Engineering molecular RNA sensors for controlling CRISPR activity *in vivo*



Oana Pelea
Wadham College
University of Oxford

Report submitted for the completion of
Doctor of Philosophy
Michaelmas 2022

Supervisors: Prof. Tatjana Sauka-Spengler, Prof. Tudor Fulga

Abstract

Cellular RNAs encode important information regarding cell identity, disease status or exposure to environmental factors. The development of CRISPR-based RNA sensors that are conditionally activated upon detection of endogenous RNA biomarkers would enable restricting CRISPR activity to target cells that were not previously targetable using tissue-specific promoters or targeted delivery approaches. My doctoral work is aiming to develop CRISPR-based RNA sensors by engineering *Streptococcus pyogenes* Cas9 single-guide RNAs (sgRNAs). sgRNAs are engineered to fold into complex secondary structures that, in the ground state, inhibit their activity. Upon recognising complementary RNA biomarkers, these sgRNAs should become activated, enabling Cas9 to perform its function.

In the first part of this project, I implemented a standard Synthetic Biology design-build-test cycle in mammalian HEK293T cells to optimise engineered sgRNA designs. Iterative design optimisations enabled development of a computational pipeline for designing engineered sgRNAs starting from the desired sgRNA and RNA biomarker sequences. Furthermore, insights into the molecular mechanism of RNA detection are also provided. After initial proof-of-concept studies in HEK293T cells, transgenic zebrafish lines were generated to enable rapid screening of different engineered sgRNA designs *in vivo*. Preliminary results suggest that engineered sgRNAs could be repurposed towards RNA detection in zebrafish embryos.

Proposed CRISPR-based sensors design could open up novel avenues towards development of new therapeutic, research and diagnostic applications. However, in order to reach its full potential, further work is still required for optimising this technology *in vivo*. We still need to explore chemical modifications that stabilise sgRNA designs *in vivo*, as well as different endogenous RNA classes that can be efficiently detected. Eukaryotic cells express many classes of transcripts of different sizes, expression levels and subcellular localisations, and these factors are expected to significantly impact dynamic ranges of sgRNA activation.

Acknowledgements

First of all, I would like to thank my supervisors Tatjana and Tudor for guiding me through this challenging DPhil project, while providing me with freedom to explore my scientific interests. I am particularly grateful to Tatjana for welcoming me into her lab and encouraging me to believe in this project and to Tudor for encouraging me to believe in my own strengths.

I would also like to thank all present and past members of the TAF and TSS lab. In particular, I would like to acknowledge Quentin Ferry for closely supervising the first year of my DPhil project, as well as for all the scientific and career advice provided. I would also like to give some special thanks to Vanessa Chong-Morrison for proofreading this thesis, helping me design zebrafish expression plasmids as well as teaching me how to perform zebrafish embryonic microinjections. I am also extremely grateful to Sarah Mayes for being a tremendous support for my project and for teaching me most of the things I know about zebrafish embryos.

I am very thankful for crossing paths with Yale Michaels and David Knapp. I will always remember those fascinating lab meetings where I learnt so much and received so much useful feedback. Martyna Lucoseviciute and Filipa Simoes were also a tremendous help for my project. I would like to thank them for helping me with embryonic injections, teaching me how to perform experiments with zebrafish embryos and for their moral support. I would also like to thank Michael Weinberger and Chloe Tubman for their great advice on how to genotype zebrafish embryos.

It's also been a great pleasure working together with Zhiyuan Hu and Evgeny Akkuratov on different collaboration projects. I would also like to thank Alice Lightowlers, Markus Toegel, Andrea Rodriguez Delherbe, Ivan Ferreira, Adi Harlianto, Ruth Williams and Perihan Ates-Kalkan for all their fantastic scientific advice and moral support. I would also like to thank my best friends from the lab-Muhammad Hanifi and Andrew Ramos. It's been fantastic sharing this journey with them, and it was so much fun working together in the lab until really late at night. I am also grateful to my students Chau Chung, Lauren Overend and Joseph Blayney for making my DPhil experience so much more enjoyable.

