9(N/C)-synVEGFR1/2

In my experiments, the VEGFA121 ligand was co-expressed from a plasmid for several reasons: (i) VEGFA121 is a protein ligand and its expression in human cells and maturation through the secretory pathways may generate a more physiologically relevant molecule compared to recombinant protein that is expressed in bacteria and subsequently purified *in vitro*; (ii) A highly specific and robust VEGFA121-dependent receptor activation was observed, demonstrating highly reproducible results between biological replicates as well as between experimental conditions; (iii) Co-delivery of VEGFA121 did not expose receptors to variable or excessive VEGFA concentrations. I investigated receptor behaviour as a function of VEGFA121 plasmid concentration and revealed a dose-dependent linear response window reflecting agonist-mediated receptor activation at non-saturating concentrations; (iv) Studies regarding cardiovascular gene therapy in various model organisms and in human clinical trials made use of CMV-driven VEGFA121 expression plasmids (Rosengart et al. 1999; Perrin et al. 2004). These treatments were well tolerated in all cases without adverse effects. Although assessing the actual VEGF concentration in these conditions would

be difficult, these experiments suggest that such promoters drive expression of VEGFA121 within a physiologically relevant range. One potential caveat of co-delivering VEGFA121 with dCas9(N/C)-synVEGFR1/2 in transient transfections is that ligand/receptor interactions can occur within the secretory pathway, triggering receptor activation before reaching the cell surface.

One area that will require further investigation is to establish the sensitivity of dCas9(N/C)-synVEGFR1/2 towards different VEGFA isoforms, such as VEGFA165 or VEGFA189 (T. T. Chen et al. 2010; Ferrara 2010), other canonical ligands, e.g., VEGFB or VEGFC (McColl et al. 2003; Dixelius et al. 2003; Anisimov et al. 2013), or non-canonical activators, such as β -galactoside-binding proteins, known as galectins (Markowska et al. 2011; Dalziel et al. 2014). Moreover, for potent endogenous VEGFR signalling and in particular for VEGFR2, a number of modulators have been reported, including VEGF-engaging co-receptors (e.g., proteoglycans and neuropilins) or non-VEGF-engaging proteins (e.g., vascular endothelial cadherin and integrins) (Simons et al. 2016). It remains to be seen if these components in complex with dCas9(N/C)-synVEGFR1/2 further improve and fine-tune its performance.

With regard to ligand selectivity, homo- and/or hetero-ectodomain combinations of VEGFR2 and VEGFR3 might be better suited to detected VEGFC or VEGFD ligands (Dixelius et al. 2003; Nilsson et al. 2010; Y. Deng et al. 2015).

5.3.2 Alternative dCas9-TF-based synthetic receptor scaffold for VEGFA detection

Similar to dCas9(N/C)-synVEGFR1/2, the synthetic receptor V-MESA-dCas9-TF facilitates VEGFA-mediated activation of dCas9-TF (Schwarz et al. 2017). Conceptually, both systems have common features that distinguish them from native

receptors (RTKs) and their associated signalling cascades. First, both chimeric receptors represent a self-contained sensing and signalling system that does not necessitate other components to function once exposed to their cognate ligands. Second, they are orthogonal to endogenous signalling pathways because they do not require native second messengers or mediators for signal transduction.

However, in contrast to dCas9(N/C)-synVEGFR1/2, analysis of V-MESA-dCas9-TF performance confirmed that tethering full length dCas9-TF to the membrane results in non-stringent OFF/ON state transitions (i.e., full activation in the uninduced state). This observation has been confirmed by other studies (Kipniss et al. 2017, Liu et al. 2017) and underscores the suboptimal behaviour of the MESA receptor scaffold for rewiring cellular input/output functions.

Chapter 6 : Implementation of therapeutically-relevant transcriptional programmes with dCas9-synRs

Most of the data presented in this chapter has been published in: Baeumler, T.A., Ahmed, A.A. & Fulga, T.A., 2017. Engineering Synthetic Signaling Pathways with Programmable dCas9-Based Chimeric Receptors. *Cell Reports*, 20(11), pp.2639–2653.

6.1 Chapter Aims

- Implementing therapeutically-relevant, non-native response programmes with chimeric dCas9-synRs in response to cognate ligand stimulation.
- Converting tumour promoting factors into therapeutic responses with opposing, tumour disfavoured properties.
- Proof of concept for engineering a glucose–insulin response signalling pathway in non-pancreatic, designer cells.

6.2 Results

To showcase the potential of redirecting the cellular information flow with programmable dCas9-synRs, I next sought to equip cells with novel, therapeutic properties. The ensuing engineered cells would elicit customised, non-native transcriptional programmes in response to a variety of disease biomarkers, thus counteracting the physiological effects of input stimuli.

6.2.1 Angiogenic signal conversion mediated by dCas9(N/C)-synVEGFR1/2

The optimised VEGFA121-responsive dCas9(N/C)-synVEGFR1/2 receptor from Chapter 5 was reprogrammed to simultaneously drive the expression of two anti-angiogenic endogenous genes: thrombospondin (*TSP-1*) and tumour necrosis factor alpha (*TNF α*) (Figure 6.1a). *TSP-1* exerts its anti-angiogenic potential directly on target cells (e.g., inhibition of migration, growth) and by antagonizing VEGF-mediated downstream signalling [Lawler:2012fa]. *TSP-1* has been shown to have tumour-inhibiting effects in a mouse model in which *TSP-1* secreting human neural stem cells impaired the progression of malignant human gliomas (van Eekelen et al. 2010). Although controversial with regard to its anti-tumour effects, locally restricted and moderate amounts of *TNF α* led to the destruction of tumour vasculature (Balkwill 2009; Burton & Libutti 2009) and promoted the production of the anti-angiogenic protein fragment angiostatin (Mauceri et al. 2002).

SAM sgRNAs for *TSP-1* and *TNF α* (Konermann et al. 2015) were co-transfected with receptor components and VEGFA121 in HEK293T cells. RT-qPCR analysis revealed robust induction of both genes simultaneously in a VEGFA121-specific manner (Figure 6.1b). Relative to the respective no-agonist control conditions, *TSP-1* was upregulated 16-fold and *TNF α* 18-fold with virtually no background activation. Importantly, VEGFA121 alone did not lead to the activation of either target gene, demonstrating a tightly controlled OFF state (Figure 6.1b).

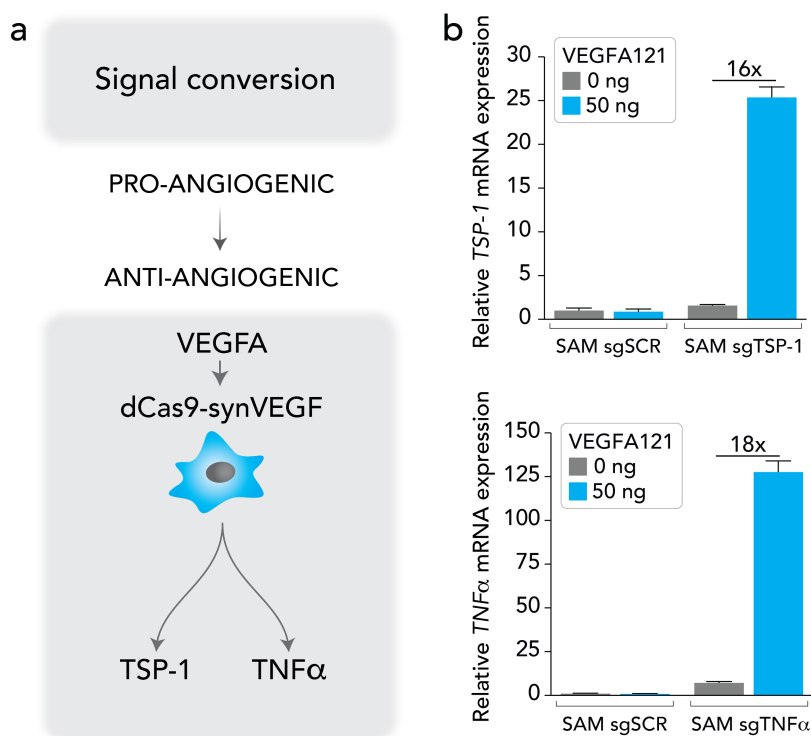


Figure 6.1 | Angiogenic signal conversion mediated by dCas9(N/C)-synVEGFR1/2. (a) Redirecting the pro-angiogenic signal of VEGFA into a user-defined, anti-angiogenic response programme by simultaneously inducing the transcriptional upregulation of the endogenous genes *TSP-1* and *TNFα*. **(b)** RT-qPCR of *TSP-1* and *TNFα* in transiently transfected HEK293T cells with dCas9(N/C)-synVEGFR1/2 and respective SAM sgRNAs. Data is represented as fold change relative to respective sgSCR no-agonist condition. N = 3 biological replicates (x3 technical replicates), mean +/- s.d.

6.2.2 Eliciting custom cytokine/chemokine programmes with dCas9-synGPCRs

Stromal cell-derived factor 1 (SDF-1) and lysophosphatidic acid (LPA) are two molecules that are highly enriched in the tumour microenvironment (Mills & Moolenaar 2003), making them ideal extracellular input signals for engineered, synthetic receptors.

The chemokine SDF-1 is the sole agonist for the CXCR4 receptor. CXCR4 is predominantly expressed on cancer cells (Zlotnik et al. 2011; Balkwill 2012) and constitutes a key receptor for metastatic dissemination from primary tumour sites to tissues with higher concentrations of SDF-1, such as lymph nodes, lungs, bone marrow, and liver (Dorsam & Gutkind 2007; Lazennec & Richmond 2010; Zlotnik et al. 2011). In addition, solid tumours display elevated levels of SDF-1 under hypoxic conditions (Dorsam & Gutkind 2007; F. Jin et al. 2012; Roccaro et al. 2014).

LPA is a single fatty acyl chain phospholipid which is secreted in high quantities by ovarian and prostate cancer cells. LPA-mediated receptor signalling promotes cancer initiation, proliferation, and metastasis formation (Y. Xie et al. 2002; Yamada et al. 2004; Mills & Moolenaar 2003; Boucharaba et al. 2004; Z. Lee et al. 2006).

For both biomarkers, I designed a non-native cytokine/chemokine programme, consisting of *IL-2*, *MIP-1 α* , and *IFN γ* that should initiate and support a strong and robust cellular immune response. Interleukin-2 (IL-2) is a versatile cytokine involved in enhanced T cell survival and expansion (W. Liao et al. 2013). Macrophage inflammatory protein 1-alpha (MIP-1 α , also known as chemokine (C-C motif) ligand 3 (CCL3)) helps with the recruitment of T cell subsets (Castellino et al. 2006). Interferon gamma (IFN γ) supports the activation of innate immune cells and potentiates the susceptibility of cancer cells to cytotoxic T cells (Zaidi & Merlino 2011; Ivashkiv & Donlin 2014; Parker et al. 2016).

In the first experiment, transiently transfected HTLA cells received dCas9(N/C)-synCXCR4 and a cocktail of previously published SAM sgRNAs for targeted expression of *IL-2*, *MIP-1 α* , and *IFN γ* (Konermann et al. 2015) (Figure 6.2a). At 0.1 μ M SDF-1 treatment, a concomitant gene upregulation of *IL-2*, *MIP-1 α* , and *IFN γ* of 61-, 6.8-, and 10-fold was detected at the mRNA level, respectively (Figure 6.2b). While maintaining low background activation in the absence of SDF-1 agonist, the scramble control conditions did not show gene upregulation. This indicates that only guided, reconstituted dCas9-TF facilitated transcriptional upregulations (Figure 6.2b).

Similarly, upon 10 μ M LPA stimulation, dCas9(N/C)-synLPAR1-engineered HTLA cells displayed robust and simultaneous target gene inductions of 89-, 16-, and 15-fold for *IL-2*, *MIP-1 α* , and *IFN γ* , respectively (Figure 6.3a,b). Again, gene upregulation was highly specific as reflected by the lack of gene induction in the SAM sgSCR control conditions (Figure 6.3b).

Taken together, these results demonstrate that a custom cytokine/chemokine programme can be induced upon the signal integration of pre-defined biomarkers. Notably, magnitudes of individual gene induction were not dependent on particular input stimuli or dCas9-synGPCR variant, demonstrating reliable and portable transcriptional gene activation across different dCas9-synGPCR receptors.

6.2.3 Glucose-induced insulin expression in designer cells with dCas9(N/C)-synT1R3

I next sought to employ the glucose-sensing dCas9(N/C)-synT1R3 receptor in HTLA cells for creating therapeutic, designer mimetic β -cells (Figure 6.4a). If successful, non-pancreatic human cells would respond to D-glucose by expressing endogenous insulin, presenting a new way for designer cells to tackle diabetes.

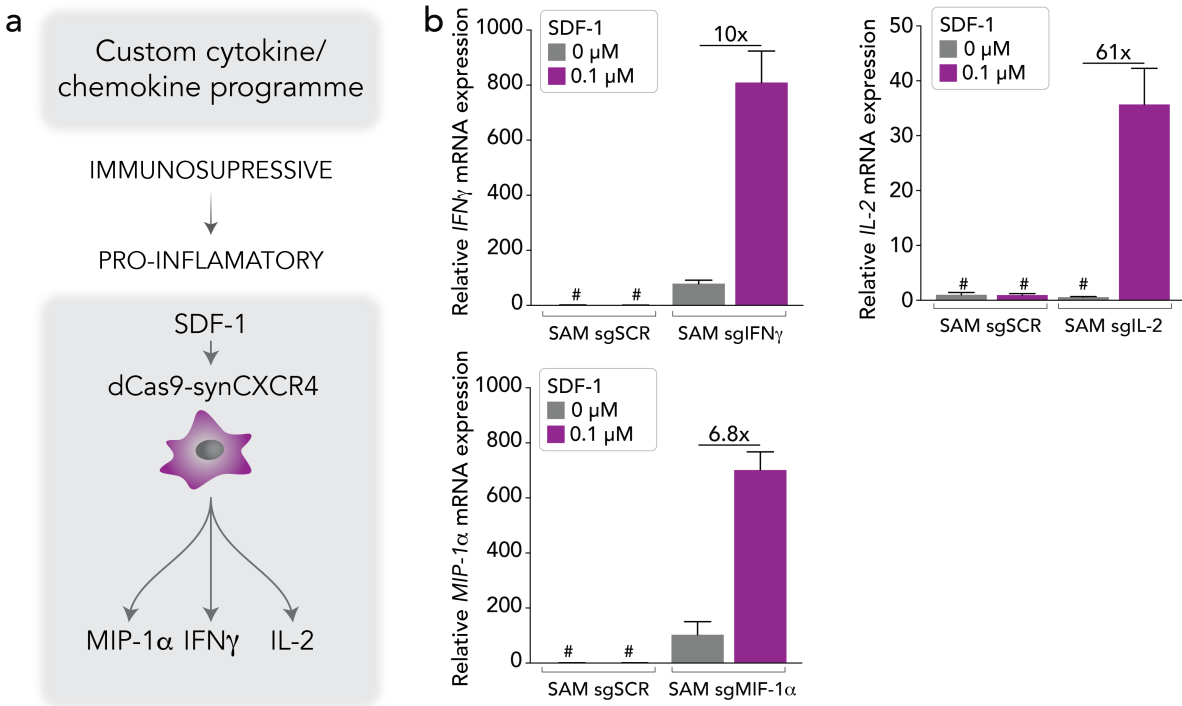


Figure 6.2 | Cytokine signal conversion mediated by dCas9(N/C)-synCXCR4. (a)

Redirecting the immunosuppressive signal of SDF-1 (CXCL12) into a user-defined, pro-inflammatory response programme by simultaneously inducing the transcriptional upregulation of the endogenous genes *MIP-1 α* , *IL-2*, and *TNF α* . **(b)** RT-qPCR of *MIP-1 α* , *IL-2*, and *TNF α* in transiently transfected HTLA cells with dCas9(N/C)-synCXCR4 and respective SAM sgRNAs. Data is represented as fold change relative to respective sgSCR no-agonist condition. N = 3 biological replicates (x3 technical replicates), mean $\pm$ s.d. #, undetermined values for the gene of interest were set to a maximum Ct = 40 cycles.

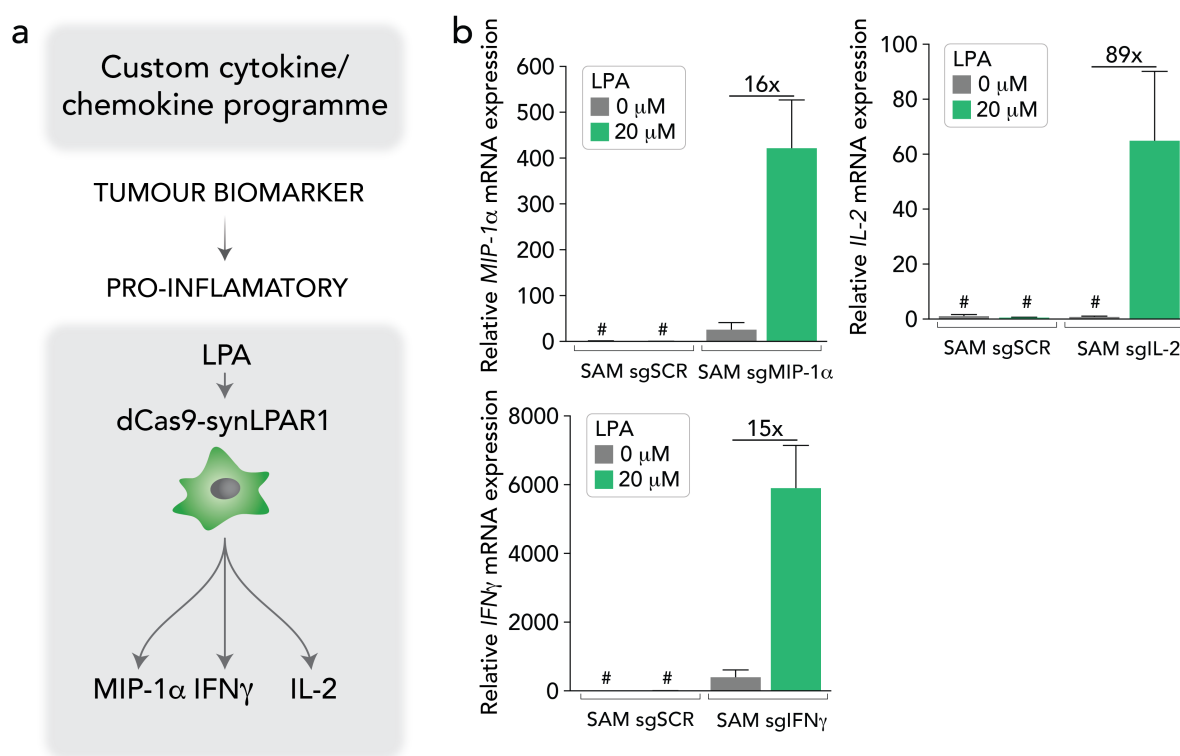


Figure 6.3 | Tumour biomarker evoked multi-factorial cytokine/chemokine response programme mediated by dCas9(N/C)-synLPA1. (a) Redirecting the tumour biomarker LPA into a user-defined, pro-inflammatory response programme by simultaneously inducing the transcriptional upregulation of the endogenous genes *MIP-1 α* , *IL-2*, and *TNF α* . (b) RT-qPCR of *MIP-1 α* , *IL-2*, and *TNF α* in transiently transfected HTLA cells with dCas9(N/C)-synLPA1 and respective SAM sgRNAs. Data is represented as fold change relative to respective sgSCR no-agonist condition. N = 3 biological replicates (x3 technical replicates), mean $\pm$ s.d. #, undetermined values for the gene of interest were set to a maximum Ct = 40 cycles.

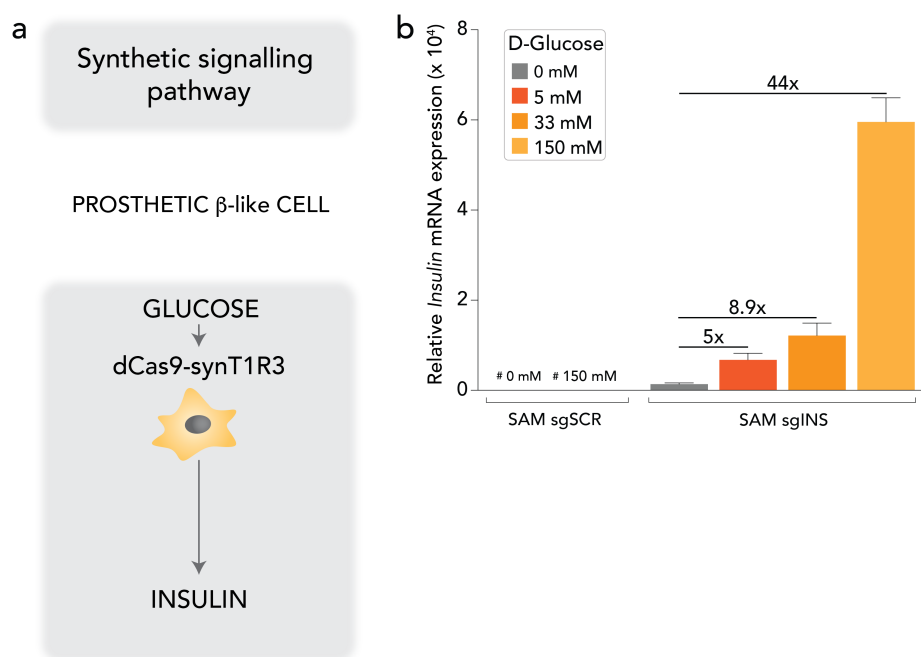


Figure 6.4 | Glucose-induced insulin expression in β -mimetic cells mediated by dCas9(N/C)-synT1R3. (a) Linking extracellular glucose levels to insulin expression by programming dCas9(N/C)-synT1R3 in engineered, non-pancreatic cells. (b) RT-qPCR of *insulin* in transiently transfected HTLA cells with dCas9(N/C)-synT1R3 and respective SAM sgRNAs. Data is represented as fold change relative to sgSCR no-agonist condition. N = 3 biological replicates (x3 technical replicates), mean $\pm$ s.d. #, undetermined values for the gene of interest were set to a maximum Ct = 40 cycles.

Diabetes mellitus is a consequence of chronically elevated blood glucose levels (hyperglycaemia) caused either by the autoimmune system-mediated destruction of pancreatic insulin-secreting β -cells (type 1 diabetes, T1D) or by the impairment of insulin action associated with the gradual and continuous exhaustion of β -cells (type 2 diabetes, T2D) (Polonsky 2012; Cornell 2015; Tuomilehto & Bahijri 2016). The former is an ideal candidate to be treated with tightly regulated insulin secreting designer cells to restore glycaemic control *in vivo*.

To this end, I transiently transfected HTLA cells with the optimised dCas9(N/C)-synT1R3 receptor and a pool of previously validated sgRNAs to target the insulin locus (Giménez et al. 2016). dCas9(N/C)-synT1R3-engineered HTLA cells were subjected to a range of D-glucose concentrations (Figure 6.4b). At the highest tested concentration of 150 mM D-glucose, a 44-fold insulin gene upregulation was observed, relative to 0 μ M D-glucose baseline. At the lower and higher end of the physiologically-relevant concentration range of blood D-glucose levels, 5 and 33 mM (Nie et al. 2005; M. Xie et al. 2016), the engineered cells displayed 5.0- and 8.9-fold elevated insulin mRNA levels, respectively (Figure 6.4b). As expected, in the absence of insulin targeting SAM sgRNAs, no insulin gene induction was observed.

To the best of my knowledge, this is the first instalment of its kind in which extracellular glucose levels have been directly linked to the gradual expression of insulin in non-pancreatic cells.

6.3 Discussion

In this chapter, I have deployed the newly designed chimeric dCas9-synRs to install therapeutically-relevant transcriptional programmes into engineered HEK293-derived cells. As a proof of concept, I demonstrated that various input signals can readily be detected and converted into customised cellular responses. For example, the multi-

layered cytokine/chemokine programme does not occur in the described combination in nature and represents a novel response relative to the input stimulus. This work illustrates an important milestone in rewiring the cellular information flow in cells by ligand-activated dCas9-synRs.

6.3.1 Converting various input signals into custom cytokine/chemokine output programmes

With regard to the proposed cytokine/chemokine programme, several constraints should be taken into account, such as the input sampling capabilities of dCas9-synRs and the output responses mediated by dCas9-synRs. To address the former, I have shown that dCas9-synRs can recognise a variety of input molecules, such as growth factor, lipids, and chemokines. However, some of these molecules are broadly expressed in the human body and have important physiological functions. SDF-1, for instance, is crucial for cerebellar development, haematopoiesis, and particularly for haematopoietic cell homing (migration towards SDF-1 gradients) (Zou et al. 1998; Egawa et al. 2001; Liles et al. 2003). To avoid adverse, potentially deleterious effects on haematopoietic cells in the bone marrow, it would be of importance that engineered cells are not restricted to recognising SDF-1 as sole input. Strategies addressing this issue may comprise of: (i) multiplexing of several dCas9-synRs; (ii) implementation of multi-input AND logic gates; (iii) combinatorial expression of multiple chimeric receptors, such as CARs or synNotch (W. A. Lim & June 2017). A similar approach could be considered for circumventing unspecific dCas9-synRs activation by other chemokines since one receptor can have several agonists and *vice versa* (Lazennec & Richmond 2010; Viola et al. 2012).

Future therapeutic designer cells could sample with a user-defined set of chimeric receptors a particular tumour microenvironment for several input factors.

These factors can include a variety of soluble molecules that are either actively secreted by cancer cells in an autocrine and paracrine fashion or by bystander host cells, such as cancer-associated fibroblasts (CAFs) and mesenchymal stem cells (MSCs) (Lazennec & Richmond 2010). Factors that support cancer cell proliferation, angiogenesis and metastasis formation are promising candidate input molecules for such engineered cells. Multi-factorial input signals may help spare non-cancerous tissues that display an incomplete set of input stimuli. Overall, an enhanced on-tumour specificity will improve the safety of utilising engineered cells.

With regard to the output programmes, it remains to be established which cocktails of cytokines/chemokines have the most potent effects *in vivo*. Activating the expression of genes counteracting the physiological input properties is an intriguing starting point (e.g., converting pro-inflammatory signals into anti-inflammatory responses). The therapeutic applicability and benefit of such cytokine conversion by engineered cells was demonstrated in an *in vivo* mouse model of inflammation (Schukur et al. 2015). Implanted HEK293T cells equipped with a closed-loop, prosthetic gene circuit detected elevated extracellular levels of pro-inflammatory tumour necrosis factor (TNF) and interleukin 22 (IL-22), and activated the secretion of transgenic anti-inflammatory cytokines IL-4 and IL-10 (Schukur et al. 2015).

6.3.2 Engineering β -mimetic cells with dCas9-synT1R3

In 2014, more than 400 million adults suffered from diabetes worldwide (NCD Risk Factor Collaboration (NCD-RisC) 2016). It is estimated that this number will increase to over 600 million by 2040 (International Diabetes Federation. IDF Diabetes Atlas, 8th edition, 2017). Untreated diabetes typically leads to further pathologies, such as cardiovascular diseases, renal failure, and metabolic syndrome (high blood pressure, increased body fat, dysregulated cholesterol level, etc.) (Grundy 2006; Rao et al. 2013;

Cleland 2012; Shulman 2014; Thomas et al. 2015; Artunc et al. 2016). Because of the sheer number of affected people and diverse diabetes-associated disorders, diabetes represents a huge economic and medical burden.

Current cell-based approaches to treat diabetes

The prevalence of diabetes has elicited extended efforts for the development of cell-based approaches aiming to safely ameliorate or cure the underlying pathology. These include (i) pancreatic cell transplantations, (ii) engineering hepatocytes with β -cells like features, (iii) redesigning mesenchymal stem cells, (iv) differentiating human embryonic stem cells (hESC), (v) converting human pluripotent stem cells (hPSC), and (vi) reprogramming iPSC.

(i) Transplantation of the whole pancreas or of microencapsulated allogeneic pancreatic islets have shown promising results but face several problems, such as longevity of transplants, sustained functionality, potential adverse immune reactions, and donor procurement (Soon-Shiong 1999; M. Qi 2014). Promising alternatives to the lack of human donor material are xenogeneic pig islets (Groth et al. 1994; Valdés-González et al. 2005) and genetically engineered pig islets (Perkel 2016; Niu et al. 2017).

(ii) Hepatocytes share the same endodermal origin as endocrine cells, are characteristically similar to pancreatic cells, and express crucial glucose-responsive elements, such as glucose transporter 2 and glucokinase (Zaret & Grompe 2008). Many groups have attempted to genetically engineer hepatocytes/hepatocellular carcinoma cell lines by incorporating transgenic insulin and/or β -cell transcription factors to render them more similar to β -cells. However, most of these studies failed to demonstrate reliable and sustained glucose-triggered insulin secretion (Vollenweider et al. 1992; Short et al. 1998; Ferber et al. 2000; Mitancher et al. 1998; Kolodka et al. 1995; Valera et al. 1994; Simpson et al. 1997; Tuch et al. 1998; Tuch et al. 2003; Gerace et al. 2015).

(iii) Non-hematopoietic stem/progenitor cells and in particular mesenchymal stem cells (MSCs) have emerged as suitable living delivery vehicles (Gerace et al. 2017). MSCs are multipotent adult stromal cells found in the adult adipose and bone marrow tissues. They are easily obtainable and transduced, possess self-renewing capacity, great expansion potential, prolonged lifespans, and present low immunogenicity in an autologous setting (Aboody et al. 2008; Sanz et al. 2012; Sherman et al. 2017).

(iv) hESCs and (v) hPSCs have been successfully converted into glucose-responsive insulin-secreting cells *ex vivo* and have been shown to ameliorate diabetes in mice (Rezania et al. 2014; Pagliuca et al. 2014; Y. Takahashi et al. 2016).

(vi) Reprogramming of human somatic cells into iPSCs holds great promise for generating a pool of unlimited, immunologically compatible, patient-derived cells (K. Takahashi et al. 2007; Kiskinis & Eggan 2010; Turner et al. 2013). In 2016, the Fussenegger laboratory designed a sophisticated, exogenously inducible, and programmable synthetic lineage control circuit capable of differentiating human iPSC-derived pancreatic progenitor cells into glucose-sensitive, insulin-secreting, β -like cells (Saxena et al. 2016). One year later, the same group demonstrated that this system could be integrated into hESC-derived pancreatic progenitor cells to generate β -like cells (Saxena et al. 2017).

Synthetic designer cell approaches to treat diabetes

Two recent studies provide important clues for addressing potential obstacles when designing β -cell mimetics from non-pancreatic human cells.

(i) In December 2016, Martin Fussenegger's laboratory published a seminal paper on engineered HEK293 cells with β -mimetic-like properties for closed-loop glycaemic control (M. Xie et al. 2016). Glucose stimulation of such designer cells led to glycolysis-mediated Ca^{2+} influx, facilitating the nuclear translocation of a calcium-

responsive transcription factor, which in turn induced the expression of an engineered insulin transgene (M. Xie et al. 2016). This analogue of insulin contains furin-cleavable sites, which ensures proper processing into mature insulin (Groskreutz et al. 1994; Hay & Docherty 2003). Eventually, processed insulin is secreted in response to the Ca^{2+} -mediated depolarisation of the cell membrane. Embedding this synthetic circuit into HEK293 cells successfully corrected hyperglycaemia in T1D mice (M. Xie et al. 2016). Intriguingly, the regulation of blood glucose levels via the transcriptional upregulation of insulin (maximum insulin secretion level reached 48 hours post glucose administration) was as effective as the protein-based direct secretion system of pancreatic β -cells (insulin secretion initiated ~20 min post glucose intake (Caumo & Luzi 2004; Mathieu et al. 2017)).

Interestingly, the authors also attempted to utilise the sweet taste receptors T1R2/T1R3 instead of the $\text{Ca}_v1.3$ channel in designer cells. However, this approach was not successful in eliciting a Ca^{2+} influx, which was deemed a key component in the signalling pathway. One explanation might be that using T1R2/T1R3 without its associated G protein, α -gustducin (McLaughlin et al. 1992; G. Q. Zhao et al. 2003; Margolskee et al. 2007; Jang et al. 2007; Tordoff et al. 2012; Masubuchi et al. 2013), impaired the endogenous downstream signalling cascade that elicits Ca^{2+} influx. Also plausible, HEK293 cells might lack other components that are crucial for functional T1R2/T1R3 signalling since T1R2/T1R3 originate from taste buds and the gut (Roper 2007; Masubuchi et al. 2013).

(ii) Also in 2016, dCas9-VP160 was programmed to upregulate the transcription of endogenous insulin (Giménez et al. 2016). Such engineered cells, however, constitutively overexpressed insulin in a non-glucose-dependent manner. This limitation will hamper potential future therapeutic applications, where inducible control of insulin is imperative. Moreover, the authors did not provide any data regarding insulin

protein expression. It remains to be seen whether HEK293T cells can produce mature insulin protein upon dCas9-TF-mediated gene activation without simultaneously co-expressing insulin processing enzymes (see below).

The future of dCas9-synT1R3-engineered designer cells

To unlock the therapeutic potential of dCas9-synT1R3 for engineering glucose-responsive artificial β -cell mimetics, at least three aspects should be taken into consideration.

First, insulin is translated as proinsulin and necessitates further peptidase-mediated processing to yield mature, bioactive insulin (Nishi & Nanjo 2011; M. Liu et al. 2014). This may require the implementation of more complex gene circuits, in which the expression of prohormone convertases (PC1/3 and PC2) and carboxypeptidase E (CPE) are co-induced (M. Liu et al. 2014; Gerace et al. 2017). This can be easily accomplished by supplying additional sgRNAs for multiplexed transcriptional gene induction, as previously demonstrated (Cong et al. 2013; Maeder et al. 2013; Cheng et al. 2013; Farzadfard et al. 2013; Gilbert et al. 2013; Kabadi et al. 2014; Chavez et al. 2015; Konermann et al. 2015; Nihongaki, Yamamoto, et al. 2015; Black et al. 2016; Gerace et al. 2017).

Second, release of bioactive insulin-containing vesicles may require Ca^{2+} signalling that could be induced by overexpressing the components of the $\text{Ca}_v1.3$ channel (D'Arco & Dolphin 2012; Ashcroft & Rorsman 2013), as previously described for HEK293 cells (M. Xie et al. 2016).

Third, engineered dCas9-synT1R3 β -cell mimetics will require shielding from the immune system, which can be achieved in principle by microencapsulation in established and clinically validated alginate microcapsules (Soon-Shiong 1999; M. Qi 2014; Schukur et al. 2015; Vegas et al. 2016; M. Xie et al. 2016).

In conclusion, dCas9-synT1R3-engineered HTLA cells represent the first implementation approach in which D-glucose sensing is directly coupled to a dose-dependent increase in endogenous insulin expression levels.

6.3.3 Choice of cell types for dCas9-synRs

Previously, HEK293 derivatives have been extensively exploited for developing new receptor technologies, such as TANGO (Barnea et al. 2008; Kroeze et al. 2015), CRISPR ChaCha (Kipniss et al. 2017), CRISPR-TANGO-dCpf1 (Y. Liu et al. 2017) and MESA (Daringer et al. 2014; Schwarz et al. 2017). In a recent study from the Fussenegger laboratory, an engineered interleukin signalling cascade was engrafted into HEK293 cells, enabling killing of cancer cells upon direct cell-to-cell contact (Kojima et al. 2018). This concept strikingly illustrates the power of synthetic receptors in rewiring cellular signalling processes and converting non-immune cells into functional immune cell mimics. However, these experimental frameworks and the ones presented in this thesis are restricted to cell lines, which prevents substantiated assumptions of *in vivo* effectiveness.

To advance the dCas9-synR platform technology towards engineering clinically-relevant, cancer-fighting designer cells, it would be desirable to select more relevant cell types. One of the most important property sought after in a candidate cell type is tumour tropic behaviour, such as in MSCs and neural stem cells (NSCs). Engineered MSCs and NSCs have been successfully utilised to trigger custom cancer killing programmes (Aboody et al. 2013; Metz et al. 2013; Matuskova et al. 2015; Nouri et al. 2015; Z. Zhang et al. 2016; Marini et al. 2017; Moradian Tehrani et al. 2017). However, primary human T cells represent the most reasonable choice because of their clinical relevance and compelling success in the context of CAR T cell clinical trials (Gill & June 2015; Kershaw et al. 2013; W. A. Lim & June 2017) and for synNotch (Roybal, J. Z.

Williams, et al. 2016). The immortalised human T cell line Jurkat is a great “intermediate” between an easy to cultivate cell line and a potent primary T cell. Jurkat cells have been successfully used to establish and characterise synNotch receptors (Morsut et al. 2016; Roybal, Rupp, et al. 2016) as well as to explore AND logic gate configurations of synNotch and CAR receptors (Roybal, Rupp, et al. 2016).

Based on these considerations, I proposed that future dCas9-synRs implementations should be first tested in Jurkat T cells and subsequently in primary T cells.

Chapter 7 : General discussion and future perspectives

In this thesis, I have described the development of a new class of dCas9-based synthetic receptors (dCas9-synR). These chimeric receptors are characterised by combining the evolutionarily optimised ligand-sensing domains of various native receptors with the highly programmable dCas9-TF-based signal transduction module. The tight OFF/ON state transition characteristics of dCas9-synRs and their versatility in redirecting the cellular information flow makes them ideally suited to engineer designer cells capable of sensing specific disease biomarkers and in turn drive customised, therapeutically-relevant, multi-gene programmes.

To attain minimal background activation in the absence of cognate ligands (OFF state) and at the same time generate a robust transcriptional activation response upon agonist stimulation (ON state), I conceived a split dCas9 core architecture through rational design and protein engineering. The customisable TEV-based signal release mechanism facilitates conditional release and subsequent reconstitution of a functional dCas9-TF (Chapter 3).

The minimal split dCas9-TF core architecture was first grafted onto the original GPCR TANGO scaffold (Kroeze et al. 2015). The generalisability of the dCas9-based chimeric receptor architecture was demonstrated by adapting several dCas9-synGPCRs, responsive to a range of ligands. dCas9-synGPCRs displayed stringent OFF/ON state transitions upon ligand challenge. Notably, dCas9-synGPCRs can induce the transcription of any user-defined, multi-gene programme in an agonist dose-dependent manner. Furthermore, I showed that Boolean AND logic gate engineering enabled the reliable integration of multi-factorial inputs (Chapter 4).

As a second example of a cell-surface receptor family, I derived and optimised a core split dCas9 architecture and alternative split TEV protease release mechanism for a dimerisation-dependent receptor tyrosine kinase, dCas9(N/C)-synVEGFR1/2. Similar to dCas9-synGPCRs, stringent OFF/ON state transitions upon cognate ligand engagement enabled robust endogenous gene upregulation (Chapter 5).

Lastly, I deployed various chimeric dCas9-synRs to activate therapeutically-relevant transcriptional programmes in mammalian cells. I demonstrated in a proof of concept that various biomarkers can serve as input stimuli for eliciting customised, non-native transcriptional response programmes (Chapter 6).

7.1 Designing the next generation of dCas9-synRs

To advance the implementation of dCas9-synRs as synthetic devices in cell-based human therapy (e.g., engineered T cells to fight cancer; prosthetic gene network-equipped designer cells to treat metabolic and autoimmune diseases), several current technical limitations will have to be addressed. These include sustained signalling capabilities of dCas9-TF upon ligand activation and the reversibility of engineered output programmes. Furthermore, the applicability to other cell types will have to be explored because all current experiments with dCas9-synRs have been conducted in HEK293-derived cells.

A board and reliably operating repertoire of dCas9-synRs will facilitate the implementation of signalling capabilities in designer cells for next-generation cell-based therapies. This may involve the adaptation of additional dCas9-synRTKs and dCas9-synGPCRs or novel dCas9-synR derivatives, such as dCas9-based synthetic Notch receptors (dCas9-synNotch, Chapter 7.1.4).

7.1.1 Duration and reversibility of dCas9-synR-mediated signalling

The precise control of receptor signalling, whether endogenous or synthetic, is of paramount importance. Synthetic receptor signalling in designer cells must follow a positive correlation with the presence of cognate ligand. If no agonist is present, the receptor should remain in an OFF state (no background activation). In the presence of a ligand, however, signalling should be robustly initiated. Following agonist removal, signalling through the receptor should diminish over time and eventually reach baseline levels. Upon repeated exposure to agonists, the receptor should again engage in robust signalling.

Well-equilibrated and reversible genetic circuits have been shown to offer a potentially curative approach for managing blood glucose levels in mice (M. Xie et al. 2016). In a prosthetic gene network, the tightly regulated insulin expression in response to varying D-glucose concentrations enabled the treatment of hyperglycaemia and concomitantly the prevention of life-threatening hypoglycaemia due to insulin overexpression (M. Xie et al. 2016). However, when these prerequisites are not met, constitutively active receptor signalling can lead to sustained expression of therapeutic programmes that can have deleterious side effects, for example in the case of adaptive T cell therapy (Kalos & June 2013; W. A. Lim & June 2017).

Within the scope of this thesis, I investigated the potential background leakiness of a candidate dCas9-synGPCR, dCas9-synBDKRB2, over a 48 hours time course (Figure 4.9b). In the no-agonist condition, I observed a transcriptional response at 48 hours. Similarly, the CRISPR ChaCha system also yielded a transcriptional output at this later time point in the absence of cognate ligands (Kipniss et al. 2017). These results imply that without agonist stimulation, dCas9-based chimeric GPCRs can produce programmed output responses over time, indicating unspecific autonomous signalling.

It remains to be seen how these synthetic receptors behave over longer periods, i.e. beyond 48 hours.

With regard to the reversibility of dCas9-based chimeric receptors, although I have not investigated this property in my thesis, Kipniss et al. (2017) determined the reversibility of dCas9-based CRISPR ChaCha receptors. Since induced gene expression is a consequence of dCas9-TF translocation into the nucleus, reversibility can only originate from a degrading pool of activated dCas9-TF. Therefore, the results by Kipniss et al. (2017) could in principle be extrapolated to dCas9-synRs.

Kipniss et al. (2017) stably integrated all CRISPR ChaCha system components into HEK392T cells. Then, they treated transduced cells with agonist for 24 hours and subsequently determined reporter gene expression over time under no-agonist conditions (Kipniss et al. 2017). This experiment revealed that initially strong agonist-induced reporter expression eventually declined to almost pre-treatment levels by day 7 (Kipniss et al. 2017). This reduction in reporter output is most likely a result of a slowly depleting pool of dCas9-TF. Taken together, this data indicates that the CRISPR ChaCha system is reversible and suggests that a similar property may apply to dCas9-synRs.

Despite these encouraging results, the next generation of dCas9-synRs should incorporate a stringent degradation system for dCas9-TF to eliminate background leakiness (beyond 48 hours) and to confer rapid reversibility (within hours rather than days). Such a strategy could provide a more refined temporal expression profile of therapeutic factors in future cell-based implementations.

A number of systems have already been established in which Cas9 and/or dCas9-TFs were fused to destabilisation domains to promote active degradation (Balboa et al. 2015; Maji et al. 2017; Senturk et al. 2017; Kleinjan et al. 2017) or induced degradation upon administration of chemical compounds (Kleinjan et al. 2017). The

former implementations completely abolished Cas9/dCas9-TF activity, which could only be restored by exogenous administration of small molecules. Nonetheless, in my opinion, it would be more desirable to employ an inherently self-sufficient, endogenous degradation system that confines transcriptional activity of reconstituted, transcriptionally active dCas9-TF in dCas9-synRs.

One such system is provided by the intracellular domain of endogenous Notch receptors (NICD). Upon ligand engagement, the Notch receptor undergoes intramembrane proteolysis which liberates NICD (Artavanis-Tsakonas et al. 1999; Kopan & Ilagan 2009). Subsequently, NICD translocates into the nucleus to drive target gene expression (Kopan & Ilagan 2009). NICD possesses two degradation sequences. On its C-terminus, there is a protein destabilising sequence, called the PEST domain (proline-, glutamic acid-, serine-, and threonine-rich region), which facilitates ubiquitin-mediated degradation of NICD in the nucleus (Fryer et al. 2002; Fryer et al. 2004; Chiang et al. 2006; O'Neil et al. 2007; B. J. Thompson et al. 2007; Moretti & Brou 2013). On its N-terminus, human Notch1 NICD contains an N1-Box that mediates degradation in the cytoplasm and nucleus (Broadus et al. 2016). Both degradation signals can work in synergy and were shown to promote rapid degradation when appended to heterologous proteins (Broadus et al. 2016). This dual degradation system presents an ideal opportunity to rapidly degrade liberated and reconstituted dCas9-TF rather than degrading the unstimulated synthetic dCas9-synR itself.

Alternatively, the PEST domain from ornithine decarboxylase can be deployed to reduce the half-life of heterologous proteins to ~2 hours (X. Li et al. 1998; Leclerc et al. 2000). Fusing dCas9-TF to this PEST domain may, however, limit the abundance of dCas9-synRs in a cell.

A recently discovered, naturally occurring mechanism to regulate the activity of CRISPR-based nucleases are phage-encoded anti-CRISPR (Acr) proteins (Bondy-

Denomy 2017; Pawluk et al. 2018). Among them, AcrIIA2, AcrIIA4, AcrIIC1, AcrIIC2, and AcrIIC3 have been shown to prevent the binding of Cas9 homologs to the DNA target site by directly binding to Cas9 *in vitro* and in mammalian cells (Pawluk et al. 2016; D. Dong et al. 2017; H. Yang & Patel 2017; Shin et al. 2017; Harrington et al. 2017; Rauch et al. 2017). Therefore, anti-CRISPR proteins represent another promising approach to actively control dCas9-synR performance at the level of dCas9-TF activity.

7.1.2 Portability and refinement of dCas9-synGPCRs' mode of activation

The GPCR TANGO-based dCas9-synGPCRs are a powerful tool to redirect the information flow of stimulated GPCRs towards novel output responses. The original GPCR TANGO architecture has been previously implemented *in vivo* (Inagaki et al. 2012; Jagadish et al. 2014; Talay et al. 2017), demonstrating portability of the system to diverse cell types. To realise the full potential of dCas9-synGPCRs for future implementations, it will be important to incorporate dCas9-synGPCRs into various cell types, such as primary T cells.

At the core of dCas9-synGPCRs signalling is the fusion protein β -arrestin2-TEV. Agonist-stimulated dCas9-synGPCRs rely on β -arrestin2-TEV recruitment for subsequent signal propagation via dCas9-TF. Initially, the dCas9-synGPCR architecture was conceived and tested in a HEK293 cell line that constitutively expresses a stably integrated β -arrestin2-TEV fusion protein at low levels (HTLA cells). It was suggested that transient transfection of β -arrestin2-TEV plasmids into HEK293 cells would not yield desirable OFF/ON state transition characteristics (Bryan L. Roth, personal communication) due to the high abundance of the fusion protein. This circumstance indicates that the absolute β -arrestin2-TEV protein level in a cell is a crucial determinant for GPCR TANGO receptor performance.

TEV protease cleavage potency can be described as a function of absolute protein level and catalytic activity. An excess of TEV protease or high catalytic activity can lead to sustained liberation of the TF, independent of GPCR TANGO receptor stimulation.

For the next generation of dCas9-synGPCRs, it will be important to not only integrate β -arrestin2-TEV for low expression but also to further investigate the cleavage properties of TEV protease. HTLA cells contain the wild-type TEV protease that undergoes intramolecular auto-inactivation, thereby rendering itself less active (Parks et al. 1995; Kapust et al. 2001). This inactivation process increases its K_M and, as a consequence, auto-inactivated TEV protease requires more time to cleave a definite amount of substrate (i.e., TCS), than non-auto-inactivated TEV protease. This property, however, is rather beneficial because it reduces the likelihood of TCS cleavage events due to simple Brownian motion and instead, favours cleavage by actively recruited β -arrestin2-TEV to agonist-stimulated synthetic GPCRs.

In support of the concept that highly active TEV protease abrogates stringent OFF/ON state transition behaviour are the studies on CRISPR ChaCha as well as MESA-dCas9-TF receptors, which display very high background activation levels in the uninduced state (Kipniss et al. 2017; Schwarz et al. 2017). Both systems utilise a highly active mutant TEV protease (S219V, TEV^{S219V}) that greatly resists auto-inactivation (Kapust et al. 2001). TEV^{S219V} is ~100-fold more stable than wild-type TEV protease (Kapust et al. 2001). Based on my studies, it is plausible that using a less active TEV protease variant might improve OFF/ON state transitions in both CRISPR ChaCha and MESA-dCas9-TF receptors.

To further reduce spontaneously occurring cleavage events in the uninduced state, a truncated TEV protease (TEV ^{Δ 219-242}) variant has been successfully employed (Gray et al. 2010; Wenjing Wang et al. 2017; D. Lee et al. 2017). This variant does not

undergo auto-inactivation, however, displays a desirable ~7-fold higher K_M than wild-type TEV protease while maintaining a similar k_{cat} value (describes the turnover rate at which an enzyme-substrate complex produces an enzyme and product) (Kapust et al. 2001).

Taken together, I conclude that the dCas9-synGPCR architecture (i) is likely to function in cells other than HEK293-derived cells; (ii) requires a low expression level of β -arrestin2-TEV; (iii) could be further improved in terms of residuals background signalling and concomitantly OFF/ON state transition characteristics by deploying the alternative TEV ^{$\Delta 219-242$} protease.

Furthermore, it is conceivable to gain improved system performance of dCas9-synGPCRs by adapting a split TEV protease release mechanism similar to dCas9-synRTKs (dCas9-synGPCR_v2, Figure 7.1a). This would render TEV protease inactive thereby mitigating unintended background cleavage of TCS sequences as described above.

A TEV(N)-dCas9(N) fusion protein could be fused downstream of the GPCR CDS-V₂-TCS sequence (GPCR-V₂-TCS-TEV(N)-dCas9(N)) and the complementary fusion protein TEV(C)-dCas9(C)^{VP64} would be attached downstream of another TCS to β -arrestin2 (Figure 7.1a). Upon ligand engagement, the GPCR actively recruits β -arrestin2-TCS-TEV(C)-dCas9(C)^{VP64}. Proximity of TEV(N) and TEV(C) promotes its reassembly into an active protease, which can then act on the TCS sequences in the GPCR and β -arrestin2 to liberate the dCas9-TF fragments (Figure 7.1a).