I wouldn't have been able to complete this project without tremendous emotional support from my family and friends. I would like to thank my mom, my dad and my grandparents for always being there for me and for believing in me. I would also like to thank my sister for encouraging me not to take things too seriously and my best friend Vanessa for providing me with fantastic life advice. Last, but not least, I would like to express my gratitude to my partner, Andrei. He was extremely supportive and encouraging and always motivated me to stay focused and fight for the things that are important to me.

I would also like to thank the funding bodies that made this project possible. My project was generously supported by the EPSRC & BBSRC Centre for Doctoral Training in Synthetic Biology, EvOX Therapeutics and Wadham College.

Authorship Statement

I declare that the submitted thesis represents my own work, unless stated otherwise. Materials from two publications as well as conference abstracts for which I am the first author were also included in this thesis.

Abstract was adapted from previous conference submissions: ***Pelea*** *O, Mayes S, Ferry Q, Fulga T, Sauka-Spengler T, Development of RNA sensors for controlling CRISPR activity in vivo.* Similar abstracts were submitted for the Keystone Precision Genome Engineering Symposia (Colorado, 2022) as well as the Oxford Genome Engineering Forum (2022). Introduction and Discussion contents were adapted from a literature review that is under process of publication in The CRISPR Journal: ***Pelea*** *O, Fulga T, Sauka-Spengler T, RNA-responsive gRNAs for controlling CRISPR activity: current advances, future directions, and potential applications.* Figures1.1-1.6 were reused from this publication. Content and figures for the Chapters 2-6, as well as Chapters 8 and 9 were adapted from a manuscript that will be submitted for publication shortly after the submission of this thesis. I would also like to acknowledge that the idea of using MREs (microRNA recognition elements) to degrade Cas9 expression in the liver (Section 7.3.1) emerged from a brainstorming session taking place at the NCCR RNA & Disease Summer School (Saas-Fee, Switzerland, 2022).

The following experiments were conducted in collaboration with others:
1. Dr Martyna Lukoseviciute and Dr Filipa Simoes performed embryonic microinjections necessary for generation of zebrafish transgenic lines (Figure 6.1).
2. Sarah Mayes helped with setting up zebrafish crosses for identifying founders and establishing first generation of transgenics (Figure 6.3, Figure 6.4).
3. Several zebrafish injection experiments were performed together with Sarah Mayes (Figure 6.8, Figure 6.9).
4. Several microscopy experiments were performed together with Prof. Tatjana Sauka-Spengler (Figure 6.7) and Sarah Mayes (Figure 6.8, Figure 6.9).

I would also like to acknowledge that Chapters 2 and parts of Chapters 3 and 4 were closely supervised by Dr. Quentin Ferry. At the beginning on my project,

Quentin already had some evidence that iSBH-sgRNAs could detect RNA triggers in HEK293T cells. Experiments presented in these chapters were carried out by me, but the concepts implemented were developed in collaboration with Quentin.

Publications and Conferences

1. ***Pelea*** *O, Fulga T, Sauka-Spengler T, RNA-responsive gRNAs for controlling CRISPR activity: current advances, future directions, and potential applications*, Literature Review (accepted for publication in The CRISPR Journal)

"CRISPR-Cas9 has emerged as a major genome manipulation tool. As Cas9 can cause off-target effects, several methods for controlling the expression of CRISPR systems were developed. Recent studies have shown that CRISPR activity could be controlled by sensing expression levels of endogenous transcripts. This is particularly interesting, as endogenous RNAs could harbour important information about the cell type, disease state, and environmental challenges cells are facing. Single-guide RNA (sgRNA) engineering played a major role in the development of RNA-responsive CRISPR systems. Following further optimisations, RNA-responsive sgRNAs could enable the development of novel therapeutic and research applications. This review introduces engineering strategies that could be employed to modify *Streptococcus pyogenes* sgRNAs with a focus on recent advances made towards the development of RNA-responsive sgRNAs. Future directions and potential applications of these technologies are also discussed."

2. ***Pelea*** *O, Mayes S, Ferry Q, Fulga T, Sauka-Spengler T, Development of RNA sensors for controlling CRISPR activity in vivo.* The following abstract was submitted for the Keystone Precision Genome Engineering Symposia (Colorado, 2022), where I was selected to give a short talk and present a poster. The same abstract was also submitted to the the Oxford Genome Engineering Forum (2022).

"Current ways of restricting the activity of CRISPR systems to cell types of interest include tissue-specific promoters or targeted delivery approaches. However, such strategies are optimised for a limited subset of cell types and tissues. The development of CRISPR-based RNA sensors that are conditionally activated upon detection of endogenous RNA biomarkers that reflect different cell identities, differentiation and disease status or exposure to environmental factors is poised to open new opportunities in research and therapeutics. Here, we use the sgRNA engineering approach to control the activity of *Streptococcus pyogenes* Cas9 (SpCas9)

in response to such endogenous inputs. sgRNAs are engineered to fold into complex secondary structures that, in the ground state, inhibit their activity. Upon recognising complementary RNA biomarkers, these sgRNAs become activated, enabling Cas9 to perform its function.

We implemented a standard Synthetic Biology design-build-test cycle in mammalian HEK293T cells to optimise engineered sgRNAs. Our iterative design optimisations improved the ability of sgRNAs to sense longer synthetic RNAs. By implementing optimised parameters, we developed a computational pipeline for designing engineered sgRNAs starting from the desired sgRNA and RNA biomarker sequences. After initial proof-of-concept studies in HEK293T cells, transgenic zebrafish lines were generated to enable rapid screening of different engineered sgRNA designs in vivo. Preliminary results suggest that engineered sgRNAs could detect synthetic RNA sequences in zebrafish embryos.

Next, we are exploring different classes of endogenous RNAs that could be sensed and used as triggers in vivo. Eukaryotic cells express many classes of transcripts of different sizes, expression levels and subcellular localisations and we expect such factors to greatly impact dynamic ranges of sgRNA activation. Our investigations use RNA sequencing databases that measure levels of different cytoplasmic and nuclear transcripts upregulated in the zebrafish neural crest. A panel of different transcript types expressed in the neural crest has been selected, and specific engineered sgRNAs were designed and delivered into zebrafish embryos by microinjection as part of our screening process. Successful designs are expected to be preferentially active in the neural crest, where complementary endogenous RNAs are expressed, whereas in non-neural crest tissues these designs would display minimal to no activity."

3. *Rajakumar T, Horos R, Jehn J, Schenz J, Muley T,* ***Pelea*** *O et al. A blood-based miRNA signature with prognostic value for overall survival in advanced stage non-small cell lung cancer treated with immunotherapy.* Npj Precis Oncol. 2022;6(1). doi:10.1038/s41698-022-00262-y

"Immunotherapies have recently gained traction as highly effective therapies in a subset of late-stage cancers. Unfortunately, only a minority of patients experience the remarkable benefits of immunotherapies, whilst others fail to respond or even come to harm through immune-related adverse events. For immunotherapies within the PD-1/PD-L1 inhibitor class, patient stratification is currently performed using tumor (tissue-based) PD-L1 expression. However, PD-L1 is an accurate predictor of response in only 30% of cases. There is pressing need for more accurate biomarkers

for immunotherapy response prediction. We sought to identify peripheral blood biomarkers, predictive of response to immunotherapies against lung cancer, based on whole blood microRNA profiling. Using three well-characterized cohorts consisting of a total of 334 stage IV NSCLC patients, we have defined a 5 microRNA risk score (miRisk) that is predictive of overall survival following immunotherapy in training and independent validation (HR 2.40, 95% CI 1.37–4.19; $P < 0.01$) cohorts. We have traced the signature to a myeloid origin and performed miRNA target prediction to make a direct mechanistic link to the PD-L1 signaling pathway and PD-L1 itself. The miRisk score offers a potential blood-based companion diagnostic for immunotherapy that outperforms tissue-based PD-L1 staining."