Such split TEV designs (Djannatian et al. 2011; D. Lee et al. 2017) as well as split TF configurations (Takakura et al. 2012; Hattori & Ozawa 2015; Kono et al. 2017)

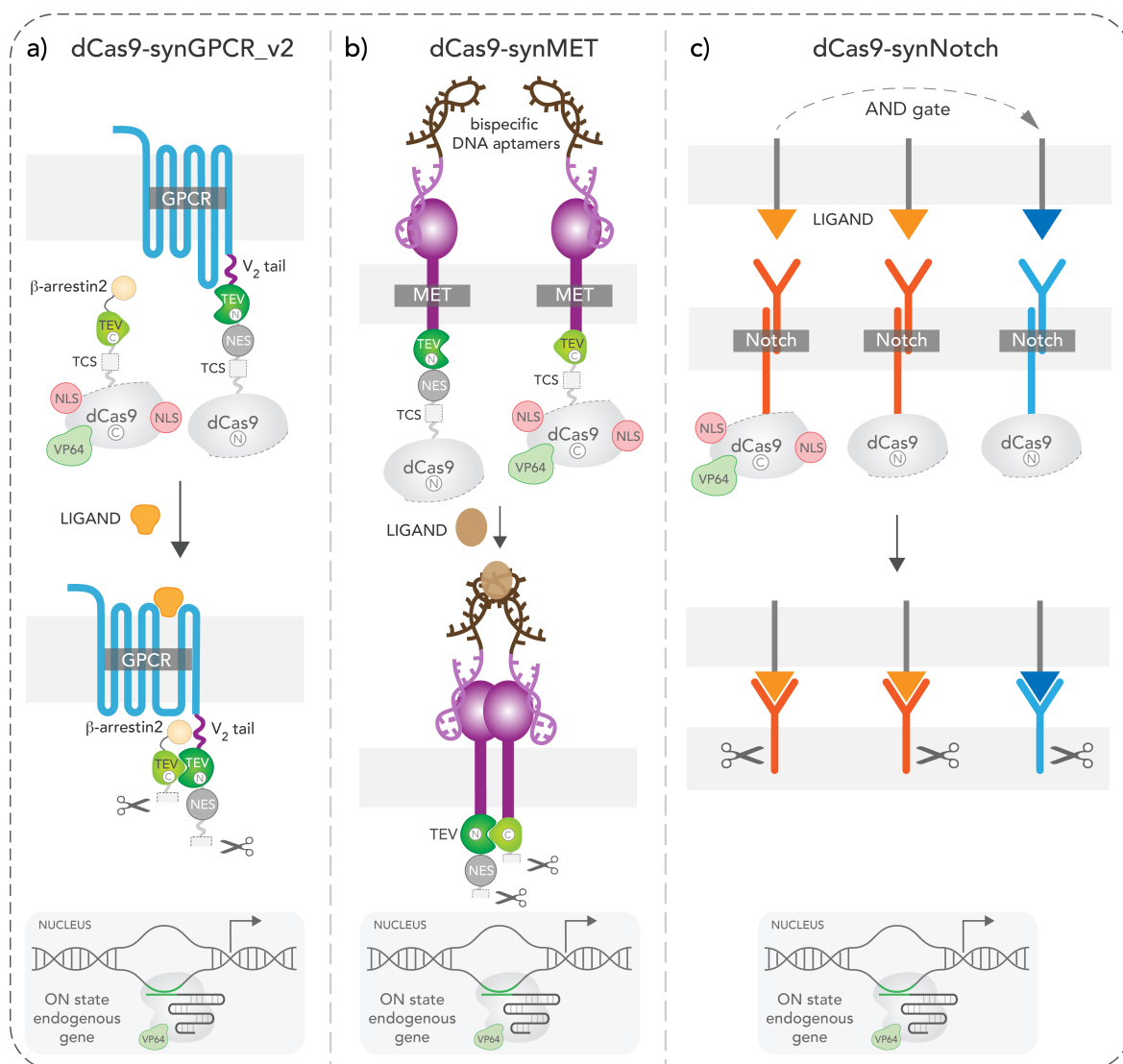


Figure 7.1 | Next-generation dCas9-synR instalments. (a) dCas9-synGPCR_v2 describes a hybrid design of dCas9-synGPCR and dCas9-synRTK, in which a dual split TEV/split dCas9-TF architecture is adapted to the original GPCR TANGO scaffold. **(b)** By analogy with dCas9(N/C)-synVEGF1/2, the MET receptor can be converted into dCas9-synMET. dCas9-synMET receptors can then be programmed to recognise any user-defined growth factor by supplying customised bispecific DNA aptamers, which provide a link between receptor and ligand. **(c)** dCas9-synNotch is based on synNotch in which the artificial TFs are replaced by dCas9 fragments. Ligand-induced intramembrane proteolysis triggers the release of dCas9 fragments. Deploying heterologous extracellular domains in dCas9-synNotch receptors can be used to create Boolean AND logic gates.

have previously been reported for GPCR TANGO scaffolds. It remains to be seen if a dual split TEV/split dCas9-TF architecture will be a viable option for dCas9-synGPCRs.

7.1.3 A universal dCas9-synRTK for growth factor signal integration

In this thesis, I created a prototype synthetic RTK, dCas9(N/C)-synVEGFR1/2, which responds to soluble VEGF. Detecting an alternative growth factor, however, will require the generation of a different dCas9-synRTK receptor.

In total, there are 20 RTK subfamilies with 58 known members (Robinson et al. 2000; Blume-Jensen & Hunter 2001). Converting all these receptors into dCas9-synRTK architectures could become a time-consuming endeavour. Furthermore, some RTKs may not be suitable for such an adaptation. One of the key requirements for the present approach is that candidate RTKs undergo homo- and/or heterodimerisation upon cognate ligand engagement (Lemmon & Schlessinger 2010; Maruyama 2014; Maruyama 2015). Based on these considerations, a universal strategy for dCas9-synRTKs would be highly desirable.

Shinsuke Sando's group has recently developed non-protein ligands for the hepatocyte growth factor (HGF) receptor (also known as MET receptor) (Ueki et al. 2016; Ueki et al. 2017). These non-protein ligands are oligonucleotide-based HGF mimetics, containing MET-specific aptamers of 100-mer single-stranded DNA (Ueki et al. 2016). Similar to endogenous HGF ligands, HGF mimetics facilitate MET dimerisation and trigger endogenous MET receptor signalling. Notably, HGF mimetics were subsequently constructed as bispecific DNA aptamers consisting of a MET receptor aptamer and a second aptamer that binds specifically to any user-defined external protein, such as platelet-derived growth factor (PDGF), which would otherwise not function as an agonist for the MET receptor (Ueki et al. 2017). Deploying such

bispecific DNA aptamers in combination with synthetic dCas9-based MET receptors (dCas9-synMET, Figure 7.1b) could be leveraged to generate a universal dCas9-synRTK with tailored ligand input specificity.

In vivo delivery of oligonucleotides, one kind of which are bispecific aptamers, has made tremendous progress in recent years and is now more conceivable for the clinic than ever before (Khvorova & Watts 2017; Khorkova & Wahlestedt 2017; Crooke et al. 2017). In a prospective therapeutic setting, designer cells equipped with dCas9-synMET receptors could then receive exogenously delivered bispecific DNA aptamers to detect any user-defined growth factor.

7.1.4 Adapting the split dCas9 core architecture to engineer a novel class of synthetic Notch receptors (dCas9-synNotch)

Wendell Lim's laboratory established a synthetic form of the cell-surface receptor Notch (synNotch) for reprogramming transgenic T cells (Morsut et al. 2016; Roybal, J. Z. Williams, et al. 2016; Roybal, Rupp, et al. 2016). The synNotch receptor with its characteristics (Chapter 1.2.3.2) would be a great addition to the repertoire of dCas9-synRs.

synNotch is activated by engaging membrane-bound endogenous and/or synthetic ligands via its variable extracellular domain. The mechanical force generated through cell-cell interaction triggers catalytic proteolysis and subsequent liberation of intracellularly fused synthetic TFs (Morsut et al. 2016). In theory, the artificial TFs could be readily replaced with split dCas9 fragments to create dCas9-synNotch (Figure 7.1c). Importantly, such a design would not necessitate a TEV protease-based release mechanism as described for dCas9-synRTKs or dCas9-synGPCRs. Since dCas9-TFs do not require the integration of transgenes and can be easily utilised for multi-gene output programmes, a dCas9-synNotch scaffold could broaden the transcriptional

response of engineered T cells. Finally, by analogy with other dCas9-synRs, the multiplexing of two different antigen-specific dCas9-synNotch receptors would allow the design of Boolean AND logic gates (Figure 7.1c).

7.2 Precautions and future applications of dCas9-synRs

The broad range of current and future dCas9-synRs can be used to implement sense and respond circuits in designer cells. The detection of disease-associated biomarkers by engineered cells and their conversion into a therapeutically-relevant, non-native output response can help tackle various disease indications. Engineering immune cells with synthetic receptors may be used for cell-based therapies to fight cancer, autoimmunity, degeneration, and inflammation (Fischbach et al. 2013; Roybal & W. A. Lim 2017). Similarly, installing prosthetic gene networks into cells can help treat metabolic and autoimmune diseases (Folcher & Fussenegger 2012; S. Ausländer et al. 2012; Bacchus et al. 2013; Heng et al. 2014; Heng et al. 2015).

7.2.1 Immunological challenges of utilising dCas9-TFs

Autologous T cells equipped with synthetic receptors pose a potentially higher risk of triggering adverse host immune responses compared to prosthetic gene networks installed in non-autologous designer cells. That is owing to the fact that the latter can be encapsulated prior to implantation into a recipient (Folcher & Fussenegger 2012; Bacchus et al. 2013; Heng et al. 2015). Biocompatible and semi-permeable packaging materials (e.g., alginate-poly-L-lysine) facilitate the diffusion of small molecules and proteins (sense and respond) while shielding the designer therapeutic cells from the host immune system (Drury & Mooney 2003).

In contrast, genetically engineered human T cells, such as CAR T cells, circulate freely in the bloodstream in search of target cells, exposing them to the host's immune

system. CARs and synNotch receptors contain scFvs as antigen recognition domains. scFvs are mostly derived from mice and therefore recognised by the human immune system as non-self antigens (Kershaw et al. 2006; Lamers et al. 2011; Turtle et al. 2016). However, immune responses arising from mouse scFvs have been addressed by deploying humanised (Westwood et al. 2005; Y. Zhao et al. 2009; Sun et al. 2014; Johnson et al. 2015) or even fully human scFvs (Morgan et al. 2012; Lanitis et al. 2012; Song et al. 2015; Sommermeyer et al. 2017).

Notably, potential immunogenicity arising from the cell-surface presentation of non-self antigens may be mitigated by dCas9-synGPCRs and dCas9-synRTKs because they are endowed with native human cell-surface sensing domains. Nonetheless, the bacterial-derived dCas9 as intracellular signal transduction module is likely to elicit MHC-mediated T cell immune responses. Recent data demonstrated that *Streptococcus pyogenes* Cas9 proteins trigger cellular and humoral immune responses in mice (Chew et al. 2016). Furthermore, pre-existing anti-Cas9 antibodies have been detected in human serum due to previous infections with common Cas9-bearing *Streptococcus pyogenes* and *Staphylococcus aureus* species (Charlesworth et al. 2018). These immunological concerns have to be addressed in order to create safe, dCas9-based designer therapeutic T cells. One conceivable approach may encompass the sensitisation of the host immune system by inducing host immune tolerance to the Cas9 protein. Autoimmune responses have already been treated by actively challenging patients with disease-causing auto-antigens (S. D. Miller et al. 2007; Sack et al. 2014; Pearson et al. 2017).

7.2.2 Multiplexing synthetic receptors for precision multi-input sensing, control, and response programmes

dCas9-synRs constitute a diverse group of novel chimeric receptors that could potentially overcome some of the problems encountered with current transgenic T cell designs. CAR-engineered T cells have been specifically created to elicit a standard, naturally encoded TCR-equivalent response programme upon recognition of a single cell-surface antigen. Single input integration can cause complications such as limited discrimination of healthy bystander tissue causing on-target, off-tumour cell killing. Furthermore, this phenomenon can also lead to continuously active and/or hyperactive CAR T cells inducing cytokine storm/cytokine release syndrome and tumour lysis syndrome with deleterious effects in patients (Kalos & June 2013; W. A. Lim & June 2017).

In addition to sense and respond complications, developing re-directed therapeutic T cells against solid tumours possess additional challenges (Sadelain et al. 2013; Joyce & Fearon 2015; Newick et al. 2016; Gross & Eshhar 2016; H. Zhang et al. 2016; Klebanoff et al. 2016; Roybal & W. A. Lim 2017; W. A. Lim & June 2017). For instance, solid tumours create immunosuppressive microenvironments by highjacking other immune cells, such as macrophages or granulocytes, which then secrete immunosuppressive compounds compromising therapeutic outcomes (Motz & Coukos 2011). In addition, solid tumours lack non-vital tissue, unlike some blood cancers. In the case of B cell malignancies, healthy and malignant CD19⁺ B cells can be completely depleted in a patient before replenishment with healthy donor material (Gross & Eshhar 2016). This unique feature underlies the success of CD19 CAR T cell therapies.

Future designer therapeutic T cells may benefit from a combination of CARs and dCas9-synRs. The sensing and integration of a broader portfolio of disease-associated

antigens in combination with biomarkers from the tumour microenvironment will create more stringent tumour-specific signatures that could be recognised by designer T cells (Teng et al. 2015; Roybal & W. A. Lim 2017). Such tailor-made, tumour-directed therapeutic T cells should display: (i) enhanced on-target specificity by increased tumour discrimination capabilities; (ii) improved safety via multi-antigen triggered activation and multi-input Boolean logic gate processing; (iii) advanced and sophisticated therapeutic programmes due to orthogonal receptor signalling and deployment of dCas9-TF. These characteristics are further discussed below in more detail.

(i) The characterisation and detection of solid tumour-specific microenvironment markers, such as the expression profile of chemokines, has become a feasible task (Di Stasi et al. 2009; Craddock et al. 2010; Moon et al. 2011). By expanding these expression profiles to other cytokines, growth factors, and tumour-specific signalling molecules, more comprehensive tumour signatures may be drafted. The greater the number of disease-associated input signals a designer T cell requires for activation, the more on-target and spatially restricted a therapeutic programme can be delivered.

A striking example of increased effectiveness in killing solid tumours are CAR T cells transduced with endogenous chemokine receptors. Such designer T cells utilised the elevated levels of cytokines produced by tumours as additional information source for directed migration and infiltration (Kershaw et al. 2002; Di Stasi et al. 2009; Craddock et al. 2010; Moon et al. 2011). This beneficial effect underscores the potential of combining receptor platforms for improved adoptive T cell therapy.

(ii) Combinatorial biomarker sensing could increase the safety of cell-based therapies not only by mitigating off-tumour killing but also by concomitantly allowing the implementation of Boolean AND logic gates. For example, CARs have been successfully used in combination with synNotch receptors in mice (Roybal, Rupp, et al.

2016). In addition, an extensive toolbox of exogenously controllable AND gate safety switches has been established in designer T cells (C.-Y. Wu, Rupp, et al. 2015; Fesnak et al. 2016; Roybal & W. A. Lim 2017). These gates and switches may also be combined with dCas9-synRs.

(iii) CARs trigger native TCR output responses in T cells. In contrast, dCas9-synRs could induce an “à la carte” programme. A customised cytokine cocktail could be compiled to more effectively eradicate a particular tumour type (Dranoff 2004). Moreover, in a temporally and spatially restricted manner, dCas9-synRs/CAR dual engineered T cells could activate the (over-)expression of additional cytokines that enhance tumour-clearing properties. Ectopic expression of IL-12, IL-15, and IL-18 in CAR T cells have been shown to support their efficacy, proliferation, and survival (Quintarelli et al. 2007; Hoyos et al. 2010; Chmielewski et al. 2011; Chmielewski & Abken 2017; B. Hu et al. 2017).

7.2.3 dCas9-synGPCRs as novel modality to fight HIV-1 infection

In principle, dCas9-synGPCRs-engineered T cells may be employed to both protect against HIV-1 infection as well as to actively kill HIV-1-infected cells.

Acquired immunodeficiency syndrome (AIDS) is caused by human immunodeficiency virus (HIV) (Barré-Sinoussi et al. 2013). In 2016, 36.7 million people were infected with HIV-1 and 1.0 million died of AIDS-related diseases (<http://www.who.int/gho/hiv/en/>, visited January 2018). HIV-1 uses its envelope protein gp120 to fuse with and infect T cells. gp120 binds to the main receptor CD4 (Dalglish et al. 1984; Klatzmann et al. 1984; Maddon et al. 1986) and to one of the two co-receptors, either CC-chemokine receptor 5 (CCR5) (Alkhatib et al. 1996; Choe et al. 1996; H. Deng et al. 1996; Doranz et al. 1996; Dragic et al. 1996) or CXCR4 (Bleul et al. 1996; Y. Feng et al. 1996). In the early course of infection, CCR5 is the predominant entry gate whereas in late-stage

infections CXCR4 is preferred, which concomitantly correlates with a more rapidly progressing disease (Connor et al. 1997).

dCas9(N/C)-synCCR5 and dCas9(N/C)-synCXCR4 engineered T cells could serve two distinct purposes. First, upon gp120 engagement, such designer T cells could induce the expression of host restriction factors to prevent the spread of HIV infection. As a second measure, engineered T cells could detect and destroy infected T cells that present gp120 proteins on their surface by activating a tailor-made antiviral therapeutic programme.

Potential antiviral host restriction factors include apolipoprotein B-editing catalytic polypeptide 3 B and G proteins (APOBEC3B, APOBEC3G), tripartite motif 5 α protein (TRIM5 α), tetherin, and sterile alpha motif and histidine-aspartate domain-containing protein 1 (SAMHD1) (Santa-Marta et al. 2013; S. Chen et al. 2018). The ectopic (over-)expression of TRIM5 α (Sayah et al. 2004; Stremlau et al. 2004), SAMHD1 (Hrecka et al. 2011; Laguette et al. 2011; Baldauf et al. 2012), and tetherin (Neil et al. 2008) were shown to interfere with virus integration, replication, and particle budding, respectively. Transcriptional upregulation of APOBEC3B and APOBEC3G by dCas9-TF was successful in interfering with viral replication (Bogerd et al. 2015). Since HIV-1 has developed mechanisms to evade the compromising effects of antiviral host effectors, a combinatorial upregulation of these genes should be explored to strengthen the antiviral defence response. Additionally, any dCas9-synGPCR-initiated programme could be complemented with existing treatments, such as antiretroviral therapy (ART) (Palella et al. 1998).

Apart from boosting the expression level of host factors, dCas9-TF could be used to upregulate a virally encoded antisense long non-coding RNA from the integrated viral genome (provirus) to suppress viral gene expression (Saayman et al. 2014). Finally, HIV-1 expression was shown to be inhibited by steric hindrance of dCas9

(Saayman et al. 2016; Gang Wang et al. 2017) or repressed with dCas9-KRAB (H.-K. Liao et al. 2015) when targeting the long terminal repeat (LTR) promoter region of the provirus.

Taken together, the extensive repertoire of programmable dCas9-synRs offer unprecedented opportunities to create novel therapeutic designer cells. The potential of combinatorial, multi-factorial sense and respond programmes will make cell-based therapies safer and more effective, promoting their wider application in modern medicine as “living therapeutics” (Kitada et al. 2018).

Chapter 8 : Bibliography

- Aboody, K.S. et al., 2013. Neural stem cell-mediated enzyme/prodrug therapy for glioma: preclinical studies. *Science translational medicine*, 5(184), pp.184ra59–184ra59.
- Aboody, K.S., Najbauer, J. & Danks, M.K., 2008. Stem and progenitor cell-mediated tumor selective gene therapy. *Gene Therapy*, 15(10), pp.739–752.
- Adamson, B. et al., 2016. A Multiplexed Single-Cell CRISPR Screening Platform Enables Systematic Dissection of the Unfolded Protein Response. *Cell*, 167(7), pp.1867–1873.e21.
- Aldrin-Kirk, P. et al., 2016. DREADD Modulation of Transplanted DA Neurons Reveals a Novel Parkinsonian Dyskinesia Mechanism Mediated by the Serotonin 5-HT₆ Receptor. *Neuron*, 90(5), pp.955–968.
- Alexander, G.M. et al., 2009. Remote control of neuronal activity in transgenic mice expressing evolved G protein-coupled receptors. *Neuron*, 63(1), pp.27–39.
- Alkhatib, G. et al., 1996. CC CKR5: a RANTES, MIP-1 α , MIP-1 β receptor as a fusion cofactor for macrophage-tropic HIV-1. *Science*, 272(5270), pp.1955–1958.
- Allen, J.A. & Roth, B.L., 2011. Strategies to discover unexpected targets for drugs active at G protein-coupled receptors. *Annual Review of Pharmacology and Toxicology*, 51, pp.117–144.
- Allgood, V.E. & Eastman, E.M., 1997. Chimeric receptors as gene switches. *Current Opinion in Biotechnology*, 8(4), pp.474–479.
- Altenschmidt, U. et al., 1996. Cytolysis of tumor cells expressing the Neu/erbB-2, erbB-3, and erbB-4 receptors by genetically targeted naive T lymphocytes. *Clinical cancer research: an official journal of the American Association for Cancer Research*, 2(6), pp.1001–1008.
- Amabile, A. et al., 2016. Inheritable Silencing of Endogenous Genes by Hit-and-Run Targeted Epigenetic Editing. *Cell*, 167(1), pp.219–232.e14.
- Anisimov, A. et al., 2013. The basis for the distinct biological activities of vascular endothelial growth factor receptor-1 ligands. *Science signaling*, 6(282), pp.ra52–ra52.
- Arber, C., Young, M. & Barth, P., 2017. Reprogramming cellular functions with engineered membrane proteins. *Current Opinion in Biotechnology*, 47, pp.92–101.
- Aricescu, A.R., Lu, W. & Jones, E.Y., 2006. A time- and cost-efficient system for high-level protein production in mammalian cells. *Acta Crystallographica Section D Biological Crystallography*, 62(10), pp.1243–1250.
- Armbruster, B.N. & Roth, B.L., 2005. Mining the receptorome. *Journal of Biological Chemistry*, 280(7), pp.5129–5132.
- Armbruster, B.N. et al., 2007. Evolving the lock to fit the key to create a family of G protein-coupled receptors potentially activated by an inert ligand. *Proceedings of the National Academy of Sciences*, 104(12), pp.5163–5168.
- Artavanis-Tsakonas, S., Rand, M.D. & Lake, R.J., 1999. Notch signaling: cell fate control and signal integration in development. *Science*, 284(5415), pp.770–776.
- Artunc, F. et al., 2016. The impact of insulin resistance on the kidney and vasculature. *Nature Publishing Group*, 12(12), pp.721–737.
- Ashcroft, F.M. & Rorsman, P., 2013. K(ATP) channels and islet hormone secretion: new insights and controversies. *Nature Publishing Group*, 9(11), pp.660–669.
- Assadi-Porter, F.M. et al., 2008. Direct NMR detection of the binding of functional ligands to the human sweet receptor, a heterodimeric family 3 GPCR. *Journal of the American Chemical Society*, 130(23), pp.7212–7213.
- Assadi-Porter, F.M. et al., 2010. Interactions between the human sweet-sensing T1R2-T1R3 receptor and sweeteners detected by saturation transfer difference NMR spectroscopy. *Biochimica et biophysica acta*, 1798(2), pp.82–86.
- Ausländer, D. et al., 2014. A Synthetic Multifunctional Mammalian pH Sensor and CO₂ Transgene-Control Device. *Molecular cell*, 55(3), pp.397–408.
- Ausländer, S. & Fussenegger, M., 2013. From gene switches to mammalian designer cells: present and future prospects. *Trends in Biotechnology*, 31(3), pp.155–168.
- Ausländer, S., Wieland, M. & Fussenegger, M., 2012. Smart medication through combination of synthetic biology and cell microencapsulation. *Metabolic engineering*, 14(3), pp.252–260.
- Bacchus, W., Aubel, D. & Fussenegger, M., 2013. Biomedically relevant circuit-design strategies in mammalian synthetic biology. *Molecular Systems Biology*, 9(1), p.691.
- Bae, K.-H. et al., 2003. Human zinc fingers as building blocks in the construction of artificial transcription factors. *Nature biotechnology*, 21(3), pp.275–280.
- Baeumler, T.A., Ahmed, A.A. & Fulga, T.A., 2017. Engineering Synthetic Signaling Pathways with Programmable dCas9-Based Chimeric Receptors. *Cell Reports*, 20(11), pp.2639–2653.
- Bai, P. et al., 2016. A synthetic biology-based device prevents liver injury in mice. *Journal of hepatology*, 65(1), pp.84–94.

- Balboa, D. et al., 2015. Stem Cell Reports. *Stem Cell Reports*, 5(3), pp.448–459.
- Baldauf, H.-M. et al., 2012. SAMHD1 restricts HIV-1 infection in resting CD4(+) T cells. *Nature medicine*, 18(11), pp.1682–1687.
- Balkwill, F., 2009. Tumour necrosis factor and cancer. *Nature Reviews Cancer*, 9(5), pp.361–371.
- Balkwill, F.R., 2012. The chemokine system and cancer. *The Journal of pathology*, 226(2), pp.148–157.
- Banaszynski, L.A., Liu, C.W. & Wandless, T.J., 2005. Characterization of the FKBP-rapamycin-FRB ternary complex. *Journal of the American Chemical Society*, 127(13), pp.4715–4721.
- Bao, Z. et al., 2017. Orthogonal Genetic Regulation in Human Cells Using Chemically Induced CRISPR/Cas9 Activators. *ACS Synthetic Biology*, 6(4), pp.686–693.
- Barnea, G. et al., 2008. The genetic design of signaling cascades to record receptor activation. *Proceedings of the National Academy of Sciences*, 105(1), pp.64–69.
- Barrangou, R., 2015. Science Direct. The roles of CRISPR. *Current Opinion in Immunology*, 32, pp.36–41.
- Barrangou, R. & Doudna, J.A., 2016. Applications of CRISPR technologies in research and beyond. *Nature biotechnology*, 34(9), pp.933–941.
- Barrangou, R. & Horvath, P., 2017. A decade of discovery: CRISPR functions and applications. *Nature Microbiology*, 2(7), p.17092.
- Barrangou, R. et al., 2007. CRISPR Provides Acquired Resistance Against Viruses in Prokaryotes. *Science*, 315(5819), pp.1709–1712.
- Barré-Sinoussi, F., Ross, A.L. & Delfraissy, J.-F., 2013. Past, present and future: 30 years of HIV research. *Nature Reviews Microbiology*, 11(12), pp.877–883.
- Barton, W.A. et al., 2006. Crystal structures of the Tie2 receptor ectodomain and the angiopoietin-2-Tie2 complex. *Nature structural & molecular biology*, 13(6), pp.524–532.
- Bayle, J.H. et al., 2006. Rapamycin analogs with differential binding specificity permit orthogonal control of protein activity. *Chemistry & Biology*, 13(1), pp.99–107.
- Beerli, R.R. & Barbas, C.F., 2002. Engineering polydactyl zinc-finger transcription factors. *Nature biotechnology*, 20(2), pp.135–141.
- Beerli, R.R. et al., 1998. Toward controlling gene expression at will: Specific regulation of the erbB-2/HER-2 promoter by using polydactyl zinc finger proteins constructed from modular building blocks. *Proceedings of the National Academy of Sciences*, 95(25), pp.14628–14633.
- Beerli, R.R., Dreier, B. & Barbas, C.F., 2000. Positive and negative regulation of endogenous genes by designed transcription factors. *Proceedings of the National Academy of Sciences*, 97(4), pp.1495–1500.
- Beerli, R.R., Schopfer, U., et al., 2000. Chemically regulated zinc finger transcription factors. *The Journal of biological chemistry*, 275(42), pp.32617–32627.
- Belshaw, P.J. et al., 1996. Controlling protein association and subcellular localization with a synthetic ligand that induces heterodimerization of proteins. *Proceedings of the National Academy of Sciences*, 93(10), pp.4604–4607.
- Beltran, A. et al., 2007. Re-activation of a dormant tumor suppressor gene maspin by designed transcription factors. *Oncogene*, 26(19), pp.2791–2798.
- Berg, J.M., 1988. Proposed structure for the zinc-binding domains from transcription factor IIIA and related proteins. *Proceedings of the National Academy of Sciences*, 85(1), pp.99–102.
- Bhattacharyya, R.P. et al., 2006. Domains, motifs, and scaffolds: The role of modular interactions in the evolution and wiring of cell signaling circuits. *Annual review of biochemistry*, 75, pp.655–680.
- Bhaya, D., Davison, M. & Barrangou, R., 2011. CRISPR-Cas Systems in Bacteria and Archaea: Versatile Small RNAs for Adaptive Defense and Regulation. *Annual Review of Genetics*, 45(1), pp.273–297.
- Black, J.B. et al., 2016. Targeted Epigenetic Remodeling of Endogenous Loci by CRISPR/Cas9-Based Transcriptional Activators Directly Converts Fibroblasts to Neuronal Cells. *Stem Cell*, pp.1–10.
- Bleul, C.C. et al., 1996. The lymphocyte chemoattractant SDF-1 is a ligand for LESTR/fusin and blocks HIV-1 entry. *Nature*, 382(6594), pp.829–833.
- Blume-Jensen, P. & Hunter, T., 2001. Oncogenic kinase signalling. *Nature*, 411(6835), pp.355–365.
- Boch, J. & Bonas, U., 2010. Xanthomonas AvrBs3 family-type III effectors: discovery and function. *Annual review of phytopathology*, 48, pp.419–436.
- Boch, J. et al., 2009. Breaking the code of DNA binding specificity of TAL-type III effectors. *Science*, 326(5959), pp.1509–1512.
- Body-Malapel, M. et al., 2008. NOD2: a potential target for regulating liver injury. *Laboratory investigation; a journal of technical methods and pathology*, 88(3), pp.318–327.
- Bogerd, H.P. et al., 2015. Specific induction of endogenous viral restriction factors using CRISPR/Cas-derived transcriptional activators. *Proceedings of the National Academy of Sciences of the United States of America*, 112(52), pp.E7249–E7256.
- Bonas, U., STALL, R.E. & STASKAWICZ, B., 1989. Genetic and structural characterization of the avirulence gene *avrBs3* from *Xanthomonas campestris* pv. *vesicatoria*. *Molecular & general genetics: MGG*, 218(1), pp.127–136.

- Bondy-Denomy, J., 2017. Protein inhibitors of CRISPR-Cas9. *ACS Chemical Biology*.
- Bordignon, C. et al., 1995. Transfer of the HSV-tk gene into donor peripheral blood lymphocytes for in vivo modulation of donor anti-tumor immunity after allogeneic bone marrow transplantation. *Human gene therapy*, 6(6), pp.813–819.
- Boucharaba, A. et al., 2004. Platelet-derived lysophosphatidic acid supports the progression of osteolytic bone metastases in breast cancer. *Journal of Clinical Investigation*, 114(12), pp.1714–1725.
- Brasemann, S., Graninger, P. & Busslinger, M., 1993. A selective transcriptional induction system for mammalian cells based on Gal4-estrogen receptor fusion proteins. *Proceedings of the National Academy of Sciences*, 90(5), pp.1657–1661.
- Bräuner-Osborne, H., Wellendorph, P. & Jensen, A.A., 2007. Structure, pharmacology and therapeutic prospects of family C G-protein coupled receptors. *Current Drug Targets*, 8(1), pp.169–184.
- Brenner, M.J. et al., 2018. Synthetic Biology: Immunotherapy by Design. *Annual Review of Biomedical Engineering*.
- Brent, R. & Ptashne, M., 1985. A eukaryotic transcriptional activator bearing the DNA specificity of a prokaryotic repressor. *Cell*, 43(3 Pt 2), pp.729–736.
- Brentjens, R.J. et al., 2013. CD19-targeted T cells rapidly induce molecular remissions in adults with chemotherapy-refractory acute lymphoblastic leukemia. *Science translational medicine*, 5(177), pp.177ra38–177ra38.
- Briggs, A.W. et al., 2012. Iterative capped assembly: rapid and scalable synthesis of repeat-module DNA such as TAL effectors from individual monomers. *Nucleic acids research*, 40(15), p.e117.
- Broadus, M.R. et al., 2016. Identification of a Paralog-Specific Notch1 Intracellular Domain Degron. *Cell Reports*, 15(9), pp.1920–1929.
- Brown, R.S., Sander, C. & Argos, P., 1985. The primary structure of transcription factor TFIIIA has 12 consecutive repeats. *FEBS Letters*, 186(2), pp.271–274.
- Bulenger, S., Marullo, S. & Bouvier, M., 2005. Emerging role of homo- and heterodimerization in G-protein-coupled receptor biosynthesis and maturation. *Trends in Pharmacological Sciences*, 26(3), pp.131–137.
- Burton, E.R. & Libutti, S.K., 2009. Targeting TNF- α for cancer therapy. *Journal of biology*, 8(9), p.85.
- Calebiro, D. et al., 2009. Persistent cAMP-signals triggered by internalized G-protein-coupled receptors. *PLoS Biology*, 7(8), p.e1000172.
- Calebiro, D., Nikolaev, V.O. & Lohse, M.J., 2010. Imaging of persistent cAMP signaling by internalized G protein-coupled receptors. *Journal of molecular endocrinology*, 45(1), pp.1–8.
- Carey, M. et al., 1990. A mechanism for synergistic activation of a mammalian gene by GAL4 derivatives. *Nature*, 345(6273), pp.361–364.
- Carroll, D. et al., 2006. Design, construction and in vitro testing of zinc finger nucleases. *Nature Protocols*, 1(3), pp.1329–1341.
- Castellino, F. et al., 2006. Chemokines enhance immunity by guiding naive CD8⁺ T cells to sites of CD4⁺ T cell-dendritic cell interaction. *Nature*, 440(7086), pp.890–895.
- Caumo, A. & Luzi, L., 2004. First-phase insulin secretion: does it exist in real life? Considerations on shape and function. *The American journal of physiology*, 287(3), pp.E371–85.
- Cesaratto, F., Burrone, O.R. & Petris, G., 2016. Tobacco Etch Virus protease: A shortcut across biotechnologies. *Journal of Biotechnology*, 231, pp.239–249.
- Chakraborty, A.K. & Weiss, A., 2014. Insights into the initiation of TCR signaling. *Nature Immunology*, 15(9), pp.798–807.
- Chandrashekar, J. et al., 2006. The receptors and cells for mammalian taste. *Nature*, 444(7117), pp.288–294.
- Chang, W.C. et al., 2007. Modifying ligand-induced and constitutive signaling of the human 5-HT₄ receptor. *PLoS ONE*, 2(12), p.e1317.
- Chang, Z.L. & Chen, Y.Y., 2017. CARs: Synthetic Immunoreceptors for Cancer Therapy and Beyond. *Trends in Molecular Medicine*, 23(5), pp.430–450.
- Changeux, J.-P. & Christopoulos, A., 2016. Allosteric Modulation as a Unifying Mechanism for Receptor Function and Regulation. *Cell*, 166(5), pp.1084–1102.
- Charlesworth, C.T. et al., 2018. Identification of Pre-Existing Adaptive Immunity to Cas9 Proteins in Humans.
- Chavez, A. et al., 2016. Comparison of Cas9 activators in multiple species. *Nature methods*, 13(7), pp.563–567.
- Chavez, A. et al., 2015. Highly efficient Cas9-mediated transcriptional programming. *Nature methods*, 12(4), pp.326–328.
- Chen, B. et al., 2013. Dynamic Imaging of Genomic Loci in Living Human Cells by an Optimized CRISPR/Cas System. *Cell*, 155(7), pp.1479–1491.
- Chen, S., Yu, X. & Guo, D., 2018. CRISPR-Cas Targeting of Host Genes as an Antiviral Strategy. *Viruses*, 10(1).
- Chen, T. et al., 2017. Chemically Controlled Epigenome Editing through an Inducible dCas9 System.

- Journal of the American Chemical Society*, 139(33), pp.11337–11340.
- Chen, T.T. et al., 2010. Anchorage of VEGF to the extracellular matrix conveys differential signaling responses to endothelial cells. *The Journal of Cell Biology*, 188(4), pp.595–609.
- Chen, Xi, Pham, E. & Truong, K., 2010. TEV protease-facilitated stoichiometric delivery of multiple genes using a single expression vector. *Protein Science*, 19(12), pp.2379–2388.
- Chen, Xin et al., 2015. The first structure-activity relationship studies for designer receptors exclusively activated by designer drugs. *ACS chemical neuroscience*, 6(3), pp.476–484.
- Cheng, A.W. et al., 2016. Casilio: A versatile CRISPR-Cas9-Pumilio hybrid for gene regulation and genomic labeling. *Cell Research*, 26(2), pp.254–257.
- Cheng, A.W. et al., 2013. Multiplexed activation of endogenous genes by CRISPR-on, an RNA-guided transcriptional activator system. *Cell Research*, 23(10), pp.1163–1171.
- Chew, W.L. et al., 2016. A multifunctional AAV-CRISPR-Cas9 and its host response. *Nature methods*, 13(10), pp.868–874.
- Chiang, M.Y. et al., 2006. Identification of a conserved negative regulatory sequence that influences the leukemogenic activity of NOTCH1. *Molecular and Cellular Biology*, 26(16), pp.6261–6271.
- Chiu, S.-C. & Yang, N.-S., 2007. Inhibition of tumor necrosis factor- α through selective blockade of Pre-mRNA splicing by shikonin. *Molecular Pharmacology*, 71(6), pp.1640–1645.
- Chmielewski, M. & Abken, H., 2017. CAR T Cells Releasing IL-18 Convert to T-Bet^{high}FoxO1^{low} Effectors that Exhibit Augmented Activity against Advanced Solid Tumors. *Cell Reports*, 21(11), pp.3205–3219.
- Chmielewski, M. et al., 2011. IL-12 release by engineered T cells expressing chimeric antigen receptors can effectively Muster an antigen-independent macrophage response on tumor cells that have shut down tumor antigen expression. *Cancer Research*, 71(17), pp.5697–5706.
- Cho, S.W. et al., 2013. Targeted genome engineering in human cells with the Cas9 RNA-guided endonuclease. *Nature biotechnology*, 31(3), pp.230–232.
- Choe, H. et al., 1996. The beta-chemokine receptors CCR3 and CCR5 facilitate infection by primary HIV-1 isolates. *Cell*, 85(7), pp.1135–1148.
- Choi, J. et al., 1996. Structure of the FKBP12-rapamycin complex interacting with the binding domain of human FRAP. *Science*, 273(5272), pp.239–242.
- Christinger, H.W. et al., 2004. The crystal structure of placental growth factor in complex with domain 2 of vascular endothelial growth factor receptor-1. *The Journal of biological chemistry*, 279(11), pp.10382–10388.
- Ciceri, F. et al., 2009. 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(5), pp.489–500.
- Claeysen, S. et al., 2003. A single mutation in the 5-HT₄ receptor (5-HT₄-R D100(3.32)A) generates a Gs-coupled receptor activated exclusively by synthetic ligands (RASSL). *The Journal of biological chemistry*, 278(2), pp.699–702.
- Claing, A. et al., 2002. Endocytosis of G protein-coupled receptors: roles of G protein-coupled receptor kinases and beta-arrestin proteins. *Progress in neurobiology*, 66(2), pp.61–79.
- Clayton, A.H.A. et al., 2005. Ligand-induced dimer-tetramer transition during the activation of the cell surface epidermal growth factor receptor-A multidimensional microscopy analysis. *The Journal of biological chemistry*, 280(34), pp.30392–30399.
- Cleland, S.J., 2012. Cardiovascular risk in double diabetes mellitus--when two worlds collide. *Nature Publishing Group*, 8(8), pp.476–485.
- Cong, L. et al., 2013. Multiplex Genome Engineering Using CRISPR/Cas Systems. *Science*, 339(6121), pp.819–823.
- Conklin, B.R. et al., 2008. Engineering GPCR signaling pathways with RASSLs. *Nature*, 5(8), pp.673–678.
- Connor, R.I. et al., 1997. Change in coreceptor use correlates with disease progression in HIV-1--infected individuals. *Journal of Experimental Medicine*, 185(4), pp.621–628.
- Copeland, M.F. et al., 2016. A transcription activator-like effector (TALE) induction system mediated by proteolysis. *Nature Publishing Group*, 12(4), pp.254–260.
- Cornell, S., 2015. Continual evolution of type 2 diabetes: an update on pathophysiology and emerging treatment options. *Therapeutics and clinical risk management*, 11, pp.621–632.
- Coward, P. et al., 1998. Controlling signaling with a specifically designed Gi-coupled receptor. *Proceedings of the National Academy of Sciences*, 95(1), pp.352–357.
- Cox, D.B.T., Platt, R.J. & Zhang, F., 2015. Therapeutic genome editing: prospects and challenges. *Nature medicine*, 21(2), pp.121–131.
- Craddock, J.A. et al., 2010. Enhanced tumor trafficking of GD2 chimeric antigen receptor T cells by expression of the chemokine receptor CCR2b. *Journal of Immunotherapy*, 33(8), pp.780–788.
- Crooke, S.T. et al., 2017. Cellular uptake and trafficking of antisense oligonucleotides. *Nature biotechnology*, 35(3), pp.230–237.

- D'Arco, M. & Dolphin, A.C., 2012. L-type calcium channels: on the fast track to nuclear signaling. *Science signaling*, 5(237), pp.pe34–pe34.
- da Costa, P.L.N. et al., 2014. The role of kinin receptors in cancer and therapeutic opportunities. *Cancer Letters*, 345(1), pp.27–38.
- Dalgleish, A.G. et al., 1984. The CD4 (T4) antigen is an essential component of the receptor for the AIDS retrovirus. *Nature*, 312(5996), pp.763–767.
- Dalziel, M. et al., 2014. Emerging principles for the therapeutic exploitation of glycosylation. *Science*, 343(6166), p.1235681.
- Daringer, N.M. et al., 2014. Modular Extracellular Sensor Architecture for Engineering Mammalian Cell-based Devices. *ACS Synthetic Biology*, 3(12), pp.892–902.
- Davila, M.L. et al., 2014. Efficacy and Toxicity Management of 19-28z CAR T Cell Therapy in B Cell Acute Lymphoblastic Leukemia. *Science translational medicine*, 6(224), pp.–224ra25.
- Davis, K.M. et al., 2015. Small molecule-triggered Cas9 protein with improved genome-editing specificity. *Nature Chemical Biology*, 11(5), pp.316–318.
- Delort, J.P. & Capecchi, M.R., 1996. TAXI/UAS: A molecular switch to control expression of genes in vivo. *Human gene therapy*.
- Deltcheva, E. et al., 2011. CRISPR RNA maturation by trans-encoded small RNA and host factor RNase III. *Nature*, 471(7340), pp.602–607.
- Deng, H. et al., 1996. Identification of a major co-receptor for primary isolates of HIV-1. *Nature*, 381(6584), pp.661–666.
- Deng, Y., Zhang, X. & Simons, M., 2015. Molecular controls of lymphatic VEGFR3 signaling. *Arteriosclerosis, thrombosis, and vascular biology*, 35(2), pp.421–429.
- Dent, C.L. et al., 2007. Regulation of endogenous gene expression using small molecule-controlled engineered zinc-finger protein transcription factors. *Gene Therapy*, 14(18), pp.1362–1369.
- Di Stasi, A. et al., 2011. Inducible Apoptosis as a Safety Switch for Adoptive Cell Therapy. *New England Journal of Medicine*, 365(18), pp.1673–1683.
- Di Stasi, A. et al., 2009. T lymphocytes coexpressing CCR4 and a chimeric antigen receptor targeting CD30 have improved homing and antitumor activity in a Hodgkin tumor model. *Blood*, 113(25), pp.6392–6402.
- Dixelius, J. et al., 2003. Ligand-induced vascular endothelial growth factor receptor-3 (VEGFR-3) heterodimerization with VEGFR-2 in primary lymphatic endothelial cells regulates tyrosine phosphorylation sites. *The Journal of biological chemistry*, 278(42), pp.40973–40979.
- Djannatian, M.S. et al., 2011. Analytical Biochemistry. *Analytical Biochemistry*, 412(2), pp.141–152.
- Dolgin, E., 2017. Epic \$12 billion deal and FDA's approval raise CAR-T to new heights. *Nature biotechnology*, 35(10), pp.891–892.
- Dominguez, A.A., Lim, W.A. & Qi, L.S., 2016. Beyond editing: repurposing CRISPR-Cas9 for precision genome regulation and interrogation. *Nature Reviews Molecular Cell Biology*, 17(1), pp.5–15.
- Dong, D. et al., 2017. Structural basis of CRISPR-SpyCas9 inhibition by an anti-CRISPR protein. *Nature*, 546(7658), pp.436–439.
- Dong, S., Rogan, S.C. & Roth, B.L., 2010. Directed molecular evolution of DREADDs: a generic approach to creating next-generation RASLS. *Nature Protocols*, 5(3), pp.561–573.
- Doranz, B.J. et al., 1996. A dual-tropic primary HIV-1 isolate that uses fusin and the beta-chemokine receptors CKR-5, CKR-3, and CKR-2b as fusion cofactors. *Cell*, 85(7), pp.1149–1158.
- Dorsam, R.T. & Gutkind, J.S., 2007. G-protein-coupled receptors and cancer. *Nature Reviews Cancer*, 7(2), pp.79–94.
- Dosch, D.D. & Ballmer-Hofer, K., 2010. Transmembrane domain-mediated orientation of receptor monomers in active VEGFR-2 dimers. *The FASEB Journal*, 24(1), pp.32–38.
- Dougherty, W.G. & Dawn Parks, T., 1989. Molecular genetic and biochemical evidence for the involvement of the heptapeptide cleavage sequence in determining the reaction profile at two tobacco etch virus cleavage sites in cell-free assays. *Virology*, 172(1), pp.145–155.
- Dougherty, W.G. et al., 1988. Biochemical and mutational analysis of a plant virus polyprotein cleavage site. *The EMBO journal*, 7(5), pp.1281–1287.
- Dragic, T. et al., 1996. HIV-1 entry into CD4+ cells is mediated by the chemokine receptor CC-CKR-5. *Nature*, 381(6584), pp.667–673.
- Dranoff, G., 2004. Cytokines in cancer pathogenesis and cancer therapy. *Nature Reviews Cancer*, 4(1), pp.11–22.
- Drury, J.L. & Mooney, D.J., 2003. Hydrogels for tissue engineering: scaffold design variables and applications. *Biomaterials*, 24(24), pp.4337–4351.
- Du, D. et al., 2017. Genetic interaction mapping in mammalian cells using CRISPR interference. *Nature methods*, 14(6), pp.577–580.
- Egawa, T. et al., 2001. The earliest stages of B cell development require a chemokine stromal cell-derived factor/pre-B cell growth-stimulating factor. *Immunity*, 15(2), pp.323–334.
- Elowitz, M. & Lim, W.A., 2010. Build life to understand it. *Nature*, 468(7326), pp.889–890.