4. ***Pelea*** *O, Mayes S, Ferry Q, Fulga T, Sauka-Spengler T, Development of RNA sensors for controlling CRISPR activity in vivo*, Research Article (manuscript in progress, to be submitted for publication shortly after submission of this thesis)

"Cellular RNAs encode important information regarding cell identity, disease status or exposure to environmental factors. The development of CRISPR-based RNA sensors that are conditionally activated upon detection of endogenous RNA biomarkers would enable restricting CRISPR activity to target cells that were not previously targetable using tissue-specific promoters or targeted delivery approaches. Here we develop CRISPR-based RNA sensors by engineering *Streptococcus pyogenes* Cas9 single-guide RNAs (sgRNAs). Engineered sgRNAs enable detection of plasmid-expressed RNA in HEK293T cells and zebrafish embryos. Iterative design optimisations allowed development of computational tools for designing sgRNAs capable of detecting RNAs of choice. Furthermore, insights into the molecular mechanism of RNA detection are also provided. We anticipate that our sensors would enable CRISPR activation in response to detection of endogenous RNA biomarkers and would open up novel opportunities towards development of research and therapeutic applications."

Patents

So far, no intellectual property arising from this project was protected. Nevertheless, I would like to retain the right of filing a patent application at a later date. Therefore, I would like to apply for a moratorium and prevent the publication of this thesis for the next 12 months.

Contents

List of Figures **xv**

List of Abbreviations **xvii**

1 Introduction **1**

1.1 CRISPR-Cas9: from bacterial immunity towards therapeutic applications . . . 1
1.2 Importance of controlling CRISPR activity . . . 6
1.3 Restricting CRISPR activity to tissues of interest *in vivo* . . . 8
1.4 The unexplored potential of RNA biomarkers for cell identity . . . 10
1.5 Progress made towards development of CRISPR-based RNA sensors 12
1.5.1 RNA sensing through engineered expression cassettes . . . 12
1.5.2 RNA sensing through engineered sgRNAs . . . 14
1.6 Purpose of this study . . . 29

2 Testing if iSBH-sgRNAs could be repurposed for RNA detection 32

2.1 Introduction to first-generation iSBH-sgRNA designs . . . 32
2.2 Introduction to the experimental pipeline for testing iSBH-sgRNAs in HEK293T cells . . . 33
2.3 First-generation iSBH-sgRNA designs detect short RNA triggers . . 36
2.4 Improving iSBH-sgRNA OFF states by changing CRISPRa reporter systems . . . 36
2.5 Activation of first-generation iSBH-sgRNA designs is trigger-specific 38
2.6 Closing remarks . . . 38

3 Further iSBH-sgRNA design optimisation and characterisation 41

3.1 Introduction to second-generation iSBH-sgRNA designs . . . 41
3.2 Testing if second-generation iSBH-sgRNA designs detect RNA triggers 42
3.3 Second-generation iSBH-sgRNA display superior performances while detecting longer RNA triggers . . . 44
3.4 Trigger features that influence second-generation iSBH-sgRNA activation . . . 46
3.4.1 Trigger length and secondary structure . . . 48

3.4.2 Mammalian expression vectors used to express RNA triggers 48
3.5 Closing remarks . . . 49

4 Designing iSBH-sgRNAs for custom CRISPR experiments 50
4.1 Limitations of second-generation iSBH-sgRNAs and need for modular iSBH-sgRNA designs . . . 50
4.2 MODesign computational pipeline . . . 51
4.3 Experimentally validating MODesign outputs . . . 53
4.4 Closing remarks . . . 56

5 Insights into the mechanism of iSBH-sgRNA activation 59
5.1 Formation of double-stranded RNA structures and its implications . 59
5.2 Measuring the size of iSBH-sgRNAs following activation . . . 61
5.3 Measuring the size of RNA triggers following activation . . . 62
5.4 Closing remarks . . . 62

6 Testing performance of iSBH-sgRNAs *in vivo* 67
6.1 Generating transgenic zebrafish lines . . . 67
6.1.1 Transgenic line design . . . 67
6.1.2 Establishment of transgenic lines . . . 68
6.1.3 Validating transgenic lines . . . 71
6.2 Testing if iSBH-sgRNAs detect synthetic RNA triggers *in vivo* . . . 75
6.3 Troubleshooting sgRNA delivery . . . 78
6.4 Testing a first version of iSBH-sgRNAs with chemical modifications *in vivo* . . . 80
6.5 Closing remarks . . . 82

7 Discussion and Further Directions 83
7.1 Results summary . . . 83
7.2 Lessons to be learnt from similar technologies . . . 86
7.2.1 Chemical modification improve probe design . . . 86
7.2.2 Probe design rules . . . 90
7.2.3 Avoiding off-target effects . . . 93
7.3 iSBH-sgRNA technology in the existing landscape of CRISPR-based RNA sensors . . . 94
7.3.1 iSBH-sgRNAs vs engineered CRISPR expression cassettes . 94
7.3.2 iSBH-sgRNAs vs other engineered sgRNAs . . . 96
7.4 Characterising endogenous RNAs that could act as triggers *in vivo* . 97
7.5 Future applications of this technology . . . 99
7.5.1 Potential iSBH-sgRNA applications within the CRISPR therapeutic space . . . 99

7.5.2 Potential iSBH-sgRNA applications in basic research 100
7.5.3 Potential iSBH-sgRNA applications for *in vitro* diagnostics . 101

8 Materials and Methods **103**
8.1 *In silico* design of iSBH-sgRNAs 103
8.2 Molecular cloning 105
8.3 Maintaining HEK293T cell lines 108
8.4 Transfecting HEK293T cell lines 108
8.5 Flow Cytometry sample preparation 109
8.6 Flow Cytometry data analysis 109
8.7 RNA circularisation assays 110
8.8 Zebrafish husbandry 111
8.9 Synthesis of mRNA for embryonic injections 111
8.10 Generation of zebrafish transgenics 111
8.11 Microinjection of zebrafish embryos using the Ac/Ds system 113
8.12 Microinjection of zebrafish embryos with chemically modified sgRNAs 113
8.13 Microinjection of zebrafish embryos with chemically modified iSBH-sgRNAs 114
8.14 Confocal microscopy 114
8.15 Code and plasmid availability 115

9 Supplementary Materials **116**

References **137**

List of Figures

1.1 Engineering sgRNA expression cassettes for RNA sensing 4
1.2 sgRNA engineering strategies 15
1.3 Conditional sgRNA designs 19
1.4 OFF-ON sgRNA switches 22
1.5 ON-OFF sgRNA switches 23
1.6 Other OFF-ON gRNA switches 26
1.7 Restricting CRISPR activity to target cells through RNA detection *in vivo* 31

2.1 Native sgRNA vs RNA-sensing first-generation iSBH-sgRNA 33
2.2 Experimental pipeline for testing iSBH-sgRNAs in HEK293T cells . 34
2.3 Expected iSBH-sgRNA system behaviour inside HEK293T cells . . 35
2.4 Testing if first-generation iSBH-sgRNAs detect short RNA triggers in HEK293T cells 37
2.5 Representative Flow Cytometry plots 39
2.6 iSBH-sgRNA activation depends on complementary RNA triggers . 40

3.1 Comparing first-generation and second-generation iSBH-sgRNA designs 42
3.2 CRISPRa reporters of choice influence ON/OFF ratios of second generation iSBH-sgRNA designs while detecting short RNA triggers 43
3.3 Second generation iSBH-sgRNAs detect longer RNA triggers in HEK293T cells 45
3.4 Assessing if RNA trigger length influences activation of second-generation iSBH-sgRNAs 46
3.5 Trigger features that influence activation of second generation iSBH-sgRNAs 47

4.1 Modular iSBH-sgRNA designs enable spatial separation of spacer and trigger-sensing sequences 51
4.2 Exploring design space for modular iSBH-sgRNA designs 52
4.3 Computational pipeline for modular iSBH-sgRNA design 54
4.4 Designing modular iSBH-sgRNAs starting from one RNA trigger and 3 sgRNAs 55

4.5 Designing modular iSBH-sgRNAs starting from 4 RNA triggers and one sgRNA sequence 57

5.1 RNA interference pathway 60
5.2 Potential scenarios for the mechanism of iSBH-sgRNA activation . . 61
5.3 Measuring the size of iSBH-sgRNA following activation 63
5.4 Measuring the size of RNA triggers following activation 64
5.5 Potential mechanisms underlying iSBH-sgRNA and RNA trigger truncation 66