- Elrod-Erickson, M. et al., 1996. Zif268 protein-DNA complex refined at 1.6 Å: a model system for understanding zinc finger-DNA interactions. *Structure (London, England: 1993)*, 4(10), pp.1171–1180.
- Emami, K.H. & Carey, M., 1992. A synergistic increase in potency of a multimerized VP16 transcriptional activation domain. *The EMBO journal*, 11(13), pp.5005–5012.
- Farzadfard, F., Perli, S.D. & Lu, T.K., 2013. Tunable and multifunctional eukaryotic transcription factors based on CRISPR/Cas. *ACS Synthetic Biology*, 2(10), pp.604–613.
- Feinstein, T.N. et al., 2013. Noncanonical control of vasopressin receptor type 2 signaling by retromer and arrestin. *The Journal of biological chemistry*, 288(39), pp.27849–27860.
- Feng, X. et al., 2017. Computational design of ligand-binding membrane receptors with high selectivity. *Nature Chemical Biology*, 426, p.145.
- Feng, Y. et al., 1996. HIV-1 entry cofactor: functional cDNA cloning of a seven-transmembrane, G protein-coupled receptor. *Science*, 272(5263), pp.872–877.
- Ferber, S. et al., 2000. Pancreatic and duodenal homeobox gene 1 induces expression of insulin genes in liver and ameliorates streptozotocin-induced hyperglycemia. *Nature medicine*, 6(5), pp.568–572.
- Ferguson, S., 2001. Evolving concepts in G protein-coupled receptor endocytosis: The role in receptor desensitization and signaling. *Pharmacological Reviews*, 53(1), pp.1–24.
- Ferguson, S. et al., 1996. Role of beta-arrestin in mediating agonist-promoted G protein-coupled receptor internalization. *Science*, 271(5247), pp.363–366.
- Ferrandon, S. et al., 2009. Sustained cyclic AMP production by parathyroid hormone receptor endocytosis. *Nature Chemical Biology*, 5(10), pp.734–742.
- Ferrara, N., 2010. Binding to the extracellular matrix and proteolytic processing: two key mechanisms regulating vascular endothelial growth factor action. *Molecular Biology of the Cell*, 21(5), pp.687–690.
- Ferrara, N. & Adamis, A.P., 2016. Ten years of anti-vascular endothelial growth factor therapy. *Nature Reviews Drug Discovery*, 15(6), pp.385–403.
- Ferré, S. et al., 2014. G protein-coupled receptor oligomerization revisited: functional and pharmacological perspectives. *Pharmacological Reviews*, 66(2), pp.413–434.
- Ferry, Q.R.V., Lyutova, R. & Fulga, T.A., 2017. Rational design of inducible CRISPR guide RNAs for de novo assembly of transcriptional programs. *Nature Communications*, 8, p.14633.
- Fesnak, A.D., June, C.H. & Levine, B.L., 2016. Engineered T cells: the promise and challenges of cancer immunotherapy. *Nature Reviews Cancer*, 16(9), pp.566–581.
- Fischbach, M.A., Bluestone, J.A. & Lim, W.A., 2013. Cell-based therapeutics: the next pillar of medicine. *Science translational medicine*, 5(179), pp.179ps7–179ps7.
- Folcher, M. & Fussenegger, M., 2012. Synthetic biology advancing clinical applications. *Current Opinion in Chemical Biology*, 16(3-4), pp.345–354.
- Fonfara, I. et al., 2016. The CRISPR-associated DNA-cleaving enzyme Cpf1 also processes precursor CRISPR RNA. *Nature*, 532(7600), pp.517–521.
- Foster, D.J. & Conn, P.J., 2017. Allosteric Modulation of GPCRs: New Insights and Potential Utility for Treatment of Schizophrenia and Other CNS Disorders. *Neuron*, 94(3), pp.431–446.
- Franco, R. et al., 2007. Basic Concepts in G-Protein-Coupled Receptor Homo- and Heterodimerization. *The Scientific World JOURNAL*, 7, pp.48–57.
- Franco, R. et al., 2016. Basic pharmacological and structural evidence for class A G-protein-coupled receptor heteromerization. *Frontiers in Pharmacology*, 7(MAR), p.76.
- Fredriksson, R. & Schiöth, H.B., 2005. The repertoire of G-protein-coupled receptors in fully sequenced genomes. *Molecular Pharmacology*, 67(5), pp.1414–1425.
- Fredriksson, R. et al., 2003. The G-protein-coupled receptors in the human genome form five main families. Phylogenetic analysis, paralogon groups, and fingerprints. *Molecular Pharmacology*, 63(6), pp.1256–1272.
- Fryer, C.J. et al., 2002. Mastermind mediates chromatin-specific transcription and turnover of the Notch enhancer complex. *Genes & Development*, 16(11), pp.1397–1411.
- Fryer, C.J., White, J.B. & Jones, K.A., 2004. Mastermind recruits CycC:CDK8 to phosphorylate the Notch ICD and coordinate activation with turnover. *Molecular cell*, 16(4), pp.509–520.
- Fuh, G. et al., 1998. Requirements for binding and signaling of the kinase domain receptor for vascular endothelial growth factor. *The Journal of biological chemistry*, 273(18), pp.11197–11204.
- Gadella, T.W. & Jovin, T.M., 1995. Oligomerization of epidermal growth factor receptors on A431 cells studied by time-resolved fluorescence imaging microscopy. A stereochemical model for tyrosine kinase receptor activation. *The Journal of Cell Biology*, 129(6), pp.1543–1558.
- Galandrin, S., Oligny-Longpré, G. & Bouvier, M., 2007. The evasive nature of drug efficacy: implications for drug discovery. *Trends in Pharmacological Sciences*, 28(8), pp.423–430.
- Gao, H. et al., 2013. Evidence of a causal relationship between adiponectin levels and insulin sensitivity: a Mendelian randomization study. *Diabetes*, 62(4), pp.1338–1344.
- Gao, X. et al., 2014. Comparison of TALE designer transcription factors and the CRISPR/dCas9 in

- regulation of gene expression by targeting enhancers. *Nucleic acids research*, 42(20), p.e155.
- Gao, Y. et al., 2016. Complex transcriptional modulation with orthogonal and inducible dCas9 regulators. *Nature methods*, 13(12), pp.1043–1049.
- Garneau, J.E. et al., 2010. The CRISPR/Cas bacterial immune system cleaves bacteriophage and plasmid DNA. *Nature*, 468(7320), pp.67–71.
- Garrett, T.P.J. et al., 2002. Crystal structure of a truncated epidermal growth factor receptor extracellular domain bound to transforming growth factor alpha. *Cell*, 110(6), pp.763–773.
- Gasiunas, G. et al., 2012. Cas9-crRNA ribonucleoprotein complex mediates specific DNA cleavage for adaptive immunity in bacteria. *Proceedings of the National Academy of Sciences of the United States of America*, 109(39), pp.E2579–E2586.
- Geissler, R. et al., 2011. Transcriptional activators of human genes with programmable DNA-specificity. S.-H. Shiu, ed. *PLoS ONE*, 6(5), p.e19509.
- Gerace, D. et al., 2017. CRISPR-targeted genome editing of mesenchymal stem cell-derived therapies for type 1 diabetes: a path to clinical success? *Stem cell research & therapy*, 8(1), p.62.
- Gerace, D. et al., 2015. The use of β -cell transcription factors in engineering artificial β cells from non-pancreatic tissue. *Gene Therapy*, 22(1), pp.1–8.
- Gilbert, L.A. et al., 2013. CRISPR-Mediated Modular RNA-Guided Regulation of Transcription in Eukaryotes. *Cell*, 154(2), pp.442–451.
- Gilbert, L.A. et al., 2014. Genome-Scale CRISPR-Mediated Control of Gene Repression and Activation. *Cell*, 159(3), pp.647–661.
- Gill, S. & June, C.H., 2015. Going viral: chimeric antigen receptor T-cell therapy for hematological malignancies. *Immunological reviews*, 263(1), pp.68–89.
- Gilman, A.G., 1987. G proteins: transducers of receptor-generated signals. *Annual review of biochemistry*, 56, pp.615–649.
- Giménez, C.A. et al., 2016. CRISPR-on system for the activation of the endogenous human INS gene. *Gene Therapy*, 23(6), pp.543–547.
- Giniger, E., Varnum, S.M. & Ptashne, M., 1985. Specific DNA binding of GAL4, a positive regulatory protein of yeast. *Cell*, 40(4), pp.767–774.
- Gitzinger, M. et al., 2009. Controlling transgene expression in subcutaneous implants using a skin lotion containing the apple metabolite phloretin. *Proceedings of the National Academy of Sciences of the United States of America*, 106(26), pp.10638–10643.
- Gitzinger, M. et al., 2012. The food additive vanillic acid controls transgene expression in mammalian cells and mice. *Nucleic acids research*, 40(5), p.e37.
- Gloriam, D.E., Fredriksson, R. & Schiöth, H.B., 2007. The G protein-coupled receptor subset of the rat genome. *BMC genomics*, 8(1), p.338.
- Gonzalez, C.B. & Figueroa, C.D., 1999. Vasopressin and bradykinin receptors in the kidney: Implications for tubular function. *Biological Research*, 32(2-3), pp.63–76.
- Gordon, W.R. et al., 2015. Mechanical Allostery: Evidence for a Force Requirement in the Proteolytic Activation of Notch. *Developmental Cell*, 33(6), pp.729–736.
- Gosmanov, A.R., Gosmanova, E.O. & Dillard-Cannon, E., 2014. Management of adult diabetic ketoacidosis. *Diabetes, metabolic syndrome and obesity: targets and therapy*, 7, pp.255–264.
- Gossen, M. & Bujard, H., 1992. Tight control of gene expression in mammalian cells by tetracycline-responsive promoters. *Proceedings of the National Academy of Sciences*, 89(12), pp.5547–5551.
- Gossen, M. et al., 1995. Transcriptional activation by tetracyclines in mammalian cells. *Science*, 268(5218), pp.1766–1769.
- Gossen, M., Bonin, A.L. & Bujard, H., 1993. Control of Gene Activity in Higher Eukaryotic Cells by Prokaryotic Regulatory Elements. *Trends in Biochemical Sciences*, 18(12), pp.471–475.
- Gray, D.C., Mahrus, S. & Wells, J.A., 2010. Activation of specific apoptotic caspases with an engineered small-molecule-activated protease. *Cell*, 142(4), pp.637–646.
- Graybill, W. et al., 2015. State of the science: Emerging therapeutic strategies for targeting angiogenesis in ovarian cancer. *Gynecologic Oncology*, 138(2), pp.223–226.
- Green, M. et al., 1995. A heat shock-responsive domain of human HSF1 that regulates transcription activation domain function. *Molecular and Cellular Biology*, 15(6), pp.3354–3362.
- Griffioen, M. et al., 2009. Retroviral transfer of human CD20 as a suicide gene for adoptive T-cell therapy. *Haematologica*, 94(9), pp.1316–1320.
- Groskreutz, D.J., SLIWKOWSKI, M.X. & GORMAN, C.M., 1994. Genetically-Engineered Proinsulin Constitutively Processed and Secreted as Mature, Active Insulin. *The Journal of biological chemistry*, 269(8), pp.6241–6245.
- Gross, G. & Eshhar, Z., 2016. Therapeutic Potential of T Cell Chimeric Antigen Receptors (CARs) in Cancer Treatment: Counteracting Off-Tumor Toxicities for Safe CAR T Cell Therapy. *Annual Review of Pharmacology and Toxicology*, 56, pp.59–83.
- Gross, G., Waks, T. & Eshhar, Z., 1989. Expression of immunoglobulin-T-cell receptor chimeric molecules as functional receptors with antibody-type specificity. *Proceedings of the National Academy of*

- Sciences*, 86(24), pp.10024–10028.
- Groth, C.G. et al., 1994. Transplantation of porcine fetal pancreas to diabetic patients. *Lancet (London, England)*, 344(8934), pp.1402–1404.
- Grundy, S.M., 2006. Drug therapy of the metabolic syndrome: minimizing the emerging crisis in polypharmacy. *Nature Reviews Drug Discovery*, 5(4), pp.295–309.
- Grusch, M. et al., 2014. Spatio-temporally precise activation of engineered receptor tyrosine kinases by light. *The EMBO journal*, 33(15), pp.1713–1726.
- Grünewald, F.S. et al., 2010. Structure-function analysis of VEGF receptor activation and the role of coreceptors in angiogenic signaling. *Biochimica et biophysica acta*, 1804(3), pp.567–580.
- Guarente, L., Yocum, R.R. & Gifford, P., 1982. A GAL10-CYC1 hybrid yeast promoter identifies the GAL4 regulatory region as an upstream site. *Proceedings of the National Academy of Sciences*, 79(23), pp.7410–7414.
- Guettier, J.-M. et al., 2009. A chemical-genetic approach to study G protein regulation of beta cell function in vivo. *Proceedings of the National Academy of Sciences of the United States of America*, 106(45), pp.19197–19202.
- Gurevich, E.V. et al., 2012. G protein-coupled receptor kinases: more than just kinases and not only for GPCRs. *Pharmacology & Therapeutics*, 133(1), pp.40–69.
- Haellman, V. & Fussenegger, M., 2017. Synthetic biology – Engineering cell-based biomedical devices. *Current Opinion in Biomedical Engineering*, 4, pp.50–56.
- Hamdan, F.F. et al., 2005. High-throughput screening of G protein-coupled receptor antagonists using a bioluminescence resonance energy transfer 1-based beta-arrestin2 recruitment assay. *Journal of Biomolecular Screening*, 10(5), pp.463–475.
- Hampsey, M., 1998. Molecular genetics of the RNA polymerase II general transcriptional machinery. *Microbiology and molecular biology reviews: MMBR*, 62(2), pp.465–503.
- Han, X. et al., 2017. Masked Chimeric Antigen Receptor for Tumor-Specific Activation. *Molecular Therapy*, 25(1), pp.274–284.
- Hanahan, D. & Weinberg, R.A., 2011. Hallmarks of Cancer: The Next Generation. *Cell*, 144(5), pp.646–674.
- Hanahan, D. & Weinberg, R.A., 2000. The hallmarks of cancer. *Cell*, 100(1), pp.57–70.
- Hanson, B.J. et al., 2009. A homogeneous fluorescent live-cell assay for measuring 7-transmembrane receptor activity and agonist functional selectivity through beta-arrestin recruitment. *Journal of Biomolecular Screening*, 14(7), pp.798–810.
- Hanyaloglu, A.C. et al., 2001. Casein Kinase II Sites in the Intracellular C-terminal Domain of the Thyrotropin-releasing Hormone Receptor and Chimeric Gonadotropin-releasing Hormone Receptors Contribute to β -Arrestin-dependent Internalization. *Journal of Biological Chemistry*, 276(21), pp.18066–18074.
- Harrington, L.B. et al., 2017. A Broad-Spectrum Inhibitor of CRISPR-Cas9. *Cell*, 170(6), pp.1224–1233.e15.
- Harris, N.C. et al., 2013. The propeptides of VEGF-D determine heparin binding, receptor heterodimerization, and effects on tumor biology. *The Journal of biological chemistry*, 288(12), pp.8176–8186.
- Hartenbach, S. et al., 2007. An engineered L-arginine sensor of Chlamydia pneumoniae enables arginine-adjustable transcription control in mammalian cells and mice. *Nucleic acids research*, 35(20), p.e136.
- Hartfield, R.M. et al., 2017. Multiplexing engineered receptors for multiparametric evaluation of environmental ligands. *ACS Synthetic Biology*.
- Hattori, M. & Ozawa, T., 2015. Bioluminescent tools for the analysis of G-protein-coupled receptor and arrestin interactions. *RSC Advances*, 5(17), pp.12655–12663.
- Haurwitz, R.E. et al., 2010. Sequence- and structure-specific RNA processing by a CRISPR endonuclease. *Science*, 329(5997), pp.1355–1358.
- Hay, C.W. & Docherty, K., 2003. Enhanced expression of a furin-cleavable proinsulin. *Journal of molecular endocrinology*, 31(3), pp.597–607.
- He, L., Huang, J. & Perrimon, N., 2017. Development of an optimized synthetic Notch receptor as an in vivo cell-cell contact sensor. *Proceedings of the National Academy of Sciences of the United States of America*, 114(21), pp.5467–5472.
- Hecht, B., Muller, G. & Hillen, W., 1993. Noninducible Tet Repressor Mutations Map From the Operator Binding Motif to the C-Terminus. *Journal of Bacteriology*, 175(4), pp.1206–1210.
- Heler, R. et al., 2015. Cas9 specifies functional viral targets during CRISPR–Cas adaptation. *Nature*, pp.1–16.
- Heng, B.C., Aubel, D. & Fussenegger, M., 2014. G protein-coupled receptors revisited: therapeutic applications inspired by synthetic biology. *Annual Review of Pharmacology and Toxicology*, 54, pp.227–249.
- Heng, B.C., Aubel, D. & Fussenegger, M., 2015. Prosthetic gene networks as an alternative to standard pharmacotherapies for metabolic disorders. *Current Opinion in Biotechnology*, 35, pp.37–45.

- Hermans, E., 2003. Biochemical and pharmacological control of the multiplicity of coupling at G-protein-coupled receptors. *Pharmacology & Therapeutics*, 99(1), pp.25–44.
- Hill, Z.B. et al., 2018. Human antibody-based chemically induced dimerizers for cell therapeutic applications. *Nature Publishing Group*, 14(2), pp.112–117.
- Hilton, I.B. et al., 2015. epigenome editing by a crisPr-cas9-based acetyltransferase activates genes from promoters and enhancers. *Nature Publishing Group*, pp.1–10.
- Hirosawa, M. et al., 2017. Cell-type-specific genome editing with a microRNA-responsive CRISPR-Cas9 switch. *Nucleic acids research*, 45(13), p.e118.
- Ho, S.N. et al., 1996. Dimeric ligands define a role for transcriptional activation domains in reinitiation. *Nature*, 382(6594), pp.822–826.
- Hopkins, A.L. & Groom, C.R., 2002. The druggable genome. *Nature Reviews Drug Discovery*, 1(9), pp.727–730.
- Horvath, P. & Barrangou, R., 2010. CRISPR/Cas, the Immune System of Bacteria and Archaea. *Science*, 327(5962), pp.167–170.
- Howard, P.L. et al., 2003. Redirecting tyrosine kinase signaling to an apoptotic caspase pathway through chimeric adaptor proteins. *Proceedings of the National Academy of Sciences*, 100(20), pp.11267–11272.
- Hoyos, V. et al., 2010. Engineering CD19-specific T lymphocytes with interleukin-15 and a suicide gene to enhance their anti-lymphoma/leukemia effects and safety. *Leukemia*, 24(6), pp.1160–1170.
- Hrecka, K. et al., 2011. Vpx relieves inhibition of HIV-1 infection of macrophages mediated by the SAMHD1 protein. *Nature*, 474(7353), pp.658–661.
- Hsiao, E.C. et al., 2008. Osteoblast expression of an engineered G(s)-coupled receptor dramatically increases bone mass. *Proceedings of the National Academy of Sciences*, 105(4), pp.1209–1214.
- Hsu, P.D., Lander, E.S. & Zhang, F., 2014. Development and applications of CRISPR-Cas9 for genome engineering. *Cell*, 157(6), pp.1262–1278.
- Hu, B. et al., 2017. Augmentation of Antitumor Immunity by Human and Mouse CAR T Cells Secreting IL-18. *CellReports*, 20(13), pp.3025–3033.
- Hu, J. et al., 2016. A G Protein-biased Designer G Protein-coupled Receptor Useful for Studying the Physiological Relevance of Gq/11-dependent Signaling Pathways. *Journal of Biological Chemistry*, 291(15), pp.7809–7820.
- Huang, T.-H., Velho, T. & Lois, C., 2016. Monitoring cell-cell contacts in vivo in transgenic animals. *Development (Cambridge, England)*, 143(21), pp.4073–4084.
- Hubbard, S.R., 2004. Juxtamembrane autoinhibition in receptor tyrosine kinases. *Nature Reviews Molecular Cell Biology*, 5(6), pp.464–471.
- Inagaki, H.K. et al., 2012. Visualizing neuromodulation in vivo: TANGO-mapping of dopamine signaling reveals appetite control of sugar sensing. *Cell*, 148(3), pp.583–595.
- Irannejad, R. et al., 2013. Conformational biosensors reveal GPCR signalling from endosomes. *Nature*, 495(7442), pp.534–538.
- Ishida, K. et al., 2018. Site-specific randomization of the endogenous genome by a regulatable CRISPR-Cas9 piggyBac system in human cells. *Scientific Reports*, 8(1), p.310.
- Ishino, Y. et al., 1987. Nucleotide sequence of the iap gene, responsible for alkaline phosphatase isozyme conversion in *Escherichia coli*, and identification of the gene product. *Journal of Bacteriology*, 169(12), pp.5429–5433.
- Ito, K. et al., 2015. Artificial human Met agonists based on macrocycle scaffolds. *Nature Communications*, 6, p.6373.
- Ivashkiv, L.B. & Donlin, L.T., 2014. Regulation of type I interferon responses. *Nature reviews. Immunology*, 14(1), pp.36–49.
- Iwamoto, M. et al., 2010. A general chemical method to regulate protein stability in the mammalian central nervous system. *Chemistry & Biology*, 17(9), pp.981–988.
- Jagadeesh, S. et al., 2014. Identifying Functional Connections of the Inner Photoreceptors in *Drosophila* using Tango-Trace. *Neuron*, 83(3), pp.630–644.
- Jain, S. et al., 2013. Chronic activation of a designer Gq-coupled receptor improves β cell function. *Journal of Clinical Investigation*, 123(4), pp.1750–1762.
- Jang, H.-J. et al., 2007. Gut-expressed gustducin and taste receptors regulate secretion of glucagon-like peptide-1. *Proceedings of the National Academy of Sciences*, 104(38), pp.15069–15074.
- Jansen, R. et al., 2002. Identification of genes that are associated with DNA repeats in prokaryotes. *Molecular Microbiology*, 43(6), pp.1565–1575.
- Jiang, P. et al., 2004. The cysteine-rich region of T1R3 determines responses to intensely sweet proteins. *The Journal of biological chemistry*, 279(43), pp.45068–45075.
- Jiang, W. et al., 2013. RNA-guided editing of bacterial genomes using CRISPR-Cas systems. *Nature biotechnology*, 31(3), pp.233–239.
- Jin, F. et al., 2012. New insight into the SDF-1/CXCR4 axis in a breast carcinoma model: hypoxia-induced endothelial SDF-1 and tumor cell CXCR4 are required for tumor cell intravasation. *Molecular Cancer*

- Research*, 10(8), pp.1021–1031.
- Jin, Z.-G. et al., 2003. Ligand-independent activation of vascular endothelial growth factor receptor 2 by fluid shear stress regulates activation of endothelial nitric oxide synthase. *Circulation Research*, 93(4), pp.354–363.
- Jinek, M. et al., 2012. A programmable dual-RNA-guided DNA endonuclease in adaptive bacterial immunity. *Science*, 337(6096), pp.816–821.
- Johnson, L.A. et al., 2015. Rational development and characterization of humanized anti-EGFR variant III chimeric antigen receptor T cells for glioblastoma. *Science translational medicine*, 7(275), pp.275ra22–275ra22.
- Johnston, M. & Davis, R.W., 1984. Sequences that regulate the divergent GAL1-GAL10 promoter in *Saccharomyces cerevisiae*. *Molecular and Cellular Biology*, 4(8), pp.1440–1448.
- Jost, M. et al., 2017. Combined CRISPRi/a-Based Chemical Genetic Screens Reveal that Rigosertib Is a Microtubule-Destabilizing Agent. *Molecular cell*, 68(1), pp.210–223.e6.
- Joung, J. et al., 2017. Genome-scale activation screen identifies a lncRNA locus regulating a gene neighbourhood. *Nature*, 548(7667), pp.343–346.
- Joyce, J.A. & Fearon, D.T., 2015. T cell exclusion, immune privilege, and the tumor microenvironment. *Science*, 348(6230), pp.74–80.
- Jusiak, B. et al., 2016. Engineering Synthetic Gene Circuits in Living Cells with CRISPR Technology. *Trends in Biotechnology*, 34(7), pp.535–547.
- Kabadi, A.M. et al., 2014. Multiplex CRISPR/Cas9-based genome engineering from a single lentiviral vector. *Nucleic acids research*, 42(19), p.e147.
- Kakidani, H. & Ptashne, M., 1988. GAL4 activates gene expression in mammalian cells. *Cell*, 52(2), pp.161–167.
- Kalderon, D. et al., 1984. A Short Amino-Acid Sequence Able to Specify Nuclear Location. *Cell*, 39(3), pp.499–509.
- Kalos, M. & June, C.H., 2013. Adoptive T cell transfer for cancer immunotherapy in the era of synthetic biology. *Immunity*, 39(1), pp.49–60.
- Kalos, M. et al., 2011. T Cells with Chimeric Antigen Receptors Have Potent Antitumor Effects and Can Establish Memory in Patients with Advanced Leukemia. *Science translational medicine*, 3(95), pp.–95ra73.
- Kang, Y. et al., 2015. Crystal structure of rhodopsin bound to arrestin by femtosecond X-ray laser. *Nature*, 523(7562), pp.561–567.
- Kapp, G.T. et al., 2012. Control of protein signaling using a computationally designed GTPase/GEF orthogonal pair. *Proceedings of the National Academy of Sciences of the United States of America*, 109(14), pp.5277–5282.
- Kapust, R.B. et al., 2002. The P1 ' specificity of tobacco etch virus protease. *Biochemical and biophysical research communications*, 294(5), pp.949–955.
- Kapust, R.B. et al., 2001. Tobacco etch virus protease: mechanism of autolysis and rational design of stable mutants with wild-type catalytic proficiency. *Protein Engineering*, 14(12), pp.993–1000.
- Karvelis, T. et al., 2013. crRNA and tracrRNA guide Cas9-mediated DNA interference in *Streptococcus thermophilus*. *RNA biology*, 10(5), pp.841–851.
- Katritch, V., Cherezov, V. & Stevens, R.C., 2012. Diversity and modularity of G protein-coupled receptor structures. *Trends in Pharmacological Sciences*, 33(1), pp.17–27.
- Kay, S. et al., 2007. A bacterial effector acts as a plant transcription factor and induces a cell size regulator. *Science*, 318(5850), pp.648–651.
- Kearns, N.A. et al., 2015. Functional annotation of native enhancers with a Cas9-histone demethylase fusion. *Nature methods*, 12(5), pp.401–403.
- Keegan, L., Gill, G. & Ptashne, M., 1986. Separation of DNA binding from the transcription-activating function of a eukaryotic regulatory protein. *Science*, 231(4739), pp.699–704.
- Kemmer, C. et al., 2010. Self-sufficient control of urate homeostasis in mice by a synthetic circuit. *Nature biotechnology*, 28(4), pp.355–360.
- Kennedy, M.J. et al., 2010. Rapid blue-light-mediated induction of protein interactions in living cells. *Nature methods*, 7(12), pp.973–975.
- Kershaw, M.H. et al., 2006. A phase I study on adoptive immunotherapy using gene-modified T cells for ovarian cancer. *Clinical cancer research: an official journal of the American Association for Cancer Research*, 12(20 Pt 1), pp.6106–6115.
- Kershaw, M.H. et al., 2002. Redirecting migration of T cells to chemokine secreted from tumors by genetic modification with CXCR2. *Human gene therapy*, 13(16), pp.1971–1980.
- Kershaw, M.H., Westwood, J.A. & Darcy, P.K., 2013. Gene-engineered T cells for cancer therapy. *Nature Reviews Cancer*, 13(8), pp.525–541.
- Khorkova, O. & Wahlestedt, C., 2017. Oligonucleotide therapies for disorders of the nervous system. *Nature biotechnology*, 35(3), pp.249–263.
- Khvorova, A. & Watts, J.K., 2017. The chemical evolution of oligonucleotide therapies of clinical utility.

- Nature biotechnology*, 35(3), pp.238–248.
- Kiani, S. et al., 2015. Cas9 gRNA engineering for genome editing, activation and repression. *Nature methods*, 12(11), pp.1051–1054.
- Kiani, S. et al., 2014. CRISPR transcriptional repression devices and layered circuits in mammalian cells. *Nature methods*, 11(7), pp.723–726.
- Kieback, E. et al., 2008. A safeguard eliminates T cell receptor gene-modified autoreactive T cells after adoptive transfer. *Proceedings of the National Academy of Sciences of the United States of America*, 105(2), pp.623–628.
- Kiel, C., Yus, E. & Serrano, L., 2010. Engineering signal transduction pathways. *Cell*, 140(1), pp.33–47.
- Kim, H. & Kim, J.-S., 2014. A guide to genome engineering with programmable nucleases. *Nature Reviews Genetics*, 15(5), pp.321–334.
- Kim, K.-M. & Caron, M.G., 2008. Complementary roles of the DRY motif and C-terminus tail of GPCRS for G protein coupling and β -arrestin interaction. *Biochemical and biophysical research communications*, 366(1), pp.42–47.
- Kim, Y. et al., 2013. A library of TAL effector nucleases spanning the human genome. *Nature biotechnology*, 31(3), pp.251–258.
- King, N., Hittinger, C.T. & Carroll, S.B., 2003. Evolution of key cell signaling and adhesion protein families predates animal origins. *Science*, 301(5631), pp.361–363.
- Kipniss, N.H. et al., 2017. Engineering cell sensing and responses using a GPCR-coupled CRISPR-Cas system. *Nature Communications*, 8(1), p.2212.
- Kiskinis, E. & Eggan, K., 2010. Progress toward the clinical application of patient-specific pluripotent stem cells. *Journal of Clinical Investigation*, 120(1), pp.51–59.
- Kitada, T. et al., 2018. Programming gene and engineered-cell therapies with synthetic biology. *Science*, 359(6376), p.eaad1067.
- Klatzmann, D. et al., 1984. T-lymphocyte T4 molecule behaves as the receptor for human retrovirus LAV. *Nature*, 312(5996), pp.767–768.
- Klebanoff, C.A., Rosenberg, S.A. & Restifo, N.P., 2016. Prospects for gene-engineered T cell immunotherapy for solid cancers. *Nature medicine*, 22(1), pp.26–36.
- Kleinjan, D.A. et al., 2017. Drug-tunable multidimensional synthetic gene control using inducible degron-tagged dCas9 effectors. *Nature Communications*, 8(1), p.1191.
- Kloss, C.C. et al., 2013. Combinatorial antigen recognition with balanced signaling promotes selective tumor eradication by engineered T cells. *Nature biotechnology*, 31(1), pp.71–75.
- Klug, A., 2010. The discovery of zinc fingers and their applications in gene regulation and genome manipulation. *Annual review of biochemistry*, 79, pp.213–231.
- Knight, S.C. et al., 2015. Dynamics of CRISPR-Cas9 genome interrogation in living cells. *Science*, 350(6262), pp.823–826.
- Kobilka, B.K., 2007. G protein coupled receptor structure and activation. *BBA - Biomembranes*, 1768(4), pp.794–807.
- Kobilka, B.K. & Deupi, X., 2007. Conformational complexity of G-protein-coupled receptors. *Trends in Pharmacological Sciences*, 28(8), pp.397–406.
- Kochenderfer, J.N. & Rosenberg, S.A., 2013. Treating B-cell cancer with T cells expressing anti-CD19 chimeric antigen receptors. *Nature Publishing Group*, 10(5), pp.267–276.
- Kochenderfer, J.N. et al., 2012. B-cell depletion and remissions of malignancy along with cytokine-associated toxicity in a clinical trial of anti-CD19 chimeric-antigen-receptor-transduced T cells. *Blood*, 119(12), pp.2709–2720.
- Kojima, R., Scheller, L. & Fussenegger, M., 2018. Nonimmune cells equipped with T-cell-receptor-like signaling for cancer cell ablation. *Nature Publishing Group*, 14(1), pp.42–49.
- Kolodka, T.M. et al., 1995. Gene-Therapy for Diabetes-Mellitus in Rats by Hepatic Expression of Insulin. *Proceedings of the National Academy of Sciences*, 92(8), pp.3293–3297.
- Konermann, S. et al., 2015. Genome-scale transcriptional activation by an engineered CRISPR-Cas9 complex. *Nature*, 517(7536), pp.583–588.
- Konermann, S. et al., 2013. Optical control of mammalian endogenous transcription and epigenetic states. *Nature*, 500(7463), pp.472–476.
- Kono, M. et al., 2017. Bioluminescence imaging of G protein-coupled receptor activation in living mice. *Nature Communications*, 8(1), p.1163.
- Kopan, R. & Ilagan, M.X.G., 2009. The canonical Notch signaling pathway: unfolding the activation mechanism. *Cell*, 137(2), pp.216–233.
- Kotowski, S.J. et al., 2011. Endocytosis promotes rapid dopaminergic signaling. *Neuron*, 71(2), pp.278–290.
- Kroeze, W.K. et al., 2015. PRESTO-Tango as an open-source resource for interrogation of the druggable human GPCRome. *Nature structural & molecular biology*, 22(5), pp.362–369.
- Kroeze, W.K., Sheffler, D.J. & Roth, B.L., 2003. G-protein-coupled receptors at a glance. *Journal of Cell Science*, 116(Pt 24), pp.4867–4869.

- Kuscu, C. et al., 2014. Genome-wide analysis reveals characteristics of off-target sites bound by the Cas9 endonuclease. *Nature biotechnology*, 32(7), pp.677–683.
- Kuwana, Y. et al., 1987. Expression of Chimeric Receptor Composed of Immunoglobulin-Derived v-Resions and T-Cell Receptor-Derived C-Regions. *Biochemical and biophysical research communications*, 149(3), pp.960–968.
- Labrie, S.J., Samson, J.E. & Moineau, S., 2010. Bacteriophage resistance mechanisms. *Nature Reviews Microbiology*, 8(5), pp.317–327.
- Lackmann, M. et al., 1998. Distinct subdomains of the EphA3 receptor mediate ligand binding and receptor dimerization. *The Journal of biological chemistry*, 273(32), pp.20228–20237.
- Laffitte, A., Neiers, F. & Briand, L., 2014. Functional roles of the sweet taste receptor in oral and extraoral tissues. *Current opinion in clinical nutrition and metabolic care*, 17(4), pp.379–385.
- Lagerström, M.C. & Schiöth, H.B., 2008. Structural diversity of G protein-coupled receptors and significance for drug discovery. *Nature Reviews Drug Discovery*, 7(4), pp.339–357.
- Laguet, N. et al., 2011. SAMHD1 is the dendritic- and myeloid-cell-specific HIV-1 restriction factor counteracted by Vpx. *Nature*, 474(7353), pp.654–657.
- Lamers, C.H.J. et al., 2011. Immune responses to transgene and retroviral vector in patients treated with ex vivo-engineered T cells. *Blood*, 117(1), pp.72–82.
- Lander, E.S., 2016. The Heroes of CRISPR. *Cell*, 164(1-2), pp.18–28.
- Lanitis, E. et al., 2013. Chimeric antigen receptor T Cells with dissociated signaling domains exhibit focused antitumor activity with reduced potential for toxicity in vivo. *Cancer immunology research*, 1(1), pp.43–53.
- Lanitis, E. et al., 2012. Redirected antitumor activity of primary human lymphocytes transduced with a fully human anti-mesothelin chimeric receptor. *Molecular Therapy*, 20(3), pp.633–643.
- Lazennec, G. & Richmond, A., 2010. Chemokines and chemokine receptors: new insights into cancer-related inflammation. *Trends in Molecular Medicine*, 16(3), pp.133–144.
- Leclerc, G.M. et al., 2000. Development of a destabilized firefly luciferase enzyme for measurement of gene expression. *BioTechniques*, 29(3), pp.590–.
- Ledford, H., 2015. CRISPR, the disruptor. *Nature*, 522(7554), pp.20–24.
- Lee, D. et al., 2017. Temporally precise labeling and control of neuromodulatory circuits in the mammalian brain. *Nature methods*, 22, p.362.
- Lee, M.S. et al., 1989. Three-dimensional solution structure of a single zinc finger DNA-binding domain. *Science*, 245(4918), pp.635–637.
- Lee, Z. et al., 2006. Lysophosphatidic acid is a major regulator of growth-regulated oncogene alpha in ovarian cancer. *Cancer Research*, 66(5), pp.2740–2748.
- Lemmon, M.A. & Schlessinger, J., 2010. Cell Signaling by Receptor Tyrosine Kinases. *Cell*, 141(7), pp.1117–1134.
- Levkaya, A. et al., 2009. Spatiotemporal control of cell signalling using a light-switchable protein interaction. *Nature*, 461(7266), pp.997–1001.
- Levy, A. et al., 2015. CRISPR adaptation biases explain preference for acquisition of foreign DNA. *Nature*, 520(7548), pp.505–510.
- Li, X. et al., 1998. Generation of destabilized green fluorescent protein as a transcription reporter. *The Journal of biological chemistry*, 273(52), pp.34970–34975.
- Li, Y. et al., 2012. Transcription activator-like effector hybrids for conditional control and rewiring of chromosomal transgene expression. *Scientific Reports*, 2, p.897.
- Li, Y.S. et al., 1996. Identification of the earliest B lineage stage in mouse bone marrow. *Immunity*, 5(6), pp.527–535.
- Liang, F.-S., Ho, W.Q. & Crabtree, G.R., 2011. Engineering the ABA plant stress pathway for regulation of induced proximity. *Science signaling*, 4(164), pp.rs2–rs2.
- Liao, H.-K. et al., 2017. In Vivo Target Gene Activation via CRISPR/Cas9-Mediated Trans-epigenetic Modulation. *Cell*, 171(7), pp.1495–1507.e15.
- Liao, H.-K. et al., 2015. Use of the CRISPR/Cas9 system as an intracellular defense against HIV-1 infection in human cells. *Nature Communications*, 6, p.6413.
- Liao, W., Lin, J.-X. & Leonard, W.J., 2013. Interleukin-2 at the Crossroads of Effector Responses, Tolerance, and Immunotherapy. *Immunity*, 38(1), pp.13–25.
- Liberles, S.D. et al., 1997. Inducible gene expression and protein translocation using nontoxic ligands identified by a mammalian three-hybrid screen. *Proceedings of the National Academy of Sciences*, 94(15), pp.7825–7830.
- Lienert, F. et al., 2014. Synthetic biology in mammalian cells: next generation research tools and therapeutics. *Nature Reviews Molecular Cell Biology*, 15(2), pp.95–107.
- Liles, W.C. et al., 2003. Mobilization of hematopoietic progenitor cells in healthy volunteers by AMD3100, a CXCR4 antagonist. *Blood*, 102(8), pp.2728–2730.
- Lim, W., Mayer, B. & Pawson, T., 2014. *Cell signaling: principles and mechanisms*, Routledge.
- Lim, W.A., 2010. Designing customized cell signalling circuits. *Nature Reviews Molecular Cell Biology*,

- 11(6), pp.393–403.
- Lim, W.A. & June, C.H., 2017. The Principles of Engineering Immune Cells to Treat Cancer. *Cell*, 168(4), pp.724–740.
- Lim, Y. et al., 2016. Structural roles of guide RNAs in the nuclease activity of Cas9 endonuclease. *Nature Communications*, 7, p.13350.
- Lin, C. et al., 2013. Nkx2-8 downregulation promotes angiogenesis and activates NF- κ B in esophageal cancer. *Cancer Research*, 73(12), pp.3638–3648.
- Lin, Y.-S. et al., 1990. How different eukaryotic transcriptional activators can cooperate promiscuously. *Nature*, 345(6273), pp.359–361.
- Liu, H. et al., 2007. Structural basis for stem cell factor-KIT signaling and activation of class III receptor tyrosine kinases. *The EMBO journal*, 26(3), pp.891–901.
- Liu, K.I. et al., 2016. A chemical-inducible CRISPR-Cas9 system for rapid control of genome editing. *Nature Publishing Group*, 12(11), pp.980–987.
- Liu, M. et al., 2014. Proinsulin entry and transit through the endoplasmic reticulum in pancreatic beta cells. *Vitamins and hormones*, 95, pp.35–62.
- Liu, P. et al., 2018. CRISPR-Based Chromatin Remodeling of the Endogenous Oct4 or Sox2 Locus Enables Reprogramming to Pluripotency. *Cell stem cell*, 22(2), pp.252–261.e4.
- Liu, P.Q. et al., 2001. Regulation of an endogenous locus using a panel of designed zinc finger proteins targeted to accessible chromatin regions. Activation of vascular endothelial growth factor A. *The Journal of biological chemistry*, 276(14), pp.11323–11334.
- Liu, Q. et al., 1997. Design of polydactyl zinc-finger proteins for unique addressing within complex genomes. *Proceedings of the National Academy of Sciences*, 94(11), pp.5525–5530.
- Liu, S.J. et al., 2017. CRISPRi-based genome-scale identification of functional long noncoding RNA loci in human cells. *Science*, 355(6320).
- Liu, X.S. et al., 2016. Editing DNA Methylation in the Mammalian Genome. *Cell*, 167(1), pp.233–247.e17.
- Liu, Y. et al., 2016. Directing cellular information flow via CRISPR signal conductors. *Nature methods*, 13(11), pp.938–944.
- Liu, Y. et al., 2017. Engineering cell signaling using tunable CRISPR-Cpf1-based transcription factors. *Nature Communications*, 8(1), p.2095.
- Lo, A. & Qi, L., 2017. Genetic and epigenetic control of gene expression by CRISPR-Cas systems. *F1000Research*, 6.
- Löffler, S. et al., 2012. Comment on "Impaired respiratory and body temperature control upon acute serotonergic neuron inhibition". *Science*, 337(6095), pp.646–author reply 646.
- Lu, J. et al., 2017. Multimode drug inducible CRISPR/Cas9 devices for transcriptional activation and genome editing. *Nucleic acids research*.
- Lu, T.K., Khalil, A.S. & Collins, J.J., 2009. Next-generation synthetic gene networks. *Nature biotechnology*, 27(12), pp.1139–1150.
- Lu, Y. et al., 2008. Targeting artificial transcription factors to the utrophin A promoter: effects on dystrophic pathology and muscle function. *The Journal of biological chemistry*, 283(50), pp.34720–34727.
- Lupo-Stanghellini, M.-T. et al., 2010. Clinical impact of suicide gene therapy in allogeneic hematopoietic stem cell transplantation. *Human gene therapy*, 21(3), pp.241–250.
- Ma, D., Peng, S. & Xie, Z., 2016. Integration and exchange of split dCas9 domains for transcriptional controls in mammalian cells. *Nature Communications*, 7, p.13056.
- Ma, J. & Ptashne, M., 1987. Deletion analysis of GAL4 defines two transcriptional activating segments. *Cell*, 48(5), pp.847–853.
- MacDonald, P.E., Joseph, J.W. & Rorsman, P., 2005. Glucose-sensing mechanisms in pancreatic β -cells. *Philosophical Transactions of the Royal Society B: Biological Sciences*, 360(1464), pp.2211–2225.
- Maddon, P.J. et al., 1986. The T4 gene encodes the AIDS virus receptor and is expressed in the immune system and the brain. *Cell*, 47(3), pp.333–348.
- Maeder, M.L. et al., 2013. CRISPR RNA-guided activation of endogenous human genes. *Nature methods*, 10(10), pp.977–979.
- Magenat, L., Schwimmer, L.J. & Barbas, C.F., 2008. Drug-inducible and simultaneous regulation of endogenous genes by single-chain nuclear receptor-based zinc-finger transcription factor gene switches. *Gene Therapy*, 15(17), pp.1223–1232.
- Mahler, S.V. et al., 2014. Designer receptors show role for ventral pallidum input to ventral tegmental area in cocaine seeking. *Nature Neuroscience*, 17(4), pp.577–585.
- Maji, B. et al., 2017. Multidimensional chemical control of CRISPR-Cas9. *Nature Publishing Group*, 13(1), pp.9–11.
- Makarova, K.S. et al., 2015. An updated evolutionary classification of CRISPR–Cas systems. *Nature Reviews Microbiology*, 13(11), pp.722–736.
- Makkerh, J., Dingwall, C. & Laskey, R.A., 1996. Comparative mutagenesis of nuclear localization signals reveals the importance of neutral and acidic amino acids. *Current biology: CB*, 6(8), pp.1025–1027.
- Mali, P., Aach, J., et al., 2013. CAS9 transcriptional activators for target specificity screening and paired

- nickases for cooperative genome engineering. *Nature biotechnology*, 31(9), pp.833–838.
- Mali, P., Esvelt, K.M. & Church, G.M., 2013. Cas9 as a versatile tool for engineering biology. *Nature methods*, 10(10), pp.957–963.
- Mali, P., Yang, L., et al., 2013. RNA-guided human genome engineering via Cas9. *Science*, 339(6121), pp.823–826.
- Mandegar, M.A. et al., 2016. CRISPR Interference Efficiently Induces Specific and Reversible Gene Silencing in Human iPSCs. *Cell stem cell*, 18(4), pp.541–553.
- Marchese, A. et al., 2003. The ins and outs of G protein-coupled receptor trafficking. *Trends in Biochemical Sciences*, 28(7), pp.369–376.
- Margolin, J.F. et al., 1994. Krüppel-associated boxes are potent transcriptional repression domains. *Proceedings of the National Academy of Sciences*, 91(10), pp.4509–4513.
- Margolskee, R.F. et al., 2007. T1R3 and gustducin in gut sense sugars to regulate expression of Na⁺-glucose cotransporter 1. *Proceedings of the National Academy of Sciences*, 104(38), pp.15075–15080.
- Marini, I. et al., 2017. Antitumor Activity of a Mesenchymal Stem Cell Line Stably Secreting a Tumor-Targeted TNF-Related Apoptosis-Inducing Ligand Fusion Protein. *Frontiers in Immunology*, 8, p.536.
- Markowska, A.I., Jefferies, K.C. & Panjwani, N., 2011. Galectin-3 protein modulates cell surface expression and activation of vascular endothelial growth factor receptor 2 in human endothelial cells. *The Journal of biological chemistry*, 286(34), pp.29913–29921.
- Marktel, S. et al., 2003. Immunologic potential of donor lymphocytes expressing a suicide gene for early immune reconstitution after hematopoietic T-cell-depleted stem cell transplantation. *Blood*, 101(4), pp.1290–1298.
- Maruyama, I.N., 2015. Activation of transmembrane cell-surface receptors via a common mechanism? The "rotation model". *BioEssays*, 37(9), pp.959–967.
- Maruyama, I.N., 2014. Mechanisms of activation of receptor tyrosine kinases: monomers or dimers. *Cells*, 3(2), pp.304–330.
- Mason, C. et al., 2011. Cell therapy industry: billion dollar global business with unlimited potential. *Regenerative medicine*, 6(3), pp.265–272.
- Masoumi Moghaddam, S. et al., 2012. Significance of vascular endothelial growth factor in growth and peritoneal dissemination of ovarian cancer. *Cancer metastasis reviews*, 31(1-2), pp.143–162.
- Masubuchi, Y. et al., 2013. A Novel Regulatory Function of Sweet Taste-Sensing Receptor in Adipogenic Differentiation of 3T3-L1 Cells X.-M. Shi, ed. *PLoS ONE*, 8(1), p.e54500.
- Mathieu, C., Gillard, P. & Benhalima, K., 2017. Insulin analogues in type 1 diabetes mellitus: Getting better all the time. *Nature Reviews Endocrinology*, 13(7), pp.385–399.
- Matuskova, M. et al., 2015. Combined enzyme/prodrug treatment by genetically engineered AT-MSC exerts synergy and inhibits growth of MDA-MB-231 induced lung metastases. *Journal of experimental & clinical cancer research: CR*, 34(1), p.33.
- Mauceri, H.J. et al., 2002. Tumor production of angiostatin is enhanced after exposure to TNF-alpha. *International journal of cancer. Journal international du cancer*, 97(4), pp.410–415.
- Maude, S.L. et al., 2014. Chimeric antigen receptor T cells for sustained remissions in leukemia. *The New England journal of medicine*, 371(16), pp.1507–1517.
- Maus, M.V. et al., 2014. Antibody-modified T cells: CARs take the front seat for hematologic malignancies. *Blood*, 123(17), pp.2625–2635.
- May, L.T. et al., 2007. Allosteric modulation of G protein-coupled receptors. *Annual Review of Pharmacology and Toxicology*, 47, pp.1–51.
- McColl, B.K. et al., 2003. Plasmin activates the lymphangiogenic growth factors VEGF-C and VEGF-D. *Journal of Experimental Medicine*, 198(6), pp.863–868.
- McLaughlin, S.K., McKinnon, P.J. & Margolskee, R.F., 1992. Gustducin Is a Taste-Cell-Specific G-Protein Closely Related to the Transducins. *Nature*, 357(6379), pp.563–569.
- Mercer, A.C. et al., 2014. Regulation of endogenous human gene expression by ligand-inducible TALE transcription factors. *ACS Synthetic Biology*, 3(10), pp.723–730.
- Metz, M.Z. et al., 2013. Neural stem cell-mediated delivery of irinotecan-activating carboxylesterases to glioma: implications for clinical use. *Stem Cells Translational Medicine*, 2(12), pp.983–992.
- Miller, J., McLachlan, A.D. & Klug, A., 1985. Repetitive zinc-binding domains in the protein transcription factor IIIA from *Xenopus* oocytes. *The EMBO journal*, 4(6), pp.1609–1614.
- Miller, J.C. et al., 2011. A TALE nuclease architecture for efficient genome editing. *Nature biotechnology*, 29(2), pp.143–148.
- Miller, S.D., Turley, D.M. & Podojil, J.R., 2007. Antigen-specific tolerance strategies for the prevention and treatment of autoimmune disease. *Nature reviews. Immunology*, 7(9), pp.665–677.
- Milligan, G., 2004. G protein-coupled receptor dimerization: function and ligand pharmacology. *Molecular Pharmacology*, 66(1), pp.1–7.
- Milligan, G. & Smith, N.J., 2007. Allosteric modulation of heterodimeric G-protein-coupled receptors. *Trends in Pharmacological Sciences*, 28(12), pp.615–620.

- Mills, G.B. & Moolenaar, W.H., 2003. The emerging role of lysophosphatidic acid in cancer. *Nature Reviews Cancer*, 3(8), pp.582–591.
- Mitanchez, D. et al., 1998. Regulated expression of mature human insulin in the liver of transgenic mice. *FEBS Letters*, 421(3), pp.285–289.
- Mitsunobu, H. et al., 2017. Beyond Native Cas9: Manipulating Genomic Information and Function. *Trends in Biotechnology*, 35(10), pp.983–996.
- Mittler, G. et al., 2003. A novel docking site on Mediator is critical for activation by VP16 in mammalian cells. *The EMBO journal*, 22(24), pp.6494–6504.
- Miyamoto, T. et al., 2012. Rapid and orthogonal logic gating with a gibberellin-induced dimerization system. *Nature Chemical Biology*, 8(5), pp.465–470.
- Mocellin, S. et al., 2003. Use of quantitative real-time PCR to determine immune cell density and cytokine gene profile in the tumor microenvironment. *Journal of immunological methods*, 280(1-2), pp.1–11.
- Mojica, F.J. et al., 1995. Long stretches of short tandem repeats are present in the largest replicons of the Archaea *Haloferax mediterranei* and *Haloferax volcanii* and could be involved in replicon partitioning. *Molecular Microbiology*, 17(1), pp.85–93.
- Mojica, F.J.M. et al., 2009. Short motif sequences determine the targets of the prokaryotic CRISPR defence system. *Microbiology*, 155(3), pp.733–740.
- Monk, B.J., Minion, L.E. & Coleman, R.L., 2016. Anti-angiogenic agents in ovarian cancer: past, present, and future. *Annals of oncology: official journal of the European Society for Medical Oncology*, 27 Suppl 1, pp.i33–i39.
- Moon, E.K. et al., 2011. Expression of a functional CCR2 receptor enhances tumor localization and tumor eradication by retargeted human T cells expressing a mesothelin-specific chimeric antibody receptor. *Clinical cancer research: an official journal of the American Association for Cancer Research*, 17(14), pp.4719–4730.
- Moore, C.A.C., Milano, S.K. & Benovic, J.L., 2007. Regulation of Receptor Trafficking by GRKs and Arrestins. *Annual Review of Physiology*, 69(1), pp.451–482.
- Moore, P.A., Ruben, S.M. & ROSEN, C.A., 1993. Conservation of transcriptional activation functions of the NF-kappa B p50 and p65 subunits in mammalian cells and *Saccharomyces cerevisiae*. *Molecular and Cellular Biology*, 13(3), pp.1666–1674.
- Moradian Tehrani, R. et al., 2017. Mesenchymal stem cells: A new platform for targeting suicide genes in cancer. *Journal of cellular physiology*, 180(10), p.1175.
- Morbiter, R. et al., 2011. Assembly of custom TALE-type DNA binding domains by modular cloning. *Nucleic acids research*, 39(13), pp.5790–5799.
- Moretti, J. & Brou, C., 2013. Ubiquitinations in the notch signaling pathway. *International Journal of Molecular Sciences*, 14(3), pp.6359–6381.
- Morgan, R.A. et al., 2012. Recognition of glioma stem cells by genetically modified T cells targeting EGFRvIII and development of adoptive cell therapy for glioma. *Human gene therapy*, 23(10), pp.1043–1053.
- Morita, S. et al., 2016. Targeted DNA demethylation in vivo using dCas9-peptide repeat and scFv-TET1 catalytic domain fusions. *Nature biotechnology*, 34(10), pp.1060–1065.
- Morsut, L. et al., 2016. Engineering Customized Cell Sensing and Response Behaviors Using Synthetic Notch Receptors. *Cell*, 164(4), pp.780–791.
- Moscou, M.J. & Bogdanove, A.J., 2009. A simple cipher governs DNA recognition by TAL effectors. *Science*, 326(5959), pp.1501–1501.
- Motz, G.T. & Coukos, G., 2011. The parallel lives of angiogenesis and immunosuppression: cancer and other tales. *Nature reviews. Immunology*, 11(10), pp.702–711.
- Muller, Y.A. et al., 1997. Vascular endothelial growth factor: crystal structure and functional mapping of the kinase domain receptor binding site. *Proceedings of the National Academy of Sciences*, 94(14), pp.7192–7197.
- Mullershausen, F. et al., 2009. Persistent signaling induced by FTY720-phosphate is mediated by internalized S1P1 receptors. *Nature Chemical Biology*, 5(6), pp.428–434.
- Nakajima, K.-I. & Wess, Juergen, 2012. Design and Functional Characterization of a Novel, Arrestin-Biased Designer G Protein-Coupled Receptor. *Molecular Pharmacology*, 82(4), pp.575–582.
- NCD Risk Factor Collaboration (NCD-RisC), 2016. Worldwide trends in diabetes since 1980: a pooled analysis of 751 population-based studies with 4.4 million participants. *Lancet (London, England)*, 387(10027), pp.1513–1530.
- Neelapu, S.S. et al., 2017. Axicabtagene Ciloleucel CAR T-Cell Therapy in Refractory Large B-Cell Lymphoma. *The New England journal of medicine*.
- Neil, S.J.D., Zang, T. & Bieniasz, P.D., 2008. Tetherin inhibits retrovirus release and is antagonized by HIV-1 Vpu. *Nature*, 451(7177), pp.425–430.
- Nelson, G. et al., 2001. Mammalian sweet taste receptors. *Cell*, 106(3), pp.381–390.
- Newick, K., Moon, E. & Albelda, S.M., 2016. Chimeric antigen receptor T-cell therapy for solid tumors. *Molecular therapy oncolytics*, 3, p.16006.