6.1 Generating zebrafish transgenic line design 69
6.2 Vision for future *in vivo* RNA detection experiments 70
6.3 Identifying founders for the dCas9-Vp64 transgenic line 72
6.4 Establishing the F1 generation of transgenics 73
6.5 Validating transgenic lines 74
6.6 Strategy for testing if iSBH-sgRNAs detect synthetic RNA triggers *in vivo* 76
6.7 Testing if iSBH-sgRNAs detect synthetic RNA triggers *in vivo* . . . 77
6.8 Troubleshooting sgRNA expression across tissues 79
6.9 Testing a first version of iSBH-sgRNA with chemical modifications . 81

7.1 Chemical modifications that could improve iSBH-sgRNA designs . . 87

List of Abbreviations

AAV . . . . . - adeno-associated virus

ADAR . . . . . - adenosine deaminase

Apo E . . . . . - apolipoprotein E

ASO . . . . . - antisense oligonucleotide

bp . . . . . - base pairs

C . . . . . - carbon

CAR-T cells . - chimeric antigen receptor T cells

CDS . . . . . -coding DNA sequence

CMV . . . . . - cytomegalovirus

CRISPRa . . . - CRISPR (transcriptional) activators

crRNA . . . . - CRISPR RNA

CTS . . . . . - CRISPR target sequence

dCas9 . . . . . - dead Cas9 with no nuclease activity

DSB . . . . . - double-strand break

dsRNA . . . . - double-stranded RNA

ECFP . . . . . - enhanced cyan fluorescent protein

EGFR . . . . . - epidermal growth factor receptor

ENCODE . . . - Encyclopaedia of DNA Elements

epegRNA . . . - engineered prime editing gRNA

F0 . . . . . - founding generation of a transgenic line

F1 . . . . . - first generation of a transgenic line

gRNA . . . . . - guide RNA

GTEx . . . . . - genotype-tissue expression

HDR . . . . . - homology-directed repair

HCR . . . . . - hybridisation-chain reaction

hPSCs . . . . . - human pluripotent stem cells

IFN . . . . . - interferon

IL6 . . . . . - interleukin 6

iSBH-sgRNA - inducible spacer-blocking hairpin sgRNA

FISH . . . . . - fluorescent *in situ* hybridisation

LNA . . . . . - locked nucleic acid

lncRNA . . . . -long non-coding RNA

LNP . . . . . - lipid nanoparticle

Me . . . . . - methyl

miR . . . . . - microRNA

miRNA . . . . - microRNA

MRE . . . . . - microRNA recognition element

mRNA . . . . - messenger RNA

HNEJ . . . . . - nonhomologous end-joining

NLS . . . . . - nuclear localisation signal

nt . . . . . - nucleotides

PAM . . . . . - palindromic associated motif

pegRNA . . . - prime editing gRNA

poly(A) . . . . - polyadenylation

PS . . . . . - phosphorothioate

O . . . . . - oxygen

RBS . . . . . - ribosomal binding site

RNP . . . . . - ribonucleoprotein

RADARS . . . - Reprogrammable ADAR Sensors

RENDR . . . - Ribozyme-ENabled Detection of RNA

RNAi . . . . . - RNA interference

rRNA . . . . . - ribosomal RNA

RT . . . . . - reverse transcription

S . . . . . - sulphur

sgRNA . . . . - single guide RNA

shRNA . . . . - short hairpin RNA

siRNA . . . . . - small interfering RNA

SSB . . . . . - single-stranded break

ssDNA . . . . - single-stranded DNA

ssRNA . . . . - single-stranded RNA

Tm . . . . . - melting temperature

tracrRNA . . - trans-activating crRNA

tRNA . . . . . - transfer RNA

TTR . . . . . - transthyretin

1 Introduction

1.1 CRISPR-Cas9: from bacterial immunity towards therapeutic applications

Recent years have seen CRISPR technology become one of the most essential tools for genome engineering[1]. CRISPR systems were initially discovered and characterised in prokaryotic systems, where they serve as a defence system against phage invaders[2]. Subsequent technology development efforts enabled repurposing of CRISPR systems towards programmable DNA cleavage[3] and gene editing applications in mammalian cells[4–6] as well as animal models[7–9]. Furthermore, recent advances enabled development of CRISPR-based therapeutic strategies for correcting disease-causing mutations in humans and several clinical trials were also carried up to date[10].

CRISPR systems were discovered in the context of bacterial and archaeal immunity against phage invaders. These systems enable prokaryotes to distinguish self from non-self-nucleic acids. In prokaryotes, CRISPR systems are encoded in arrays consisting of DNA repeat sequences separated by variable sequences as well as *cas* genes. Variable sequences, also known as spacers, have their origin in viral or plasmid sequences. *Cas* genes are a large family of effector proteins that act on

nucleic acids, including nucleases, nucleic-acid binding proteins, polymerases or helicases[2].

The general mechanism of action for CRISPR-Cas systems is maintained between different prokaryotic species and involves three main steps. Firstly, spacer sequences are acquired from invading nucleic acid species and integrated into the host genome. This process is carried out by Cas1 and Cas2, two proteins that are conserved amongst different CRISPR systems. The second step involves transcription of the CRISPR arrays containing direct repeats and their processing into functional CRISPR-RNAs (crRNA). Thirdly, mature crRNAs interact with Cas effector proteins and target them to complementary nucleic acid species, resulting in either DNA or RNA cutting[11].

Based on the mechanism governing crRNA processing as well as the number of Cas effector proteins involved in RNA/DNA interference, CRISPR systems could be divided into two classes. In class 1 CRISPR systems, crRNA processing is typically carried out by Cas6, while effector modules consist of multiple Cas proteins. In class 2 CRISPR systems, crRNA processing is carried out by bacterial RNase III or by the Cas effector proteins themselves. Class 2 systems are also characterised by the presence of single, large effector protein. Depending on the effector protein, class 2 systems could be further divided into 3 different types. Type II systems have Cas9 as the main effector, type V systems rely on Cas12 effectors, while type VI systems have Cas13 as their main effector. While Cas9 and Cas12 mediate targeted DNA cutting, Cas13 is involved in RNA cleavage[12].

Type II CRISPR systems were the first to be biochemically characterised[3, 13] and to gain popularity in the genome editing field. Important components of *Streptococcus pyogenes* CRISPR systems are the SpCas9, crRNA and the trans-activating crRNA (tracrRNA)[13]. Within this complex, crRNA plays an essential role in programming Cas9 to recognise a desired DNA sequence, while the tracrRNA

is involved in mediating molecular interactions with Cas9 (Figure 1.1.A).

For downstream technology developments, the number of CRISPR components could be further reduced by using single-guide RNAs (sgRNAs) obtained by fusing the crRNA and the tracrRNA (Figure 1.1.B). Changing the spacer sequence of the crRNA enables targeting Cas9 to desired DNA sequences, where it creates a double-strand break (DSB). DNA sequences that could be targeted by Cas9 are restricted to the presence of an NGG palindromic associated motif (PAM). This short DNA sequence is involved in R-loop formation, and it is essential for the Cas9's ability to cleave double-stranded DNA[3].

Bacterial and *in vitro* work suggested that CRISPR-Cas9 could become a useful tool for programmable genome editing[3]. Soon after these findings, several labs made progress towards implementing CRISPR systems in mammalian cells[4–6]. Cas9 was found to be able to introduce DNA DSBs in mammalian cells, with endogenous DNA repair machineries involved in fixing such breaks.

Depending on the DNA repair pathway of choice, CRISPR-Cas9 enabled introduction of disruptive frameshift mutations in coding genes or could facilitate precise genome editing. In the absence of a provided repair template for homology-directed repair (HDR), nonhomologous end-joining (NHEJ) is the pathway of choice for repairing DSBs. This is an error-prone mechanism where small deletions or insertions are introduced during break repair, resulting in frameshift mutations. Alternatively, supplying cells with DNA templates for HDR could result in precise DNA editing at desired genomic locations[6].