- Nguyen, D.P. et al., 2016. Ligand-binding domains of nuclear receptors facilitate tight control of split CRISPR activity. *Nature Communications*, 7, p.12009.
- Nie, Y. et al., 2005. Distinct Contributions of T1R2 and T1R3 Taste Receptor Subunits to the Detection of Sweet Stimuli. *Current Biology*, 15(21), pp.1948–1952.
- Nihongaki, Y. et al., 2017. CRISPR-Cas9-based photoactivatable transcription systems to induce neuronal differentiation. *Nature methods*, 14(10), pp.963–966.
- Nihongaki, Y., Kawano, F., et al., 2015. Photoactivatable CRISPR-Cas9 for optogenetic genome editing. *Nature biotechnology*, 33(7), pp.755–760.
- Nihongaki, Y., Otabe, T. & Sato, M., 2018. Emerging Approaches for Spatiotemporal Control of Targeted Genome with Inducible CRISPR-Cas9. *Analytical Chemistry*, 90(1), pp.429–439.
- Nihongaki, Y., Yamamoto, S., et al., 2015. CRISPR-Cas9-based photoactivatable transcription system. *Chemistry & Biology*, 22(2), pp.169–174.
- Nilsson, I. et al., 2010. VEGF receptor 2/3 heterodimers detected in situ by proximity ligation on angiogenic sprouts. *The EMBO journal*, 29(8), pp.1377–1388.
- Nishi, M. & Nanjo, K., 2011. Insulin gene mutations and diabetes. *Journal of diabetes investigation*, 2(2), pp.92–100.
- Nissim, L. et al., 2014. Multiplexed and Programmable Regulation of Gene Networks with an Integrated RNA and CRISPR/Cas Toolkit in Human Cells. *Molecular cell*, 54(4), pp.698–710.
- Niu, D. et al., 2017. Inactivation of porcine endogenous retrovirus in pigs using CRISPR-Cas9. *Science*, 357(6357), pp.1303–1307.
- Nolen, B., Taylor, S. & Ghosh, G., 2004. Regulation of protein kinases; controlling activity through activation segment conformation. *Molecular cell*, 15(5), pp.661–675.
- Nouri, F.S., Wang, X. & Hatefi, A., 2015. Genetically engineered theranostic mesenchymal stem cells for the evaluation of the anticancer efficacy of enzyme/prodrug systems. *Journal of controlled release: official journal of the Controlled Release Society*, 200, pp.179–187.
- Nuñez, J.K., Harrington, L.B., et al., 2015. Foreign DNA capture during CRISPR–Cas adaptive immunity. *Nature*, pp.1–13.
- Nuñez, J.K., Lee, A.S.Y., et al., 2015. Integrase-mediated spacer acquisition during CRISPR–Cas adaptive immunity. *Nature*, pp.1–17.
- O'Geen, H. et al., 2015. A genome-wide analysis of Cas9 binding specificity using ChIP-seq and targeted sequence capture. *Nucleic acids research*, 43(6), pp.3389–3404.
- O'Neil, J. et al., 2007. FBW7 mutations in leukemic cells mediate NOTCH pathway activation and resistance to gamma-secretase inhibitors. *Journal of Experimental Medicine*, 204(8), pp.1813–1824.
- Oakes, B.L. et al., 2016. Profiling of engineering hotspots identifies an allosteric CRISPR-Cas9 switch. *Nature Publishing Group*, 34(6), pp.646–651.
- Oakley, R.H. et al., 1999. Association of beta-arrestin with G protein-coupled receptors during clathrin-mediated endocytosis dictates the profile of receptor resensitization. *The Journal of biological chemistry*, 274(45), pp.32248–32257.
- Oakley, R.H. et al., 2000. Differential affinities of visual arrestin, beta arrestin1, and beta arrestin2 for G protein-coupled receptors delineate two major classes of receptors. *The Journal of biological chemistry*, 275(22), pp.17201–17210.
- Oakley, R.H. et al., 2001. Molecular Determinants Underlying the Formation of Stable Intracellular G Protein-coupled Receptor-β-Arrestin Complexes after Receptor Endocytosis. *Journal of Biological Chemistry*, 276(22), pp.19452–19460.
- Ogiso, H. et al., 2002. Crystal structure of the complex of human epidermal growth factor and receptor extracellular domains. *Cell*, 110(6), pp.775–787.
- Olsson, A.-K. et al., 2006. VEGF receptor signalling - in control of vascular function. *Nature Reviews Molecular Cell Biology*, 7(5), pp.359–371.
- Ossovskaia, V. et al., 2008. FAK nuclear export signal sequences. *FEBS Letters*, 582(16), pp.2402–2406.
- Ottow, M.K. et al., 2014. The helminth *Trichuris suis* suppresses TLR4-induced inflammatory responses in human macrophages. *Genes and immunity*, 15(7), pp.477–486.
- Pagliuca, F.W. et al., 2014. Generation of functional human pancreatic β cells in vitro. *Cell*, 159(2), pp.428–439.
- Palella, F.J. et al., 1998. Declining morbidity and mortality among patients with advanced human immunodeficiency virus infection. HIV Outpatient Study Investigators. *New England Journal of Medicine*, 338(13), pp.853–860.
- Papworth, M. et al., 2003. Inhibition of herpes simplex virus 1 gene expression by designer zinc-finger transcription factors. *Proceedings of the National Academy of Sciences*, 100(4), pp.1621–1626.
- Parker, B.S., Rautela, J. & Hertzog, P.J., 2016. Antitumour actions of interferons: Implications for cancer therapy. *Nature Reviews Cancer*, 16(3), pp.131–144.
- Parks, T.D. et al., 1995. Expression and Purification of a Recombinant Tobacco Etch Virus Nla Proteinase - Biochemical Analyses of the Full-Length and a Naturally-Occurring Truncated Proteinase Form.

- Virology*, 210(1), pp.194–201.
- Parks, T.D. et al., 1994. Release of proteins and peptides from fusion proteins using a recombinant plant virus proteinase. *Analytical Biochemistry*, 216(2), pp.413–417.
- Passananti, C. et al., 2010. Transgenic mice expressing an artificial zinc finger regulator targeting an endogenous gene. *Methods in molecular biology (Clifton, N.J.)*, 649, pp.183–206.
- Pavletich, N.P. & Pabo, C.O., 1991. Zinc finger-DNA recognition: crystal structure of a Zif268-DNA complex at 2.1 Å. *Science*, 252(5007), pp.809–817.
- Pawluk, A. et al., 2016. Naturally Occurring Off-Switches for CRISPR-Cas9. *Cell*, 167(7), pp.1829–1838.e9.
- Pawluk, A., Davidson, A.R. & Maxwell, K.L., 2018. Anti-CRISPR: discovery, mechanism and function. *Nature Reviews Microbiology*, 16(1), pp.12–17.
- Pawson, T., 2004. Specificity in signal transduction: from phosphotyrosine-SH2 domain interactions to complex cellular systems. *Cell*, 116(2), pp.191–203.
- Pawson, T. & Nash, P., 2003. Assembly of cell regulatory systems through protein interaction domains. *Science*, 300(5618), pp.445–452.
- Pearson, R.M. et al., 2017. In vivo reprogramming of immune cells: Technologies for induction of antigen-specific tolerance. *Advanced Drug Delivery Reviews*, 114, pp.240–255.
- Pei, Y. et al., 2008. Engineered GPCRs as tools to modulate signal transduction. *Physiology*, 23(6), pp.313–321.
- Pennisi, E., 2013. The CRISPR craze. *Science*, 341(6148), pp.833–836.
- Perez-Pinera, P. et al., 2013. RNA-guided gene activation by CRISPR-Cas9-based transcription factors. *Nature methods*, 10(10), pp.973–976.
- Perkel, J.M., 2016. Xenotransplantation makes a comeback. *Nature biotechnology*, 34(1), pp.3–4.
- Perrin, L.A. et al., 2004. Increased revascularization efficacy after administration of an adenovirus encoding VEGF121. *Gene Therapy*, 11(6), pp.512–521.
- Phan, J. et al., 2002. Structural basis for the substrate specificity of tobacco etch virus protease. *The Journal of biological chemistry*, 277(52), pp.50564–50572.
- Philip, B. et al., 2014. A highly compact epitope-based marker/suicide gene for easier and safer T-cell therapy. *Blood*, 124(8), pp.1277–1287.
- Pierce, K.L., Premont, R.T. & Lefkowitz, R.J., 2002. Seven-transmembrane receptors. *Nature Reviews Molecular Cell Biology*, 3(9), pp.639–650.
- Pollock, R. & Clackson, T., 2002. Dimerizer-regulated gene expression. *Current Opinion in Biotechnology*, 13(5), pp.459–467.
- Polonsky, K.S., 2012. The past 200 years in diabetes. *The New England journal of medicine*, 367(14), pp.1332–1340.
- Polstein, L.R. & Gersbach, C.A., 2015. A light-inducible CRISPR-Cas9 system for control of endogenous gene activation. *Nature Chemical Biology*, 11(3), pp.198–200.
- Polstein, L.R. et al., 2015. Genome-wide specificity of DNA binding, gene regulation, and chromatin remodeling by TALE- and CRISPR/Cas9-based transcriptional activators. *Genome Research*, 25(8), pp.1158–1169.
- Porter, D.L. et al., 2011. Chimeric Antigen Receptor–Modified T Cells in Chronic Lymphoid Leukemia. *The New England journal of medicine*, 365(8), pp.725–733.
- Premont, R.T. & Gainetdinov, R.R., 2007. Physiological Roles of G Protein–Coupled Receptor Kinases and Arrestins. *Annual Review of Physiology*, 69(1), pp.511–534.
- Ptashne, M., 1986. Gene regulation by proteins acting nearby and at a distance. *Nature*, 322(6081), pp.697–701.
- Qi, M., 2014. Transplantation of Encapsulated Pancreatic Islets as a Treatment for Patients with Type 1 Diabetes Mellitus. *Advances in medicine*, 2014, p.429710.
- Quintarelli, C. et al., 2007. Co-expression of cytokine and suicide genes to enhance the activity and safety of tumor-specific cytotoxic T lymphocytes. *Blood*, 110(8), pp.2793–2802.
- Ramirez, C.L. et al., 2008. Unexpected failure rates for modular assembly of engineered zinc fingers. *Nature methods*, 5(5), pp.374–375.
- Ran, F.A. et al., 2013. Genome engineering using the CRISPR-Cas9 system. *Nature Protocols*, 8(11), pp.2281–2308.
- Ran, F.A. et al., 2015. In vivo genome editing using Staphylococcus aureus Cas9. *Nature*, 520(7546), pp.186–191.
- Rao, P.M., Kelly, D.M. & Jones, T.H., 2013. Testosterone and insulin resistance in the metabolic syndrome and T2DM in men. *Nature Reviews Endocrinology*, 9(8), pp.479–493.
- Rauch, B.J. et al., 2017. Inhibition of CRISPR-Cas9 with Bacteriophage Proteins. *Cell*, 168(1-2), pp.150–158.e10.
- Ray, M. et al., 2015. Quantitative Tracking of Protein Trafficking to the Nucleus Using Cytosolic Protein Delivery by Nanoparticle-Stabilized Nanocapsules. *Bioconjugate Chemistry*, 26(6), pp.1004–1007.
- Reardon, S., 2016. Welcome to the CRISPR zoo. *Nature*, 531(7593), pp.160–163.

- Rebar, E.J. et al., 2002. Induction of angiogenesis in a mouse model using engineered transcription factors. *Nature medicine*, 8(12), pp.1427–1432.
- Redchuk, T.A. et al., 2017. Near-infrared optogenetic pair for protein regulation and spectral multiplexing. *Nature Chemical Biology*, pp.1–9.
- Redfern, C.H. et al., 1999. Conditional expression and signaling of a specifically designed Gi-coupled receptor in transgenic mice. *Nature biotechnology*, 17(2), pp.165–169.
- Reiter, E. & Lefkowitz, R.J., 2006. GRKs and beta-arrestins: roles in receptor silencing, trafficking and signaling. *Trends in Endocrinology & Metabolism*, 17(4), pp.159–165.
- Restifo, N.P. et al., 1996. Loss of functional beta 2-microglobulin in metastatic melanomas from five patients receiving immunotherapy. *JNCI Journal of the National Cancer Institute*, 88(2), pp.100–108.
- Reynolds, L. et al., 2003. Repression of the HIV-1 5' LTR promoter and inhibition of HIV-1 replication by using engineered zinc-finger transcription factors. *Proceedings of the National Academy of Sciences*, 100(4), pp.1615–1620.
- Reyon, D. et al., 2012. FLASH assembly of TALENs for high-throughput genome editing. *Nature biotechnology*, 30(5), pp.460–465.
- Rezania, A. et al., 2014. Reversal of diabetes with insulin-producing cells derived in vitro from human pluripotent stem cells. *Nature biotechnology*, 32(11), pp.1121–1133.
- Richter, F. et al., 2016. Engineering of temperature- and light-switchable Cas9 variants. *Nucleic acids research*, 44(20), pp.10003–10014.
- Richter, F. et al., 2017. Switchable Cas9. *Current Opinion in Biotechnology*, 48, pp.119–126.
- Ritter, S.L. & Hall, R.A., 2009. Fine-tuning of GPCR activity by receptor-interacting proteins. *Nature Reviews Molecular Cell Biology*, 10(12), pp.819–830.
- Rivera, V.M. et al., 1996. A humanized system for pharmacologic control of gene expression. *Nature medicine*, 2(9), pp.1028–1032.
- Roberts, M.R. et al., 1994. Targeting of human immunodeficiency virus-infected cells by CD8+ T lymphocytes armed with universal T-cell receptors. *Blood*, 84(9), pp.2878–2889.
- Robinson, D.R., Wu, Y.M. & Lin, S.F., 2000. The protein tyrosine kinase family of the human genome. *Oncogene*, 19(49), pp.5548–5557.
- Roccaro, A.M. et al., 2014. SDF-1 inhibition targets the bone marrow niche for cancer therapy. *Cell Reports*, 9(1), pp.118–128.
- Roemer, K. & Friedmann, T., 1993. Modulation of cell proliferation and gene expression by a p53-estrogen receptor hybrid protein. *Proceedings of the National Academy of Sciences*, 90(20), pp.9252–9256.
- Roper, S.D., 2007. Signal transduction and information processing in mammalian taste buds. *Pflügers Archiv - European Journal of Physiology*, 454(5), pp.759–776.
- Rosenberg, S.A. & Restifo, N.P., 2015. Adoptive cell transfer as personalized immunotherapy for human cancer. *Science*, 348(6230), pp.62–68.
- Rosenbluh, J. et al., 2017. Complementary information derived from CRISPR Cas9 mediated gene deletion and suppression. *Nature Communications*, 8, p.15403.
- Rosenfeld, N. et al., 2005. Gene regulation at the single-cell level. *Science*, 307(5717), pp.1962–1965.
- Rosengart, T.K. et al., 1999. Angiogenesis gene therapy - Phase I assessment of direct intramyocardial administration of an adenovirus vector expressing VEGF121 cDNA to individuals with clinically significant severe coronary artery disease. *Circulation*, 100(5), pp.468–474.
- Roth, B.L., Sheffler, D.J. & Kroeze, W.K., 2004. Magic shotguns versus magic bullets: selectively non-selective drugs for mood disorders and schizophrenia. *Nature Reviews Drug Discovery*, 3(4), pp.353–359.
- Roybal, K.T. & Lim, W.A., 2017. Synthetic Immunology: Hacking Immune Cells to Expand Their Therapeutic Capabilities. *Annual Review of Immunology*, 35, pp.229–253.
- Roybal, K.T., Rupp, L.J., et al., 2016. Precision Tumor Recognition by T Cells With Combinatorial Antigen-Sensing Circuits. *Cell*, 164(4), pp.770–779.
- Roybal, K.T., Williams, J.Z., et al., 2016. Engineering T Cells with Customized Therapeutic Response Programs Using Synthetic Notch Receptors. *Cell*, 167(2), pp.419–432.e16.
- Römer, P. et al., 2007. Plant pathogen recognition mediated by promoter activation of the pepper Bs3 resistance gene. *Science*, 318(5850), pp.645–648.
- Rössger, K., Charpin-El Hamri, G. & Fussenegger, M., 2013a. A closed-loop synthetic gene circuit for the treatment of diet-induced obesity in mice. *Nature Communications*, 4, p.2825.
- Rössger, K., Charpin-El Hamri, G. & Fussenegger, M., 2013b. Reward-based hypertension control by a synthetic brain-dopamine interface. *Proceedings of the National Academy of Sciences of the United States of America*, 110(45), pp.18150–18155.
- Ruben, S.M. et al., 1991. Isolation of a Rel-Related Human Cdna That Potentially Encodes the 65-Kd Subunit of Nf-Kappa-B. *Science*, 251(5000), pp.1490–1493.
- Ruch, C. et al., 2007. Structure of a VEGF-VEGF receptor complex determined by electron microscopy. *Nature structural & molecular biology*, 14(3), pp.249–250.
- Saayman, S. et al., 2014. An HIV-encoded antisense long noncoding RNA epigenetically regulates viral

- transcription. *Molecular Therapy*, 22(6), pp.1164–1175.
- Saayman, S.M. et al., 2016. Potent and Targeted Activation of Latent HIV-1 Using the CRISPR/dCas9 Activator Complex. *Molecular Therapy*, 24(3), pp.488–498.
- Sack, B.K. et al., 2014. Development of Gene Transfer for Induction of Antigen-specific Tolerance. *Molecular therapy. Methods & clinical development*, 1, p.14013.
- Sadelain, M., Brentjens, R. & Rivière, I., 2013. The basic principles of chimeric antigen receptor design. *Cancer Discovery*, 3(4), pp.388–398.
- Sadelain, M., Rivière, I. & Riddell, S., 2017. Therapeutic T cell engineering. *Nature*, 545(7655), pp.423–431.
- Sander, J.D. et al., 2011. Selection-free zinc-finger-nuclease engineering by context-dependent assembly (CoDA). *Nature methods*, 8(1), pp.67–69.
- Santa-Marta, M. et al., 2013. Host factors and HIV-1 replication: clinical evidence and potential therapeutic approaches. *Frontiers in Immunology*, 4.
- Sanz, L. et al., 2012. Non-hematopoietic stem cells as factories for in vivo therapeutic protein production. *Gene Therapy*, 19(1), pp.1–7.
- Sasaki, T. et al., 2006. Structural basis for Gas6-Axl signalling. *The EMBO journal*, 25(1), pp.80–87.
- Sato, T. et al., 2007. Engineered human tmpr/AZT as a novel enzyme/prodrug axis for suicide gene therapy. *Molecular Therapy*, 15(5), pp.962–970.
- Saxena, P. et al., 2016. A programmable synthetic lineage-control network that differentiates human iPSCs into glucose-sensitive insulin-secreting beta-like cells. *Nature Communications*, 7, pp.11247–14.
- Saxena, P. et al., 2017. Generation of glucose-sensitive insulin-secreting beta-like cells from human embryonic stem cells by incorporating a synthetic lineage-control network. *Journal of Biotechnology*, 259, pp.39–45.
- Sayah, D.M. et al., 2004. Cyclophilin A retrotransposition into TRIM5 explains owl monkey resistance to HIV-1. *Nature*, 430(6999), pp.569–573.
- Schellenberger, V. et al., 2009. A recombinant polypeptide extends the in vivo half-life of peptides and proteins in a tunable manner. *Nature biotechnology*, 27(12), pp.1186–1190.
- Schiöth, H.B., Nordström, K.J.V. & Fredriksson, R., 2007. Mining the gene repertoire and ESTs for G protein-coupled receptors with evolutionary perspective. *Acta physiologica (Oxford, England)*, 190(1), pp.21–31.
- Schlessinger, J. & Lemmon, M.A., 2003. SH2 and PTB domains in tyrosine kinase signaling. *Science's STKE: signal transduction knowledge environment*, 2003(191), pp.RE12–re12.
- Schmid-Burgk, J.L. et al., 2013. A ligation-independent cloning technique for high-throughput assembly of transcription activator-like effector genes. *Nature biotechnology*, 31(1), pp.76–81.
- Schmitz, M.L. & Baeuerle, P.A., 1991. The p65 subunit is responsible for the strong transcription activating potential of NF-kappa B. *The EMBO journal*, 10(12), pp.3805–3817.
- Scholten, D.J. et al., 2012. Pharmacological modulation of chemokine receptor function. *British Journal of Pharmacology*, 165(6), pp.1617–1643.
- Schukur, L. & Fussenegger, M., 2016. Engineering of synthetic gene circuits for (re-)balancing physiological processes in chronic diseases. *Wiley interdisciplinary reviews. Systems biology and medicine*, 8(5), pp.402–422.
- Schukur, L. et al., 2015. Implantable synthetic cytokine converter cells with AND-gate logic treat experimental psoriasis. *Science translational medicine*, 7(318), pp.318ra201–318ra201.
- Schuster, S.J. et al., 2017. Chimeric Antigen Receptor T Cells in Refractory B-Cell Lymphomas. *The New England journal of medicine*, 377(26), pp.2545–2554.
- Schwarz, K.A. et al., 2017. Rewiring human cellular input-output using modular extracellular sensors. *Nature Publishing Group*, 13(2), pp.202–209.
- Schwimmer, L.J., Gonzalez, B. & Barbas, C.F., 2012. Benzoate X receptor zinc-finger gene switches for drug-inducible regulation of transcription. *Gene Therapy*, 19(4), pp.458–462.
- Scofield, M.D. et al., 2015. Gq-DREADD Selectively Initiates Glial Glutamate Release and Inhibits Cue-induced Cocaine Seeking. *Biological psychiatry*, 78(7), pp.441–451.
- Seipel, K., Georgiev, O. & Schaffner, W., 1992. Different Activation Domains Stimulate Transcription From Remote (Enhancer) and Proximal (Promoter) Positions. *The EMBO journal*, 11(13), pp.4961–4968.
- Senturk, S. et al., 2017. Rapid and tunable method to temporally control gene editing based on conditional Cas9 stabilization. *Nature Communications*, 8, p.14370.
- Sera, T., 2009. Zinc-finger-based artificial transcription factors and their applications. *Advanced Drug Delivery Reviews*, 61(7-8), pp.513–526.
- Serafini, M. et al., 2004. Characterization of CD20-transduced T lymphocytes as an alternative suicide gene therapy approach for the treatment of graft-versus-host disease. *Human gene therapy*, 15(1), pp.63–76.
- Shalem, O. et al., 2014. Genome-Scale CRISPR-Cas9 Knockout Screening in Human Cells. *Science*, 343(6166), pp.84–87.

- Shao, J. et al., 2017. Smartphone-controlled optogenetically engineered cells enable semiautomatic glucose homeostasis in diabetic mice. *Science translational medicine*, 9(387), p.ea12298.
- Shechner, D.M. et al., 2015. Multiplexable, locus-specific targeting of long RNAs with CRISPR-Display. *Nature methods*, 12(7), pp.664–670.
- Shenoy, S.K. & Lefkowitz, R.J., 2003. Multifaceted roles of beta-arrestins in the regulation of seven-membrane-spanning receptor trafficking and signalling. *The Biochemical journal*, 375(Pt 3), pp.503–515.
- Shenoy, S.K. & Lefkowitz, R.J., 2005. Seven-transmembrane receptor signaling through beta-arrestin. *Science's STKE: signal transduction knowledge environment*, 2005(308), pp.cm10–cm10.
- Shenoy, S.K. et al., 2006. beta-arrestin-dependent, G protein-independent ERK1/2 activation by the beta 2 adrenergic receptor. *The Journal of biological chemistry*, 281(2), pp.1261–1273.
- Sherman, L.S. et al., 2017. Mesenchymal stromal/stem cells in drug therapy: New perspective. *Cytotherapy*, 19(1), pp.19–27.
- Shi, Y. et al., 1991. Transcriptional repression by YY1, a human GLI-Krüppel-related protein, and relief of repression by adenovirus E1A protein. *Cell*, 67(2), pp.377–388.
- Shin, J. et al., 2017. Disabling Cas9 by an anti-CRISPR DNA mimic. *Science Advances*, 3(7), p.e1701620.
- Short, D.K. et al., 1998. Adenovirus-mediated transfer of a modified human proinsulin gene reverses hyperglycemia in diabetic mice. *The American journal of physiology*, 275(5), pp.E748–E756.
- Shrimp, J.H. et al., 2018. Chemical Control of a CRISPR-Cas9 Acetyltransferase. *ACS Chemical Biology*.
- Shukla, A.K. et al., 2014. Visualization of arrestin recruitment by a G-protein-coupled receptor. *Nature*, 512(7513), pp.218–222.
- Shulman, G.I., 2014. Ectopic fat in insulin resistance, dyslipidemia, and cardiometabolic disease. *The New England journal of medicine*, 371(12), pp.1131–1141.
- Simeonov, D.R. et al., 2017. Discovery of stimulation-responsive immune enhancers with CRISPR activation. *Nature*, 549(7670), pp.111–115.
- Simons, M., Gordon, E. & Claesson-Welsh, L., 2016. Mechanisms and regulation of endothelial VEGF receptor signalling. *Nature Reviews Molecular Cell Biology*, 17(10), pp.611–625.
- Simpson, A.M. et al., 1997. Gene therapy of diabetes: glucose-stimulated insulin secretion in a human hepatoma cell line (HEP G2ins/g). *Gene Therapy*, 4(11), pp.1202–1215.
- Singh, D. et al., 2016. Real-time observation of DNA recognition and rejection by the RNA-guided endonuclease Cas9. *Nature Communications*, 7, p.12778.
- Siuda, E.R. et al., 2015. Spatiotemporal control of opioid signaling and behavior. *Neuron*, 86(4), pp.923–935.
- Smith-Garvin, J.E., Koretzky, G.A. & Jordan, M.S., 2009. T cell activation. *Annual Review of Immunology*, 27, pp.591–619.
- Sommermeier, D. et al., 2017. Fully human CD19-specific chimeric antigen receptors for T-cell therapy. *Leukemia*, 31(10), pp.2191–2199.
- Song, D.-G. et al., 2015. A fully human chimeric antigen receptor with potent activity against cancer cells but reduced risk for off-tumor toxicity. *Oncotarget*, 6(25), pp.21533–21546.
- Soon-Shiong, P., 1999. Treatment of type I diabetes using encapsulated islets. *Advanced Drug Delivery Reviews*, 35(2-3), pp.259–270.
- Sprinzak, D. et al., 2010. Cis-interactions between Notch and Delta generate mutually exclusive signalling states. *Nature*, 465(7294), pp.86–90.
- Srivastava, S. & Riddell, S.R., 2015. Engineering CAR-T cells: Design concepts. *Trends in Immunology*, 36(8), pp.494–502.
- Stachniak, T.J., Ghosh, A. & Sternson, S.M., 2014. Chemogenetic synaptic silencing of neural circuits localizes a hypothalamus→midbrain pathway for feeding behavior. *Neuron*, 82(4), pp.797–808.
- Stanley, S.A. et al., 2012. Radio-Wave Heating of Iron Oxide Nanoparticles Can Regulate Plasma Glucose in Mice. *Science*, 336(6081), pp.604–608.
- Stanley, S.A. et al., 2014. Remote regulation of glucose homeostasis in mice using genetically encoded nanoparticles. *Nature medicine*, 21(1), pp.92–98.
- Stepper, P. et al., 2017. Efficient targeted DNA methylation with chimeric dCas9-Dnmt3a-Dnmt3L methyltransferase. *Nucleic acids research*, 45(4), pp.1703–1713.
- Sternberg, S.H. et al., 2014. DNA interrogation by the CRISPR RNA-guided endonuclease Cas9. *Nature*, 507(7490), pp.62–67.
- Stremlau, M. et al., 2004. The cytoplasmic body component TRIM5alpha restricts HIV-1 infection in Old World monkeys. *Nature*, 427(6977), pp.848–853.
- Struhl, G. & Adachi, A., 1998. Nuclear access and action of notch in vivo. *Cell*, 93(4), pp.649–660.
- Sun, M. et al., 2014. Construction and evaluation of a novel humanized HER2-specific chimeric receptor. *Breast cancer research: BCR*, 16(3), p.R61.
- Superti-Furga, G. et al., 1991. Hormone-dependent transcriptional regulation and cellular transformation by Fos-steroid receptor fusion proteins. *Proceedings of the National Academy of Sciences*, 88(12), pp.5114–5118.

- Szczepek, M. et al., 2014. Crystal structure of a common GPCR-binding interface for G protein and arrestin. *Nature Communications*, 5, p.4801.
- Tak, Y.E. et al., 2017. Inducible and multiplex gene regulation using CRISPR-Cpf1-based transcription factors. *Nature methods*, 14(12), pp.1163–1166.
- Takahashi, K. et al., 2007. Induction of pluripotent stem cells from adult human fibroblasts by defined factors. *Cell*, 131(5), pp.861–872.
- Takahashi, Y., Takebe, T. & Taniguchi, H., 2016. Engineering pancreatic tissues from stem cells towards therapy. *Regenerative Therapy*, 3, pp.15–23.
- Takakura, H. et al., 2012. Visualization and quantitative analysis of g protein-coupled receptor- β -arrestin interaction in single cells and specific organs of living mice using split luciferase complementation. *ACS Chemical Biology*, 7(5), pp.901–910.
- Takeda, S. et al., 2002. Identification of G protein-coupled receptor genes from the human genome sequence. *FEBS Letters*, 520(1-3), pp.97–101.
- Talay, M. et al., 2017. Transsynaptic Mapping of Second-Order Taste Neurons in Flies by trans-Tango. *Neuron*, 96(4), pp.783–795.e4.
- Tanenbaum, M.E. et al., 2014. A Protein-Tagging System for Signal Amplification in Gene Expression and Fluorescence Imaging. *Cell*, pp.1–12.
- Tang, W., Hu, J.H. & Liu, D.R., 2017. Aptazyme-embedded guide RNAs enable ligand-responsive genome editing and transcriptional activation. *Nature Communications*, 8, pp.1–8.
- Teixeira, A.P. & Fussenegger, M., 2017. Synthetic biology-inspired therapies for metabolic diseases. *Current Opinion in Biotechnology*, 47, pp.59–66.
- Teng, M.W.L. et al., 2015. Classifying Cancers Based on T-cell Infiltration and PD-L1. *Cancer Research*, 75(11), pp.2139–2145.
- Tey, S.-K. et al., 2007. Inducible caspase 9 suicide gene to improve the safety of allogeneic T cells after haploidentical stem cell transplantation. *Biology of blood and marrow transplantation: journal of the American Society for Blood and Marrow Transplantation*, 13(8), pp.913–924.
- Thakore, P.I. et al., 2015. Highly specific epigenome editing by CRISPR-Cas9 repressors for silencing of distal regulatory elements. *Nature methods*, 12(12), pp.1143–1149.
- Thomas, M.C. et al., 2015. Diabetic kidney disease. *Nature reviews. Disease primers*, 1, p.15018.
- Thompson, B.J. et al., 2007. The SCFFBW7 ubiquitin ligase complex as a tumor suppressor in T cell leukemia. *Journal of Experimental Medicine*, 204(8), pp.1825–1835.
- Thompson, M.D., Burnham, W.M. & Cole, D.E.C., 2005. The G protein-coupled receptors: pharmacogenetics and disease. *Critical reviews in clinical laboratory sciences*, 42(4), pp.311–392.
- Thomsen, A.R.B. et al., 2016. GPCR-G Protein- β -Arrestin Super-Complex Mediates Sustained G Protein Signaling. *Cell*, 166(4), pp.907–919.
- Tiberghien, P. et al., 2001. Administration of herpes simplex-thymidine kinase-expressing donor T cells with a T-cell-depleted allogeneic marrow graft. *Blood*, 97(1), pp.63–72.
- Tordoff, M.G. et al., 2012. T1R3: a human calcium taste receptor. *Scientific Reports*, 2, p.496.
- Tran, E., Longo, D.L. & Urba, W.J., 2017. A Milestone for CAR T Cells. *The New England journal of medicine*, 377(26), pp.2593–2596.
- Travis, J., 2015. Making the cut. *Science*, 350(6267), pp.1456–1457.
- Triezenberg, S.J., Kingsbury, R.C. & McKnight, S.L., 1988. Functional dissection of VP16, the trans-activator of herpes simplex virus immediate early gene expression. *Genes & Development*, 2(6), pp.718–729.
- Tsai, S.Q. et al., 2014. Dimeric CRISPR RNA-guided FokI nucleases for highly specific genome editing. *Nature biotechnology*, 32(6), pp.569–576.
- Tuch, B.E. et al., 2003. Function of a genetically modified human liver cell line that stores, processes and secretes insulin. *Gene Therapy*, 10(6), pp.490–503.
- Tuch, B.E. et al., 1998. Transplantation of genetically engineered insulin-producing hepatocytes into immunoincompetent mice. *Transplantation proceedings*, 30(2), p.473.
- Tuomilehto, J. & Bahijri, S., 2016. Epidemiology: Lifetime risk of diabetes mellitus--how high? *Nature Reviews Endocrinology*, 12(3), pp.127–128.
- Turner, M. et al., 2013. Toward the development of a global induced pluripotent stem cell library. *Cell stem cell*, 13(4), pp.382–384.
- Turtle, C.J. et al., 2016. CD19 CAR-T cells of defined CD4⁺:CD8⁺ composition in adult B cell ALL patients. *Journal of Clinical Investigation*, 126(6), pp.2123–2138.
- Tzima, E. et al., 2005. A mechanosensory complex that mediates the endothelial cell response to fluid shear stress. *Nature*, 437(7057), pp.426–431.
- Ueki, R. et al., 2017. Nongenetic Reprogramming of the Ligand Specificity of Growth Factor Receptors by Bispecific DNA Aptamers. *Journal of the American Chemical Society*, 139(19), pp.6554–6557.
- Ueki, R. et al., 2016. Oligonucleotide-Based Mimetics of Hepatocyte Growth Factor. *Angewandte Chemie International Edition*, 55(2), pp.579–582.
- Uhlén, M. et al., 2015. Proteomics. Tissue-based map of the human proteome. *Science*, 347(6220),

- p.1260419.
- Ullrich, A. & Schlessinger, J., 1990. Signal transduction by receptors with tyrosine kinase activity. *Cell*, 61(2), pp.203–212.
- Urban, D.J. & Roth, B.L., 2015. DREADDs (designer receptors exclusively activated by designer drugs): chemogenetic tools with therapeutic utility. *Annual Review of Pharmacology and Toxicology*, 55, pp.399–417.
- Urban, J.D. et al., 2007. Functional selectivity and classical concepts of quantitative pharmacology. *The Journal of pharmacology and experimental therapeutics*, 320(1), pp.1–13.
- Valdés-González, R.A. et al., 2005. Xenotransplantation of porcine neonatal islets of Langerhans and Sertoli cells: a 4-year study. *European journal of endocrinology*, 153(3), pp.419–427.
- Valera, A. et al., 1994. Regulated expression of human insulin in the liver of transgenic mice corrects diabetic alterations. *FASEB journal: official publication of the Federation of American Societies for Experimental Biology*, 8(6), pp.440–447.
- van Eekelen, M. et al., 2010. Human stem cells expressing novel TSP-1 variant have anti-angiogenic effect on brain tumors. *Oncogene*, 29(22), pp.3185–3195.
- Vardy, E. et al., 2015. A New DREADD Facilitates the Multiplexed Chemogenetic Interrogation of Behavior. *Neuron*, 86(4), pp.936–946.
- Vegas, A.J. et al., 2016. Long-term glycemic control using polymer-encapsulated human stem cell-derived beta cells in immune-competent mice. *Nature medicine*, 22(3), pp.306–311.
- Vegeto, E. et al., 1992. The Mechanism of Ru486 Antagonism Is Dependent on the Conformation of the Carboxy-Terminal Tail of the Human Progesterone-Receptor. *Cell*, 69(4), pp.703–713.
- Venter, J.C. et al., 2001. The sequence of the human genome. *Science*, 291(5507), pp.1304–1351.
- Vilardaga, J.-P. et al., 2002. Internalization determinants of the parathyroid hormone receptor differentially regulate beta-arrestin/receptor association. *The Journal of biological chemistry*, 277(10), pp.8121–8129.
- Vincent, A., 1986. TFIIIA and homologous genes. The “finger” proteins. *Nucleic acids research*, 14(11), pp.4385–4391.
- Viola, A. et al., 2012. The pros and cons of chemokines in tumor immunology. *Trends in Immunology*, 33(10), pp.496–504.
- Violin, J.D. & Lefkowitz, R.J., 2007. Beta-arrestin-biased ligands at seven-transmembrane receptors. *Trends in Pharmacological Sciences*, 28(8), pp.416–422.
- Vogler, I. et al., 2010. An improved bicistronic CD20/tCD34 vector for efficient purification and in vivo depletion of gene-modified T cells for adoptive immunotherapy. *Molecular Therapy*, 18(7), pp.1330–1338.
- Vojnic, E. et al., 2011. Structure and VP16 binding of the Mediator Med25 activator interaction domain. *Nature structural & molecular biology*, 18(4), pp.404–409.
- Vojta, A. et al., 2016. Repurposing the CRISPR-Cas9 system for targeted DNA methylation. *Nucleic acids research*, p.gkw159.
- Vollenweider, F. et al., 1992. Processing of proinsulin by transfected hepatoma (FAO) cells. *The Journal of biological chemistry*, 267(21), pp.14629–14636.
- Vrecl, M. et al., 2004. Development of a BRET2 screening assay using beta-arrestin 2 mutants. *Journal of Biomolecular Screening*, 9(4), pp.322–333.
- Vrecl, M. et al., 2009. β -Arrestin-based bRET2 screening assay for the “nong” β -arrestin binding CB1 receptor. *Journal of Biomolecular Screening*, 14(4), pp.371–380.
- Wang, Gang et al., 2017. CRISPR-Cas based antiviral strategies against HIV-1. *Virus Research*.
- Wang, Jiuyu et al., 2015. Structural and Mechanistic Basis of PAM-Dependent Spacer Acquisition in CRISPR-Cas Systems. *Cell*, pp.1–15.
- Wang, Tim et al., 2014. Genetic Screens in Human Cells Using the CRISPR-Cas9 System. *Science*, 343(6166), pp.80–84.
- Wang, Wenjing et al., 2017. A light- and calcium-gated transcription factor for imaging and manipulating activated neurons. *Nature biotechnology*, 35(9), pp.864–871.
- Wang, Xiuli et al., 2011. A transgene-encoded cell surface polypeptide for selection, in vivo tracking, and ablation of engineered cells. *Blood*, 118(5), pp.1255–1263.
- Wang, Xue, Chen, X. & Yang, Y., 2012. Spatiotemporal control of gene expression by a light-switchable transgene system. *Nature methods*, 9(3), pp.266–269.
- Wang, Y et al., 1994. A regulatory system for use in gene transfer. *Proceedings of the National Academy of Sciences*, 91(17), pp.8180–8184.
- Wang, Y et al., 1997. Positive and negative regulation of gene expression in eukaryotic cells with an inducible transcriptional regulator. *Gene Therapy*, 4(5), pp.432–441.
- Wang, Yuedi et al., 2017. New Chimeric Antigen Receptor Design for Solid Tumors. *Frontiers in Immunology*, 8(DEC), p.1934.
- Ward, C.W. et al., 2007. The insulin and EGF receptor structures: new insights into ligand-induced receptor activation. *Trends in Biochemical Sciences*, 32(3), pp.129–137.

- Weber, W. & Fussenegger, M., 2009. Engineering of synthetic mammalian gene networks. *Chemistry & Biology*, 16(3), pp.287–297.
- Weber, W. & Fussenegger, M., 2010. Synthetic gene networks in mammalian cells. *Current Opinion in Biotechnology*, 21(5), pp.690–696.
- Webster, N. et al., 1988. The yeast UASG is a transcriptional enhancer in human HeLa cells in the presence of the GAL4 trans-activator. *Cell*, 52(2), pp.169–178.
- Webster, N.J. et al., 1988. The hormone-binding domains of the estrogen and glucocorticoid receptors contain an inducible transcription activation function. *Cell*, 54(2), pp.199–207.
- Wehr, M.C. et al., 2006. Monitoring regulated protein-protein interactions using split TEV. *Nature methods*, 3(12), pp.985–993.
- Wei, H. et al., 2003. Independent beta-arrestin 2 and G protein-mediated pathways for angiotensin II activation of extracellular signal-regulated kinases 1 and 2. *Proceedings of the National Academy of Sciences*, 100(19), pp.10782–10787.
- Werthmann, R.C. et al., 2012. Persistent cAMP signaling by internalized TSH receptors occurs in thyroid but not in HEK293 cells. *The FASEB Journal*, 26(5), pp.2043–2048.
- Wess, Jürgen, 2016. Use of Designer G Protein-Coupled Receptors to Dissect Metabolic Pathways. *Trends in endocrinology and metabolism: TEM*, 27(9), pp.600–603.
- Wess, Jürgen, Nakajima, K. & Jain, S., 2013. Novel designer receptors to probe GPCR signaling and physiology. *Trends in Pharmacological Sciences*, 34(7), pp.385–392.
- West, R.W., Yocum, R.R. & Ptashne, M., 1984. Saccharomyces cerevisiae GAL1-GAL10 divergent promoter region: location and function of the upstream activating sequence UASG. *Molecular and Cellular Biology*, 4(11), pp.2467–2478.
- Westwood, J.A. et al., 2005. Adoptive transfer of T cells modified with a humanized chimeric receptor gene inhibits growth of Lewis-Y-expressing tumors in mice. *Proceedings of the National Academy of Sciences*, 102(52), pp.19051–19056.
- Whelan, J. & Miller, N., 1996. Generation of estrogen receptor mutants with altered ligand specificity for use in establishing a regulatable gene expression system. *The Journal of Steroid Biochemistry and Molecular Biology*, 58(1), pp.3–12.
- Wieland, M. & Fussenegger, M., 2012. Engineering molecular circuits using synthetic biology in mammalian cells. *Annual Review of Chemical and Biomolecular Engineering*, 3, pp.209–234.
- Wiesmann, C. et al., 1997. Crystal structure at 1.7 Å resolution of VEGF in complex with domain 2 of the Flt-1 receptor. *Cell*, 91(5), pp.695–704.
- Wiesmann, C. et al., 1999. Crystal structure of nerve growth factor in complex with the ligand-binding domain of the TrkA receptor. *Nature*, 401(6749), pp.184–188.
- Wilkie, S. et al., 2012. Dual targeting of ErbB2 and MUC1 in breast cancer using chimeric antigen receptors engineered to provide complementary signaling. *Journal of clinical immunology*, 32(5), pp.1059–1070.
- Williams, D.J., Puhl, H.L. & Ikeda, S.R., 2009. Rapid modification of proteins using a rapamycin-inducible tobacco etch virus protease system. *PLoS ONE*, 4(10), p.e7474.
- Woese, C.R., Kandler, O. & Wheelis, M.L., 1990. Towards a Natural System of Organisms - Proposal for the Domains Archaea, Bacteria, and Eucarya. *Proceedings of the National Academy of Sciences*, 87(12), pp.4576–4579.
- Wooten, D., Christopoulos, A. & Sexton, P.M., 2013. Emerging paradigms in GPCR allostery: implications for drug discovery. *Nature Reviews Drug Discovery*, 12(8), pp.630–644.
- Wright, D.A. et al., 2006. Standardized reagents and protocols for engineering zinc finger nucleases by modular assembly. *Nature Protocols*, 1(3), pp.1637–1652.
- Wu, C.-Y., Roybal, K.T., et al., 2015. Remote control of therapeutic T cells through a small molecule-gated chimeric receptor. *Science*, 350(6258), p.aab4077.
- Wu, C.-Y., Rupp, L.J., et al., 2015. Synthetic biology approaches to engineer T cells. *Current Opinion in Immunology*, 35, pp.123–130.
- Wu, X. et al., 2014. Genome-wide binding of the CRISPR endonuclease Cas9 in mammalian cells. *Nature biotechnology*, pp.1–9.
- Xie, M. et al., 2014. Antagonistic control of a dual-input mammalian gene switch by food additives. *Nucleic acids research*, 42(14), p.e116.
- Xie, M. et al., 2016. β -cell-mimetic designer cells provide closed-loop glycemic control. *Science*, 354(6317), pp.1296–1301.
- Xie, Y. et al., 2002. Role for 18:1 lysophosphatidic acid as an autocrine mediator in prostate cancer cells. *The Journal of biological chemistry*, 277(36), pp.32516–32526.
- Xie, Z. et al., 2011. Multi-Input RNAi-Based Logic Circuit for Identification of Specific Cancer Cells. *Science*, 333(6047), pp.1307–1311.
- Xu, G.L. & Bestor, T.H., 1997. Cytosine methylation targetted to pre-determined sequences. *Nature Genetics*, 17(4), pp.376–378.
- Xu, L. et al., 2001. A versatile framework for the design of ligand-dependent, transgene-specific

- transcription factors. *Molecular Therapy*, 3(2), pp.262–273.
- Xu, X. et al., 2016. A CRISPR-based approach for targeted DNA demethylation. *Nature Publishing Group*, pp.1–12.
- Xu, Y. et al., 2014. Optogenetic control of chemokine receptor signal and T-cell migration. *Proceedings of the National Academy of Sciences of the United States of America*, 111(17), pp.6371–6376.
- Yamada, T. et al., 2004. Lysophosphatidic acid (LPA) in malignant ascites stimulates motility of human pancreatic cancer cells through LPA1. *The Journal of biological chemistry*, 279(8), pp.6595–6605.
- Yang, F.J. et al., 2004. The activator-recruited cofactor/Mediator coactivator subunit ARC92 is a functionally important target of the VP16 transcriptional activator. *Proceedings of the National Academy of Sciences*, 101(8), pp.2339–2344.
- Yang, H. & Patel, D.J., 2017. Inhibition Mechanism of an Anti-CRISPR Suppressor AcrIIA4 Targeting SpyCas9. *Molecular cell*, 67(1), pp.117–127.e5.
- Yao, L.-C. et al., 2014. Pulmonary lymphangiectasia resulting from vascular endothelial growth factor-C overexpression during a critical period. *Circulation Research*, 114(5), pp.806–822.
- Yarranton, G.T., 1992. Inducible vectors for expression in mammalian cells. *Current Opinion in Biotechnology*, 3(5), pp.506–511.
- Yazawa, M. et al., 2009. Induction of protein-protein interactions in live cells using light. *Nature biotechnology*, 27(10), pp.941–945.
- Ye, H. et al., 2011. A synthetic optogenetic transcription device enhances blood-glucose homeostasis in mice. *Science*, 332(6037), pp.1565–1568.
- Ye, H. et al., 2017. Self-adjusting synthetic gene circuit for correcting insulin resistance. *Nature Biomedical Engineering*, 1(1), p.0005.
- Ye, H., Aubel, D. & Fussenegger, M., 2013. Synthetic mammalian gene circuits for biomedical applications. *Current Opinion in Chemical Biology*, 17(6), pp.910–917.
- Ye, H., Charpin-El Hamri, G., et al., 2013. 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(1), pp.141–146.
- Yip, A. & Webster, R.M., 2018. The market for chimeric antigen receptor T cell therapies. *Nature Publishing Group*.
- Zaidi, M.R. & Merlino, G., 2011. The two faces of interferon- γ in cancer. *Clinical cancer research: an official journal of the American Association for Cancer Research*, 17(19), pp.6118–6124.
- Zalatan, J.G. et al., 2015. Engineering complex synthetic transcriptional programs with CRISPR RNA scaffolds. *Cell*, 160(1-2), pp.339–350.
- Zaret, K.S. & Grompe, M., 2008. Generation and regeneration of cells of the liver and pancreas. *Science*, 322(5907), pp.1490–1494.
- Zetsche, B. et al., 2017. Multiplex gene editing by CRISPR-Cpf1 using a single crRNA array. *Nature biotechnology*, 35(1), pp.31–34.
- Zetsche, B., Gootenberg, J.S., et al., 2015. Cpf1 is a single RNA-guided endonuclease of a class 2 CRISPR-Cas system. *Cell*, 163(3), pp.759–771.
- Zetsche, B., Volz, S.E. & Zhang, F., 2015. A split-Cas9 architecture for inducible genome editing and transcription modulation. *Nature biotechnology*, 33(2), pp.139–142.
- Zhai, P. et al., 2005. Cardiac-specific overexpression of AT1 receptor mutant lacking G alpha q/G alpha i coupling causes hypertrophy and bradycardia in transgenic mice. *Journal of Clinical Investigation*, 115(11), pp.3045–3056.
- Zhang, E., Gu, J. & Xu, H., 2018. Prospects for chimeric antigen receptor-modified T cell therapy for solid tumors. *Molecular cancer*, 17(1), p.7.
- Zhang, F. et al., 2011. Efficient construction of sequence-specific TAL effectors for modulating mammalian transcription. *Nature biotechnology*, 29(2), pp.149–U90.
- Zhang, H. et al., 2016. New Strategies for the Treatment of Solid Tumors with CAR-T Cells. *International journal of biological sciences*, 12(6), pp.718–729.
- Zhang, L. et al., 2000. Synthetic zinc finger transcription factor action at an endogenous chromosomal site. Activation of the human erythropoietin gene. *The Journal of biological chemistry*, 275(43), pp.33850–33860.
- Zhang, Y. et al., 2003. Coding of sweet, bitter, and umami tastes: different receptor cells sharing similar signaling pathways. *Cell*, 112(3), pp.293–301.
- Zhang, Z. et al., 2016. Probe-Based Confocal Laser Endomicroscopy for Imaging TRAIL-Expressing Mesenchymal Stem Cells to Monitor Colon Xenograft Tumors In Vivo. *PLoS ONE*, 11(9), p.e0162700.
- Zhao, G.Q. et al., 2003. The receptors for mammalian sweet and umami taste. *Cell*, 115(3), pp.255–266.
- Zhao, Y. et al., 2009. A herceptin-based chimeric antigen receptor with modified signaling domains leads to enhanced survival of transduced T lymphocytes and antitumor activity. *Journal of immunology (Baltimore, Md: 1950)*, 183(9), pp.5563–5574.
- Zheng, Y. et al., 2018. CRISPR interference-based specific and efficient gene inactivation in the brain.

- Nature Neuroscience*, 77, pp.1–8.
- Zhong, G. et al., 2017. Cpf1 proteins excise CRISPR RNAs from mRNA transcripts in mammalian cells. *Nature Chemical Biology*, pp.1–5.
- Zhou, X.X. et al., 2017. A Single-Chain Photoswitchable CRISPR-Cas9 Architecture for Light-Inducible Gene Editing and Transcription. *ACS Chemical Biology*, p.acschembio.7b00603.
- Ziemke, F. & Mantzoros, C.S., 2010. Adiponectin in insulin resistance: lessons from translational research. *The American journal of clinical nutrition*, 91(1), pp.258S–261S.
- Zlotnik, A., Burkhardt, A.M. & Homey, B., 2011. Homeostatic chemokine receptors and organ-specific metastasis. *Nature Publishing Group*, 11(9), pp.597–606.
- Zou, Y.R. et al., 1998. Function of the chemokine receptor CXCR4 in haematopoiesis and in cerebellar development. *Nature*, 393(6685), pp.595–599.

Appendix

Appendix A and B represent an extended version of the supplemental information reported in: Baeumler, T.A., Ahmed, A.A. & Fulga, T.A., 2017. Engineering Synthetic Signaling Pathways with Programmable dCas9-Based Chimeric Receptors. Cell Reports, 20(11), pp.2639–2653.

A: Amino acid sequence of constructs

A:1 dCas9^{VP64}, Cas9, and associated constructs

NLS(+)-dCas9^{VP64} used in *pNLS(+)-HA-dCas9m4-VP64*

MA**PKKKRKVG**IGH**VPAA**GGG**STS**YPYDV**PDYA**GG**STG**MDKKYSIGLAIGTNSVGWAVITDEYKVPSSKKFKVLGNTDRHSIKKNLIGALLFD
SGETAETRLKRTARRRYTRRNRIQCYLQEIFSNEMAKVDDSFHRLSEESFLVEEDKKHERHPIFGNIVDEVAYHEKYPTIYHLRKKLVDS
TDKADLRLLIYLALAHMIKFRGHFLIEGDLNPDNSDVKLFILQVQTYNQLFEENPINASGVDAKAILSARLSKSRRLLENLIAQLPGEKKNG
LFGNLIALSLGLTPNFKSNDLAEDAKLQLSKDTYDDDLNLLAQIGDQYADFLAAKNLSDAILLSDILRVNTEITKAPLSASMIKRYDE
HHQDLTLLKALVRQQLPEKYKEIFFDQSKNGYAGYIDGGASQEEFYKFIKPILEKMDGTEELLVKLNREDLLRKQRTFDNGSIPHQIHLGE
LHAILRRQEDFYFPLKDNREKIEKILTFRIPIYVGPLARGNSRFAMWTRKSEETITPWNFEVVVDKGASAQSFIERMTNFDKNLPNEKVL
KHSLLYEYFTVYNELTKVKYVTEGMRKPAFLSGEQKKAIVDLLFKTNRKVTVKQLKEDYFKKIECFDSVEISGVEDRFNASLGTYHDLK
IKDKDFLDNEENEDILEDIVLTTLTFEDREMIEERLKYAHLFDDKVMKQLKRRRYTGWGRLSRKLINGIRDQSGKTILDFLKSDGFANR
NFMQLIHDDSLTFKEDIQKAQVSGQDLSLHEHIANLAGSPAIIKKGILQTVKVVDELVKVMGRHKPENIVEMARENQTTQKGQKNSRERMK
RIEEGIKELGSQILKEHPVENTQLQNEKLYLYLQNGRDMYVDQELDNRLSDYDVAIVPQSFLKDDSIDNKVLTSDKARGKSDNVPSE
EVVKKMKNYWRQLLNAKLITQRKFDNLTKAERGGLSELDKAGFIKRLVETRQITKHVAQILDSRMNTKYDENDKLIREVKVITLKSCLVS
DFRKDFQFYKREINNYHHAHDAYLNAVGTALIKKYPKLESEFVYGDYKVDVRKMIKSEQEI GKATAKYFFYSNIMNFFKTEITLANG
EIRKRPLIETNGETGEIVWDKGRDFATVRKVLSPQVNIKKTEVQTGGFSKESILPKRNSDKLIARKKDWDPKKYGGFDSPTVAYSVLV
AKVEKGSKKLKSVKELLGITIMERSSEFKNPIDFLEAKGYKEVKDLIKLPKYSLELENGRKRMLASAGELQKGNELALPSKYVNFY
LASHYEKLKSPEDNEQQLFVEQHKHYLDEIEQISEFSKRVLADANLDKVL SAYNKHDKPIREQAENIIHLFTLTNLGAPAAFKYFD
TTIDRKRYTSTKEVLDTLIHQSIITGLYETRIDLSQLGGDSRADPKKKRKVEAS**GSGRADALDDFDLMDLGS**DALDDFDLMDLGS
DALDDFDLMDLGSDALDDFDLMDLGS

NN NLS(+), SV40 nuclear localisation sequence

NN GGGS linker

NN HA, hemagglutinin A epitope

NN dCas9m4, nuclease deficient *Streptococcus pyogenes* Cas9

NN NLS, SV40 nuclear localisation sequence

NN VP64, transcriptional activator domain

NLS-dCas9^{VP64} used in *pNLS-HA-dCas9m4-VP64*

MAS**PKKKRKVG**GGG**STS**YPYDV**PDYA**GG**STG**MDKKYSIGLAIGTNSVGWAVITDEYKVPSSKKFKVLGNTDRHSIKKNLIGALLFD
SGETAETRLKRTARRRYTRRNRIQCYLQEIFSNEMAKVDDSFHRLSEESFLVEEDKKHERHPIFGNIVDEVAYHEKYPTIYHLRKKLVDS
TDKADLRLLIYLALAHMIKFRGHFLIEGDLNPDNSDVKLFILQVQTYNQLFEENPINASGVDAKAILSARLSKSRRLLENLIAQLPGEKKNG
LFGNLIALSLGLTPNFKSNDLAEDAKLQLSKDTYDDDLNLLAQIGDQYADFLAAKNLSDAILLSDILRVNTEITKAPLSASMIKRYDE
HHQDLTLLKALVRQQLPEKYKEIFFDQSKNGYAGYIDGGASQEEFYKFIKPILEKMDGTEELLVKLNREDLLRKQRTFDNGSIPHQIHLGE
LHAILRRQEDFYFPLKDNREKIEKILTFRIPIYVGPLARGNSRFAMWTRKSEETITPWNFEVVVDKGASAQSFIERMTNFDKNLPNEKVL
KHSLLYEYFTVYNELTKVKYVTEGMRKPAFLSGEQKKAIVDLLFKTNRKVTVKQLKEDYFKKIECFDSVEISGVEDRFNASLGTYHDLK
IKDKDFLDNEENEDILEDIVLTTLTFEDREMIEERLKYAHLFDDKVMKQLKRRRYTGWGRLSRKLINGIRDQSGKTILDFLKSDGFANR
NFMQLIHDDSLTFKEDIQKAQVSGQDLSLHEHIANLAGSPAIIKKGILQTVKVVDELVKVMGRHKPENIVEMARENQTTQKGQKNSRERMK
RIEEGIKELGSQILKEHPVENTQLQNEKLYLYLQNGRDMYVDQELDNRLSDYDVAIVPQSFLKDDSIDNKVLTSDKARGKSDNVPSE
EVVKKMKNYWRQLLNAKLITQRKFDNLTKAERGGLSELDKAGFIKRLVETRQITKHVAQILDSRMNTKYDENDKLIREVKVITLKSCLVS
DFRKDFQFYKREINNYHHAHDAYLNAVGTALIKKYPKLESEFVYGDYKVDVRKMIKSEQEI GKATAKYFFYSNIMNFFKTEITLANG
EIRKRPLIETNGETGEIVWDKGRDFATVRKVLSPQVNIKKTEVQTGGFSKESILPKRNSDKLIARKKDWDPKKYGGFDSPTVAYSVLV
AKVEKGSKKLKSVKELLGITIMERSSEFKNPIDFLEAKGYKEVKDLIKLPKYSLELENGRKRMLASAGELQKGNELALPSKYVNFY
LASHYEKLKSPEDNEQQLFVEQHKHYLDEIEQISEFSKRVLADANLDKVL SAYNKHDKPIREQAENIIHLFTLTNLGAPAAFKYFD
TTIDRKRYTSTKEVLDTLIHQSIITGLYETRIDLSQLGGDSRADPKKKRKVEAS**GSGRADALDDFDLMDLGS**DALDDFDLMDLGS
DALDDFDLMDLGSDALDDFDLMDLGS

NN NLS, SV40 nuclear localisation sequence

NN GGGS linker

NN HA, hemagglutinin A epitope

Appendix

NN dCas9m4, nuclease deficient *Streptococcus pyogenes* Cas9
NN VP64, transcriptional activator domain

dCas9^{VP64} used in *pHA-dCas9m4-VP64*

MASYPYDVPDYAGGSTGMDKKYSIGLAIGTNSVGWAVITDEYKVPSSKKFKVLGNTDRHSIKKNLIGALLFDSGETAEATRLKRTARRRYTR
RKNRICYLQEIFSNEMAKVDDSFHRLSEESFLVEEDKKHERHPIFGNIVDEVAYHEKYPTIYHLRKKLVDSTDKADLRLIYLALAHMIKFR
GHFLIEGDLNPDNSDVKLFIQLVQTYNQLFEEENPINASGVDAKAILSARLSKSRRLENLIAQLPGEKKNGLFGNLIALSLGLTPNFKSNF
DLAEDAKLQLSKDTYDDDLNLLAQIGDQYADFLAAKNLSDAILLSDILRVNTEITKAPLSASMIKRYDEHHQDLTLLKALVRQQLPEKY
KEIFFDQSKNGYAGYIDGGASQEEFYKFIKPILEKMDGTEELLVKLNREDLLRKQRTFDNGSIPHQIHLGELHAILRRQEDFYFPLKDNRE
KIEKILTFRIPIYYVGPLARGNSRFAWMTRKSEETITPWNFEVVVDKGASAQSFIERMTNFDKNLPNEKVLPHKSLLEYFTVYNELTKVKY
VTEGMRKPAFLSGEQKKAIVDLLFKTNRKVTVKQLKEDYFKKIECFDSVEISGVEDRFNASLGTYHDLKIIKDKDFLDNEENEDILEDIV
LTLTLFEDREMIEERLKYAHLFDDKVMKQLKRRRYTGWGRLSRKLINGIRDQSGKTILDFLKSDGFANRNFQMQLIHDDSLTFKEDIQKA
QVSGQGDSLHEHIANLAGSPAIAKKGILQTVKVVDELVKVMGRHKPENIVIEMARENQTTQKGQKNSRERMKRIEEGIKELGSQILKEHPVE
NTQLQNEKLYLYLQNGRDMYVDQELDINRLSDYDVAIVPQSFLKDDSIDNKVLTRSDKARGKSDNVPSEEVVKKMKNYWRQLLNAKLIT
QRKFDNLTKAERGGLSELDKAGFIKRLVETROITKHVAQILDSRMNTKYDENDKLIREVKVITLKSCLVSDFRKDFQFYKVIENNYHHA
HDAYLNAVGTALIKKYPKLESEFVYGDYKVDVRKMIKSEQEI GKATAKYFFYSNIMNFFKTEITLANGEIRKRPLIETNGETGEIVWD
KGRDFATVRKVLSPQVNIIVKKTEVQTGGFSKESILPKRNSDKLIARKKDWDPKKYGGFDSPTVAYSVLVVAKEVGKSKKLKSVKELLGI
TIMERSSEFKNPIDFLEAKGYEVKKDLIKLPKYSLFELNGRKRMLASAGELQKGNELALPSKYVNFYLASHYKLGSPEDNEQKQL
FVEQHKHYLDEIIIEQISEFSKRVILADANLDKVL SAYNKHDKPIREQAENI IHLFTLTNLGAPAAFKYFDTTIDRKRYTSTKEVLDATLI
HQSITGLYETRIDLSQLGGDGGGSQ LGGGS**GSGRADALDDFDLDMLGSDALDDFDLDMLGSDALDDFDLDMLGSDALDDFDLDMLGSDA**
LDDFDLDMLINSR*

MN HA, hemagglutinin A epitope
NN dCas9m4, nuclease deficient *Streptococcus pyogenes* Cas9
NN (GGGS)₂ linker
NN VP64, transcriptional activator domain

NES-dCas9^{VP64} used in *pNES-HA-dCas9m4-VP64*

MASLDLASLILGGGSASYPYDVPDYAGGSTGMDKKYSIGLAIGTNSVGWAVITDEYKVPSSKKFKVLGNTDRHSIKKNLIGALLFDSGETA
EATRLKRTARRRYTRRKNRICYLQEIFSNEMAKVDDSFHRLSEESFLVEEDKKHERHPIFGNIVDEVAYHEKYPTIYHLRKKLVDSTDKAD
LRLIYLALAHMIKFRGHFLIEGDLNPDNSDVKLFIQLVQTYNQLFEEENPINASGVDAKAILSARLSKSRRLENLIAQLPGEKKNGLFGNL
IALSLGLTPNFKSNFDLAEDAKLQLSKDTYDDDLNLLAQIGDQYADFLAAKNLSDAILLSDILRVNTEITKAPLSASMIKRYDEHHQDL
TLLKALVRQQLPEKYKEIFFDQSKNGYAGYIDGGASQEEFYKFIKPILEKMDGTEELLVKLNREDLLRKQRTFDNGSIPHQIHLGELHAIL
RRQEDFYFPLKDNREKIEKILTFRIPIYYVGPLARGNSRFAWMTRKSEETITPWNFEVVVDKGASAQSFIERMTNFDKNLPNEKVLPHKSL
YEFYTVYNELTKVKYVTEGMRKPAFLSGEQKKAIVDLLFKTNRKVTVKQLKEDYFKKIECFDSVEISGVEDRFNASLGTYHDLKIIKDKD
FLDNEENEDILEDIVLTLTLFEDREMIEERLKYAHLFDDKVMKQLKRRRYTGWGRLSRKLINGIRDQSGKTILDFLKSDGFANRNFQMQL
IHDDSLTFKEDIQKAQVSGQGDSLHEHIANLAGSPAIAKKGILQTVKVVDELVKVMGRHKPENIVIEMARENQTTQKGQKNSRERMKRIE
IKELGSQILKEHPVENTQLQNEKLYLYLQNGRDMYVDQELDINRLSDYDVAIVPQSFLKDDSIDNKVLTRSDKARGKSDNVPSEEVVKK
MKNYWRQLLNAKLITQRKFDNLTKAERGGLSELDKAGFIKRLVETROITKHVAQILDSRMNTKYDENDKLIREVKVITLKSCLVSDFRKD
FQFYKVIENNYHHAHDAYLNAVGTALIKKYPKLESEFVYGDYKVDVRKMIKSEQEI GKATAKYFFYSNIMNFFKTEITLANGEIRKR
PLIETNGETGEIVWDKGRDFATVRKVLSPQVNIIVKKTEVQTGGFSKESILPKRNSDKLIARKKDWDPKKYGGFDSPTVAYSVLVVAKEV
GKSKKLKSVKELLGITIMERSSEFKNPIDFLEAKGYEVKKDLIKLPKYSLFELNGRKRMLASAGELQKGNELALPSKYVNFYLASHY
EKLKSPEDNEQKQLFVEQHKHYLDEIIIEQISEFSKRVILADANLDKVL SAYNKHDKPIREQAENI IHLFTLTNLGAPAAFKYFDTTIDR
KRYTSTKEVLDATLIHQSITGLYETRIDLSQLGGDGGGSQ LGGGS**GSGRADALDDFDLDMLGSDALDDFDLDMLGSDALDDFDLDMLGSDA**
LDDFDLDMLINSSAGGSLDLASLILSR*

NN NES, nuclear export signal from human protein tyrosine kinase 2
MN HA, hemagglutinin A epitope
NN dCas9m4, nuclease deficient *Streptococcus pyogenes* Cas9
NN GGGS linker
NN VP64, transcriptional activator domain

A:2 TMt-dCas9^{VP64}, TMt-Cas9, and associated constructs

TMt-NLS-dCas9^{VP64} used in *pTMt_TCS(Q'G)_NLS-HA-dCas9m4-VP64*

METDTLLLWVLLWVPGSTGDHSGGSGGGSGRQEQKLISEEDLNAVQDQTQEVIVVPHSLPFKVVVISAILALVVLTIISLIILIMLWQK
KPRQLQSGSETPGTSESATPESASHVDHAAANENLYFQGPKKKRKVGGGSTSYPYDVPDYAGGSTGMDKKYSIGLAIGTNSVGWAVITDEYKV
PSKKFKVLGNTDRHSIKKNLIGALLFDSGETAEATRLKRTARRRYTRRKNRICYLQEIFSNEMAKVDDSFHRLSEESFLVEEDKKHERHPI
FGNIVDEVAYHEKYPTIYHLRKKLVDSTDKADLRLIYLALAHMIKFRGHFLIEGDLNPDNSDVKLFIQLVQTYNQLFEEENPINASGVDAK
AILSARLSKSRRLENLIAQLPGEKKNGLFGNLIALSLGLTPNFKSNFDLAEDAKLQLSKDTYDDDLNLLAQIGDQYADFLAAKNLSDAI
LLSDILRVNTEITKAPLSASMIKRYDEHHQDLTLLKALVRQQLPEKYKEIFFDQSKNGYAGYIDGGASQEEFYKFIKPILEKMDGTEELLV
KLNREDLLRKQRTFDNGSIPHQIHLGELHAILRRQEDFYFPLKDNREKIEKILTFRIPIYYVGPLARGNSRFAWMTRKSEETITPWNFEVV
DKGASAQSFIERMTNFDKNLPNEKVLPHKSLLEYFTVYNELTKVKYVTEGMRKPAFLSGEQKKAIVDLLFKTNRKVTVKQLKEDYFKKIE
CFDSVEISGVEDRFNASLGTYHDLKIIKDKDFLDNEENEDILEDIVLTLTLFEDREMIEERLKYAHLFDDKVMKQLKRRRYTGWGRLSR
KLINGIRDQSGKTILDFLKSDGFANRNFQMQLIHDDSLTFKEDIQKAQVSGQGDSLHEHIANLAGSPAIAKKGILQTVKVVDELVKVMGRHK
PENIVIEMARENQTTQKGQKNSRERMKRIEEGIKELGSQILKEHPVENTQLQNEKLYLYLQNGRDMYVDQELDINRLSDYDVAIVPQSFL
LKDDSIDNKVLTRSDKARGKSDNVPSEEVVKKMKNYWRQLLNAKLITQRKFDNLTKAERGGLSELDKAGFIKRLVETROITKHVAQILDS
RMNTKYDENDKLIREVKVITLKSCLVSDFRKDFQFYKVIENNYHHAHDAYLNAVGTALIKKYPKLESEFVYGDYKVDVRKMIKSEQEI
IGKATAKYFFYSNIMNFFKTEITLANGEIRKRPLIETNGETGEIVWDKGRDFATVRKVLSPQVNIIVKKTEVQTGGFSKESILPKRNSDKL
IARKKDWDPKKYGGFDSPTVAYSVLVVAKEVGKSKKLKSVKELLGITIMERSSEFKNPIDFLEAKGYEVKKDLIKLPKYSLFELNGR

Appendix

[KRMLASAGELQKGNELALPSKYVNFLYLASHYEKLKGGSPEDNEQKQLFVEQHKHYLDEIIEQISEFSKRVI LADANLDKVL SAYNKHRDKP
IREQAE NIIHLFTLTNLGAPAAFKYFDTTIDRKRYTSTKEVLDATLIHQ SITGLYETRIDLSQLGGD](#)[SRADPKKKRKVEAS](#)[GSGRADALDD](#)
[FDLMDLGS DALDDFDLMDLGS DALDDFDLMDLGS DALDDFDLMDL IN](#)^{SR*}

NN Igκ, murine Immunoglobulin kappa-chain signal peptide

NN (GGGS)₂ linker

NN Myc epitope tag

NN TM (PDGFR), transmembrane domain from platelet derived growth factor receptor

NN XTEN linker

NN efficient TCS; tobacco etch virus (TEV) protease cleavage sequence

NN NLS, SV40 nuclear localisation sequence

NN HA, hemagglutinin A epitope

NN dCas9m4, nuclease deficient *Streptococcus pyogenes* Cas9

NN VP64, transcriptional activator domain

TMt^{ΔTCS}-NLS-dCas9^{VP64} used in *pTMt_NLS-HA-dCas9m4-VP64*

[METDTLLLWVLLWVPGSTGDHSGGGSGGGSGRQEQKLISEEDLNAVQDQTQEVIVVPHSLPFKVVVISAILALVVLTIISLIILIMLWQK
KPRLQSGSETPGTSESATPES](#)[ASHVDHAAAGGSGPKKKRKVGGGSTS](#)[YPYDVDPDYAGGSTGMDKKYSIGLAIGTNSVGWAVITDEYKVPSPK
KFKVLGNTDRHSIKKNLIGALLFDSGETAEATRLKRTARRRYTRRKNRICYLQEIFS](#)[NEMAKVDDSF](#)[FHRLEESFLVEEDKKHERHPIFGN
IVDEVAYHEKYPTIYHLRKKLVDS](#)[TDKADLRLIY LALAHMIKFRGHFLIEGDLNPDNSDVKLF](#)[IQLVQTYNQLF](#)[EENPINASGVDAKAIL
SARLSKSRLENLIAQLPGEKKNGLF](#)[GNLIALSLGLTPNFKS](#)[NFDLAEDAKLQLSKDTYDDDLNLLAQIGDQYADLFLAAKNLSDAILLS
DILRVNTEITKAPLSASMIKRYDEHHQDLTLKALVRQQLPEKYKEIFFDQSKNGYAGYIDGGASQEEFYKFIKPILEKMDGTEELLVKLN
REDLRLKQRTFDNGSIPHQIHLGELHAILRRQEDFY](#)[PFLKDNREKIEKILTFRIPYVVGPLARGNSRFAWMTRKSEETITPWNFEEVVDKG
ASAQSFIERMTNFDKNLPNEKVLPHKSLLYEYFTVYNELTKVKYVTEGMRKPAFLS](#)[GEQKKAIVDLLFKTNRKVT](#)[VKQLKEDYFKKIECFD
SVEISGVEDRFNASLGTYHDLKIIKDKDFLDNEENEDILEDIVLTTLTFEDREMIEERLKYAHLFDDKVMKQLKRRRYTGWGRLSRKLI
NGIRDKQSGKTI](#)[LDFLKSDGFANRNFMQLIHDDSLTFKEDIQKAQVSGQGDSLHEHIANLAGSPA](#)[IKKGILQTVKVVD](#)[ELVKVMGRHKPEN
IVIEMARENQTTQKGQKNSRERMKRIE](#)[EGIKELGSQILKEHPVENTQLONEKLYLYYLQNGRDMYVDQELDINRLSDYDVAAIVPQSFLKD
DSIDNKVLTRSDKARGKSDNVPSEEVVKKMKNYWRQLLNAKLITQRKFDNLT](#)[KAERGGLSEL](#)[DKAGFIKRQLVET](#)[RQITKHVAQILDSRMN
TKYDENDKLI](#)[REVKVI](#)[TLKSKLVSDFRKDFQFYK](#)[VREINNYHHAHDAYLNAVVG](#)[TALIKKYPKLESEFVYGDYKVYDVRKMI](#)[AKSEQEIGK
ATAKYFFYSINIMNFFKTEIT](#)[LANGEIRKRPLIETNGETGEIVWDKGRDFATVRKVL](#)[SMPQVNI](#)[VKKTEVQ](#)[TGGFSKESILPKRNSDKLIAR
KKDWDPKKYGGFDSPTVAYS](#)[VLVAKVEKGSKKLKSVKELLGITIMERS](#)[SFEKNPIDFLEAKGYKEVKKDLI](#)[IKLPKYS](#)[LFELENGRKRML
LASAGELQKGNELALPSKYVNFLYLASHYEKLKGGSPEDNEQKQLFVEQHKHYLDEIIEQISEFSKRVI LADANLDKVL SAYNKHRDKP
IREQAE NIIHLFTLTNLGAPAAFKYFDTTIDRKRYTSTKEVLDATLIHQ SITGLYETRIDLSQLGGD](#)[SRADPKKKRKVEAS](#)[GSGRADALDDFDL
DMLGSDALDDFDLMDLGS DALDDFDLMDLGS DALDDFDLMDL IN](#)^{SR*}

NN Igκ, murine Immunoglobulin kappa-chain signal peptide

NN (GGGS)₂ linker

NN Myc epitope tag

NN TM (PDGFR), transmembrane domain from platelet derived growth factor receptor

NN XTEN linker

NN NLS, SV40 nuclear localisation sequence

NN HA, hemagglutinin A epitope

NN dCas9m4, nuclease deficient *Streptococcus pyogenes* Cas9

NN VP64, transcriptional activator domain

TMt-dCas9^{VP64} used in *pTMt_TCS(Q'G)_HA-dCas9m4-VP64*

[METDTLLLWVLLWVPGSTGDHSGGGSGGGSGRQEQKLISEEDLNAVQDQTQEVIVVPHSLPFKVVVISAILALVVLTIISLIILIMLWQK
KPRLQSGSETPGTSESATPES](#)[ASHVDHAAENLYFQGGGGSTS](#)[YPYDVDPDYAGGSTGMDKKYSIGLAIGTNSVGWAVITDEYKVPSPKFKV
LGNTDRHSIKKNLIGALLFDSGETAEATRLKRTARRRYTRRKNRICYLQEIFS](#)[NEMAKVDDSF](#)[FHRLEESFLVEEDKKHERHPIFGNIVDE
VAYHEKYPTIYHLRKKLVDS](#)[TDKADLRLIY LALAHMIKFRGHFLIEGDLNPDNSDVKLF](#)[IQLVQTYNQLF](#)[EENPINASGVDAKAIL
SARLSKSRLENLIAQLPGEKKNGLF](#)[GNLIALSLGLTPNFKS](#)[NFDLAEDAKLQLSKDTYDDDLNLLAQIGDQYADLFLAAKNLSDAILLS
VNTTEITKAPLSASMIKRYDEHHQDLTLKALVRQQLPEKYKEIFFDQSKNGYAGYIDGGASQEEFYKFIKPILEKMDGTEELLVKLNREDL
LRKQRTFDNGSIPHQIHLGELHAILRRQEDFY](#)[PFLKDNREKIEKILTFRIPYVVGPLARGNSRFAWMTRKSEETITPWNFEEVVDKGASAQ
SFIERMTNFDKNLPNEKVLPHKSLLYEYFTVYNELTKVKYVTEGMRKPAFLS](#)[GEQKKAIVDLLFKTNRKVT](#)[VKQLKEDYFKKIECFDSVEI
SGVEDRFNASLGTYHDLKIIKDKDFLDNEENEDILEDIVLTTLTFEDREMIEERLKYAHLFDDKVMKQLKRRRYTGWGRLSRKLINGIR
DKQSGKTI](#)[LDFLKSDGFANRNFMQLIHDDSLTFKEDIQKAQVSGQGDSLHEHIANLAGSPA](#)[IKKGILQTVKVVD](#)[ELVKVMGRHKPENIVIE
MARENQTTQKGQKNSRERMKRIE](#)[EGIKELGSQILKEHPVENTQLONEKLYLYYLQNGRDMYVDQELDINRLSDYDVAAIVPQSFLKDDSID
NKVLTRSDKARGKSDNVPSEEVVKKMKNYWRQLLNAKLITQRKFDNLT](#)[KAERGGLSEL](#)[DKAGFIKRQLVET](#)[RQITKHVAQILDSRMNTKYD
ENDKLIREVKVITL](#)[KSKLVSDFRKDFQFYK](#)[VREINNYHHAHDAYLNAVVG](#)[TALIKKYPKLESEFVYGDYKVYDVRKMI](#)[AKSEQEIGKATAK
YFFYSINIMNFFKTEIT](#)[LANGEIRKRPLIETNGETGEIVWDKGRDFATVRKVL](#)[SMPQVNI](#)[VKKTEVQ](#)[TGGFSKESILPKRNSDKLIAR
KKDWDPKKYGGFDSPTVAYS](#)[VLVAKVEKGSKKLKSVKELLGITIMERS](#)[SFEKNPIDFLEAKGYKEVKKDLI](#)[IKLPKYS](#)[LFELENGRKRMLASA
GELQKGNELALPSKYVNFLYLASHYEKLKGGSPEDNEQKQLFVEQHKHYLDEIIEQISEFSKRVI LADANLDKVL SAYNKHRDKP
IREQAE NIIHLFTLTNLGAPAAFKYFDTTIDRKRYTSTKEVLDATLIHQ SITGLYETRIDLSQLGGD](#)[GGGSQLGGG](#)[GSGRADALDDFDLMDLGS DAL
DDFDLMDLGS DALDDFDLMDLGS DALDDFDLMDL IN](#)^{SR*}

NN Igκ, murine Immunoglobulin kappa-chain signal peptide

NN (GGGS)₂ linker

NN Myc epitope tag

NN TM (PDGFR), transmembrane domain from platelet derived growth factor receptor

NN XTEN linker

NN efficient TCS; tobacco etch virus (TEV) protease cleavage sequence

NN HA, hemagglutinin A epitope

NN dCas9m4, nuclease deficient *Streptococcus pyogenes* Cas9

Appendix

NN VP64, transcriptional activator domain

TMt-NES-dCas9^{VP64} used in *pTmt_TCS(Q'G)_NES-HA-dCas9m4-VP64*

METDTLLWVLLWVPGSTGDHSGGSGGGSGRQEQKLISEEDLNAV**QGDTQE**VIVVPHSLPFKVVVISAILALVVL**TIISLIILIMLWQK**
KPRLQSGSETPGTSESATPESASHVDHAA**ENLYFQGGGGSTS****LDLASLI**IGGGSASYPYDV**PDYAGG**STGTGMDKKYSIGLAIGTNSVG
WAVITDEYKVPSSKKFKVLGNTDRHSIKKNLIGALLFDSGETAEATRLKRTARRRYTRRKNRICYLQEIFSNEMAKVDDSFHRLSEESFLVE
EDKKHERHPIFGNIVDEVAYHEKYPTIYHLRKKLVSDTDKADRLRIYLALAHMIKFRGHFLIEGDLNPDNSDVDFLFIQLVQTYNQLFEEFN
PINASGVDAKAILSARLSKSRRLLENLIAQLPGEKKNGFLGNLIALSLGLTPNFKSNFDLAEDAKLQLSKDTYDDDLNLLAQIGDQYADLF
LAAKNLSDAILLSDILRVNTEITKAPLSASMIKRYDEHHQDLTLLKALVRQQLPEKYKEIFFDQSKNGYAGYIDGGASQEEFYKFIKPILE
KMDGTEELLVKLNREDLLRKQRTFDNGSIHQIHLGELHAILRRQEDFYFPLKDNREKIEKILTFRIPIYVGPLARGNSRFAMWTRKSEET
ITPWNFEEVVDKGASAQSFIERMTNFDKNLPNEKVLPHKSHLLYEYFTVYNELTKVKYVTEGMRKPAFLSGEQKKAIVDLLFKTNRKVTVKQ
LKEDYFKKIECFDSVEISGVEDRFNASLGTYHDLLKI IKDKDFLDNEENEDILEDIVLTTLTFEDREMIEERLKYAHLFDDKVMKQLKRR
RYTGWGRLSRKLINGIRDKQSGKTI LDFLKSDFANRNFQMQLIHDDSLTFKEDIQKAQVSGQGDLSLHEHIANLAGSPAIKKGILQTVKVVD
ELVKVMGRHKPENIVIAMARENQTTQKGQKNSRERMKRIEEGIKELGSQILKEHPVENTQLQNEKLYLYLQNGRDMYVDQELDINRLSDY
DVAAIVPQSFLLKDDSIDNKVLTNRSDKARGKSDNVPSEEVVKKMKNYWRQLLNAKLITQRFKNLTKAERGGLSELDKAGFIKRLVETRQI
TKHVAQILDSRMNTKYDENDKLIREVKVITLKSCLVSDFRKDFQFYKVINNYHHAHDAYLNAVVGTAIIKKYPKLESEFVYGDYKVYDV
RKMIKSEQEIGKATAKYFFYSNIMNFFKTEITLANGEIRKRPLIETNGETGEIVWDKGRDFATVRKVLSPQVNIKKTEVQTGGFSKES
ILPKRNSDKLIARNDWDPKKYGGFDSPTVAYSVLVAKVEKGSKKLKSVKELLGITIMERSFFEKNPIDFLEAKGYKEVKDLIIKLPK
YSLFELENGRKRMLASAGELQKGNELALPSKYVNFYLLASHYEKLGKSPEDNEQKQLFVEQHKHYLDEIIIEQISEFSKRVLADANLDKVL
SAYNKHRRDKPIREQAENIIHLFTLTNLGAPAAFKYFDTTIDRKRYTSTKEVLDATLIHQSIITGLYETRIDLSQLGGDGGGSQLGGGS**GSGR**
ADALDDFDLMLGSDALDDFDLMLGSDALDDFDLMLGSDALDDFDLMLINSSAGGGSLD**LASLI**SR*

NN Igk, murine Immunoglobulin kappa-chain signal peptide

NN (GGGS)₂ linker

NN Myc epitope tag

NN TM (PDGFR), transmembrane domain from platelet derived growth factor receptor

NN XTEN linker

NN NES, nuclear export signal form human protein tyrosine kinase 2

NN efficient TCS; tobacco etch virus (TEV) protease cleavage sequence

NN HA, hemagglutinin A epitope

NN dCas9m4, nuclease deficient *Streptococcus pyogenes* Cas9

NN VP64, transcriptional activator domain

TMt^{NES}-dCas9^{VP64} used in *pTmt_NES_TCS(Q'G)_HA-dCas9m4-VP64*

METDTLLWVLLWVPGSTGDHSGGSGGGSGRQEQKLISEEDLNAV**QGDTQE**VIVVPHSLPFKVVVISAILALVVL**TIISLIILIMLWQK**
KPRLQSGSETPGTSESATPESAS**LDLASLI**LGKLG**ENLYFQGGGGSTS**YPYDV**PDYAGG**STGMDKKYSIGLAIGTNSVGWAVITDEYKVPSS
KKFKVLGNTDRHSIKKNLIGALLFDSGETAEATRLKRTARRRYTRRKNRICYLQEIFSNEMAKVDDSFHRLSEESFLVEEDKKHERHPIFG
NIVDEVAYHEKYPTIYHLRKKLVSDTDKADRLRIYLALAHMIKFRGHFLIEGDLNPDNSDVDFLFIQLVQTYNQLFEEFNPINASGVDAKAI
LSARLSKSRRLLENLIAQLPGEKKNGFLGNLIALSLGLTPNFKSNFDLAEDAKLQLSKDTYDDDLNLLAQIGDQYADLFLLAAKNLSDAILL
SDILRVNTEITKAPLSASMIKRYDEHHQDLTLLKALVRQQLPEKYKEIFFDQSKNGYAGYIDGGASQEEFYKFIKPILEKMDGTEELLVKLN
NREDLLRKQRTFDNGSIHQIHLGELHAILRRQEDFYFPLKDNREKIEKILTFRIPIYVGPLARGNSRFAMWTRKSEETITPWNFEEVVDK
GASAQSFIERMTNFDKNLPNEKVLPHKSHLLYEYFTVYNELTKVKYVTEGMRKPAFLSGEQKKAIVDLLFKTNRKVTVKQLKEDYFKKIECF
DSVEISGVEDRFNASLGTYHDLLKI IKDKDFLDNEENEDILEDIVLTTLTFEDREMIEERLKYAHLFDDKVMKQLKRRRYTGWGRLSRKL
INGIRDKQSGKTI LDFLKSDFANRNFQMQLIHDDSLTFKEDIQKAQVSGQGDLSLHEHIANLAGSPAIKKGILQTVKVVDDELVKVMGRHKPE
NIVIAMARENQTTQKGQKNSRERMKRIEEGIKELGSQILKEHPVENTQLQNEKLYLYLQNGRDMYVDQELDINRLSDYDVAAIVPQSFLLK
DDSIDNKVLTNRSDKARGKSDNVPSEEVVKKMKNYWRQLLNAKLITQRFKNLTKAERGGLSELDKAGFIKRLVETRQITKHVAQILDSRM
NTKYDENDKLIREVKVITLKSCLVSDFRKDFQFYKVINNYHHAHDAYLNAVVGTAIIKKYPKLESEFVYGDYKVYDVRKMIKSEQEIG
KATAKYFFYSNIMNFFKTEITLANGEIRKRPLIETNGETGEIVWDKGRDFATVRKVLSPQVNIKKTEVQTGGFSKESILPKRNSDKLIA
RKWDWDPKKYGGFDSPTVAYSVLVAKVEKGSKKLKSVKELLGITIMERSFFEKNPIDFLEAKGYKEVKDLIIKLPKYSLELENGRKR
MLASAGELQKGNELALPSKYVNFYLLASHYEKLGKSPEDNEQKQLFVEQHKHYLDEIIIEQISEFSKRVLADANLDKVL SAYNKHRRDKPIR
EQAENIIHLFTLTNLGAPAAFKYFDTTIDRKRYTSTKEVLDATLIHQSIITGLYETRIDLSQLGGDGGGSQLGGGS**GSGRADALDDFDLML**
GSDALDDFDLMLGSDALDDFDLMLGSDALDDFDLMLINSR*

NN Igk, murine Immunoglobulin kappa-chain signal peptide

NN GGGS linker

NN Myc epitope tag

NN TM (PDGFR), transmembrane domain from platelet derived growth factor receptor

NN XTEN linker

NN NES, nuclear export signal form human protein tyrosine kinase 2

NN efficient TCS; tobacco etch virus (TEV) protease cleavage sequence

NN HA, hemagglutinin A epitope

NN dCas9m4, nuclease deficient *Streptococcus pyogenes* Cas9

NN VP64, transcriptional activator domain

TMt^{NESATCS}-dCas9^{VP64} used in *pTmt_NES_HA-dCas9m4-VP64*

METDTLLWVLLWVPGSTGDHSGGSGGGSGRQEQKLISEEDLNAV**QGDTQE**VIVVPHSLPFKVVVISAILALVVL**TIISLIILIMLWQK**
KPRLQSGSETPGTSESATPESASLDLASLIIGKLG**GGGGSTS**YPYDV**PDYAGG**STGMDKKYSIGLAIGTNSVGWAVITDEYKVPSSKKFKVLG
NTDRHSIKKNLIGALLFDSGETAEATRLKRTARRRYTRRKNRICYLQEIFSNEMAKVDDSFHRLSEESFLVEEDKKHERHPIFGNIVDEVA
YHEKYPTIYHLRKKLVSDTDKADRLRIYLALAHMIKFRGHFLIEGDLNPDNSDVDFLFIQLVQTYNQLFEEFNPINASGVDAKAILSARLSK
SRRLLENLIAQLPGEKKNGFLGNLIALSLGLTPNFKSNFDLAEDAKLQLSKDTYDDDLNLLAQIGDQYADLFLLAAKNLSDAILLSDILRVN
TEITKAPLSASMIKRYDEHHQDLTLLKALVRQQLPEKYKEIFFDQSKNGYAGYIDGGASQEEFYKFIKPILEKMDGTEELLVKLNREDLLR

Appendix

KQRTFDNGSIPHQIHLGELHAILRRQEDFYFLKDNREKIEKILTFRIPIYYVGPLARGNSRFAWMTRKSEETITPWNFEEVVDKGASAQSF
IERMTNFDKNLPNEKVLPKHSLLEYFTVYNELTKVKYVTEGMRKPAFLSGEQKKAIVDLLFKTNRKVTQVKLKEDYFKKIECFDSVEISG
VEDRFNASLGTYHDLKI IKDKDFLDNEENEDILEDIVLTLTLFEDREMIEERLKYAHLFDDKVMKQLKRRRYTGWGRLSRKLINGIRDK
QSGKTIILDFLKSDFANRNFMQLIHDDSLTFKEDIQKAQVSGQGDSLHEHIANLAGSPAIKKGILQTVKVVDLVKVMGRHKPENIVIEMA
RENQTTQKGQKNSRERMKRIE EGIKELGSQILKEHPVENTQLQNEKLYLYYLQNGRDMYVDQELDINRLSDYDVA AIVPQSFLKDDSIDNK
VLTRSDKARGKSDNVPSEEVVKMKMNYWRQLLNAKLITQRKFDNLTKAERGGLSELDKAGFIKRQLVETRQITKHVAQILDSRMNTKYDEN
DKLIREVKVITLTKSLVSDFRKDFQFYKVR EINNYHHAHDAYLNAVVG TALIKKYPKLESEFVYGDYKVYDVRKMI AKSEQEIGKATAKYF
FYSNIMNFFKTEITLANGEIRKRPLIETNGETGEIVWDKGRDFATVRKVL SMPQVNI VKKTEVQTGGFSKESILPKRNSDKLIARKKDWD
KKYGGFDSPTVAYSVLVVAKEVGKSKKLKSVKELLGITIMERSSFEKNPIDFLEAKGYKEVKKDLIIKLPKYSLFELENGRKRMLASAGE
LQKGNELALPSKYVNFYLYLASHYEKLGKSPEDNEQQLFVEQHKHYLDEIIIEQISEFSKRVILADANLDKVL SAYNKH RDKPIREQAENII
HLFTLTNLGAPAAFKYFDTTIDRKRYTSTKEVL DATLIHQ SITGLYETRIDLSQLGGDGGGSQ LGGSGSGRADALDDFDLMDLGS DALDD
FDLMDLGS DALDDFDLMDLGS DALDDFDLMDLINSR*

NN Igk, murine Immunoglobulin kappa-chain signal peptide

NN GGGS linker

NN Myc epitope tag

NN TM (PDGFR), transmembrane domain from platelet derived growth factor receptor

NN XTEN linker

NN NES, nuclear export signal form human protein tyrosine kinase 2

NN HA, hemagglutinin A epitope

NN dCas9m4, nuclease deficient *Streptococcus pyogenes* Cas9

NN VP64, transcriptional activator domain

NLS-Cas9-mCherry used in *pFLAG-NLS-Cas9-XTEN-mCherry-NLS*

MDYKDHGDGYKDHDIDYKDDDDK**MAPKKRKVGVIHGVPAA**DKKYSIGLDIGTNSVGWAVITDEYKVP SKKFKVLGNTDRHSIKKNLIGALL
FDSGETAEATRLKRTARRRYTRRKNRICY LQEIFSNEMAKVDDSFHRL EESFLVEEDKKHERHPIFGNIVDEVAYHEKYPTIYHLRKKLV
DSTDKADLRLIY LALAHMIKFRGHFLIEGDLNPDNSDV DKLFIQLVQTYNQ LFEENPINASGVDAKAIL SARLSKSRLENLIAQLPGEKK
NGLFNGNIALSLGLTPNFKSNFDLAEDAKLQ LSKD TYDDDLNLLAQIGDQYADLFLAAKNLSDAILLSDILRVNTEITKAPLSASMIKRY
DEHHQDLTLLKALVRQQLPEKYKEIFFDQSKNGYAGYIDGGASQEEFYKFIKPILEKMDGTEELLV KLNREDLLRKQRTFDNGSIPHQIHL
GELHAILRRQEDFYFLKDNREKIEKILTFRIPIYYVGPLARGNSRFAWMTRKSEETITPWNFEEVVDKGASAQSFIERMTNFDKNLPNEKV
LPKHSLLEYFTVYNELTKVKYVTEGMRKPAFLSGEQKKAIVDLLFKTNRKVTQVKLKEDYFKKIECFDSVEISGVEDRFNASLGTYHDL
KI IKDKDFLDNEENEDILEDIVLTLTLFEDREMIEERLKYAHLFDDKVMKQLKRRRYTGWGRLSRKLINGIRDKQSGKTIILDFLKSDF
ANRNFMQLIHDDSLTFKEDIQKAQVSGQGDSLHEHIANLAGSPAIKKGILQTVKVVDLVKVMGRHKPENIVIEMARENQTTQKGQKNSR
MKRIE EGIKELGSQILKEHPVENTQLQNEKLYLYYLQNGRDMYVDQELDINRLSDYDV DHI VQP SFLKDDSIDNKVLTRSDKNRGKSDNVP
SEEVVKMKMNYWRQLLNAKLITQRKFDNLTKAERGGLSELDKAGFIKRQLVETRQITKHVAQILDSRMNTKYDENDKLIREVKVITLTKSL
VSDFRKDFQFYKVR EINNYHHAHDAYLNAVVG TALIKKYPKLESEFVYGDYKVYDVRKMI AKSEQEIGKATAKYFFYSNIMNFFKTEITL
ANGEIRKRPLIETNGETGEIVWDKGRDFATVRKVL SMPQVNI VKKTEVQTGGFSKESILPKRNSDKLIARKKDWDPKKYGGFDSPTVAYS
VLVVAKEVGKSKKLKSVKELLGITIMERSSFEKNPIDFLEAKGYKEVKKDLIIKLPKYSLFELENGRKRMLASAGELQKGNELALPSKYV
NLYLASHYEKLGKSPEDNEQQLFVEQHKHYLDEIIIEQISEFSKRVILADANLDKVL SAYNKH RDKPIREQAENIIHLFTLTNLGAPAAFKY
FDTTIDRKRYTSTKEVL DATLIHQ SITGLYETRIDLSQLGGDKRPAATKKAGL**SGSETPGTSESATPES**GRMVSKGEEDNMAIIKEFMRFK
VHMEGSVNGHEFEIEGEGEGRPYEGTQTAKLKVT KGGPLPFAWDILSPQFMYGSKAYVKHPADIPDY LKLSFPEGFKWERVMNFEDGGVV
TQTQSSLDGGEFIYKVKLRGTNFPSDGPVMQKKTMGWEASSERMPEDGALKEIKQRLKLKDGGHYDAEVKTTYKAKKPVQLPGAYNVNI
KLDITSHNEDYTIVEQYERAEGRHSTGGMDELYKKRPAATKKAGQAKKKK*

NN FLAG epitope tag

NN NLS(+), SV40 nuclear localisation sequence

NN catalytically active Cas9

NN NLS, nucleoplasmin NLS

NN XTEN linker

NN mCherry

Cas9-mCherry used in *pCas9-XTEN-mCherry*

MDKKYSIGLDIGTNSVGWAVITDEYKVP SKKFKVLGNTDRHSIKKNLIGALLFDSGETAEATRLKRTARRRYTRRKNRICY LQEIFSNEMA
KVDDSFHRL EESFLVEEDKKHERHPIFGNIVDEVAYHEKYPTIYHLRKKLV DSTDKADLRLIY LALAHMIKFRGHFLIEGDLNPDNSDV
KLFIQLVQTYNQ LFEENPINASGVDAKAIL SARLSKSRLENLIAQLPGEKKNGLFNGNIALSLGLTPNFKSNFDLAEDAKLQ LSKD TYDD
DLNLLAQIGDQYADLFLAAKNLSDAILLSDILRVNTEITKAPLSASMIKRYDEHHQDLTLLKALVRQQLPEKYKEIFFDQSKNGYAGYID
GGASQEEFYKFIKPILEKMDGTEELLV KLNREDLLRKQRTFDNGSIPHQIHLGELHAILRRQEDFYFLKDNREKIEKILTFRIPIYYVGPL
ARGNSRFAWMTRKSEETITPWNFEEVVDKGASAQSFIERMTNFDKNLPNEKVLPKHSLLEYFTVYNELTKVKYVTEGMRKPAFLSGEQK
KAIVDLLFKTNRKVTQVKLKEDYFKKIECFDSVEISGVEDRFNASLGTYHDLKI IKDKDFLDNEENEDILEDIVLTLTLFEDREMIEERL
KYAHLFDDKVMKQLKRRRYTGWGRLSRKLINGIRDKQSGKTIILDFLKSDFANRNFMQLIHDDSLTFKEDIQKAQVSGQGDSLHEHIANLA
GSPAIKKGILQTVKVVDLVKVMGRHKPENIVIEMARENQTTQKGQKNSRERMKRIE EGIKELGSQILKEHPVENTQLQNEKLYLYYLQNG
RDMYVDQELDINRLSDYDV DHI VQP SFLKDDSIDNKVLTRSDKNRGKSDNVPSEEVVKMKMNYWRQLLNAKLITQRKFDNLTKAERGGL
SELDKAGFIKRQLVETRQITKHVAQILDSRMNTKYDENDKLIREVKVITLTKSLVSDFRKDFQFYKVR EINNYHHAHDAYLNAVVG TALIKK
PKLESEFVYGDYKVYDVRKMI AKSEQEIGKATAKYFFYSNIMNFFKTEITLANGEIRKRPLIETNGETGEIVWDKGRDFATVRKVL SMPQV
NIVKKTEVQTGGFSKESILPKRNSDKLIARKKDWDPKKYGGFDSPTVAYSVLVVAKEVGKSKKLKSVKELLGITIMERSSFEKNPIDFLE
AKGYKEVKKDLIIKLPKYSLFELENGRKRMLASAGELQKGNELALPSKYVNFYLYLASHYEKLGKSPEDNEQQLFVEQHKHYLDEIIIEQIS
EFSKRVILADANLDKVL SAYNKH RDKPIREQAENIIHLFTLTNLGAPAAFKYFDTTIDRKRYTSTKEVL DATLIHQ SITGLYETRIDLSQL
GGDAAS**SGSETPGTSESATPES**GRMVSKGEEDNMAIIKEFMRFKVHMEGSVNGHEFEIEGEGEGRPYEGTQTAKLKVT KGGPLPFAWDILSP
QFMYGSKAYVKHPADIPDY LKLSFPEGFKWERVMNFEDGGVVTVTQDSSLDGGEFIYKVKLRGTNFPSDGPVMQKKTMGWEASSERMPED
GALKGEIKQRLKLKDGGHYDAEVKTTYKAKKPVQLPGAYNVNIKLDITSHNEDYTIVEQYERAEGRHSTGGMDELYK*

NN catalytically active Cas9

NN XTEN linker

NN mCherry

TMt-NLS-Cas9-mCherry used in *pTMt_TCS(Q'G)_FLAG-NLS-Cas9-XTEN-mCherry-NLS*

METDTLLLWVLLLWVPGSTGDHS **GGGSGGGSGRQE**QKLISEEDL**NAV**QDQTQEVIVVPHSLPFKVVVISAILALVVLTIISLIILIMLWQK
KPRLQSGSETPGTSESATPESASHVDHAAAE**NLY**FQGTGATMDYKDHDGDYKDHDIDYKDDDDK**MAPKKR**KRVGIHGVPAADKKYSIGLDI
GTNSVGWAVITDEYKVP**SKKFKVL**GN**TD**RHS**IKKNLI**GALL**FD**SG**ETAE**AT**RLKRTARRRYTRRKNRI**CYL**QE**IF**SNEMAKVDD**SFFH**RL**E
ESFL**VEED**KK**HERHPI**FG**NI**VDE**VAYHEKYPTI**YHL**RKKLV**D**STDKAD**LR**LIY**LALAHMI**KFRGHFL**IEG**DLNPD**NSD**VDKLF**IQL**VQ**TY**N**
QLFEEN**PIN**AS**GVDAKAILSARLSKSRRLEN**LIAQL**PGEKK**NGL**FGN**LIALSLGL**TPNF**KS**NFD**LAEDAK**LQLSKD**TYDD**DLN**LLAQ**IGD**
QYADLF**LAAK**NLS**DAILL**SD**ILRVN**TEIT**KAPLS**ASMI**KRYDEHHQD**L**TLLKAL**V**RQQLPEKYKE**IFF**DQ**SKNGYAGYID**GGASQ**EEFY**KF**
IKPILE**KMDG**TEELL**VKLN**RED**LLRQRT**FD**NGSI**PH**QI**HL**GELHAILRRQ**EDFY**PFLK**DN**REKIEKIL**TF**RI**PYY**VG**PL**ARGNSR**FAW**MT**
RKSEETIT**PN**WFEE**VVDK**G**ASAQ**S**FIER**MT**NF**DK**NLPNEK**VLP**KHS**LLY**EYFTVY**NEL**TKVY**VTE**GM**R**KPA**FL**S**GE**QK**KA**IVD**LL**F**KT**N**R
KVTV**KQL**KEDY**FKKIE**CFDS**VEIS**GV**ED**RF**NAS**L**GTYH**DL**LKIIK**DK**DF**L**DNE**ENED**I**LED**IVL**TL**TL**LF**ED**REMI**EE**RL**KTYA**HL**F**DD**KV**M
KQLKRR**RYT**GW**GR**LS**RKL**ING**IR**DK**QSG**KT**IL**DF**LK**SD**G**FAN**RNF**Q**LI**HDD**SL**TF**KED**I**Q**KA**Q**VS**GQ**GD**SL**HEH**IAN**LAG**S**PA**IKK**G**IL**Q
TVK**V**DE**LV**K**V**MGR**HK**PEN**IV**IE**MARE**N**QTTQ**KG**Q**KS**R**ERM**KRI**EE**G**IK**EL**GS**QIL**KE**HP**VENT**QLQ**NE**KLY**LY**YLQ**NGR**DMY**VD**Q**EL**D**I
NRLSDY**D**VD**H**IV**PQ**S**FL**KDD**S**ID**NK**VL**TR**SD**K**NR**GK**SD**N**VP**SE**EV**VK**M**K**NY**WR**QL**NA**KL**ITQ**R**K**FD**N**L**T**KA**ERG**GL**S**EL**D**KAG**F**IK**R**Q**L**
VETR**QIT**K**H**VA**Q**IL**D**SR**M**NT**KY**DEND**KLIRE**V**KVIT**L**K**SK**L**VS**D**FR**KD**F**Q**FY**K**VR**E**IN**NY**HH**A**DAY**L**NA**V**VG**T**AL**IK**KY**P**KL**E**S**E**FY**G**D
YKYD**V**VR**K**MI**A**K**SE**QE**IG**K**A**T**A**KY**FF**YS**NIM**N**FF**K**TEIT**L**ANGE**IR**K**R**PL**IE**T**NG**ET**GE**IV**WD**K**GR**D**FAT**VR**K**V**LS**MP**Q**V**N**IV**K**K**TE**V**Q**T**G
GFS**K**ES**IL**PK**R**NS**D**K**L**IAR**K**DK**D**W**D**PK**Y**GG**F**D**SPT**V**A**YS**V**LV**V**AK**VE**G**K**SK**L**KS**V**K**ELL**G**IT**IM**ER**SS**F**E**K**NP**ID**FL**E**AK**Y**KE**V**KK**D**L
IIK**L**PK**Y**SL**FE**LE**NG**R**K**R**M**L**A**S**A**G**E**L**Q**K**G**N**E**L**AL**PS**KY**V**N**FLY**L**ASHY**E**KL**G**SP**E**D**NE**Q**K**QL**F**VE**Q**HK**H**YL**DEI**IE**Q**IS**E**FS**K**R**V**IL**A**D**A**
NL**D**K**V**LS**A**YN**K**HR**D**K**P**IRE**Q**A**ENI**I**HL**FT**L**T**N**L**G**AP**A**AF**KY**FD**TTID**R**K**RY**T**ST**K**EV**L**D**A**T**LI**H**Q**S**IT**GLY**E**TR**ID**LS**Q**LG**G**D**KR**PA**AT**KK**AGL**
AGL**SGSETPGTSESATPES**GRMV**SK**GEED**NMAI**I**KE**FM**R**FK**V**HMEGS**V**NG**HE**FE**IE**GE**G**EG**R**PY**E**GT**Q**AK**L**K**V**T**K**GG**P**LP**FA**WD**IL**SP**Q**F
MYGSKAYVKHPADIPDYLKLSFPEGFKWERMNFEDGGVVTVTQDSSLQDGEFTYKVKLRGNTNFPSDGPVMQKKTMGWEASSERMPEDGALKGEIK
QRLKLKDGGHYDAEVKTTYKAKKPVQLPGAYNVNIKLDITSHNEDYTIVEQYERAEGRHSTGGMDELYKKRPAATKKAGQAKKKK
 *

NN Igk, murine Immunoglobulin kappa-chain signal peptide

NN (GGGS)₂ linker

NN Myc epitope tag

NN TM (PDGFR), transmembrane domain from platelet derived growth factor receptor

NN XTEN linker

NN efficient TCS; tobacco etch virus (TEV) protease cleavage sequence

NN FLAG epitope tag

NN NLS(+), SV40 nuclear localisation sequence

NN catalytically active Cas9

NN NLS, nucleoplasmin NLS

NN XTEN linker

NN mCherry

TMt^{ΔTCS}-NLS-Cas9-mCherry used in *pTMt_FLAG-NLS-Cas9-XTEN-mCherry-NLS*

METDTLLLWVLLLWVPGSTGDHS **GGGSGGGSGRQE**QKLISEEDL**NAV**QDQTQEVIVVPHSLPFKVVVISAILALVVLTIISLIILIMLWQK
KPRL**QSGSETPGTSESATPES**ASHVDHAAAT**GAT**MDYKDHDGDYKDHDIDYKDDDD**KMAPKKR****KRVGIHGVPAADKKYSIGLDIGTNSV**
AVITDEY**KVP**SK**KFKVL**GN**TD**RHS**IKKNLI**GALL**FD**SG**ETAE**AT**RLKRTARRRYTRRKNRI**CYL**QE**IF**SNEMAKVDD**SFFH**RL**E**S**FL**VEE**
DKK**HERHPI**FG**NI**VDE**VAYHEKYPTI**YHL**RKKLV**D**STDKAD**LR**LIY**LALAHMI**KFRGHFL**IEG**DLNPD**NSD**VDKLF**IQL**VQ**TY**N**QL**FEEN**P
INAS**GVDAKAILSARLSKSRRLEN**LIAQL**PGEKK**NGL**FGN**LIALSLGL**TPNF**KS**NFD**LAEDAK**LQLSKD**TYDD**DLN**LLAQ**IGDQYAD**LF**L**
AAKNLS**DAILL**SD**ILRVN**TEIT**KAPLS**ASMI**KRYDEHHQD**L**TLLKAL**V**RQQLPEKYKE**IFF**DQ**SKNGYAGYID**GGASQ**EEFY**KF**IK**PILE**K
MDGTEELL**VKLN**RED**LLRQRT**FD**NGSI**PH**QI**HL**GELHAILRRQ**EDFY**PFLK**DN**REKIEKIL**TF**RI**PYY**VG**PL**ARGNSR**FAW**MT**R**K**SEETI
TPWN**FE**EV**VDK**G**ASAQ**S**FIER**MT**NF**DK**NLPNEK**VLP**KHS**LLY**EYFTVY**NEL**TKVY**VTE**GM**R**KPA**FL**S**GE**QK**KA**IVD**LL**F**KT**N**R**KV**TV**K**Q**L**
KEDY**FKKIE**CFDS**VEIS**GV**ED**RF**NAS**L**GTYH**DL**LKIIK**DK**DF**L**DNE**ENED**I**LED**IVL**TL**TL**LF**ED**REMI**EE**RL**KTYA**HL**F**DD**KV**M**K**Q**L**KRRR
YTGW**GR**LS**RKL**ING**IR**DK**QSG**KT**IL**DF**LK**SD**G**FAN**RNF**Q**LI**HDD**SL**TF**KED**I**Q**KA**Q**VS**GQ**GD**SL**HEH**IAN**LAG**S**PA**IKK**G**IL**Q**TV**K**V**DE
LVK**V**MGR**HK**PEN**IV**IE**MARE**N**QTTQ**KG**Q**KS**R**ERM**KRI**EE**G**IK**EL**GS**QIL**KE**HP**VENT**QLQ**NE**KLY**LY**YLQ**NGR**DMY**VD**Q**EL**D**IN**RL**SDY**D**
VDH**IV**P**Q**S**FL**KDD**S**ID**NK**VL**TR**SD**K**NR**GK**SD**N**VP**SE**EV**VK**M**K**NY**WR**QL**NA**KL**ITQ**R**K**FD**N**L**T**KA**ERG**GL**S**EL**D**KAG**F**IK**R**Q**L**V**ET**R**QIT**
KHVA**Q**IL**D**SR**M**NT**KY**DEND**KLIRE**V**KVIT**L**K**SK**L**VS**D**FR**KD**F**Q**FY**K**VR**E**IN**NY**HH**A**DAY**L**NA**V**VG**T**AL**IK**KY**P**KL**E**S**E**FY**G**DY**K**YD**V**R
KMI**A**K**SE**QE**IG**K**A**T**A**KY**FF**YS**NIM**N**FF**K**TEIT**L**ANGE**IR**K**R**PL**IE**T**NG**ET**GE**IV**WD**K**GR**D**FAT**VR**K**V**LS**MP**Q**V**N**IV**K**K**TE**V**Q**T**GG**F**SK**ES**I
LPK**R**NS**D**K**L**IAR**K**DK**D**W**D**PK**Y**GG**F**D**SPT**V**A**YS**V**LV**V**AK**VE**G**K**SK**L**KS**V**K**ELL**G**IT**IM**ER**SS**F**E**K**NP**ID**FL**E**AK**Y**KE**V**KK**D**L**I**IK**L**PK**Y**
SLFE**LE**NG**R**K**R**M**L**A**S**A**G**E**L**Q**K**G**N**E**L**AL**PS**KY**V**N**FLY**LASHY**E**KL**G**SP**E**D**NE**Q**K**QL**F**VE**Q**HK**H**YL**DEI**IE**Q**IS**E**FS**K**R**V**IL**A**D**A**N**L**D**K**VLS
AYN**K**HR**D**K**P**IRE**Q**A**ENI**I**HL**FT**L**T**N**L**G**AP**A**AF**KY**FD**TTID**R**K**RY**T**ST**K**EV**L**D**A**T**LI**H**Q**S**IT**GLY**E**TR**ID**LS**Q**LG**G**D**KR**PA**AT**KK**AGL****SGSE**
TPGTSESATPESGRMV**SK**GEED**NMAI**I**KE**FM**R**FK**V**HMEGS**V**NG**HE**FE**IE**GE**G**EG**R**PY**E**GT**Q**AK**L**K**V**T**K**GG**P**LP**FA**WD**IL**SP**Q**F**MYG**SK**A**Y
VK**H**PA**D**IP**DY**L**K**LS**F**PE**G**FK**W**ERM**N**F**E**D**G**GV**V**TV**TQ**D**S**SL**Q**D**G**E**F**TY**K**V**K**LR**G**NT**N**FP**S**D**G**P**V**M**Q**K**K**T**M**G**W**E**A**S**S**ER**M**P**E**D**G**AL**K**GE**I**K**Q**
RLKL**K**D**G**GHY**DA**EV**K**TTY**K**AK**K**P**V**QL**P**G**A**Y**N**V**N**IK**L**D**IT**SH**N**ED**YT**IVE**Q**Y**E**R**A**E**G**R**H**ST**G**G**M**D**E**LY**K**K**R**PA**AT**KK**AG**Q**AK**KK**K***

NN Igk, murine Immunoglobulin kappa-chain signal peptide

NN (GGGS)₂ linker

NN Myc epitope tag

NN TM (PDGFR), transmembrane domain from platelet derived growth factor receptor

NN XTEN linker

NN FLAG epitope tag

NN NLS(+), SV40 nuclear localisation sequence

NN catalytically active Cas9

NN NLS, nucleoplasmin NLS

NN XTEN linker

NN mCherry

Tmt-Cas9-mCherry used in *pTmt_TCS(Q'G)_Cas9-XTEN-mCherry*

METDTLLLLWVLLWVPGSTGDHSGGGSGGGSGRQEQKLISEEDLNAVQDQTQEVIVVPHSLPFKVVVISAILALVVLTIISLIILIMLWQK
KPRLQSGSETPGTSESATPESASHVDHAAAENLYFQGTGMDKKYSIGLDIGTNSVGWAVITDEYKVPSSKKFVLGNDRHSIKKNLIGALL
FDSGTEAETRLKRTARRRYTRRKNRICYLQEIFSNEMAKVDDSFHRLEESFLVEEDKKHERHPIFGNIVDEVAYHEKYPTIYHLRKKLIV
DSTDKADRLRIYLALAHMIKFRGHFLIEGDLNPDNSDVDKLFIQLVQTYNQLFEEENPINASGVDAKAILSARLSKSRRLENLIAQLPGKEK
NGLFGNLIASLSGLTPNFKSNFDLAEDAKLQLSKDITYDDDLNLLAQIGDQYADLFLAAKNLSDAILLSDILRVNTEITKAPLSASMIKRY
DEHHQDITLLKALVRQQLPEKYKEIFFDQSKNGYAGYIDGGASQEEFYKFIKPILEKMDGTEELLVKLNREDLLRKQRTFDNGSIPHQIHL
GELHAILRRQEDFYFPLKDNREKIEKILTFRIPIYYVGPLARGNSRFAMWTRKSEETITPWNFEVVVDKGASAQSFIERMTNFDKNLPNEKV
LPKHSLLYEYFTVYNELTKVKYVTEGMRKPAFLSGEQKKAIVDLLFKTNRKVTVKQLKEDYFKKIECFDSVEISGVEDRFNASLGTYHDL
KI IKDKDFLDNEENEDILEDIVLTLTLFEDREMIEERLKYAHLFDDKVMQQLKRRRYTGWGRLSRKLINGIRDKQSGKTI LDFLKS
NRNFMQLIHDDSLTFKEDIQKAQVSGQGDSLHEHIANLAGSPAIAKKGILQTVKVVDELVKVMGRHKPENIVEMARENQTTQKGQKNSRER
MKRIEIEGKELGSQLI KEHPVENTQLQNEKLYLYYLQNGRDMYVDQELDINRLSDYDVVDHIVPQSFLKDDSIDNKVLRSDKNRGKSDNVP
SEEVVKMKMNYWRQLLNAKLITQRKFDNLTKAERGGLSELDKAGFIKRQLVETRQITKHVAQILDSRMNTKYDENDKLIREVKVITLKS
VSDFRKDFQFYKVRINNYHHAHDAYLNAVVGTA LIKKYPKLESEFVYGDYKVDVRKMIKSEQIEGKATAKYFFYSNIMNFFKTEITLA
NGEIRKRP LIETNGETGEIVWDKGRDFATVRKVL SMPQVNI VKKTEVQTGGFSKESILPKRNSDKLIARKKDWDPKKYGGFDSPTVAYS
VVAKVEKGKSKKLKSVKELLGITIMERSSEFKNPIDFLEAKGYKEVKKDLI IKLPKYSLFELENGRKRMLASAGELQKGNELALPSKYVNF
LYLASHYEKLGKSPEDNEQQLFVEQHKHYLDEII EQISEFSKRVLADANL DKVLSAYNKHHRDKPIREQAENI IHLFTLTNLGAPAAFKY
FDTTIDRKRYTSTKEVLDATLIHQ SITGLYETRIDLSQLGGDASGSETPGTSESATPESGRMVSKGEEDNMAI IKFEMRFKVMHEGVSNG
 HEFEIEGEGEGRPYEGTQTA KLKVTGKGPLPFAWDILSPQFMYGSKAYVKHPADIPDYLKLSFPEGFKWERVMNFEDGGVVTVTQDSSLQD
 GEFIYKVKLRGTNFPSDGPMQKKTMGWEASSERMPEDGALKGEIKQRLKLDGGHYDAEVKTTYAKKPVQLPGAYNVNIKLDITSHNE
 DYTIVEQYERAEGRHSTGGMDELYK*

NN Igk, murine Immunoglobulin kappa-chain signal peptide

NN (GGGS)₂ linker

NN Myc epitope tag

NN TM (PDGFR), transmembrane domain from platelet derived growth factor receptor

NN XTEN linker

NN efficient TCS; tobacco etch virus (TEV) protease cleavage sequence

NN catalytically active Cas9

NN mCherry

A:3 Split Tmt-dCas9^{VP64} and associated constructsNLS(+)-dCas9(C)^{VP64}-FKBP used in *pNLS(+)-FKBP-dCas9(C)-VP64*

MAPPKKRKVGIHGVPAAGVQVETISPGDGRFTPKRGQTCVHYTGMLLEDGKKKFDSSRDNRNPKPFKMLGKQEVIRGWEEGVAQMSVGQRAKL
TISPDYAYGATGHGPIIPPHATLVFDVLELLKLEGGGSGGGSGGGGSKPAFLSGEQKKAIVDLLFKTNRKVTVKQLKEDYFKKIECFDSV
EISGVEDRFNASLGTYHDLKI IKDKDFLDNEENEDILEDIVLTLTLFEDREMIEERLKYAHLFDDKVMQQLKRRRYTGWGRLSRKLING
IRDKQSGKTI LDFLKSDGFANRNMQLIHDDSLTFKEDIQKAQVSGQGDSLHEHIANLAGSPAIAKKGILQTVKVVDELVKVMGRHKPENIV
EMARENQTTQKGQKNSRERMKRIEIEGKELGSQLI KEHPVENTQLQNEKLYLYYLQNGRDMYVDQELDINRLSDYDVVDHIVPQSFLKDD
IDNKVLRSDKARGKSDNVPSEEVVKMKMNYWRQLLNAKLITQRKFDNLTKAERGGLSELDKAGFIKRQLVETRQITKHVAQILDSRMNTK
YDENDKLIREVKVITLKSKLVSDFRKDFQFYKVRINNYHHAHDAYLNAVVGTA LIKKYPKLESEFVYGDYKVDVRKMIKSEQIEGKAT
AKYFFYSNIMNFFKTEITLANGEIRKRP LIETNGETGEIVWDKGRDFATVRKVL SMPQVNI VKKTEVQTGGFSKESILPKRNSDKLIARKK
DWDPKKYGGFDSPTVAYSSVLVVAKVEKGKSKKLKSVKELLGITIMERSSEFKNPIDFLEAKGYKEVKKDLI IKLPKYSLFELENGRKRMLA
SAGELQKGNELALPSKYVNFLYLASHYEKLGKSPEDNEQQLFVEQHKHYLDEII EQISEFSKRVLADANL DKVLSAYNKHHRDKPIREQA
ENI IHLFTLTNLGAPAAFKYFDTTIDRKRYTSTKEVLDATLIHQ SITGLYETRIDLSQLGGDSPKKRRKVEASGRADALDDFDLMLGSDA
LDDFDLMLGSDALDDFDLMLGSDALDDFDLMLIN*

NN NLS(+), SV40 nuclear localisation sequence

NN FKBP12, FK506 binding protein 12

NN (GGGS)₃ linker

NN dCas9(C), C-terminal moiety of human codon-optimized *Streptococcus pyogenes* Cas9

NN NLS, SV40 nuclear localisation sequence

NN VP64, transcriptional activator domain

Tmt-NLS-dCas9(C)^{VP64}-FKBP used in *pTmt_TCS(Q'G)_NLS-HA-FKBP-dCas9(C)-VP64*

METDTLLLLWVLLWVPGSTGDHSGGGSGGGSGRQEQKLISEEDLNAVQDQTQEVIVVPHSLPFKVVVISAILALVVLTIISLIILIMLWQK
KPRLQSGSETPGTSESATPESASHVDHAAAENLYFQGPKKRKVGGGSTSYPYDVPDYAGGSTGGVQVETISPGDGRFTPKRGQTCVHYT
GMLLEDGKKKFDSSRDNRNPKPFKMLGKQEVIRGWEEGVAQMSVGQRAKLITISPDYAYGATGHGPIIPPHATLVFDVLELLKLEGGGSGGGGSG
GGGSKPAFLSGEQKKAIVDLLFKTNRKVTVKQLKEDYFKKIECFDSVEISGVEDRFNASLGTYHDLKI IKDKDFLDNEENEDILEDIVLT
LTLFEDREMIEERLKYAHLFDDKVMQQLKRRRYTGWGRLSRKLINGIRDKQSGKTI LDFLKSDGFANRNMQLIHDDSLTFKEDIQKAQV
SGQDLSLHEHIANLAGSPAIAKKGILQTVKVVDELVKVMGRHKPENIVEMARENQTTQKGQKNSRERMKRIEIEGKELGSQLI KEHPVENT
QLQNEKLYLYYLQNGRDMYVDQELDINRLSDYDVVDHIVPQSFLKDDSIDNKVLRSDKARGKSDNVPSEEVVKMKMNYWRQLLNAKLITQR
KFDNLTKAERGGLSELDKAGFIKRQLVETRQITKHVAQILDSRMNTKYDENDKLIREVKVITLKSKLVSDFRKDFQFYKVRINNYHHAHD
AYLNAVVGTA LIKKYPKLESEFVYGDYKVDVRKMIKSEQIEGKATAKYFFYSNIMNFFKTEITLANGEIRKRP LIETNGETGEIVWDKGR
DFATVRKVL SMPQVNI VKKTEVQTGGFSKESILPKRNSDKLIARKKDWDPKKYGGFDSPTVAYSSVLVVAKVEKGKSKKLKSVKELLGITI
MERSSEFKNPIDFLEAKGYKEVKKDLI IKLPKYSLFELENGRKRMLASAGELQKGNELALPSKYVNFLYLASHYEKLGKSPEDNEQQLFV
EQHKHYLDEII EQISEFSKRVLADANL DKVLSAYNKHHRDKPIREQAENI IHLFTLTNLGAPAAFKYFDTTIDRKRYTSTKEVLDATLIHQ
SITGLYETRIDLSQLGGDSPKKRRKVEASGRADALDDFDLMLGSDALDDFDLMLGSDALDDFDLMLGSDALDDFDLMLIN*

NN Igk, murine Immunoglobulin kappa-chain signal peptide

NN GGS linker

NN Myc epitope tag

Appendix

NN TM (PDGFR), transmembrane domain from platelet derived growth factor receptor

NN XTEN linker

NN efficient TCS; tobacco etch virus (TEV) protease cleavage sequence

NN NLS, SV40 nuclear localisation sequence

NN HA, hemagglutinin A epitope

NN FKBP12, FK506 binding protein 12

NN (GGGS)₃ linker

NN dCas9(C), C-terminal moiety of human codon–optimized *Streptococcus pyogenes* Cas9

NN VP64, transcriptional activator domain

NES-dCas9(N)-FRB used in *pNES-dCas9(N)-FRB*

MLDLASLILDKKYSIGLAIGTNSVGWAVITDEYKVPSSKKFKVLGNTDRHSIKKNLIGALLFDSGETAEATRLKRTARRRYTRRKNRICYLQEIFSNEMAKVDDSDFFHRLEESFLVEEDKKHERHPIFGNIVDEVAYHEKYPTIYHLRKKLV DSTDKADLRLIYLALAHMIKFRGHFLIEGDLNPDNSDVKLFIQLVQTYNQLFEEENPINASGVDAKAILSARLSKSRRLLENLIAQLPGEKKNGLFGNLIASLGLTPNFKSNFDLAEDAKLQLSKDTYDDDLNLLAQIGDQYADLFLAAKNLSDAILLSDILRVNTEITKAPLSASMIKRYDEHHQDLTLLKALVRQQLPEKYKEIFFDQSKNGYAGYIDGGASQEEFYKFIKPILEKMDGTEELLVKLNREDLLRKQRTFDNGSIPHQIHLGELHAILRRQEDFYFPLKDNREKIEKILTFRIPYYVGPLARGNSRFAMWTRKSEETITPWNFEEVVDKGASAQSFIERMTNFDKNLPNEKVLPKHSLLYEYFTVYNELTKVKYVTEGMRGGGGGGGGGGGSEMWHEGLEEASRLYFGERNVKGMEFVLEPLHAMMERGPQTLKETSFNQAYGRDLMEAQEWCRKYMKSGNVKDLTQAWDLYYHVFRRISKQLDLASLIL*

NN NES, nuclear export signal form human protein tyrosine kinase 2

NN dCas9(N), N-terminal moiety of human codon–optimized *Streptococcus pyogenes* Cas9

NN (GGGS)₃ linker

NN FKBP rapamycin binding domain (FRB)

dCas9(N)-FRB used in *pdCas9(N)-FRB*

MDDKYSIGLAIGTNSVGWAVITDEYKVPSSKKFKVLGNTDRHSIKKNLIGALLFDSGETAEATRLKRTARRRYTRRKNRICYLQEIFSNEMAKVDDSDFFHRLEESFLVEEDKKHERHPIFGNIVDEVAYHEKYPTIYHLRKKLV DSTDKADLRLIYLALAHMIKFRGHFLIEGDLNPDNSDVKLFIQLVQTYNQLFEEENPINASGVDAKAILSARLSKSRRLLENLIAQLPGEKKNGLFGNLIASLGLTPNFKSNFDLAEDAKLQLSKDTYDDDLNLLAQIGDQYADLFLAAKNLSDAILLSDILRVNTEITKAPLSASMIKRYDEHHQDLTLLKALVRQQLPEKYKEIFFDQSKNGYAGYIDGGASQEEFYKFIKPILEKMDGTEELLVKLNREDLLRKQRTFDNGSIPHQIHLGELHAILRRQEDFYFPLKDNREKIEKILTFRIPYYVGPLARGNSRFAMWTRKSEETITPWNFEEVVDKGASAQSFIERMTNFDKNLPNEKVLPKHSLLYEYFTVYNELTKVKYVTEGMRGGGGGGGGGGGSEMWHEGLEEASRLYFGERNVKGMEFVLEPLHAMMERGPQTLKETSFNQAYGRDLMEAQEWCRKYMKSGNVKDLTQAWDLYYHVFRRISKQ*

NN dCas9(N), N-terminal moiety of human codon–optimized *Streptococcus pyogenes* Cas9

NN (GGGS)₃ linker

NN FKBP rapamycin binding domain (FRB)

TMt^{NES}-dCas9(N)-FRB used in *pTMt_{NES}_TCS(Q'G)_HA-dCas9(N)-FRB*

METDTLLLWVLLWVPGSTGDHSGGGSGGGSGRQEQKLISEEDLNAVQDQTQEVIVVPHSLPFKVVVISAILALVVLTIISLIILIMLWQK KPRLQSGSETPGTSESATPESASLDLASLILGKLGENLYFQGGGGSTSYPYDVDPYAGGGSGSDKKYSIGLAIGTNSVGWAVITDEYKVPSSKKFKVLGNTDRHSIKKNLIGALLFDSGETAEATRLKRTARRRYTRRKNRICYLQEIFSNEMAKVDDSDFFHRLEESFLVEEDKKHERHPIFGNIVDEVAYHEKYPTIYHLRKKLV DSTDKADLRLIYLALAHMIKFRGHFLIEGDLNPDNSDVKLFIQLVQTYNQLFEEENPINASGVDAKAILSARLSKSRRLLENLIAQLPGEKKNGLFGNLIASLGLTPNFKSNFDLAEDAKLQLSKDTYDDDLNLLAQIGDQYADLFLAAKNLSDAILLSDILRVNTEITKAPLSASMIKRYDEHHQDLTLLKALVRQQLPEKYKEIFFDQSKNGYAGYIDGGASQEEFYKFIKPILEKMDGTEELLVKLNREDLLRKQRTFDNGSIPHQIHLGELHAILRRQEDFYFPLKDNREKIEKILTFRIPYYVGPLARGNSRFAMWTRKSEETITPWNFEEVVDKGASAQSFIERMTNFDKNLPNEKVLPKHSLLYEYFTVYNELTKVKYVTEGMRGGGGGGGGGGGSEMWHEGLEEASRLYFGERNVKGMEFVLEPLHAMMERGPQTLKETSFNQAYGRDLMEAQEWCRKYMKSGNVKDLTQAWDLYYHVFRRISKQ*

NN Igκ, murine Immunoglobulin kappa-chain signal peptide

NN GGGS linker

NN Myc epitope tag

NN TM (PDGFR), transmembrane domain from platelet derived growth factor receptor

NN XTEN linker

NN efficient TCS; tobacco etch virus (TEV) protease cleavage sequence

NN HA, hemagglutinin A epitope

NN NES, nuclear export signal form human protein tyrosine kinase 2

NN dCas9(N), N-terminal moiety of human codon–optimized *Streptococcus pyogenes* Cas9

NN (GGGS)₃ linker

NN FKBP rapamycin binding domain (FRB)

TMt^{NES}-dCas9(N) used in *pTMt_{NES}_TCS(Q'G)_HA-dCas9(N)_P2A-Puro-WPRE*

METDTLLLWVLLWVPGSTGDHSGGGSGGGSGRQEQKLISEEDLNAVQDQTQEVIVVPHSLPFKVVVISAILALVVLTIISLIILIMLWQK KPRLQSGSETPGTSESATPESASLDLASLILGKLGENLYFQGGGGSTSYPYDVDPYAGGGSGSDKKYSIGLAIGTNSVGWAVITDEYKVPSSKKFKVLGNTDRHSIKKNLIGALLFDSGETAEATRLKRTARRRYTRRKNRICYLQEIFSNEMAKVDDSDFFHRLEESFLVEEDKKHERHPIFGNIVDEVAYHEKYPTIYHLRKKLV DSTDKADLRLIYLALAHMIKFRGHFLIEGDLNPDNSDVKLFIQLVQTYNQLFEEENPINASGVDAKAILSARLSKSRRLLENLIAQLPGEKKNGLFGNLIASLGLTPNFKSNFDLAEDAKLQLSKDTYDDDLNLLAQIGDQYADLFLAAKNLSDAILLSDILRVNTEITKAPLSASMIKRYDEHHQDLTLLKALVRQQLPEKYKEIFFDQSKNGYAGYIDGGASQEEFYKFIKPILEKMDGTEELLVKLN

Appendix

[REDLLRKQRTFDNGSIPHQIHLGELHAILRRQEDFYFPLKDNREKIEKILTFRIPYYVGPLARGNSRFAMWTRKSEETITPWNFEFVVDKG](#)
[ASAQSFIERMTNFDKNLPNEKVLPHKSLLYEYFTVYNELTKVKYVTEGMRGGGSGTGSGATNFSLLKQAGDVEENPGPEFMTFYEYKPTVRL](#)
[ATRDDVPRAVRTLAAAFADYPATRHVTDPDRHIERTVTELQELFLTRVGLDIGKVVWADDGAAVAVWTTPEVSVEAGAVFAEIGPRMAELSGS](#)
[RLAAQQQMEGLLAPHRPKEPAWFLATVGVSPDHQGGKGLGSVVLPGEAAERAGVPAFLETSAPRNLPFYERLGFVTADVEVEPEGPRTWC](#)
[MTRKG*](#)

NN Igκ, murine Immunoglobulin kappa-chain signal peptide

NN (GGGS)₂ linker

NN Myc epitope tag

NN TM (PDGFR), transmembrane domain from platelet derived growth factor receptor

NN XTEN linker

NN NES, nuclear export signal form human protein tyrosine kinase 2

NN efficient TCS; tobacco etch virus (TEV) protease cleavage sequence

NN HA, hemagglutinin A epitope

NN dCas9(N), N-terminal moiety of human codon–optimized *Streptococcus pyogenes* Cas9

NN P2A, 2A self-cleaving peptide from porcine teschovirus-1

NN Puromycin resistance protein

TMt-NLS-dCas9(C)^{VP64} used in *pTmt_TCS(Q'G)_NLS-HA-dCas9(C)-VP64_T2A_MCP-P65-HSF1*

[METDTLLLWVLLWVPGSTGDSGGSGSGSGRQEQKLISEEDLNAVQDQTQEVIVVPHSLPFKVVVISAILALVVLTIISLIILIMLWQK](#)
[KPRLQSGSETPGTSESATPESASHVDHAAENLYFQCPKKRRKVGGGSTSYYPDVDPDYAGGSGSGSGGSKPAFLSGEQKKAIVDLLFKTNRK](#)
[VTVKQLKEDYFKKIECFDSVEISGVEDRFNLSLGTYHDLKIKDKDFLDNEENEDILEDIVLTLTLFEDREMIEERLKYAHLFDDKVMK](#)
[QLKRRRYTGWGRLSRKLINGIRDKQSGKTILDFLKSDFGANRNFQMQLIHDDSLTFKEDIQKAQVSGQGDSLHEHIANLAGSPAIKKGILQT](#)
[VKVVDLVKVMGRHKPENIVEMARENQTTQKGQKNSRERMKRIEIGIKELGSQILKEHPVENTQLQNEKLYLYYLQNGRDMYVDQELDIN](#)
[RLSDYDVHIVPQSFLKDDSIDNKVLTRSDKARGKSDNVPSEEVVKMKMNYWRQLLNAKLITQRKFDNLTKAERGGLSELDKAGFIKRQLV](#)
[ETRQITKHVAQILDSRMNTKYDENDKLIREVKVITLKSCLVSDFRKDFQFYKVRINNYHHAHDAYLNAVVGITALIKKYPKLESEFVYGDY](#)
[KVYDVRKMIKSEQEI GKATAKYFFYSNIMNFFKTEITLANGEIRKRPLIETNGETGEIVWDKGRDFATVRKVLSPMPQVNVKKTEVQTGG](#)
[FSKESILPKRNSDKLIARKKDWDPKYGGGFSPTVAYSVLVAKVEKGKSKLKSVKELLGITIMERSSSFENPIDFLEAKGYKEVKKDLI](#)
[IKLPKYSLFELNGRKRMLASAGELQKGNELALPSKYVNFYLLASHYEKLGKSPEDNEQKQLFVEQHKHYLDEIIIEQISEFSKRVLADAN](#)
[LDKVL SAYNKHRRDKPIREQAENI IHLFTLTNLGAPAAFKYFDTTIDRKRYTSTKEVLDTLHQSTITGLYETRIDLSQLGGDSPKKRRKVE](#)
[ASGRADALDDFDLMDLGS DALDDFDLMDLGS DALDDFDLMDLGS DALDDFDLMDLIN](#)[GTASGSGEGRGSLLTCGDVEENPGPVSKLMASNF](#)
[TQFVLVDNGGTGDVTVAPSNFANGVAEWISSNSRSQAYKVTCSVRQSSAQKRKYTIKVEVPKVATQTVGGVELPVAAWRSYLNMELTIPIF](#)
[ATNSDCELI VKAMQGLLKDGNPISAI AANS GIYSAGGGSGGGSGGGSGGSGPKKKRKVAAAGSPSGQISNQALALAPSSAPVLAQTMVPS](#)
[SAMVPLAQPPAPVLTTPGPPQSL SAPVPKSTQAGECTLSEALLHLQFDAEDL GALLGNSTDPGVFTDLASVDNSEFQQLLNQGVSMSSH](#)
[TAEPMLMEYPEAITRLVTGSQRPPDPAPTPLGTSGLPNGLSGDEDFSSIADMDFSALLSQISSSGQGGGSGFSVDTSALLDLFSPSVTVP](#)
[DMSLPDLSSLASIQELLS PQEPPRPPEAENS SPDSGKQLVHYTAQPLFLLDPGSVDTGSNDLPVL FELGEGSYFSEGDGFAEDPTISLLT](#)
[GSEPPKAKDPTVS*](#)

NN Igκ, murine Immunoglobulin kappa-chain signal peptide

NN (GGGS)₂ linker

NN Myc epitope tag

NN TM (PDGFR), transmembrane domain from platelet derived growth factor receptor

NN XTEN linker

NN efficient TCS; tobacco etch virus (TEV) protease cleavage sequence

NN NLS, SV40 nuclear localisation sequence

NN HA, hemagglutinin A epitope

NN dCas9(C), C-terminal moiety of human codon–optimized *Streptococcus pyogenes* Cas9

NN VP64, transcriptional activator domain

NN T2A, 2A self-cleaving peptide from thosea asigna virus

NN MCP (MS2 protein), MS2 bacteriophage coat protein

NN P65, activation domain from human NF-κB trans-activating subunit p65

NN HSF1, activation domains from human heat-shock factor 1

NLS-dCas9(C)^{VP64} used in *pNLS-HA-dCas9(C)-VP64_T2A_MCP-P65-HSF1*

[MPKKRRKVGGGSTSYYPDVDPDYAGGSGSGGGSKPAFLSGEQKKAIVDLLFKTNRKVTVKQLKEDYFKKIECFDSVEISGVEDRFNLSLGT](#)
[YHDLKIKDKDFLDNEENEDILEDIVLTLTLFEDREMIEERLKYAHLFDDKVMKQLKRRRYTGWGRLSRKLINGIRDKQSGKTILDFLKS](#)
[DGFANRNFQMQLIHDDSLTFKEDIQKAQVSGQGDSLHEHIANLAGSPAIKKGILQTVKVVDELVKVMGRHKPENIVEMARENQTTQKGQKN](#)
[SRERMKRIEIGIKELGSQILKEHPVENTQLQNEKLYLYYLQNGRDMYVDQELDINRLSDYDVHIVPQSFLKDDSIDNKVLTRSDKARGKS](#)
[DNVPSEEVVKMKMNYWRQLLNAKLITQRKFDNLTKAERGGLSELDKAGFIKRQLVETRQITKHVAQILDSRMNTKYDENDKLIREVKVITL](#)
[KSKLVSDFRKDFQFYKVRINNYHHAHDAYLNAVVGITALIKKYPKLESEFVYGDYKVYDVRKMIKSEQEI GKATAKYFFYSNIMNFFKTE](#)
[ITLANGEIRKRPLIETNGETGEIVWDKGRDFATVRKVLSPMPQVNVKKTEVQTGGFSKESILPKRNSDKLIARKKDWDPKKYGGFDSPTVA](#)
[YSVLVAKVEKGKSKLKSVKELLGITIMERSSSFENPIDFLEAKGYKEVKKDLI IKLPKYSLFELNGRKRMLASAGELQKGNELALPSK](#)
[YVNFYLLASHYEKLGKSPEDNEQKQLFVEQHKHYLDEIIIEQISEFSKRVLADANLDKVL SAYNKHRRDKPIREQAENI IHLFTLTNLGAP](#)
[AFKYFDTTIDRKRYTSTKEVLDTLHQSTITGLYETRIDLSQLGGDSPKKRRKVEASGRADALDDFDLMDLGS DALDDFDLMDLGS DALDD](#)
[FDLMDLGS DALDDFDLMDLIN](#)[GTASGSGEGRGSLLTCGDVEENPGPVSKLMASNF](#)
[TQFVLVDNGGTGDVTVAPSNFANGVAEWISSNSRSQAYKVTCSVRQSSAQKRKYTIKVEVPKVATQTVGGVELPVAAWRSYLNMELTIPIF](#)
[ATNSDCELI VKAMQGLLKDGNPISAI AANS GIYSAGGGSGGGSGGGSGGSGPKKKRKVAAAGSPSGQISNQALALAPSSAPVLAQTMVPS](#)
[SAMVPLAQPPAPVLTTPGPPQSL SAPVPKSTQAGECTLSEALLHLQFDAEDL GALLGNSTDPGVFTDLASVDNSEFQQLLNQGVSMSSH](#)
[TAEPMLMEYPEAITRLVTGSQRPPDPAPTPLGTSGLPNGLSGDEDFSSIADMDFSALLSQISSSGQGGGSGFSVDTSALLDLFSPSVTVP](#)
[DMSLPDLSSLASIQELLS PQEPPRPPEAENS SPDSGKQLVHYTAQPLFLLDPGSVDTGSNDLPVL FELGEGSYFSEGDGFAEDPTISLLT](#)
[GSEPPKAKDPTVS*](#)

NN NLS, SV40 nuclear localisation sequence

NN HA, hemagglutinin A epitope

Appendix

NN dCas9(C), C-terminal moiety of human codon–optimized *Streptococcus pyogenes* Cas9
NN VP64, transcriptional activator domain
NN T2A, 2A self-cleaving peptide from thosea asigna virus
NN MCP (MS2 protein), MS2 bacteriophage coat protein
NN P65, activation domain from human NF-κB trans-activating subunit p65
NN HSF1, activation domains from human heat-shock factor 1

NLS(+)-dCas9(C)VP64 used in pNLS(+)-HA-dCas9(C)-VP64_T2A_MCP-P65-HSF1

MA**PKKKRKVG**I**HGVPAA**GGGSGTSYPYDVDPDYAGGGSGGGGSKPAFLSGEQKKAIVDLLFKTNRKVTVKQLKEDYFKKIECFDSVEISGVED
RFNASLGTYYHDLKKIKDKDFLDNEENEDILEDIVLTTLTFEDREMIEERLKYAHLFDDKVMKQLKRRRYTGWGRLSRKLINGIRDKQSG
KTILDFLKSDFANRNFMLHDDSLTFKEDIQKAQVSGQDLSLHEHIANLAGSPAIKKGILQTVKVVDELVKVMGRHKPENIVIEAREN
OTTQKGQKNSRERMKRIEELGKELGSQLKEHPVENTQLQNEKLYLYLLQNGRDMYVDQELDINRLSDYDVDHIVPQSFLKDDSIDNKVLT
RSDKARGKSDNVPSEEVVKKMKNYWRQLLNAKLITQRKFDNLTKAERGGLSELDKAGFIKRQLVETRQITKHVAQILDSRMNTKYDENDKL
IREVKVITLKSCLVSDFRKDFQFYKVIENNYHHAHDAYLNAVVGTLIKKYPKLESEFVYGDYKVYDVRKMIKAKSEQEIGKATAKYFFYS
NIMNFFKTEITLANGEIRKRPLIETNGETGEIVWDKGRDFATVRKVLSPQVNIKKTEVQTTGGFSKESILPKRNSDKLIARKDWDPKKY
GGFDSPTVAYSVLVAKVEKGSKKLKSVELLGITIMERSSEFKNPIDFLEAKGYKEVKKDLIIKLPKYSLELENGRKRMLASAGELQK
GNELALPSKYVNFYLLASHYEKLKGSPEDEQKQLFVEQHKHYLDEIIIEQISEFSKRVLADANLKDVLISAYNKHDKPIREQAENIHLF
TLTNLGAFAAFKYFDTTIDRKRYTSTKEVLDATLIHQISITGLYETRIDLSQLGGDSPKKKRKVEAS**GRADALDDFDLMDLGS****DALDDFDL**
MLGSDALDDFDLMDLGS**DALDDFDLMDL****ING**TASGSG**EGRGSIL****TCGDVEEN****PGPV**SKLMAENFTQFVLVDNGGTGDTVAPSNFANGVAE
WISSNSRSQAYKVTCSVRQSSAQKRYTIKVEVPKVATQTVGGVELPVAAWRSYLNMEITPIFATNSDCELIKAMQGLLKDGNPISAI
AANSGLIYSAAGGGSGGGSGGGSGGSG**PKKKRKV**AAAGS**PSGQIS****NQALALAPSSAPVLAQTMVPS****SAMVPLAQPPAPAPVLT****P****GPPQ****SL****SAP**
VPKSTQAG**EDLSEALLHLQF****DADEDLGALLGN****STDPGVFTDLASVDN****SEFQQLLNQGVSM****SHSTAEP****MLMEY****PEAITRLV****TGSQR****PPDPA**
PTPLGT**SGLPNGLSG****DEDFSS****ADMF****SALLS****QISSGQ****GGGSG****GF****SVDT****SALLDL****FS****PSVT****VPDMS****LPD****LD****SS****SLAS****IQELL****SPQE****PPRPP**
EAENSS**PD****SG****Q****LVHYTAQ****PL****FL****DP****GS****VDT****GS****ND****LP****VL****FEL****GE****SY****F****SE****GD****GF****AE****DP****T****IS****LL****TG****SE****PP****KA****KD****PT****VS*******

NN NLS(+), SV40 nuclear localisation sequence

NN HA, hemagglutinin A epitope

NN dCas9(C), C-terminal moiety of human codon–optimized *Streptococcus pyogenes* Cas9

NN VP64, transcriptional activator domain

NN T2A, 2A self-cleaving peptide from thosea asigna virus

NN MCP (MS2 protein), MS2 bacteriophage coat protein

NN P65, activation domain from human NF-κB trans-activating subunit p65

NN HSF1, activation domains from human heat-shock factor 1

TMt^{NES}-FKBP-TEV(N)-dCas9(N) used in pTMt_FKBP_TEV(N)_NES_TCS(Q'G)_HA-dCas9(N)_P2A-Puro-WPRE

METDTLLLLWVLLWVPGSTGDHSGGGSGGGSGRQEQKLISEEDLNAVQD**TQEVIVVPHSLPFKVVVISAILALV****VTIISLIILIMLWQK**
KPRLQSGSETPGTSESATPESASMGVQVETISPGDGRFTFKRGQTCCVHYTGMLDGGKKFDSSDRNKPFFKMLGKQEVIRGWEEGVAQMS
VGQRAKLITISPDYAYGATGHPGIIPPHATLVFDVLLKLEGGSGGGSGGESLFKGPDRYNPISSTICHLTNESDGHNTSLYIGIGFGPFIIT
NKHLFRRNNGTLLVQSLHGVFKVKNNTTLQOHLIDGRDMIIRMPKDFPPFPQKLKFRFPQREERICLVTTNFQTGGGSLDLASLILGKLG
ENLYFQGGGGSTSYPYDVDPDYAGGGSGDKKYSIGLAIGTNSVGVAVITDEYKVPSSKKFKVLGNTDRHSIKKNLIGALLFDSGETAEATRLK
RTARRRYTRRKNRICYLOEIFSNEMAKVDDSFHRLSEESFLVEEDKKHERHPIFGNIVDEVAYHEKYPTIYHLRKKLVSDTDKADLRILIY
ALAHMIKFRGHFLIEGDLNPDNSDVDFLFIQLVQTYNQLFEENPINASGVDAKAILSARLSKSRRLENLIAQLPGEKKNGLFGNLIALSLG
LTPNFKSNFDLAEDAKLQLSKDTYDDDLNLLAQIGDQYADFLAAKNLSDAILLSDILRVNTEITKAPLSASMIKRYDEHHQDLTLLKAL
VRQQLPEKYKEIFFDQSKNGYAGYIDGGASQEEFYKFIKPILEKMDGTEELLVKLNREDLLRKQRTFDNGSIPHQIHGELHAILRRQEDF
YPFLKDNREKIEKILTRIPYVVGPLARGNSRFAMWTRKSEETITPWNFEFVVDKGASAQSFIERMTNFDKNLPNEKVLPHKSHLLYEYFTV
YNELTKVYVTEGMRGGGSGTGSG**GATN****FSL****LKQAGD****VEEN****PGPE****F****MTE****Y****K****P****T****V****R****L****A****T****R****D****D****V****P****R****A****V****R****T****L****A****A****F****A****D****Y****P****A****T****R****H****T****V****D****P****D****R****H****I****E**
VTELQELF**FLTRVGLD****IGK****VW****ADDGA****AV****W****TT****P****S****VE****AG****AV****FAE****I****G****P****R****MA****EL****S****G****S****R****L****A****A****Q****Q****ME****GL****L****A****P****H****R****P****K****E****P****A****W****F****L****A****T****V****G****V****S****P****D****H****Q**
KGLG**SA****V****L****P****GV****E****A****A****R****AG****V****PA****LE****T****S****A****P****R****N****L****P****F****Y****E****R****L****G****F****T****V****T****A****D****V****E****P****E****G****P****R****T****W****C****M****T****R****K*******

NN Igκ, murine Immunoglobulin kappa-chain signal peptide

NN (GGGS)_x linker

NN Myc epitope tag

NN TM (PDGFR), transmembrane domain from platelet derived growth factor receptor

NN XTEN linker

NN FKBP12, FK506 binding protein 12

NN N-TEV

NN NES, nuclear export signal from human protein tyrosine kinase 2

NN efficient TCS; tobacco etch virus (TEV) protease cleavage sequence

NN HA, hemagglutinin A epitope

NN dCas9(N), N-terminal moiety of human codon–optimized *Streptococcus pyogenes* Cas9

NN P2A, 2A self-cleaving peptide from porcine teschovirus-1

NN Puromycin resistance protein

TMt-FKBP-TEV(N)-NLS-dCas9(C)^{VP64} used in pTMt_FKBP_TEV(N)_TCS(Q'G)_NLS-HA-dCas9(C)-VP64_T2A_MCP-P65-HSF1

METDTLLLLWVLLWVPGSTGDHSGGGSGGGSGRQEQKLISEEDLNAVQD**TQEVIVVPHSLPFKVVVISAILALV****VTIISLIILIMLWQK**
KPRLQSGSETPGTSESATPESASMGVQVETISPGDGRFTFKRGQTCCVHYTGMLDGGKKFDSSDRNKPFFKMLGKQEVIRGWEEGVAQMS
VGQRAKLITISPDYAYGATGHPGIIPPHATLVFDVLLKLEGGSGGGSGGESLFKGPDRYNPISSTICHLTNESDGHNTSLYIGIGFGPFIIT
NKHLFRRNNGTLLVQSLHGVFKVKNNTTLQOHLIDGRDMIIRMPKDFPPFPQKLKFRFPQREERICLVTTNFQTGGGSENLYFQ**GPKKKR**

Appendix

KVGGGSTSYPYDVDPYAGGSGSGGGSKPAFLSGEQKKAIVDLLFKTNRKVTVKQLKEDYFKKIECFDSVEISGVEDRFNASLGTYHDLKI
IKDKDFLDNEENEDILEDIVLTLTLFEDREMIEERLKYAHLFDDKVMKQLKRRRYTGWGRLSRKLINGIRDKQSGKTILDFLKSDGFANR
NFMQLIHDDSLTFKEDIQKAQVSGQGDSLHEHIANLAGSPAIIKKGILQTVKVVDELVKVMGRHKPENIVIAMARENQTTQKGQKNSRERMK
RIEEGIKELGSQILKEHPVENTQLQNEKLYLYLQNGRDMYVDQELDINRLSDYDVHIVPQSFLKDDSIDNKVLTRSDKARGKSDNVPSE
EVVKKMKKNYWRQLLNAKLITQRKFDNLTKAERGGLSELDKAGFIKRLVETRQITKHVAQILDSRMNTKYDENDKLIREVKVITLKSCLVS
DFRKDFQFYKVRINNYHHAHDAYLNAVVGTAIIKKYPKLESEFVYGDYKVDVRKMIKSEQEI GKATAYFFYSNIMNFFKTEITLANG
EIRKRPLIETNGETGEIVWDKGRDFATVRKVLSPQVNIIVKKTEVQTGGFSKESILPKRNSDKLIARKKDWDPKKYGGFDSPTVAYSVLVV
AKVEKGSKKLKSVKELLGITIMERSSEFKNPIDFLEAGYKEVKKDLIIKLPKYSLFELENGRKRMLASAGELQKGNELALPSKYVNFY
LASHYEKLKSGPEDNEQKQLFVEQHKHYLDEIIIEQISEFSKRVLADANLDKVL SAYNKHDKPIREQAENI IHLFTLTNLGAPAAFKYFD
TTIDRKRYTSTKEVLDTLIHQSIITGLYETRIDLSQLGGDSPKKRRKVEASGRADALDDFDLMDLGS DALDDFDLMDLGS DALDDFDLMDL
GSDALDDFDLMDLINGTASGSGEGRGSLTTCGDVEENPGPVSKLMASNFTQFVLVDNGGTGDVTVAPSNFANGVAEWISSNSRSQAYKVTC
SVRQSSAQKRKYTIKVEVPKVATQTVGGVELPVAAWRSYLNMEITPIFATNSDCELVKAMQGLLKDGNIIPSAIAANSNGIYSAGGGGSG
GGSGGGGSGPKKKRKVAAGSPSGQISNQALALAPSSAPVLAQTMVPSSAMVPLAQPPAPAPVLTGPQQSLAPVPKSTQAGEGTLSEA
LLHLQFDAQEDLGALLGNSTDPGVFTDLASVDNSEFQQLLNQGVSMSSHSTAEPMLMEYPEAITRLVTGSQRPPDPAPTPLGTSGLPNGLSG
DEDFSSIADMDFSALLSQISSSGQGGGSGFSVDTSALLDLFSPSVTVPDMSLPDLSSLASIQELLSQEPFRPPPEAENS SPDSGKQLVH
YTAQPLFLDPGSDVTGSDNLPVLFELGEGSYFSEGDGFAEDPTIISLLTGSEPPKADPTVS*

NN Igk, murine Immunoglobulin kappa-chain signal peptide

NN (GGGS)₂ linker

NN Myc epitope tag

NN TM (PDGFR), transmembrane domain from platelet derived growth factor receptor

NN XTEN linker

NN FKBP12, FK506 binding protein 12

NN N-TEV

NN efficient TCS; tobacco etch virus (TEV) protease cleavage sequence

NN NLS, SV40 nuclear localisation sequence

NN HA, hemagglutinin A epitope

NN dCas9(C), C-terminal moiety of human codon-optimized *Streptococcus pyogenes* Cas9

NN VP64, transcriptional activator domain

NN T2A, 2A self-cleaving peptide from thossea asigna virus

NN MCP (MS2 protein), MS2 bacteriophage coat protein

NN P65, activation domain from human NF-κB trans-activating subunit p65

NN HSF1, activation domains from human heat-shock factor 1

TEV protease used in *pcDNA3.1_TEV*

MHHHHHHHHHGESLFKGRPDYNPISSTICHLTNESDGHTTSLYGIGFGPFIITNKHLFRRNNGTLLVQSLHGVFKVKNTTTLQQHLIDGRD
MI IIRMPKDFPPFPQKLKFREPQREERICLVTTNFQTKSMSSMVSDTSCTFPSSDGI FWKHWIQT KDGCQSPLVSTRDGFIVGIHSASNF
TNTNNYFTSVPKNFMELLTNQEAQQWVSGWRLNADSVLWGGHKVFMVKPEEPFQPVKEATQLMNRRRRP*

NN His9x tag

NN TEV protease sequence

TEV mutant protease (C151A) used in *pcDNA3.1_TEV^{C151A}*

MHHHHHHHHHGESLFKGRPDYNPISSTICHLTNESDGHTTSLYGIGFGPFIITNKHLFRRNNGTLLVQSLHGVFKVKNTTTLQQHLIDGRD
MI IIRMPKDFPPFPQKLKFREPQREERICLVTTNFQTKSMSSMVSDTSCTFPSSDGI FWKHWIQT KDGCQSPLVSTRDGFIVGIHSASNF
TNTNNYFTSVPKNFMELLTNQEAQQWVSGWRLNADSVLWGGHKVFMVKPEEPFQPVKEATQLMNRRRRP*

NN His9x tag

NN TEV protease sequence

NN C151A mutation

FKBP-TEV(N) used in *pcDNA3.1_FKBP-TEV(N)*

MGVQVETISPGDGRTFPKRGQTCVVHYTGMLDGGKFDSSSRDNKPFKFMKGQEVIRGWEEGVAQMSVQRAKLTISPDIYAGATGHPGI
IPPHATLVFVDELLKLEGGSGGGSGESLFKGRPDYNPISSTICHLTNESDGHTTSLYGIGFGPFIITNKHLFRRNNGTLLVQSLHGVFKV
KNTTTLQQHLIDGRDMI IIRMPKDFPPFPQKLKFREPQREERICLVTTNFQTKLESRGPV*

NN FKBP12, FK506 binding protein 12

NN (GGGS)₂ linker

NN N-TEV

FRB-TEV(C) used in *pcDNA3.1_FRB-TEV(C)*

MEMWHEGLEEASRLYFCERNVKGMEFVLEPLHAMMERGPQTLKETSFNQAYGRDLMEAEQWCRKYMKSGNVKDLTQAWDLYYHVFRISKQ
GGSGGGSGKSMSSMVSDTSCTFPSSDGI FWKHWIQT KDGCQSPLVSTRDGFIVGIHSASNF TNTNNYFTSVPKNFMELLTNQEAQQWVSG
WRLNADSVLWGGHKVFMVKPEEPFQPVKEATQLMNRRRRP*

NN FKBP rapamycin binding domain (FRB)

NN (GGGS)₂ linker

NN C-TEV

Appendix

NN inefficient TCS; tobacco etch virus (TEV) protease cleavage sequence

NN NLS, SV40 nuclear localisation sequence

NN GGGS linker

NN HA, hemagglutinin A epitope

NN dCas9m4, nuclease deficient *Streptococcus pyogenes* Cas9

NN VP64, transcriptional activator domain

dCas9(C)-synBDKRB2 used in pBDKRB2_TCS (Q'G)_NLS-HA-dCas9(C)-VP64_T2A_MCP-P65-HSF1

MKTIIALSYIFCLVFADYKDDDDASIDMFSPWKISMFLSVREDSVPTTASFADMLNVTLQGPTLNGTFAQSKCPQVEWLGLNLTIQPPFL
WVFLVFLATLENIFVLVSFCLHKSSCTVAEIIYLGNLAAADLILACGLPFWAITISNNFDWLFGETLCRVVNAIISMNLYSSICFLMLVSI
YLALVKTMSMGRMRGVRWAKLYSLVIWGTLLSSPMLVFRMTKEYSDEGHNVACVISYPSLIWEVFTNMLNVLNVGFLPLSVITFTCTMQ
IMQVLRNNEMQKFKEIQTERRATVLLVLLVLLFIICWLPFQISTFDTLHRLGILSSCQDERIIDVITQIASFMAYSNSCLNPLVYVIVGK
RFRKKSWEVYQGVQCQKGGCRSEPIQMENSMTGLRTSISVERQIHKLQDWAGSRQIDTGGRTPPSLGPQDESCTASSSLAKDTSSTGENLY
FQGPKKKKRVGGGSGSYPYDVPDYAGGSGSGGSGSKPAFLSGEQKKAIVDLLPKTNRKVTVKQLKEDYFKKIECFDSVEISGVEDRFNASLG
TYHDLKIIKDKDFLDNEENEDILEDIVLTTLTFEDREMIEERLKYAHLFDDKVMKQLKRRRYTGWGRLSRKLINGIRDKQSGKTIIDFL
KSDGFANRNFQQLIHDDSLTFKEDIQKAQVSGQDLSHEHIANLAGSPAIAKKGILQTVKVVDELVKVMGRHKPENIVEMARENQTTQKQ
KNSRERMKRIEIEGKELGSQLKEHPVENTQLQNEKLYLYLQNGRDMYVDQELDINRLSDYDVDHIVPQSFLKDDSIDNKVLTNRSDKARG
KSDNVPSEEVVKKMKNYWRQLLNAKLITQRKFDNLTKAERGGLSELDKAGFIKRQLVETRQITKHVAQILDSRMNTKYDENDKLIREVKVI
TLKSLVSDFRKDFQFYKREINNYHHAADAYLNAVGTALIKKYPKLESEFVYGDYKVDVRKMIKSESEQEIGKATAKYFFYSNIMNFFK
TEITLANGEIRKRPLIETNGETGEIVWDKGRDFATVRKVLSPQVNIIVKKTEVQTGGFSKESILPKRNSDKLIARKKDWDPKKYGGFDSPT
VAYSVLVVAKEGKSKKLKSVKELGITIMERSSFEKNPIDFLEAKGYKEVKKDLIIKLPKYSLFELENGRKRMLASAGELQKGNELALP
SKYVNFYLYLASHYEKLKSPEDNEQKQLFEVQHKHYLDEIIEQISEFSKRVLADANLKDVLKSAYNKHRDKPIREQAENIIHLFTLTNLGA
PAAFKYFDTTIDRKRYTSTKEVLDATLIHQISITGLYETRIDLSLQGGDSPKKKRVKVEASGRADALDDFDLMDLGSALDDFDLMDLGSAL
DDFDLMDLGSALDDFDLMDLNGTASGSGEGRGSLLTCGDVEENPGPVSKLMASNFTQFVLVDNGGTGDTVAPSNFANGVAEWISSNR
SQAYKVTCSVRQSSAQKRKYTIKVEVPKVATQTVGGVELPVAAWRSYLNMLTIPIFATNSDCELVKAMQGLLKDGNPIPSAIAANSNGIY
SAGGGGSGGGGSGGGGSGPKKKRKYAAAGSPSGQISNQALALAPSSAPVLAQTMVPSAMVPLAQPPAPAPVLTGPPQSLAPVPKSTQA
GEGTISEALHLQFDEADLGLLGNSTDPGVFTDLASVDNSEFQQLNQGVSMHSTAEPMLMEYPEAITRLVTGSRPPDPAPTPLGTS
GLPNGLSGDEDFSSIADMDFSALLSQISSGQGGGSGGFSVDTSALLDLFSPSVTVPDMSLPDLSSLASIQELLSPQEPPEPPPEAENSSP
DSGKQLVHYTAQPLFLDPGSVDTGSNDLPVLFELGEGSYFSEGDFEAEDPTISLLTGSEPPKAKDPTVS*

NN membrane localisation signal sequence and Flag epitope tag

NN BDKRB2 coding sequence

NN V₂-tail

NN efficient TCS; tobacco etch virus (TEV) protease cleavage sequence

NN NLS, SV40 nuclear localisation sequence

NN GGGS linker

NN HA, hemagglutinin A epitope

NN dCas9(C), C-terminal moiety of human codon-optimized *Streptococcus pyogenes* Cas9

NN VP64, transcriptional activator domain

NN T2A, 2A self-cleaving peptide from thosea asigna virus

NN MCP (MS2 protein), MS2 bacteriophage coat protein

NN P65, activation domain from human NF-κB trans-activating subunit p65

NN HSF1, activation domains from human heat-shock factor 1

Note: For TEV cleavage optimization, the TCS sequences were iteratively replaced with ENLYFQG, ENLYFQY and ENLYFQL.

dCas9(N)-synBDKRB2 used in pBDKRB2_TCS(Q'L)_HA-dCas9(N)_P2A-Puro-WPRE

MKTIIALSYIFCLVFADYKDDDDASIDMFSPWKISMFLSVREDSVPTTASFADMLNVTLQGPTLNGTFAQSKCPQVEWLGLNLTIQPPFL
WVFLVFLATLENIFVLVSFCLHKSSCTVAEIIYLGNLAAADLILACGLPFWAITISNNFDWLFGETLCRVVNAIISMNLYSSICFLMLVSI
YLALVKTMSMGRMRGVRWAKLYSLVIWGTLLSSPMLVFRMTKEYSDEGHNVACVISYPSLIWEVFTNMLNVLNVGFLPLSVITFTCTMQ
IMQVLRNNEMQKFKEIQTERRATVLLVLLVLLFIICWLPFQISTFDTLHRLGILSSCQDERIIDVITQIASFMAYSNSCLNPLVYVIVGK
RFRKKSWEVYQGVQCQKGGCRSEPIQMENSMTGLRTSISVERQIHKLQDWAGSRQIDTGGRTPPSLGPQDESCTASSSLAKDTSSTGENLY
FQLTSYPYDVPDYAGGSGSDKKYSIGLAIGTNSVGWAVITDEYKVPSSKKFKVLGNTDRHSIKKNLIGALLFDSGETAEATRLKRTARRRYT
RRKNRICYLQEIFSNEMAKVDDSFHRLSEESFLVEEDKKHERHPIFGNIIVDEVAYHEKYPTIYHLRKKLVDSDDKADLRLLIYALAHMIKF
RGHFLIEGDLNPDNSVDVKLFIQLVQTYNQLFEEENPINASGVDAKAILSARLSKSRLENLIAQLPGEKKNGLFGNLIALLSLGLTPNFKSN
FDLAEDAKLQLSKDTYDDDLNLLAQIGDQYADFLAAKNLSDAILSDILRVNTEITKAPLSASMIKRYDEHHQDLTLLKALVRQQLPEK
YKEIFFDQSKNGYAGYIDGGASQEEFYKFIKPILEKMDGTEELLVKLNREDLLRKQRTFDNGSIPHQIHLGELHAILRRQEDFYFPLKDNR
EKIEKILTFRIPIYYVGPLARGNSRFAWMTRKSEETITPNWFEEVVDKGSASQSFIERMTNFDKNLPNEKVLPKHSLLEYEFTVYNELTKVK
YVTEGMRGGGSGTGS~~GATNFSLLKQAGDVEENPGPE~~FMTEYKPTVRLATRDDVPRAVRTLAAAFADYPATRHTVDPDRHIERVTELQELF
LTRVGLDIGKVVVADGAAVVWTTPESEAGAVFAEIGPRMAELSGSRLAAQQQMEGLLAPHRPKEPAWFLATVGVSPDHQKGLGSAVV
LPGVEAAERAGVPAFLETAPRNLPFYERLGTFTVADVEVEGPRTWCMTRKG*

NN membrane localisation signal sequence and Flag epitope tag

NN BDKRB2 coding sequence

NN V₂-tail

NN inefficient TCS (Q'L); tobacco etch virus (TEV) protease cleavage sequence

NN HA, hemagglutinin A epitope

NN dCas9(N), N-terminal moiety of human codon-optimized *Streptococcus pyogenes* Cas9

NN P2A, 2A self-cleaving peptide from porcine teschovirus-1

NN Puromycin resistance protein

Appendix

Note: For the linker optimization of dCas9(N)-synBDKRB2 in Figure 4.3, the following amino acid sequences were inserted between the V₂ tail and the TCS: for b) GGS and for c) EFLDLASLILGKLG. For TEV cleavage optimization, the TCS in dCas9(N)-synBDKRB2 were consecutively replaced with ENLYFQY (for construct in d) and ENLYFQG (for construct in e). For the extracellular NES optimisation in Figure 4.4, in one case, dCas9(N)-synBDKRB2 received a NES sequence that was inserted downstream of the membrane localisation signal and FLAG tag and upstream of the BDKRB2 coding sequence, to generate pNES-BDKRB2_TCS(Q'L)_HA-dCas9(N)_P2A-Puro-WPRE. In the other case, the NES sequence was flanked by GGS short linker sequences was inserted in the same location, to generate pNES(GGS)-BDKRB2_TCS(Q'L)_HA-dCas9(N)_P2A-Puro-WPRE.

dCas9(N)-synBDKRB2^{DHFR} used in pBDKRB2_TCS(Q'L)_HA-dCas9(N)_DHFR

MKTIIALSYIFCLVFADYKDDDDASIDMFSPWKISMFLSVREDSVPTTASFADMLNVTLQGPTLNGTFAQSKCPQVEWLGWLNITIQPPFL
WVLFVLATLENI FVLVSFCLHKSSCTVAEIIYLGNLAAADLLACGLPFWAITISNNFDWLFGETLCRVVNAIISMNLYSSICFLMLVSI
YLALVKTMSMGRMRGVRWAKLYSLVIWGCTLLSSPMLVFRMTKEYSDEGHNVTAACVISYPSLIWEVFTNMLNVLVGGFLPLSVITFCTMQ
IMQVLRNNEMQKFKEIQTERRATVLLVLLFFIICWLPFQISTFLDTHLRLGILSSCQDERIIDVITQIASFMAYSNSCLNPLVYVIVGK
RFRKKSWEVYQGVCKQGGCRSEPIQMENSMTGLRTSISVERQIHKLQDWAGSRQIDTGGRTPPSLGPDQESCTASSSLAKDTSSTGENLY
FQLTSYYPYDVPDYAGGSGSKKYSIGLAIGTNSVGWAVITDEYKVPSSKKFKVLGNTDRHSIKKNLIGALLFDSGETAEATRLKRTARRRYT
RRKNRICYLQEIFSNEMAKVDDSFHRLSEESFLVEEDKKHERHPIFGNIVDEVAYHEKYPTIYHLRKKLVDSITKADLRLLIYLAHAHMIKF
RGHFLIEGDLNPDNSDVKLFQVLQVQTYNQLFEEENPINASGVDAKAILSLARLSKSRLENLIAQLPGEKKNGLFGNLIALSLGLTPNFKSN
FDLAEDAKLQLSKDTYDDDLNLLAQIGDQYADFLAAKNLSDAILLSDILRVNTEITKAPLSASMIKRYDEHHQDLTLLKALVRQQLPEK
YKEIFFDQSKNGYAGYIDGGASQEEFYKFIKPILEKMDGTEELLVKLNREDLLRKQRTFDNGSIPHQIHLGELHAILRRQEDFYFPLKDN
EKIEKILTFRIPIYVGLPARGNSRFAMWTRKSEETITPWNFEVVDKGASQSFIERMTNFDKNLPNEKVLPKHSLLEYFTVYNELTKVK
YVTEGMRGGGSGTGSMISLIAALAVDYVIGMENAMPWNLPADLAWFKRNTLNKPVIMGRHTWESIGRPLPGRKNIIILSSQSTDDRVTW
KSVDEATAACGDVPEIMVIGGGRVIEQFLPKAQKLYLTHIDAEEVEGDTHFPDYEPPDWESVSEFHDADAQNHSYCFEILERR*

NN membrane localisation signal sequence and Flag epitope tag

NN BDKRB2 coding sequence

NN V₂-tail

NN inefficient TCS (Q'L); tobacco etch virus (TEV) protease cleavage sequence

NN HA, hemagglutinin A epitope

NN dCas9(N), N-terminal moiety of human codon-optimized *Streptococcus pyogenes* Cas9

NN *E. coli* dihydrofolate reductase

dCas9(N)-synBDKRB2^{Rl} used in pBDKRB2_TCS(Q'L)_FKBP-HA-dCas9(N)_P2A_Puro-WPRE

MKTIIALSYIFCLVFADYKDDDDASIDMFSPWKISMFLSVREDSVPTTASFADMLNVTLQGPTLNGTFAQSKCPQVEWLGWLNITIQPPFL
WVLFVLATLENI FVLVSFCLHKSSCTVAEIIYLGNLAAADLLACGLPFWAITISNNFDWLFGETLCRVVNAIISMNLYSSICFLMLVSI
YLALVKTMSMGRMRGVRWAKLYSLVIWGCTLLSSPMLVFRMTKEYSDEGHNVTAACVISYPSLIWEVFTNMLNVLVGGFLPLSVITFCTMQ
IMQVLRNNEMQKFKEIQTERRATVLLVLLFFIICWLPFQISTFLDTHLRLGILSSCQDERIIDVITQIASFMAYSNSCLNPLVYVIVGK
RFRKKSWEVYQGVCKQGGCRSEPIQMENSMTGLRTSISVERQIHKLQDWAGSRQIDTGGRTPPSLGPDQESCTASSSLAKDTSSTGENLY
FQLTSYYPYDVPDYAGGSGSKKYSIGLAIGTNSVGWAVITDEYKVPSSKKFKVLGNTDRHSIKKNLIGALLFDSGETAEATRLKRTARRRYT
RRKNRICYLQEIFSNEMAKVDDSFHRLSEESFLVEEDKKHERHPIFGNIVDEVAYHEKYPTIYHLRKKLVDSITKADLRLLIYLAHAHMIKF
RGHFLIEGDLNPDNSDVKLFQVLQVQTYNQLFEEENPINASGVDAKAILSLARLSKSRLENLIAQLPGEKKNGLFGNLIALSLGLTPNFKSN
FDLAEDAKLQLSKDTYDDDLNLLAQIGDQYADFLAAKNLSDAILLSDILRVNTEITKAPLSASMIKRYDEHHQDLTLLKALVRQQLPEK
YKEIFFDQSKNGYAGYIDGGASQEEFYKFIKPILEKMDGTEELLVKLNREDLLRKQRTFDNGSIPHQIHLGELHAILRRQEDFYFPLKDN
EKIEKILTFRIPIYVGLPARGNSRFAMWTRKSEETITPWNFEVVDKGASQSFIERMTNFDKNLPNEKVLPKHSLLEYFTVYNELTKVK
YVTEGMRGGGSGTGSMISLIAALAVDYVIGMENAMPWNLPADLAWFKRNTLNKPVIMGRHTWESIGRPLPGRKNIIILSSQSTDDRVTW
KSVDEATAACGDVPEIMVIGGGRVIEQFLPKAQKLYLTHIDAEEVEGDTHFPDYEPPDWESVSEFHDADAQNHSYCFEILERR*

NN membrane localisation signal sequence and Flag epitope tag

NN BDKRB2 coding sequence

NN V₂-tail

NN inefficient TCS (Q'L); tobacco etch virus (TEV) protease cleavage sequence

NN FKBP12, FK506 binding protein 12

NN HA, hemagglutinin A epitope

NN dCas9(N), N-terminal moiety of human codon-optimized *Streptococcus pyogenes* Cas9

NN P2A, 2A self-cleaving peptide from porcine teschovirus-1

NN Puromycin resistance protein

FRB-VP64 used in pcDNA3.1_FRB-VP64

EMWHEGLEEASRLYFGERNVKGMEFVLEPLHAMMERGPQTLKETSFNQAYGRDLMEAEQWCRKYMKSNGVKDLTQAWDLYYHVFRRIKQGG
GSGSGSGSGRADALDDFDLDMLGSDALDDFDLDMLGSDALDDFDLDMLGSDALDDFDLDMLIN*

NN FKBP rapamycin binding domain (FRB)

NN (GGGS)₂ linker

NN VP64, transcriptional activator domain

dCas9(C)-synBDKRB2^{-VP64} used in pBDKRB2_TCS(Q'G)_NLS-HA-dCas9(C)_T2A_DHFR-PP7-VP64

MKTIIALSYIFCLVFADYKDDDDASIDMFSPWKISMFLSVREDSVPTTASFADMLNVTLQGPTLNGTFAQSKCPQVEWLGWLNITIQPPFL
WVLFVLATLENI FVLVSFCLHKSSCTVAEIIYLGNLAAADLLACGLPFWAITISNNFDWLFGETLCRVVNAIISMNLYSSICFLMLVSI

YALAVKTM~~SMGRMR~~GVRWAKLYSLV~~IWGCTLL~~SSPMLV~~FRTMKEYS~~DEGHNV~~TACVISY~~PSLIWEV~~FTNMLLN~~VGV~~FLLPLSVIT~~CTMQ
IMQVL~~RNRNE~~YQK~~FKEIQT~~TERRATVLVLV~~LLLFI~~CWLP~~FPQIST~~FLDTL~~HLRGLISSCQ~~DERI~~IDVIQT~~QAS~~FMAYSNS~~CLNP~~LYV~~VYVIGK
RFRK~~NSWEVYQ~~QV~~QKCGK~~SEPIQ~~MENSMGL~~TRTSIS~~VERQI~~HKLD~~WAGSRQID~~TGGR~~TPPSLPQ~~DESCT~~ASSLAKD~~STSG~~ENLY~~
FYQ~~PKKKRKV~~GGGTSY~~PPDVDPY~~AGGS~~GGGGG~~SK~~PAFLSE~~GQKKAIV~~DLDF~~LFT~~NKRNV~~TV~~QKLKEDY~~FKKIE~~CFDSVET~~SGV~~EDFR~~NASLG
TH~~DDLKI~~IK~~KDKD~~FL~~DNE~~ENE~~DILEDIV~~LT~~TLF~~ED~~REMI~~EER~~LKTYA~~HL~~FDDK~~V~~MQK~~L~~KRRRY~~TG~~WGR~~L~~SRK~~L~~INGI~~R~~DQ~~S~~GKT~~IL~~D~~FL
KSD~~GFAN~~R~~NFM~~QLI~~HDD~~SL~~TFKEDI~~Q~~KAQ~~V~~SGQ~~GD~~SL~~HE~~HI~~AN~~LAG~~SPAI~~KKG~~IL~~QTV~~K~~VVD~~EL~~V~~K~~V~~M~~GR~~H~~KPEN~~IV~~IE~~MA~~REN~~Q~~TTQ~~K~~Q~~
K~~NSR~~EM~~KRIE~~GI~~KEL~~GS~~QILKE~~H~~PVENTQ~~L~~QNE~~K~~LYL~~Y~~LQ~~N~~GRD~~MY~~VQ~~DEL~~DIR~~L~~NSD~~Y~~QID~~H~~IV~~Q~~PS~~FL~~KDD~~S~~DN~~K~~V~~L~~TR~~S~~KD~~KA~~R~~
K~~SD~~NP~~SE~~EV~~VKKM~~K~~KNY~~Q~~LLNA~~K~~LIT~~Q~~RK~~FD~~NLT~~KA~~ERG~~SEL~~DKAG~~F~~IK~~Q~~L~~RV~~ET~~Q~~IT~~K~~HA~~V~~Q~~IL~~DS~~RM~~TKY~~D~~END~~K~~L~~IR~~EV~~K~~V~~
TL~~KS~~L~~V~~S~~DFR~~Q~~DFY~~K~~V~~Y~~RE~~IN~~NY~~HA~~HD~~AY~~LN~~AV~~VG~~TAL~~IK~~K~~YP~~KL~~SE~~F~~YV~~G~~DY~~K~~V~~Y~~VD~~R~~K~~MI~~A~~Q~~SE~~Q~~ET~~G~~K~~AT~~AKY~~FF~~YS~~NI~~M~~N~~FF~~K
TE~~IT~~LAN~~GE~~IR~~KR~~PLI~~ET~~NG~~ET~~GEI~~V~~WD~~K~~GR~~D~~FAT~~VR~~K~~V~~LS~~M~~PP~~Q~~VNI~~V~~KK~~TE~~V~~Q~~T~~GG~~F~~S~~KE~~S~~IL~~P~~K~~R~~NS~~D~~K~~L~~I~~A~~R~~K~~K~~D~~W~~D~~PK~~Y~~GG~~F~~DS~~P~~T
V~~A~~YS~~V~~LV~~V~~AK~~VE~~KG~~SK~~KL~~S~~V~~K~~ELL~~GT~~TE~~IM~~ERS~~S~~FE~~K~~NP~~ID~~FL~~E~~AK~~G~~Y~~K~~EV~~K~~DL~~I~~IK~~L~~P~~K~~YS~~L~~FB~~E~~NG~~R~~K~~R~~ML~~A~~S~~A~~GE~~L~~Q~~K~~NE~~L~~AL~~P~~
SK~~YV~~N~~F~~LY~~L~~ASH~~Y~~E~~K~~L~~G~~S~~P~~ED~~NE~~Q~~KL~~TV~~E~~Q~~H~~K~~HY~~L~~DE~~IE~~Q~~IS~~E~~FS~~K~~RV~~I~~AD~~AN~~LD~~K~~VS~~A~~YN~~K~~H~~R~~DK~~P~~IRE~~Q~~AE~~NI~~I~~H~~LT~~F~~LT~~N~~LG~~A~~
PAA~~FY~~FD~~TT~~DI~~DR~~K~~RY~~TS~~KE~~VD~~L~~AT~~L~~I~~H~~QS~~IT~~GL~~Y~~ET~~R~~LD~~S~~Q~~L~~GD~~S~~PK~~K~~RR~~V~~E~~ASS~~SG~~S~~E~~GR~~SL~~TC~~GD~~VEN~~GP~~M~~IS~~L~~IA~~LA~~V
DY~~V~~IG~~M~~EN~~A~~MP~~W~~N~~L~~PAD~~L~~AW~~F~~K~~R~~NT~~L~~N~~K~~P~~V~~IM~~G~~R~~H~~T~~W~~ES~~I~~GR~~P~~L~~P~~GR~~K~~NI~~I~~LS~~S~~Q~~P~~ST~~D~~DR~~V~~T~~V~~W~~K~~S~~V~~DE~~A~~IA~~A~~AC~~G~~D~~V~~PE~~I~~M~~V~~IG~~G~~GR~~V~~IE
Q~~F~~L~~P~~KA~~Q~~K~~L~~Y~~L~~TH~~I~~DA~~E~~VE~~G~~D~~TH~~FP~~D~~YE~~P~~DD~~W~~ES~~V~~FSE~~F~~HD~~A~~DAQ~~N~~SH~~S~~Y~~C~~FE~~I~~LER~~R~~GG~~G~~SG~~G~~GG~~G~~SG~~G~~GG~~S~~SK~~T~~IV~~L~~SV~~G~~EAT~~R~~TL~~T~~ET~~I~~
Q~~ST~~ADR~~Q~~I~~F~~E~~E~~K~~V~~GL~~V~~GR~~L~~RL~~T~~AS~~L~~R~~Q~~NG~~A~~K~~T~~AY~~R~~N~~L~~K~~L~~DQ~~A~~D~~V~~VS~~D~~GL~~P~~K~~V~~RY~~T~~Q~~V~~W~~S~~H~~D~~V~~T~~IV~~A~~NS~~T~~E~~A~~SR~~K~~SL~~Y~~DL~~T~~K~~S~~L~~V~~ATS~~Q~~V
ED~~L~~V~~N~~LV~~P~~LG~~R~~SAG~~G~~GG~~S~~GG~~G~~GG~~S~~GG~~G~~SG~~S~~PK~~K~~RR~~K~~VAA~~A~~GS~~G~~RA~~D~~AL~~D~~DF~~D~~LD~~M~~LG~~S~~DAL~~D~~DF~~D~~LD~~M~~LG~~S~~DAL~~D~~DF~~D~~LD~~M~~LG~~S~~DAL~~D~~DF~~D~~LD~~M~~LG~~S~~DAL~~D~~DF~~D~~
DL~~D~~ML~~N~~I^{*}

NN membrane localisation signal sequence and Flag epitope tag

NN BDKRB2 coding sequence

NN V_2 -tail

NN efficient TCS; tobacco etch virus (TEV) protease cleavage sequence

NN NLS, SV40 nuclear localisation sequence

NN GGGS linker

NN HA, hemagglutinin A epitope

NN dCas9(C), C-terminal moiety of human codon-optimized *Streptococcus pyogenes* Cas9

NN T2A, 2A self-cleaving peptide from thossea asigna virus

NN *E. coli* dihydrofolate reductase

NN PP7

NN VP64, transcriptional activator domain

dCas9(C)-synAVPR2 used in *pAVPR2_TCS (Q'G)_NLS-HA-dCas9(C)-VP64_T2A_MCP-P65-HSF1*

MKTTIALSYIFCLVFADYKDDDDASIDMLMASTTSAPVGHPSLPSLPSNSSQERPLDRDPLLARAEALLSIVFVAVALSNGVLVAALAR
RGRRRGWAPIHVFIGHCLADLAVALFQVLPQLAWKATDRFRGPDALCRAYKVLQVMGYASSYMLAMTDLRRHAICRPLAYRHGSGAH
WNRPVLVAWAFSLLSLLQLFLFAIQNRVEGGSGVTDWCACFAEPWGRRTYVTWIALVMFVAPLTGIAACQVLFREIHASLVPSERPGRG
RRRRRTGSGPEGGAHVSAAVAKTVMRTLIVVVVVLWCAPFFLVQLWAWDPEALGEPFVLLMLLASLNSCTNPWISYASFSSVSSSELF
SLLCCARGRTPPSLGPQDESCTTASSSLAKDTSSIDTGGRTPPSLGPQDESCTTASSSLAKDTSSTGENLYFQCGKKKKRVGGGSTSYPYD
VPDYAGGGSGGGGSKPAFLSGEQKKAIVDLLFKTNRKVTVKQLKEDYFKKIECFDSVEISGVEDRFNASLGTYHDLKIKIKDKDFLDNEEN
EFILEDVLTLTLTFEDREMIERLKLTYAHLFDDKVMKQLKRRRYTGWGLRSKRLINGIRDKQSGKTLIDFLKSDGRFARNRMFQIHDDSLT
FKEDIQAQVSGQGDLSHEHIANLAGSAPAKKGLQTVKVVDVLKVMGRHKPENVIEMARENQTTQCGKNSRANKRIEELKELGSG
ILKEHPVENTQIQNEKLYLYYLQNGRDMYVDQELDINRLSDYDVHIVPQSFLKDDSIDNKVLTBSDKARGKSDNPVSEEVKKMKNYWRQ
LLNAKLITQRKFDNLTAKERGGSELDKAGFIKRLQVETQITKHVAQILDSRMNTYDENDKLIREVKVITLKSGLVSDFRKDFQFYKVR
SINNYHHAHDYLNNAVGTALIKKYPKLESEFVYGDYKVDVRMKIAKSEQEIGKATAKYFFYSINIMNFKTEITLANGEIRKPLIETNG
ETGEIIVWDKGRDFATVRKVLSPMQVNIKKTEVTQGGFSKESILPKRNSDKLIRKKDDPDKKYGGFDSPTVAYSVLVVAKEKGSKKLK
SVKELGITIMERSFSEKNPIDFLEAKGYKEVKKDLIKLPKYSLFELENGRKRMLASAGELQKGNELALPSKYVNFVLYLASHYEKLGKSP
EDNEQKQLFVEQHKHYLDEIIEQISEFSKRVILADANLDKVL SAYNKHDKPIREQAENIIHLFTLTNLGAPAAFKYFDTTIDRKRYTSTK
EVLDTALIHQSTITGLYTERIDLSQLGDSPKKKRKVEASGRADALDDFDLMDLGSDFALDDFDLMDLGSDFALDDFDLMDLGSDFALDDFDLMD
LINGTASGSGEGRSLLTCGDVEENPGPVSYNMLTMSASNFQTVFLVDNGGTDVTVAPSNDFFVAEWISSNRSQAYKVTCSVRQSSAQKKRY
TIKEVPKVATQTVGGVELPVAWRSVLNMELTPIFATNSDCELLIVKAMQGLLDGKNPISAI AANS GIYSAAGGGSGGGSGGGSGGSGPK
KKRKVAAAGSPSGQISNQALALAPSSAPVLAQTMVPSSAMVPLAQPPAPAPVLTGPPQSLSAPVPKSTQAGEGTLSEALLHQFDADEDL
GALLNSTDPGVFTDLASVDNSEFQQLLNQGVSMSSHSTAEPMLMEYEAETIRLVTGSQRPPDPAPTPLGTSGLPNGLSGDEDFSSIADMDF
SALLSQISSSGGGSGFSVDTALLDLFSPSVTPVDMSLPDLDDSLAIQELSSQEPPEPRPPEAENSSPDGSKQLVHYTAQPLFLDDPG
SVDTGNSNLDPLVLELGGSGVFSFGDGFAEDPTISILLTGSEPPKAKDPTVS*

NN membrane localisation signal sequence and Flag epitope tag

NN AVPR2 coding sequence

NN V_2 -tail

NN efficient TCS; tobacco etch virus (TEV) protease cleavage sequence

NN NLS, SV40 nuclear localisation sequence

NN GGGS linker

NN HA, hemagglutinin A epitope

NN dCas9(C), C-terminal moiety of human codon-optimized *Streptococcus pyogenes* Cas9

NN VP64, transcriptional activator domain

NN T2A, 2A self-cleaving peptide from thosea asigna virus

■ MCP (MS2 protein). MS2 bacteriophage coat protein

NN P65, activation domain from human NF- κ B trans-activating subunit p65

NN HSF1, activation domains from human heat-shock factor 1

dCas9(N)-synAVPR2 used in *pAVPR2_TCS(Q'L)_HA-dCas9(N)_P2A-Puro-WPRE*

MKTIIALSYIFCLVFADYKDDDDASIDMLMASTTSVAVPGHPSLPSLPSNSSQERPLDTRDPLLARAELALLSIVFVAVALSNGLVLAALAR
RGRRGHWAPIHVFIHGLCLADLAVALFQVLPQLAWKATDRFRGPDALCRAVKYLQVMGYASSYMLAMTLDRHRAICRPMLAYRHGSGAH
WNRPLVAVAFSLLLSLPQLFI FAQRNVEGSGVTDCAWCAFAEPWGRRTYVWIALMVFAPTLGIACQVLI FREIHASLVPGPSEPPGG
RRRGRTGSPGEGAHVSAVAKTVMRLTVV VVYVLCWAPFVLVQLWAAWDEAPLEGAPFVLLMLLASLNSCTNPWIYASFSSSVSSEL
SLCCARGRTPPSLGPQDESCCTASSSLAKDTSSIDTGGRTPPSLGPQDESCCTASSSLAKDTSSSTGENLYFQLTSYPYDVPDYAGGSGSD
KKYSIGLAIGTNSVGWAVITDEYKVPSSKKFKVLGNTDRHSIKKNLIGALLFDSGETAEATRLKRTARRRYTRKKNRICYLQEIFSNEMAKV
DDSFHRLSESLVEEDKKHERHPIFGNIVDEVAYHEKYPTIYHLRKKLVSDTKADLRILIYLAHAMIKFRGHFLIEGDLNPDNSDVKL
FIQLVQTYNQLFEEENPINASGVDKAILSARLSKSRLENLIAQLPGEKKNGLFGNLIALSLGLTPNFKSNFDLAEDAKLQLSKDTYDDDL
DNLLAQIGDYADLFLLAANKLSDAILSDILRVNTEITKAPLSASMIKRYDEHHQDLTLLKALVRQQLPEKYKEIFFDQSKNGYAGYIDGG
ASQEEFYKFIKPILEKMDGTEELLVKNREDLLRKQRTFDNGSIPHQIHLGELHAILRRQEDFYFPLKDNREKIEKILTRIPYVYVGLAR
GNSRFAWMTRKSEETITPWNFEVVDKGASQSFIERMTNFDKNLPNEKVLPHKSLLEYEFTVYNELTKVKYVTEGMRGGGSGTGS**GATN**
FSLKQAGDVEENPGPEFMTYKPTVRLATRDDVPRVARTLAAAFADYPATRH TVDPDRHIERVTELQELFLTRVGLDIGKVVVADDGAAV
AVWTTPEVSVEAGAVFAEIGPRMAELSGSRLAAQQQMEGLLAPHRKPEFAWFLATVGVS PDHQGKGLGSAVVLPGVEAAERAGVPAFLETS
ARNLPFYERLGTFTVTADVEVPEGPRTWCMTRKG*

NN membrane localisation signal sequence and Flag epitope tag

NN AVPR2 coding sequence

NN V₂-tail

NN inefficient TCS (Q'L); tobacco etch virus (TEV) protease cleavage sequence

NN HA, hemagglutinin A epitope

NN dCas9(N), N-terminal moiety of human codon-optimized *Streptococcus pyogenes* Cas9

NN P2A, 2A self-cleaving peptide from porcine teschovirus-1

NN Puromycin resistance protein

Note: For the extracellular NES optimisation in Figure 4.5, dCas9(N)-synAVPR2 received a NES sequence flanked by GGS short linker sequences, which was inserted downstream of the membrane localisation signal and FLAG tag and upstream of the AVPR2 coding sequence, to generate pNES(GGS)-AVPR2_TCS(Q'L)_HA-dCas9(N)_P2A-Puro-WPRE.

dCas9(C)-synCXCR4 used in *pCXCR4_TCS (Q'G)_NLS-HA-dCas9(C)-VP64_T2A_MCP-P65-HSF1*

MKTIIALSYIFCLVFADYKDDDDASIDMEGISIYTS DNYTEEMSGDYDSMKEPCFREANANFNKIFLPTIYSIIFLTGIVGNGLVILVMG
YQKKLRSM TDYRLHLSVADLLFVITL PFWAVDAVANWYFGNFLCKAVHVIYTVNLYSSVLILAFISLDRLAIVHATNSQRPRKLLAEKV
VYGVVWIPALLTIPDFIFANVSEADDRYICDRFYPN DLWVVFQFQHIMVGLILPGIVILSCYCI IISKLSSHSGHQRKALKTTVILIL
AFFACWL PYYIGISIDSFILEI IKQGEFENTVHKWISITEALAFFHCCLNPILYAF LGA FKTSQAHALTSVSRGSSLKILSKGKRGGH
SSVSTESSESSFHSSIDTGGRTPPSLGPQDESCCTASSSLAKDTSSSTGENLYFQGPKKKKRVGGGSTSYPYDVPDYAGGSGSGGSGSKPAFL
SGEQKKAIVDLLFKTNRKVTVKQLKEDYFKKIECFDSVEISGVEDRFNASLGTYHDLKIIKDKDFLDNEENEDILEDIVLTTLTFEDREM
TEERLKYAHLFDDKVMQQLKRRRYTGWGRLSRKLINGIRDQSGKTILDFLKS DGFANRNFMLIHDDSLTFKEDIQKAQVSGQGDLSHE
HIANLAGSPA KKGILQTVKVVDELVKVMGRHKPENIVIE MARENQTQKGQKNSRERMKRIE EGIKELGSQLKEHPVENTQLQNEKLYL
YYLQNGRDMYVDQELDINRLSDYDV DHI V PQSFLKDDSIDNKVLT RSDKARGKSDNPVSEEVVKKMKNYWRQLLNAKLITQRKFDNLTKAE
RGGLSELDKAGFIKRLVETRQITKHVAQILDSRMNTKYDENDKLIREVKVITLKS KLVSDFRKDFQFYKVRINNYHHAHDAYLNAVVG
ALIKKYPKLESEFVYGDYKVDVRKMIKSEQEIGKATAKYFYFSNIMNFFKTEITLANGEIRKRPLIETNGETGEIVWDKGRDFATVRKV
LSMPQVNI VKKTEVQTGGFSKESILPKRNSDKLIARKKDWDPKKYGGFDSPTVAYSVLVAKVEKGKSKKLKSVKELLGITIMERSSFEKN
PIDFLEAKGYKEVKDLI IKLPKYSLFEL ENGRKRMLASAGELQKGNELALPSKYVNFYLYLASHYEKLKGPS PEDNEQQLFVEQHKHYLDE
IIEQISEFSKRVLADANLDKVL SAYNKH RDKPIREQAENIIHLFTLTNLGAPAAFKYFDTTIDRKRYTSTKEVLDATLIHQSI TGLYETR
IDLSQLGGDSPKKKKRVEASGRADALDDFDLDMLGSDALDDFDLDMLGSDALDDFDLDMLINGTASGSGEGRGSLLTC
GDVEENPGPVSKLMASNFTQFVLVDNGGTGDVTVAPS NFANGVAEWISSNSRSQAYKVTCSVRQSSAQKRKYTIKVEVPKVATQTVGGVEL
PVAAWRSYLNMELTIPIFATNSDCELI V KAMQGLLDGNPIPSAIAANSIGY SAGGGSGGGSGGGSGSGPKKKRKYAAAGSPSGQISNQA
LALAPSSAPVLAQT MVPSSAMVPLAQPPAPAPVLT PGPQSL SAPVPKSTQAGEGTLSEALLHLQFDAEDLGALLGNSTDPGVFTDLASV
DNSEFQQLNQGVSM SHSTAEPMLMEYPEAITRLVTGSQRPPDPAPTPLGTSGLPNGLSGDEDFSSIA DMDFSALLSQISSSGQGGGSGF
SVDTSALLDLFSPSVTPVDMSLPDL DSSLASIQELLSQEPPEPRPPEAENSSPD SGKQLVHYTAQPLFLDPGSVD TGSNDLPVLFELGEGS
YFSEGDGFAEDPTISLLTGSEPPKADPTVS*

NN membrane localisation signal sequence and Flag epitope tag

NN CXCR4 coding sequence

NN V₂-tail

NN efficient TCS; tobacco etch virus (TEV) protease cleavage sequence

NN NLS, SV40 nuclear localisation sequence

NN GGS linker

NN HA, hemagglutinin A epitope

NN dCas9(C), C-terminal moiety of human codon-optimized *Streptococcus pyogenes* Cas9

NN VP64, transcriptional activator domain

NN T2A, 2A self-cleaving peptide from thosea asigna virus

NN MCP (MS2 protein), MS2 bacteriophage coat protein

NN P65, activation domain from human NF-κB trans-activating subunit p65

NN HSF1, activation domains from human heat-shock factor 1

dCas9(C)-synCXCR4^(-VP64) used in *pCXCR4_TCS(Q'G)_NLS-HA-dCas9(C)*

MKTIIALSYIFCLVFADYKDDDDASIDMEGISIYTS DNYTEEMSGDYDSMKEPCFREANANFNKIFLPTIYSIIFLTGIVGNGLVILVMG
YQKKLRSM TDYRLHLSVADLLFVITL PFWAVDAVANWYFGNFLCKAVHVIYTVNLYSSVLILAFISLDRLAIVHATNSQRPRKLLAEKV
VYGVVWIPALLTIPDFIFANVSEADDRYICDRFYPN DLWVVFQFQHIMVGLILPGIVILSCYCI IISKLSSHSGHQRKALKTTVILIL
AFFACWL PYYIGISIDSFILEI IKQGEFENTVHKWISITEALAFFHCCLNPILYAF LGA FKTSQAHALTSVSRGSSLKILSKGKRGGH
SSVSTESSESSFHSSIDTGGRTPPSLGPQDESCCTASSSLAKDTSSSTGENLYFQGPKKKKRVGGGSTSYPYDVPDYAGGSGSGGSGSKPAFL

Appendix

SGEQKKAIVDLLFKTNRKVTQKQLKEDYFKKIECFDSVEISGVEDRFNASLGTYHDLKI IKDKDFLDNEENEDILEDIVLTLTLFEDREM
IEERLKTYAHLFDDKVMKQLKRRRYTGWGRLSRKLINGIRDKQSGKTI LDFLKS DGFANRNFQMQLIHDDSLTFKEDIQKAQVSGQGDSLHE
HIANLAGSPAIIKGILOTQVKVVDLVKVMGRHKPENIVIEMARENQTTQKGQKNSRERMKRIEEGIKELGSQLKEHPVENTQLQNEKLYL
YYLQNGRDMYVDQELDINRLSDYDVHIVPQSFLKDDSIDNKVLT RSDKARGKSDNPVSEEVVKKMKNYWRQLLNAKLITQRKFDNLTKEA
RGGLSELDKAGFIKRQLVETRQITKHVAQILDSRMNTKYDENDKLIREVKVITLKS KLVSDFRKDFQFYKVREINNYHHAHDAYLNAVVG
ALIKKYPKLESEFVYGDYKVDVRKMIKSEQEIGKATAKYFFYSNIMNFFKTEITLANGEIRKRPLIETNGETGEIVWDKGRDFATVRKV
LSMPQVNIVKKTEVQTGGFSKESILPKRNSDKLIARKKDWDPKKYGGFDSPTVAYSVLVVAKVEKGSKKLKSVKELLGITIMERSSFEKN
PIDFLEAKGYKEVKKDLI IKLPKYSLFELENGRKRMLASAGELQKGNELALPSKYVNFLYLASHYEKLGKSPEDNEQQLFVEQHKHYLDE
ITIEQISEFSKRVLADANLDKVL SAYNKH RDKPIREQAENI IHLFTLTNLGAPAAFKYFDTTIDRKRYTSTKEVLDATLIHQSIITGLYETR
IDLSQLGGDSPKKKRKVEASGGGS*

NN membrane localisation signal sequence and Flag epitope tag

NN CXCR4 coding sequence

NN V₂-tail

NN efficient TCS; tobacco etch virus (TEV) protease cleavage sequence

NN NLS, SV40 nuclear localisation sequence

NN GGGS linker

NN HA, hemagglutinin A epitope

NN dCas9(C), C-terminal moiety of human codon-optimized *Streptococcus pyogenes* Cas9

dCas9(N)-synCXCR4 used in *pCXCR4_TCS(Q'L)_HA-dCas9(N)_P2A-Puro-WPRE*

MKTIIALSYIFCLVFADYKDDDDASIDMEGISIYTS DNYTEEMSGDYDSMKEPCFREANANFNKIFLPTIYSIIIFLTGIVGNGLVILVMG
YQKKLRSMTPKYRLHLSVADLLFVITLPFWAVDAVANWYFGNFLCKAVHVIYTVNLYSSVLILAFISLDRYLAI VHATNSQRPRKLLAEKV
VYVGWVIPALLTIPDFIFANVSEADRYICDRFYPNDLVVVVFQFQHIMVGLILPGIVILSCYCI IISKLSSHKGHKRKA LKTTVILIL
AFFACWLPPYYIGISIDSFILLEI IKQGEFENTVHKWISITEALAFFHCCLNPILYAF LGAKFKTSAQHALTSVSRGSSKLKLSKGKRGGH
SSVSTESSESSFHSSIDTGGRTPPSLGPQDESCTTASSSLAKDTSSTGENLYFQGLTS YPYDVPDYAGGSGS DKKYSIGLAIGTNSVGWAVI
TDEYKVPKSKFKVLGNTDRHSI KKNLIGALLFDSGETAEATRLKRTARRRYTRRKNRICYLQEIFSNEMAKVDDSFHRLEESFLVEEDKK
HERHPIFGNIVDEVAYHEKYPTIYHLRKKLV DSTDKADRLIYLALAHMIKFRGHFLIEGDLNPDNSDVKLF IQLVQTYNQLFEENPINA
SGVDAAKAILSARLSKSRRLENLIAQLPGEKKNLFGNLIALSLGLTPNFKSNFDLAEDAKLQLSKD TYDDDLNLLAQIGDQYADFLAAK
NLSDAILLSDILRVNTEITKAPLSASMIKRYDEHHQDLTLLKALVRQQLPEKYKEIFFDQSKNGYAGYIDGGASQEEFYKFIKPILEKMDG
TEELLVKLNREDLLRKQRTFDNGSIPHQIHLGELHAILRRQEDFY PFLKDNREKIEKILTFRIPYYVGPLARGNSRFAMWTRKSEETITPW
NFEEVVDKGAASAQSFIERMTNFDKNLPNEKVL PKHSLLYEYFTVYNELTKVKYVTEGMRGGGSGTGS GATNFSLKQAGDVEENPGPEFM
TEYKPTVRLATRDDVPRAVRTLAAAFADYPATRHVTDPDRHIERVTE LQELFLTRVGLDIGKVVWADDGAAVAVWTTPE SVEAGAVFAEIG
PRMAELSGSRLAAQQQMEGLLAPHRPKEPAWFLATVGVS PDHQGKGLGS AVVLPGV EAAERAGVPAFLETSAPRNL PFYERLGFTVTADVE
VPEGPRTWCMTRKG*

NN membrane localisation signal sequence and Flag epitope tag

NN CXCR4 coding sequence

NN V₂-tail

NN inefficient TCS (Q'L); tobacco etch virus (TEV) protease cleavage sequence

NN HA, hemagglutinin A epitope

NN dCas9(N), N-terminal moiety of human codon-optimized *Streptococcus pyogenes* Cas9

NN P2A, 2A self-cleaving peptide from porcine teschovirus-1

NN Puromycin resistance protein

dCas9(C)-synLPAR1 used in *pLPAR1_TCS(Q'G)_NLS-HA-dCas9(C)-VP64_T2A_MCP-P65-HSF1*

MKTIIALSYIFCLVFADYKDDDDASIDMAAISTSIPIVISQPQTAMNEPQC FYNESIAFFYNRSGKHLATEWNTVSKLVMGLGITVCIFIM
LANLIVMVAIYVNRPHFPIYYLMANLAAADFFAGLAYFYLMFNTGPNTRRLTVSTWLLRQGLIDTSLTASVANLLAIAIERHITVFRMQL
HTRMSNRVRVVVVIWITMAIVMGAIPSVGWNCICDIENCNMAPLYSDSYLVFWAIFNLVTFVVMVVL YAHIFGYVRQTRMRSRHSSGP
RRNRDTMMSLLKTVVLVLGAFIICWTPGLVLLLDVCCPQCDVLAYEKFFLLLAEFNSAMNPIIYSYRDKEMSATFRQILCCQRSENP TGP
TEGSDRSASSLNHTILAGVHSNDHSVVIDTGGRTPPSLGPQDESCTTASSSLAKDTSSTGENLYFQGP KKKRKVGGGSTS YPYDVPDYAGG
SGSGGGSKPAPFLSGEQKKAIVDLLFKTNRKVTQKQLKEDYFKKIECFDSVEISGVEDRFNASLGTYHDLKI IKDKDFLDNEENEDILEDI
VLTTLTLFEDREMIEERLKTYAHLFDDKVMKQLKRRRYTGWGRLSRKLINGIRDKQSGKTI LDFLKS DGFANRNFQMQLIHDDSLTFKEDIQK
AQVSGQDSLSHEHIANLAGSPAIIKGILOTQVKVVDLVKVMGRHKPENIVIEMARENQTTQKGQKNSRERMKRIEEGIKELGSQLKEHPV
ENTQLQNEKLYLYYLLQNGRDMYVDQELDINRLSDYDVHIVPQSFLKDDSIDNKVLT RSDKARGKSDNPVSEEVVKKMKNYWRQLLNAKL I
TQRKFDNLTKAERGGLSELDKAGFIKRQLVETRQITKHVAQILDSRMNTKYDENDKLIREVKVITLKS KLVSDFRKDFQFYKVREINNYHH
AHDAYLNAVVG TALIKKYPKLESEFVYGDYKVDVRKMIKSEQEIGKATAKYFFYSNIMNFFKTEITLANGEIRKRPLIETNGETGEIVW
DKGRDFATVRKVL SMPQVNIVKKTEVQTGGFSKESILPKRNSDKLIARKKDWDPKKYGGFDSPTVAYSVLVVAKVEKGSKKLKSVKELLG
ITIMERSSFEKNPIDFLEAKGYKEVKKDLI IKLPKYSLFELENGRKRMLASAGELQKGNELALPSKYVNFLYLASHYEKLGKSPEDNEQKQ
LFVEQHKHYLDEIIEQISEFSKRVLADANLDKVL SAYNKH RDKPIREQAENI IHLFTLTNLGAPAAFKYFDTTIDRKRYTSTKEVLDATL
IHQSIITGLYETRIDLSQLGGDSPKKKRKVEASGRADALDDFDLMDLGS DALDDFDLMDLGS DALDDFDLMDLGS DALDDFDLMDL INGTAS
GSGEGRGSLITCGDVEENPGVPVSKLMA SNFTQFVLVDNGGTGDVTVA PSNFANGVAEWISSNSRSQYKVTCSVRQSSAQKRKYTIKVEVP
KVATQTVGGVELPVAAWRSYLNMELTIPIFATNSDCELIVKAMQGLLKDGNPIPSAIAANS GIYSAGGGGGGGGGGGGGSGPKKKRKVAA
AGSPSGQISNQALALAPSSAPVLAQTMVPSAMVPLAQPPAPAPVLTGPPQSL SAPVPKSTQAGEGTLSEALLHLQFDAEDLGALLGNS
TDPGVFTDLASVDNSEFQQLNQGVSMSSHSTAEPMLMEYPEA ITRLVTGSQRPPDPAPTPLGTSGLPNGLSGDEDFSSIADMDFSALLSQI
SSSGGGGGGFSVDT SALLDLFSPSVTVPDMSLDPLDSSLSIQELLS PQEPPRPPEAENSSPD SGKQLVHYTAQPLFLLDPGSVD TGSN
DLPLVLFELGEGSYFSEGDGFAEDPTISLLTGSEPPKAKDPTVS*

NN membrane localisation signal sequence and Flag epitope tag

NN LPAR1 coding sequence

NN V₂-tail

NN efficient TCS; tobacco etch virus (TEV) protease cleavage sequence

NN NLS, SV40 nuclear localisation sequence

NN GGGS linker

Appendix

- HA**, hemagglutinin A epitope
dCas9(C), C-terminal moiety of human codon–optimized *Streptococcus pyogenes* Cas9
VP64, transcriptional activator domain
T2A, 2A self-cleaving peptide from thesaccharin virus
MCP (MS2 protein), MS2 bacteriophage coat protein
P65, activation domain from human NF-κB trans-activating subunit p65
HSF1, activation domains from human heat-shock factor 1

dCas9(C)-synLPAR1^(-VP64) used in *pLPAR1_TCS(Q'G)_NLS-HA-dCas9(C)*

MKTIIALSYIFCLVFADYKDDDDASIDMAAISTSIPIVISQPQFTAMNEPQCIFYNESIAFFYNRSKGHLATEWNTVSKLVMGLGITVCIFIM
LANLLVMVAIYVNRHFHPIYYLMANLAAADFFAGLAYFYLMFNTGPNTRRLTVSTWLLRQGLIDTSLTASVANLLAIAIERHITVFRMQL
HTRMSNRRVVVIVVWITMAIVMGAIPSVGWNCICDIENCSNMPLYSYSYLVFWAIFNLVTFVVMVVLVYAHIFGYVRQRTMRMRHSSGP
RRNRDTMMSLLKTIVIVLGAFIICWTPGLVLLLLDVCCPQCDVLAYEKFFLLAEFNSAMNPIIYSYRDKEMSATFRQILCCQRSENPTGP
TEGSDRSASSLNHTILAGVHSNDHVSVIDTGGRTPPSLGPDQESCTTASSSLAKDTSSTGENLYFQGPKKKKRVGGGSTSYPYDVPDYAGG
SGSGGSKPAFLSGEQKKAIVDLLFKTNRKVTVKQLKEDYFKKIECFDSVEISGVEDRFNASLGTYHDLKKIKDKDFLDNEENEDILEDI
VLTTLTFEDREMIEERLKYAHLFDDKVMQKLRRTYTGWGRLSRKLINGIRDKQSGKTILDFLKSDFANRNFQMQLIHDDSLTFKEDIQK
AQVSGQGDLSLHEHIANLAGSPAIIKKGILQTVKVVDELVKVMGRHKPENIVIMARENQTTQKGQKNSRERMKRIEEGIKELGSQILKEHPV
ENTQLQNEKLYLYLQNGRDMYVDQELDINRLSDYDVDHIVPQSFLKDDSIDNKVLRSDKARGKSDNPVSEEVVKMKNYWRQLLNAKLI
TQRKFDNLTKAERGGSELKAGFIKRLVETRQITKHVAQILDSRMNTKYDENDKLIREVKVITLKSCLVDFRDKDFQFYKVRINNYHH
AHDAYLNAVVGTAIIKKYPKLESEFVYGDYKVDVVRKMIKSEQEI GKATAKYFFYSNIMNFFKTEITLANGEIRKRPLIETNGETGEIVW
DKGRDFATVRKVLSPQVNIYVKKTEVQTGGFSKESILPKRNSDKLIARKKDWDPKKYGGFDSPTVAYSVLVVAKEGKSKKLKSVKELLG
ITIMERSSFEKNPIDFLEAGKYKEVKDLIKLPKYSLEFENGRKRLASAGELQKGNELALPSKYVNFYLAHYEKLKSGPEDNEQKQ
LFVEQHKHYLDEIIIEQISEFSKRVLADANLDKLSAYNKHDKPIREQAENIIHLFTLTNLGAPAAFKYFDTTIDRKRYTSTKEVLDATL
IHQSITGLYETRIDLSQLGGDSPKKKKRVEASGGS*

- membrane localisation signal sequence and Flag epitope tag**
LPAR1 coding sequence
V₂-tail
efficient TCS; tobacco etch virus (TEV) protease cleavage sequence
NLS, SV40 nuclear localisation sequence
GGGS linker
HA, hemagglutinin A epitope
dCas9(C), C-terminal moiety of human codon–optimized *Streptococcus pyogenes* Cas9

dCas9(N)-synLPAR1 used in *pLPAR1_TCS(Q'L)_HA-dCas9(N)_P2A-Puro-WPRE*

MKTIIALSYIFCLVFADYKDDDDASIDMAAISTSIPIVISQPQFTAMNEPQCIFYNESIAFFYNRSKGHLATEWNTVSKLVMGLGITVCIFIM
LANLLVMVAIYVNRHFHPIYYLMANLAAADFFAGLAYFYLMFNTGPNTRRLTVSTWLLRQGLIDTSLTASVANLLAIAIERHITVFRMQL
HTRMSNRRVVVIVVWITMAIVMGAIPSVGWNCICDIENCSNMPLYSYSYLVFWAIFNLVTFVVMVVLVYAHIFGYVRQRTMRMRHSSGP
RRNRDTMMSLLKTIVIVLGAFIICWTPGLVLLLLDVCCPQCDVLAYEKFFLLAEFNSAMNPIIYSYRDKEMSATFRQILCCQRSENPTGP
TEGSDRSASSLNHTILAGVHSNDHVSVIDTGGRTPPSLGPDQESCTTASSSLAKDTSSTGENLYFQGLTSYPYDVPDYAGGSGSKKYSIGL
AIGTNSVGWAVITDEYKVPKFKVLGNTDRHSIKKNLIGALLFDSGETAEATRLKRTARRRYTRRKNRICYLQEIFSNEMAKVDDSFHHR
LEESFLVEEDKKHERHPIFGNIVDEVAYHEKYPTIYHLRKLKLVSDTKADRLIYLAALAHMIKFRGHFLIEGDLNPDNSDVKLFQIQLVQT
YNQLFEENPINASGVDAKAILSARLSKSRLENLIAQLPGEKKNGLFPGNLIALSLGLTPNFKSNFDLAEDAKLQLSKDQYDDBDLNLLAQI
GDQYADFLAAKNLSDAILSDILRVNTEITKAPLSAMIKRYDEHHQDLTLKALVRQQLPEKYKEIFFDQSKNGYAGYIDGGASQEEFY
KFIKPILEKMDGTIELLVKLNREDLLRKQRTFDNGSIPHQIHLGELHAILRRQEDFYFPLKDNREKIEKILTFRIPIYVYVGLARGNSRFAW
MTRKSEETITPWNFEVVVDKGASAQSFIERMTNFDKNLPNEKVLPHKSLLEYFTVYNELTKVKYVTEGMRGGGSGTSGATNFSLLKQA
GDVEENPGPEFMTEYKPTVRLATRDVPRVRTLAAAFADYPATRHVTDPDRHIERVTELQELFLTRVGLDIGKVVWADDGAAVAVWTTPE
SEAGAVFAEIGPRMAELSGSRLAAQQQMEGLLAPHRPKEPAWFLATVGVSPDHQKGKLGSAVVLPGVEAAERAGVPAPLETSAAPRNLFFY
ERLGTFTVADVEVPEGPRTWCMTRKG*

- membrane localisation signal sequence and Flag epitope tag**
LPAR1 coding sequence
V₂-tail
inefficient TCS (Q'L); tobacco etch virus (TEV) protease cleavage sequence
HA, hemagglutinin A epitope
dCas9(N), N-terminal moiety of human codon–optimized *Streptococcus pyogenes* Cas9
P2A, 2A self-cleaving peptide from porcine teschovirus-1
Puromycin resistance protein

dCas9(C)-synLPAR2 used in *pLPAR2_TCS(Q'G)_NLS-HA-dCas9(C)-VP64_T2A_MCP-P65-HSF1*

MKTIIALSYIFCLVFADYKDDDDASIDMVMGQCYYNETIGFFYNNSGKELSSHWRPKDVVVVALGLTVSVLVLLTNLLVIAAIAASNRHFH
QPIYYLLGNLAAADLFAGVAYFLMFHTGPRATRLSLEGWFLRQGLDTSLSATVATLLAIAVERHRSVMAVQLHSRLPRGRVVMILVGVW
VAALGLGLLPAHSHWCLCALDRCSRMAPLLSRSYLAVWALSSLLVFLLMVAVYTRIFFYVRRRVQRMAEHVSCHPRYRETTLSLVKTIVV
LGAFVVCWTGQVVLVLLDGLGCESCNVLAKEYFLLAEANSLVNAAVYSCRDAEMRRTFRLLCCACLRQSTRESVHYTSSAQGGASTRI
MLPENGLPLMDSTLIDTGGRTPPSLGPDQESCTTASSSLAKDTSSTGENLYFQGPKKKKRVGGGSTSYPYDVPDYAGGSGSGGSKPAFLS
GEQKKAIVDLLFKTNRKVTVKQLKEDYFKKIECFDSVEISGVEDRFNASLGTYHDLKKIKDKDFLDNEENEDILEDIVLTTLTFEDREMIE
ERLKYAHLFDDKVMQKLRRTYTGWGRLSRKLINGIRDKQSGKTILDFLKSDFANRNFQMQLIHDDSLTFKEDIQKQVSGQGDLSLHEH
IANLAGSPAIIKKGILQTVKVVDELVKVMGRHKPENIVIMARENQTTQKGQKNSRERMKRIEEGIKELGSQILKEHPVENTQLQNEKLYLY
YLQNGRDMYVDQELDINRLSDYDVDHIVPQSFLKDDSIDNKVLRSDKARGKSDNPVSEEVVKMKNYWRQLLNAKLITQRKFDNLTKAER
GGLSELDKAGFIKRLVETRQITKHVAQILDSRMNTKYDENDKLIREVKVITLKSCLVDFRDKDFQFYKVRINNYHHAHDAYLNAVVGTA
LIKKYPKLESEFVYGDYKVDVVRKMIKSEQEI GKATAKYFFYSNIMNFFKTEITLANGEIRKRPLIETNGETGEIVWDKGRDFATVRKVL

Appendix

NI SFYITDVPNGFHVNLKEMPTGEDLKL SCTVNKFLYRDVTWILLRTVNNRTMHYSISKQKMAITKEHSITLNLTIMNVSLQDSGTYACR
ARNVYTGEETLQKKIEITIRDOEAPYLLRNLS DHTVAISSSTLDC HANGVPEPQITWFKNNHKIQQEPGII LGPGSSSTLFIERVTEEDG
YHCKATNQKGSVSSAYLTVQGTSDKPNLELITLTCTCVAATLFWLLLTFLF I GGGSGGSGSKSMSSMVSDTSCTFPSSDGI FWKHWIQT
KDG QCGSLVSTRDGFI VGIHSAFNFTNTNNYFTSVPKNFMELLTNFEAQQWVSGWRNLNADSVLWGGHKVFMVKPEEPFQPVKEATQLMNR
RRR PGGGS ENLYFQGPKKKRKVGGGSTS YPYDVPDYAGGSGSGGSGSKPAPLSGEQKKAIVDLLFKTNRKVTVKQLKEDYFKKIECFDS
VEISGV EDRFNASLGTYHDLKIIKDKDFLDNEENEDILEDIVLTTLTFEDREMI EERLKYAHLFDDKVMQQLKRRRYTGWGRLSRKLINGIR
DKQ SGKTILDFLKSDGFANRNFMLIHDDSLTFKEDIQKAQVSGQGSLSHEHIANLAGSPAIIKKGILQTVKVVDELVKVMGRHKPENIVIE
MAR ENQTTQKQKNSRERMKRIEIEGKELGSQLKEHPVENTQLQNEKLYLYYLQNGRDMYVDQELDINRLSDYDVHDHIVPQSFLKDDSD
DNKV LTRSDKARGKSDNVPSEEVVKMKNYWRQLNAKLITQRKFDNLTKAERGGSELDAKAGFIKRQLVETRQITKHVAQILDSRMNTKYD
END KLIREVKVITLKSCLVSDFRKDFQFYKVRINNYYHHAHDAYLNAVVGTAIIKKYPKLESEFVYGDYKVYDVRKMIKSEQEIGKATAKY
FF YSNIMNFFKTEITLANGEIRKRPLIETNGETGEIVWDKGRDFATVRKVL SMPQVNI VKKTEVQTGGFSKESILPKRNSDKLIARKKDW
DPK KYGGFDSPTVAYSVLVAKVEKGSKKLKSVKELLGITIMERSSFEKNPIDFLEAKGYKEVKKDLIIKLPKYSLFELENGRKRMLAS
GEL QKGNELALPSKYVNFYLLASHYEKLGKSPEDNEQQLFVEQHKHYLDEIIEQISEFSKRVIADANLDKVL SAYNKHDKPIREQAENI
IH LFTLTNLGAPAAFKYFDTTIDRKRYTSTKEVLDATLIHQSI TGLYETRIDLSQLGGDSPKKKRKVEAS GRADALDDFDLMDLGS
DALDDFD LMDLGS DALDDFDLMDLGS DALDDFDLMDL INGTASGSGEGRGSLLTCGDVEENPGPVSKLMASNFTQFVLVDN
GGTGDVTVAPSNFANGV AEWISSNSRSQAYKVC SVRQSSAQKRKYTIKVEVPKVATQTVGGVELPVAWRLSYLNMELTPIFATNS
DCELPIDVKAQQGLLKDGNIIPS AIAANS GIY SAGGGSGGGSGGGSGSGPKKKRKVAAGS PSQG I SNQALALAPSSAPVLAQTM
VPSSAMVLAQPPAPAPVLTGPPQSL SAPVPKSTQAGEGTLSEALLHLQF DAD ELDGALLGNSTDPGVFTDLASVDNSEFQQLN
QGVSMSSHSTAEPMLMEYPAIATRLVTGSGQRPPD PAPTPLGTSGLPNGLSGDEDFSSIADMDFSALLSQISSSG
GGGSGS GFSVDT SALLDLFSPSVTVPDMSLPDLSSLASIQELLS PQEPPR PPEAENSPPDSGKQVHYTAQPLFLPD
PGSVDTGSNDLPVLFLEGEYSFSEGDGFAEDPTISLLTGSEPPKAKDP TVS*

NN VEGFR1 leader peptide, the extracellular domains and transmembrane domain

NN C-TEV

NN efficient TCS; tobacco etch virus (TEV) protease cleavage sequence

NN NLS, SV40 nuclear localisation sequence

NN HA, hemagglutinin A epitope

NN dCas9(C), C-terminal moiety of human codon-optimized *Streptococcus pyogenes* Cas9

NN VP64, transcriptional activator domain

NN T2A, 2A self-cleaving peptide from thossea asigna virus

NN MCP (MS2 protein), MS2 bacteriophage coat protein

NN P65, activation domain from human NF-κB trans-activating subunit p65

NN HSF1, activation domains from human heat-shock factor 1

Note: For the optimization of dCas9-synVEGFR dimer combinations, VEGFR1 and VEGFR2 were interchangeably swapped to generate alternative plasmid pVEGFR2_TEV(C)_TCS(Q'G)_NLS-HA-dCas9(C)-VP64_T2A_MCP-P65-HSF1. For TEV cleavage optimization the TCS in dCas9(C)-synVEGFR-1 was iteratively replaced by ENLYFQG, ENLYFQY and ENLYFQL.

dCas9(N)-synVEGFR-2 used in pVEGFR2_TEV(N)_NES_TCS(Q'L)_HA-dCas9(N)_P2A_Puro-WPRE

MQSKVLLAVALWLVCVETRAASVGLPSVSLDLPRLSIQKDILT IKANTTLQITCRGQRDLWLWPNNQSGSEQRVEVTECSDGLFCKTLTIP
KVIGNDTGAYKCFYRETDLASVIYVYVQDYRSPFIASVSDQHGVVYITENKNKT VVIPCLGSI SNLNVSLCARYPEKRFVDPGNRISWDSK
KGFTTIPSYMTSYAGVMVFCEAKINDESYQSIMYIVVVVGYRIYDVVLSPSHGIELSVGEKLVNCTARTELNVDIFDNWEYPPSKHQHKKLV
NRDLKTQSGSEMKKFLSTLTIDGITRSDQGLYTCAASSGLMTKKNSTFVRVHEKPFVAFGSGMESLVEATVGERVRIPAKYLGYPPEIKW
YKNGIPLESNHTIKAGHVLTMEVSEDTGNVTILTNPISEKQSHVVS LVVYVPPQIGEKSLISPVDSYQGTQTTLCTVYAI PPHH
IHWWYQLEEECANEPSQAVSVTNPYPCEEWRSVEDFQGGNKIEVNKNQFALIEGKNKT VSTLVIQAANVSALYKCEAVNKVGRGERVISFH
VTRGPEITLQPDMPTEQESVSLWCTADRSTFENLTWYKLGFPQLPIHVGELPTPVCKNLDTLWKLNATMFNSNTNDILIMELKNASLQDQ
GDYVCLAQDRKTKKRHCVVRLTVLERVAPTITGNLENQTLTTSIGESIEVSC TASGNPPQIMWFKDNETLVEDSGIVLKDGNRNLTI RRV
KEDEGLYTQACSVLGC AKVEAFFIIEGAQEKTNLEI IILVGTAVIAMFVWLLVII GGGSGGSGSGLFKGPRDYNPISSTICHLTNE
SD GHTTSLYIGIFGFPFIITNKHLFRNRNGTLLVQSLHGVFKVKNTTTTLQOHLIDGRDMI IIRMPKDFPPFPQKLKFRPQREERICLV
TTFNQ TGGGSLDLASLILGKLG ENLYFQLGGGSTS YPYDVPDYAGGSGS DKKYSIGLAIGTNSVGWAVITDEYKVP SKKFKVLGNTDRHS
IKKNLI GALLFDSGETAEATRLKRTARRRYTRKRNRI CYLQEIFSNEMAKVDDSFHRL EESFLVEEDKKHERHPIFGNIVDEVAYHEKYPTI
YHLR KKLVDSTDKADRLIYLALAHMIKFRGHFLIEGDLNPDNSVDKDLFIQLVQTYNQLFEE NPINASGVDAKAILSARLSKSRRL
ENLIAQLP GEKKNGLFGNLIALSLGLTPNFKS NFDLAEDAKLQLSKDTYDDDLNLLAQIGDQYADLFLAAKNLSDAILSDILRVNTEIT
KAPLSASM IKRYDEHHQDLTLLKALVRQQLPEKYKEIFFDQSKNGYAGYIDGGASQEEFYKFIKPILEKMDGTEELLVKLNREDLLR
KQRTFDNGSI PH QIHLGELHAILRRQEDFY PFLKDNREKIEKILTFRI PYYVGPLARGNSRFAWMTRKSEETITPWNFEVV
VDKGASAQSFIERMTNFDKNLP NEKVLPKHSLLHYEYFTVYNELTKVKYVTEGMR GGGSGGTSGS GATNFSLLKQAGDVEENPGPE
FMTEYKPTVRLATRDDVPRAVRTLAAAFADYPATRKHTVDPDRHIERVTELQELFLTRVGLDIGKVVADGGA AVAVWTTPEVS
EAGAVFAEIGPRMAELSGSRLAAQQQMEGLLAPHRPK EPAWFLATVGSPDHQGKGLGSAVVLPGVEAAERAGVPAFLETS
APRNL PFYERLGT VTDADVEVPEGPRTWCMT RKG*

NN VEGFR2 leader peptide, the extracellular domains and transmembrane domain

NN N-TEV

NN NES, nuclear export signal form human protein tyrosine kinase 2

NN inefficient TCS (Q'L); tobacco etch virus (TEV) protease cleavage sequence

NN HA, hemagglutinin A epitope

NN dCas9(N), N-terminal moiety of human codon-optimized *Streptococcus pyogenes* Cas9

NN P2A, 2A self-cleaving peptide from porcine teschovirus-1

NN Puromycin resistance protein

Note: For the optimization of dCas9-synVEGFR dimer combinations, VEGFR1 and VEGFR2 were interchangeably swapped to generate alternative plasmid pVEGFR1_TEV(N)_NES_TCS(Q'G)_HA-dCas9(N)_P2A_Puro-WPRE. For TEV cleavage optimization the TCS in dCas9(N)-synVEGFR-2 was iteratively replaced by ENLYFQG, ENLYFQY and ENLYFQL.

dCas9(C)-synVEGFR1^{RI} used in *pVEGFR1_TEV(C)_TCS(Q'G)_NLS-HA-dCas9(C)*

MVSYWDTGVLLCALLSCLLLTGSSSGSKLKDPELSLKGTQHIMQAGQTLHLQCRGEAAHKWSLPEMVSKESERLSITKSACGRNGKQFCST
 LTLNTAQANHTGFYSCYKYLAVPTSKKKETESAIIYIFISDTGRPFVEMYSEIPEIIHMTGRELVI PCRVTSPNITVTLKKFPLDTLIPDGK
 RI IWD SRKGFI ISNATYKEIGLLTCEATVNGHLYKTNLYLTHRQTNTIIDVQISTPRPVKLLRGHTLVNLCTATTPLNTRVQMTWSYPDEKN
 KRASVRRRIDQSNSHANIFYSVLTIDKMQNKDKGLYTCRVRSGPSFKSVNTSVHIYDKAFITVKHRKQOVLETVAGKRSYRLSMKVKAFFPS
 PEVVLWKDGLPATEKSARYLTRGYSLLIKDVTEEDAGNYTILLSIKQSNVFNLTATLIVNVKPKIYEKAVSSFPDPALYPLGSRQILTCT
 AYGIPQPTIKWFHPCNNHSEARCDFCNNNEESFILDADSNMGNRIEITQRMALIEGKNKMASTLVVADSRISGIYICIASNKGVTGVR
 NISFYITDVPNGFHVNLKEMPTEGEDLKLSTVKNFLYRDVTWILLRTVNNRTMHYSISKQKMAITKEHSITLNLTIMNVSLQDSGTYACR
 ARNVYTGEELQKKEITIRDQEAAPYLLRNLSDHVTVAISSSTLTDCHANGVPEPQITWFKNNHKIQQEPGIIILGPGSSTLFIERVTEEDEGV
 YHCKATNQKGSVESSAYLTVQGTSDKPNLELITLTCTCVAATLFWLLLTLLFI GGGSGGSGSKSMSSMVS DTSCTFPSSDGI FWKHWIQT KD G
 QCGSPLVSTRDGFIVGIHSASNTNTNNYFTSVPKNFEMELLTNQEAQQVSGWRLNADSVLWGGHKVFMVKPEEPFQPVKEATQLMNRRRR
 PGGGS^{ENLYFQGP}KKKRKVGGGSTS^{YPYDVPDYAGGSGSGGSGSKPAFLS}GEGKKAIVDLLFKTNRKVTVKQLKEDYFKKIECFDSVEISGV
 EDRFNASLGTYHDLKI IKDKDFLDNEENEDILEDIVLTLLTFEDREMI EERLKYAHLFDDKVMQKLRRTYTGWGRLSRKLINGIRDKQ
 SGKTILDFLKS DGFANRNFQLIHDDSLTFKEDIQKAQVSGQGSLSLHEHIANLAGSPAIIKKGILQTVKVVDELVKVMGRHKPENIVIEMAR
 ENQTTQKQKNSRERMKRIEIEGKELGSQILKEHPVENTQLQNEKLYLYLQNGRDMYVDQELDINRLSDYDVHIVPQSFLKDDSIDNKV
 LTRSDKARGKSDNPSEEVVKMKMYWRQLNAKLITQRKFDNLTKAERGGSELDAKAGFIKRLVETRQITKHVAQILDSRMNTKYDEND
 KLIREVKVITLKS KLVSDFRKDFQFYKVRINNYYHHAHDAYLNAVVG TALIKY PKLESEFVYGDYKVYDVRKMI AKSEQEIGKATAKYFF
 YSNIMNFFKTEITLANGEIRKRPLIETNGETGEIVWDKGRDFATVRKVLSPQVNI VKKTEVQTGGFSKESILPKRNSDKLIARKKDWDPK
 KYGGFDSPTVAYSVLVVAKEGKSKKLKSVKELLGITIMERSSFEKNPI DFLQNGRDMYVDQELDINRLSDYDVHIVPQSFLKDDSIDNKV
 QKGNELALPSKYVNFYLLASHYEKLGKSPEDNEQKQLFVEQHKHYLDEIIEQISEFSKRVI LADANL DKVLSAYNKH RDKPIREQAENIIH
 LFTLTLNLGAPAAFKYFDTTIDRKRYTSTKEVLDATLIHQSI TGLYETRIDLSQLGGDSPKKKRKVEASGGS*

NN VEGFR1 leader peptide, the extracellular domains and transmembrane domain

NN C-TEV

NN efficient TCS; tobacco etch virus (TEV) protease cleavage sequence

NN NLS, SV40 nuclear localisation sequence

NN HA, hemagglutinin A epitope

NN dCas9(C), C-terminal moiety of human codon-optimized *Streptococcus pyogenes* Cas9

dCas9(N)-synVEGFR2^{RI} used in *pVEGFR2_TEV(N)_NES_TCS(Q'L)_FKBP-HA-dCas9(N)_P2A_Puro-WPRE*

MQSKVLLAVALWLCVETRAASVGLPSVSLDLPLRLSIQKDILTIKANTTLQITCRGQRDLDWLWPNNSGSEQRVEVTECSDGLFCKTLTIP
 KVIGNDTGAYKCFYRETDLASVIYVYVQDYRSPFIASVSDQHGVVYITENKNKTIVIPCLGSI SNLNVSLCARYPEKRFVDPGNRISWDSK
 KGFTTIPSYMISYAGVMVFCEAKINDESYQSIMYIVVVVGYRIYDVVLSPSHGIELSVGEKLVNLCTARTELVNGIDFNWEYPSKHKHKLIV
 NRDLKTQSGSEMKKFLSTLTIDGVTRSDQGLYTCAASSGLMTKKNSTFVRVHEKPFVAFGSGMESLVEATVGERVRIPAKYLGYPPEIKW
 YKNGIPLSNHTIKAGHVLTIMEVSRDGTNYTVILTNPISKEKQSHVFLSVVYVPPQIGEKSLISPVDSYQYGTQTTLTCTVYAI PPPHH
 IHWYQLEEECANEPSQASVVTNPYPCEEWRSDVDFQGGNKIEVNKNVLFALIEGKNKTVSTLVIQAANVSALYKCEAVNKVGRGERVISFH
 VTRGPEITLQPDMPTEQESVSLWCTADRSTFENLTWYKLGQPLPIHVGELEPTPVCKNLDTLWKLNATMFSNSTNDILIMELKNASLQDQ
 GDYVCLAQDRKTKKRHCVVQLTIVLERVAPTITGNLENQTTSIGESI EVSCTASGNPPPQIMWFKDNETLVEDSGIVLKDGNRNLTRRVR
 KEDEGLYTKQACSVLGCACVAAFFIIEBGAQKTNLEIIILVGTAVIAMFFWLLLVII GGGSGGSGSGESLFGKPRDYNPISSTICHLTNESD
 GHTTSLYIGIFGPGFIITNKLHFRNRNGTLLVQSLHGVFKVKNNTTLLQOHLIDGRDMI IIRMPKDFPPFPQKLKFRPQREERICLVTTNFQ
 TGGGSLDLASLILGLKENLYFQLGGGSTSGVQVETISPGDGRTPFKRGQTCVVHYTGMLEDGKKFDSSRDNRNPKFKMLGKQEVIRGWEE
 GVAQMSVGQRAKLTISPDIYAGATGHPGIIIPPHATLVDFVELLKLKETS^{YPYDVPDYAGGSGSDKKYSIGLAIGTNSVGWAVITDEYKVP}SK
 KFKVLGNTDRHSIKKNLIGALLFDSGETAEATRLKRTARRRYTRRNKRICYLCQEIFSNEMAKVDDSFHRLSESFLEEDKKHERHPIFGN
 IVDEVAYHEKYPTIYHLRKLVDSTDKADRLIYLALAHMKIVYGFLEIEGDLNPDNSDVKLFIQLVQTYNQLFEENPINASGVDKAIL
 SARLSKSRLENLIAQLPGEKKNGLFGNLIALSLGLTPNFKSNFDLAEDAKLQLSKDTYDDDLNLLAQIGDQYADFLAAKNLSDAILLS
 DILRVNTEITKAPLSASMIKRYDEHHQDLTLLKALVRQQLPEKYKEIFFDQSKNGYAGYIDGGASQEEFYKFIKPILEKMDGTEELLVKLN
 REDLRKQRTFDNGSIPHQIHLGELHAILRRQEDFYFLKDNREKIEKILTRIPYVVGPLARGNSRFAWMTRKSEETITPWNFEFVVDKG
 ASAQSFIERMTNFDKNLPNEKVLPHKSLLYEYFTVYNELTKVYVTEGMRGGGSGTGS^{GATNFSLLKQAGVDEENPGPEFMT}EYKPTVRL
 ATRDDVPRAVRTLAAAFADYPATRHTVDPDRHIERVTELQELFLTRVGLDIGKVWVADDGAAVAVWTTPESEVAGAVFAEIGPRMAELSGS
 RLAAQQQMEGLLAPHRPKEPAWFLATVGVSPDHQKGKLGSAVVLPGVAAERAGVPAFLETSA PRNLPFYERLGFTVTADVEVPEGPRTW
 MTRKG*

NN VEGFR2 leader peptide, the extracellular domains and transmembrane domain

NN N-TEV

NN NES, nuclear export signal form human protein tyrosine kinase 2

NN inefficient TCS (Q'L); tobacco etch virus (TEV) protease cleavage sequence

NN FKBP12, FK506 binding protein 12

NN HA, hemagglutinin A epitope

NN dCas9(N), N-terminal moiety of human codon-optimized *Streptococcus pyogenes* Cas9

NN P2A, 2A self-cleaving peptide from porcine teschovirus-1

NN Puromycin resistance protein

VEGFR1-TEV(C) used in *pVEGFR1_TEV(C)*

MVSYWDTGVLLCALLSCLLLTGSSSGSKLKDPELSLKGTQHIMQAGQTLHLQCRGEAAHKWSLPEMVSKESERLSITKSACGRNGKQFCST
 LTLNTAQANHTGFYSCYKYLAVPTSKKKETESAIIYIFISDTGRPFVEMYSEIPEIIHMTGRELVI PCRVTSPNITVTLKKFPLDTLIPDGK
 RI IWD SRKGFI ISNATYKEIGLLTCEATVNGHLYKTNLYLTHRQTNTIIDVQISTPRPVKLLRGHTLVNLCTATTPLNTRVQMTWSYPDEKN
 KRASVRRRIDQSNSHANIFYSVLTIDKMQNKDKGLYTCRVRSGPSFKSVNTSVHIYDKAFITVKHRKQOVLETVAGKRSYRLSMKVKAFFPS
 PEVVLWKDGLPATEKSARYLTRGYSLLIKDVTEEDAGNYTILLSIKQSNVFNLTATLIVNVKPKIYEKAVSSFPDPALYPLGSRQILTCT
 AYGIPQPTIKWFHPCNNHSEARCDFCNNNEESFILDADSNMGNRIEITQRMALIEGKNKMASTLVVADSRISGIYICIASNKGVTGVR
 NISFYITDVPNGFHVNLKEMPTEGEDLKLSTVKNFLYRDVTWILLRTVNNRTMHYSISKQKMAITKEHSITLNLTIMNVSLQDSGTYACR
 ARNVYTGEELQKKEITIRDQEAAPYLLRNLSDHVTVAISSSTLTDCHANGVPEPQITWFKNNHKIQQEPGIIILGPGSSTLFIERVTEEDEGV

Appendix

YHCKATNQKGSVESSAYLTVQGTSDKPNLELITLTCTCVAATLFWLLLTIFI GGGSGGGSKSMSSMVSDTSCTFPSSDGI FWKHWIQT KDG
QCGSPLVSTRDGFIVGIHSASNFTNTNNYFTSVPKNFME LLTNQEAQQWVSGWRLNADSVLWGGHKVFMVKPEEPFQPVKEATQLMNRRRR
P*
NN VEGFR1 leader peptide, the extracellular domains and transmembrane domain
NN C-TEV

VEGFA121 used in *pcDNA3.1_VEGFA121*

MNFLLSVHWWSLALLLYLHHAKWSQAAPMAEGGQNHHEVVKFMDVYQRSYCHPIETLV DIFQEYPDEIEYIFKPSCVPLMR CGGCCNDEG
LECVPTESNITMQIMRIKPHQGQHIGEMSFLQHNKCECRPKKDRARQEKCDKPRR*

B: Transfection schematics for all experiments

--- (next page) ---

Appendix

Figure 3.1	Assay	Plate (wells)	Assay length (hours)	Transfected plasmid (ng / well)	
				dCas9 ^{VP64} variant	pcDNA3.1
	Imaging	12	24	500ng	500ng

Figure 3.2	Assay	Plate (wells)	Assay length (hours)	Transfected plasmid (ng / well)			
				EYFP reporter	sgSCR or sgEYFP	NLS-dCas9-VP64 or pcDNA3.1	EYFP / mCherry / pcDNA3.1
b, d	Fluorescence reporter	24	48	200ng	200ng	200ng	-
e	Fluorescence compensation	24	48	-	-	-	2 x 250ng as indicated

Figure 3.3	Assay	Plate (wells)	Assay length (hours)	Transfected plasmid (ng / well)					
				EYFP reporter	sgEYFP	SAM sgASCL1 (pool of 4 sgRNAs) or SAM sgSCR	SAM MCP	dCas9 ^{VP64} variant	pcDNA3.1
a-e	Fluorescence reporter	24	48	200ng	200ng	-	-	200ng	-
f	Endo. gene activation	24	44	-	-	250ng	250ng	250ng	250ng

Figure 3.5	Assay	Plate (wells)	Assay length (hours)	Transfected plasmid (ng / well)	
				TMt-dCas9 ^{VP64} variant	TEV or pcDNA3.1
	Imaging	12	24	500ng	500ng

Figure 3.6	Assay	Plate (wells)	Assay length (hours)	Transfected plasmid (ng / well)	
				TMt-dCas9 ^{VP64} variant or dCas9 ^{VP64}	TEV or pcDNA3.1
	Imaging	12	24	500ng	500ng

Appendix

Figure 3.7	Assay	Plate (wells)	Assay length (hours)	Transfected plasmid (ng / well)			
				EYFP reporter	sgEYFP	TMt-dCas9 ^{VP64} variant	TEV or pcDNA3.1
	Fluorescence reporter	24	48	200ng	200ng	200ng each	200ng

Figure 3.8	Assay	Plate (wells)	Assay length (hours)	Transfected plasmid (ng / well)	
				TMt-dCas9 ^{VP64} variant	TEV / TEV ^{C151A} / pcDNA3.1
	Imaging	12	24	500ng	500ng

Figure 3.9	Assay	Plate (wells)	Assay length (hours)	Transfected plasmid (ng / well)			
				EYFP reporter	sgEYFP	TMt-NLS-dCas9 ^{VP64} variant	TEV / TEV ^{C151A} / pcDNA3.1
	Fluorescence reporter	24	48	200ng	200ng	200ng	200ng

Figure 3.10	Assay	Plate (wells)	Assay length (hours)	Transfected plasmid (ng / well)			
				EYFP reporter	sgEYFP	TMt ^{NES} -dCas9 ^{VP64} variant	TEV / TEV ^{C151A} / pcDNA3.1
	Fluorescence reporter	24	48	200ng	200ng	200ng	200ng

Figure 3.11	Assay	Plate (wells)	Assay length (hours)	Transfected plasmid (ng / well)			
				EYFP reporter	sgSCR or sgEYFP	TMt-NLS-dCas9 ^{VP64} variant	TEV or pcDNA3.1
	Fluorescence reporter	24	72	200ng	200ng	stably integrated	200ng

Figure 3.12	Assay	Plate (wells)	Assay length (hours)	Transfected plasmid (ng / well)			
				TMt-Cas9-mCherry variant	TEV or pcDNA3.1	Cas9-mCherry variant	pcDNA3.1
a, b	Imaging	12	24	500ng	500ng	-	-
c, d	Imaging	12	24	-	-	500ng	500ng

Appendix

Figure 3.14 - 3.16	Assay	Plate (wells)	Assay length (hours)	Transfected plasmid (ng / well)							
				mCherry	Cas9-mCherry with sgSCR or sgEGFP	NLS-Cas9-mCherry with sgSCR or sgEGFP	TMt-Cas9-mCherry with sgSCR or sgEGFP	TMt-NLS-Cas9-mCherry with sgSCR or sgEGFP	TEV or pcDNA3.1	TMt ^{ΔTCS} -NLS-Cas9-mCherry with sgSCR or sgEGFP	sgSCR or sgEGFP
3.14a,b + 3.15b-d	dEGFP disruption	24	72	400ng	-	-	-	-	400ng	-	-
				-	400ng	-	-	-	400ng	-	-
				-	-	400ng	-	-	400ng	-	-
				-	-	-	400ng	-	400ng	-	-
				-	-	-	-	400ng	400ng	-	-
3.16a,b	dEGFP disruption	24	72	-	-	400ng	-	-	400ng	-	-
				-	-	-	-	-	400ng	400ng	-
				-	-	-	-	-	400ng	-	400ng

Figure 3.17	Assay	Plate (wells)	Assay length (hours)	Transfected plasmid (ng / well)				
				EYFP reporter	sgEYFP	NLS-dCas9(C) ^{VP64} -FKBP variant	dCas9(N)-FRB variant	TEV or pcDNA3.1
b	Fluorescence reporter	24	48	200ng	200ng	200ng	200ng	200ng

Figure 3.18	Assay	Plate (wells)	Assay length (hours)	Transfected plasmid (ng / well)				
				EYFP reporter	sgSCR or sgEYFP	NLS-dCas9(C) ^{VP64} -FKBP variant	TMt ^{NES} -dCas9(N)-FRB	TEV or pcDNA3.1
b,c	Fluorescence reporter	24	48	200ng	200ng	200ng	200ng	200ng

Figure 3.20	Assay	Plate (wells)	Assay length (hours)	Transfected plasmid (ng / well)					
				EYFP reporter	sgEYFP	NLS-dCas9(C) ^{VP64} variant	TMt ^{NES} -dCas9(N)	SAM sgSCR	TEV or pcDNA3.1
a	Imaging	12	24	-	-	250ng	250ng	250ng	250ng
b, c	Fluorescence reporter	24	48	200ng	200ng	200ng	200ng	-	200ng
d, e	Fluorescence reporter	24	48	200ng	200ng	200ng	200ng	-	200ng

Appendix

Figure 3.21 - 3.22	Assay	Plate (wells)	Assay length (hours)	Transfected plasmid (ng / well)							
				EYFP reporter	sgSCR or sgEYFP	SAM sgSCR or sgASCL1 guide#1	NLS- dCas9(C) ^{VP64} variant	Tmt ^{NES} - dCas9(N)	FKBR- TEV(N) /FRB- TEV(C)	pcDNA3.1	TEV
3.21c,d	Fluorescence reporter	24	44	200ng	200ng	-	200ng	200ng	200ng each	-	-
3.21e	Fluorescence reporter	24	44	200ng	200ng	-	200ng	200ng	200ng each	-	-
3.22	Endo. gene activation	24	44	-	-	250ng	10ng	10ng	250ng each	230ng	-
				-	-	250ng	10ng	10ng	-	230ng	250ng

Figure 3.24 - 3.25	Assay	Plate (wells)	Assay length (hours)	Transfected plasmid (ng / well)							
				EYFP reporter	sgSCR or sgEYFP	SAM sgSCR or sgASCL1 guide#1	Tmt-FKBP- TEV(N)-NLS- dCas9(C) ^{VP64} or NLS- dCas9(C) ^{VP64}	Tmt ^{NES} - FKBP- TEV(N)- dCas9(N)	FKBR- TEV(N) or FRB- TEV(C) or pcDNA3.1	NLS(+)- dCas9(C)VP64- FKBP / NES- dCas9(N)-FRB	pcDNA3.1
3.24a,b	Fluorescence reporter	24	48	200ng	200ng	-	200ng	200ng	200ng	-	-
				200ng	200ng	-	-	-	-	200ng each	200ng
3.25	Endo. gene activation	24	44	-	-	250ng	250ng	250ng	250ng	-	-

Figure 4.2	Assay	Plate (wells)	Assay length (hours)	Transfected plasmid (ng / well)		
				EYFP reporter	sgSCR or sgEYFP	BDKRB2- dCas9 ^{VP64} variants
b, c	Fluorescence reporter	24	44	200ng	200ng	200ng

Figure 4.3	Assay	Plate (wells)	Assay length (hours)	Transfected plasmid (ng / well)			
				EYFP reporter	sgEYFP	dCas9(N)- synBDKRB2 variants	NLS- dCas9(C) ^{VP64}
b, c	Fluorescence reporter	24	44	200ng	200ng	200ng	200ng

Appendix

Figure 4.4	Assay	Plate (wells)	Assay length (hours)	Transfected plasmid (ng / well)			
				EYFP reporter	sgSCR or sgEYFP	dCas9(N)-synBDKRB2 variants	NLS-dCas9(C) ^{VP64}
b, c	Fluorescence reporter	24	44	200ng	200ng	200ng	200ng
d	Fluorescence reporter	24	44	200ng	200ng	200ng	200ng

Figure 4.5	Assay	Plate (wells)	Assay length (hours)	Transfected plasmid (ng / well)			
				EYFP reporter	sgSCR or sgEYFP	dCas9(N)-synAVPR2 variants	NLS-dCas9(C) ^{VP64}
b, c	Fluorescence reporter	24	44	200ng	200ng	200ng	200ng

Figure 4.6	Assay	Plate (wells)	Assay length (hours)	Transfected plasmid (ng / well)						
				SAM sgASCL1 (pool of 4 sgRNAs) or SAM sgSCR	SAM sgHBG1 (pool of 4 sgRNAs) or SAM sgSCR	SAM sgIL1B (pool of 4 sgRNAs) or SAM sgSCR	dCas9(N)-synBDKRB2	dCas9(N)-synAVPR2	NLS-dCas9(C) ^{VP64}	pcDNA3.1
a	Endo. gene activation	24	44	10ng	-	-	100ng	-	100ng	790ng
b	Endo. gene activation	24	44	-	10ng	-	100ng	-	100ng	790ng
c	Endo. gene activation	24	44	-	-	10ng	100ng	-	100ng	790ng
d	Endo. gene activation	24	44	10ng	-	-	-	100ng	100ng	790ng

Figure 4.7	Assay	Plate (wells)	Assay length (hours)	Transfected plasmid (ng / well)				
				EYFP reporter	sgSCR or sgEYFP	dCas9(N)-synBDKRB2 variants	TMT ^{NES} -dCas9(N)	TEV
b, c	Fluorescence reporter	24	44	200ng	200ng	200ng	200ng	-
e, f	Fluorescence reporter	24	44	200ng	200ng	200ng	200ng	200ng

Appendix

Figure 4.8	Assay	Plate (wells)	Assay length (hours)	Transfected plasmid (ng / well)		
				EYFP reporter	sgSCR or sgEYFP	dCas9(N/C)-synBDKBR2
b - f	Fluorescence reporter	24	44	200ng	200ng	200ng each

Figure 4.9	Assay	Plate (wells)	Assay length (hours)	Transfected plasmid (ng / well)					
				EYFP reporter	sgSCR or sgEYFP	dCas9(N/C)-synBDKBR2	SAM sgASCL1 (pool of 4 sgRNAs) or SAM sgSCR	SAM sgGOI (pool of 4 sgRNAs) or SAM sgSCR	pcDNA3.1
a	Fluorescence reporter	24	44	200ng	200ng	200ng each	-	-	-
b	Fluorescence reporter	24	as indicated	200ng	200ng	200ng each	-	-	-
c	Endo. gene activation	12	44	-	-	100ng each	10ng	-	790ng
d	Endo. gene activation	12	44	-	-	100ng each	-	10ng each sgRNA pool	770ng

Figure 4.10	Assay	Plate (wells)	Assay length (hours)	Transfected plasmid (ng / well)			
				EYFP reporter	sgEYFP	dCas9(N/C)-synAVPR2	dCas9(N/C)-synCXCR4
a	Fluorescence reporter	24	44	200ng	200ng	200ng each	-
b	Fluorescence reporter	24	44	200ng	200ng	-	200ng each

Figure 4.11	Assay	Plate (wells)	Assay length (hours)	Transfected plasmid (ng / well)		
				EYFP reporter	sgEYFP	dCas9(N/C)-synLPA variants
a, b	Fluorescence reporter	24	44	200ng	200ng	200ng each
c	Fluorescence reporter	24	44	200ng	200ng	200ng each

Appendix

Figure 4.12	Assay	Plate (wells)	Assay length (hours)	Transfected plasmid (ng / well)		
				EYFP reporter	sgEYFP	dCas9(N/C)-synT1R variants
a, b	Fluorescence reporter	24	44	200ng	200ng	200ng each
c	Fluorescence reporter	24	44	200ng	200ng	200ng each
d	Fluorescence reporter	24	72	200ng	200ng	200ng each

Figure 4.13	Assay	Plate (wells)	Assay length (hours)	Transfected plasmid (ng / well)				
				EYFP reporter	sgEYFP	dCas9(N/C)-synBDKBR2	dCas9(N/C)-synAVPR2	dCas9(N/C)-synCXCR4
b, d	Fluorescence reporter	24	44	200ng	200ng	200ng either	200ng either	-
f, h	Fluorescence reporter	24	44	200ng	200ng	200ng either	-	200ng either

Figure 4.14	Assay	Plate (wells)	Assay length (hours)	Transfected plasmid (ng / well)			
				EYFP reporter	sgSCR or sgEYFP	scSCR ^{PP7} or scEYFP ^{PP7}	dCas9(N/C)-synBDKBR2 variant
b, c	Fluorescence reporter	24	44	200ng	200ng	-	200ng each
e, f	Fluorescence reporter	24	44	200ng	-	200ng	100ng each

Figure 4.15	Assay	Plate (wells)	Assay length (hours)	Transfected plasmid (ng / well)					
				EYFP reporter	sgEYFP	dCas9(N)-synBDKBR2 ^{Ri}	dCas9(C)-synCXCR4 ⁽⁻⁾ _{VP64}	dCas9(N/C)-synLPAR1 ⁽⁻⁾ _{VP64}	FRB-VP64
b	Fluorescence reporter	24	44	200ng	200ng	200ng	200ng	-	200ng
d	Fluorescence reporter	24	44	200ng	200ng	200ng	-	200ng	200ng

Appendix

Figure 5.1	Assay	Plate (wells)	Assay length (hours)	Transfected plasmid (ng / well)			
				EYFP reporter	sgEYFP or sgSCR	dCas9(N/C)-synVEGFR variant	VEGFA121 or pcDNA3.1
b, c	Fluorescence reporter	24	44	200ng	200ng	100ng each	50ng

Figure 5.2	Assay	Plate (wells)	Assay length (hours)	Transfected plasmid (ng / well)			
				EYFP reporter	sgEYFP or sgSCR	dCas9(N/C)-synVEGFR variant	VEGFA121 or pcDNA3.1
b, c	Fluorescence reporter	24	44	200ng	200ng	100ng each	50ng

Figure 5.3	Assay	Plate (wells)	Assay length (hours)	Transfected plasmid (ng / well)		
				dCas9(N/C)-synVEGFR variant	pcDNA3.1	EGFP (pCSGW)
	Direct flow cytometry	24	44	200ng either	200ng	200ng
				200ng each	-	200ng

Figure 5.4	Assay	Plate (wells)	Assay length (hours)	Transfected plasmid (ng / well)					
				EYFP reporter	sgEYFP or sgSCR	SAM sgASCL1 (pool of 4 sgRNAs) or SAM sgSCR	dCas9(N/C)-synVEGFR variant	VEGFA121 or pcDNA3.1	pcDNA3.1
b - f	Fluorescence reporter	24	44	200ng	200ng	-	100ng each	50ng	-
g	Fluorescence reporter	24	44	200ng	200ng	-	100ng each	as indicated (to 100ng)	-
h	Endo. gene activation	12	44	-	-	10ng	100ng each	as indicated (to 50ng)	740ng

Figure 5.5	Assay	Plate (wells)	Assay length (hours)	Transfected plasmid (ng / well)							
				EYFP reporter	sgEYFP or sgSCR	dCas9(N/C)-synVEGFR1/2 ^{RC}	dCas9(N/C)-synVEGFR1/2 ^{RI}	VEGFA121 or pcDNA3.1	FRB-TEV(C).	VEGFR1-TEV(C).	FRB-VP64 or pcDNA3.1
b, c	Fluorescence reporter	24	44	200ng	200ng	200ng each	-	50ng	200ng	200ng	-
e	Fluorescence reporter	24	44	200ng	200ng	-	200ng each	50ng	-	-	200ng

Appendix

Figure 5.6	Assay	Plate (wells)	Assay length (hours)	Transfected plasmid (ng / well)							
				EYFP reporter	sgEYFP or sgSCR	dCas9(N/C)-synVEGFR variant	dCas9 ^{VP64} (Mali <i>et al.</i> , 2013)	MESA-dCas9 (V1 or V2)	MESA-TEV (V1 or V2) or pcDNA3.1	pcDNA3.1	VEGFA121 or pcDNA3.1
b, c	Fluorescence reporter	24	44	200ng	200ng	100ng each	-	-	-	-	50ng
						-	100ng	-	-	100ng	
						-	-	100ng	10ng	90ng	

Figure 6.1	Assay	Plate (wells)	Assay length (hours)	Transfected plasmid (ng / well)		
				SAM sgGOI (pool of 3 sgRNAs) or SAM sgSCR	dCas9(N/C)-synVEGFR1/2	VEGFA121 or pcDNA3.1
b	Endo. gene activation	12	44	50ng each sgRNA pool	400ng each	50ng

Figure 6.2	Assay	Plate (wells)	Assay length (hours)	Transfected plasmid (ng / well)		
				SAM sgGOI (pool of 3 sgRNAs) or SAM sgSCR	dCas9(N/C)-synCXCR4	pcDNA3.1
b	Endo. gene activation	12	44	50ng each sgRNA pool	100ng each	650ng

Figure 6.3	Assay	Plate (wells)	Assay length (hours)	Transfected plasmid (ng / well)		
				SAM sgGOI (pool of 3 sgRNAs) or SAM sgSCR	dCas9(N/C)-synLPAR1	pcDNA3.1
b	Endo. gene activation	12	44	50ng each sgRNA pool	100ng each	650ng

Figure 6.4	Assay	Plate (wells)	Assay length (hours)	Transfected plasmid (ng / well)		
				SAM sgINS (pool of 4 sgRNAs) or SAM sgSCR	dCas9(N/C)-synT1R3	pcDNA3.1
b	Endo. gene activation	12	44	100ng	400ng each	100ng