Soon after showing that CRISPR-Cas9 could edit DNA in mammalian cells, evidence was also provided that CRISPR could edit DNA *in vivo*. Editing of mutations associated with Duchenne Muscular Dystrophy in mouse germline cells

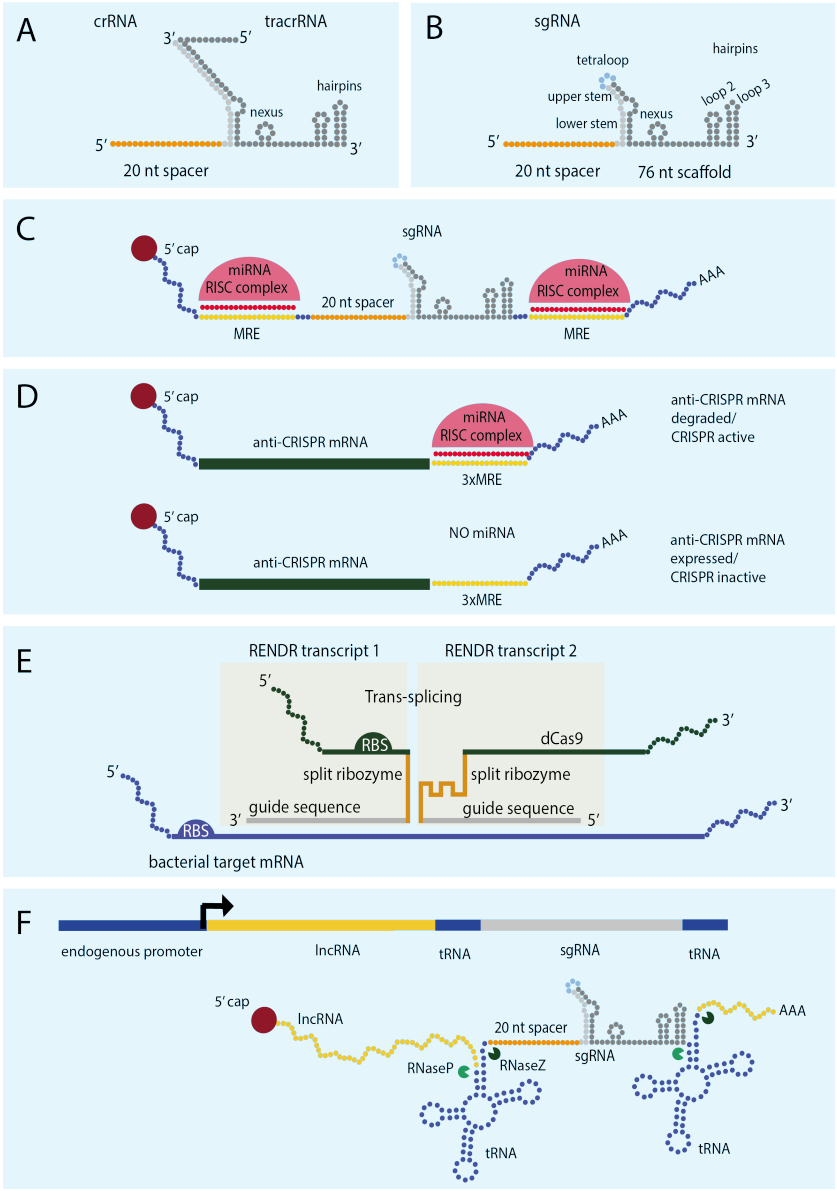


Figure 1.1: Engineering sgRNA expression cassettes for RNA sensing

A. CRISPR RNA (crRNA) base-pairs with the trans-activating CRISPR RNA (tracrRNA), forming a functional gRNA complex[3]. **B.** Single-guide RNAs (sgRNAs) are generated by fusing crRNA and tracrRNAs while employing a minimal tetraloop structure (cyan)[3]. The sgRNA sequence is composed of 6 functional modules: spacer, lower stem, upper stem, nexus as well as two hairpins[14]. **C.** MicroRNA-recognition elements (MREs) enable sgRNA processing from polymerase II transcripts[15]. **D.** Anti-CRISPR proteins are employed for microRNA sensing. In tissues expressing microRNAs of interest, anti-CRISPR genes are silenced, thus derepressing CRISPR activity[16, 17]. **E.** Bacterial mRNA targets could be detected using RENDR technology. Upon recognition of the target mRNA, two RENDR transcripts are located in close spatial proximity, resulting in ribozyme reconstitution. dCas9 expression is enabled by the fusion between RBS (ribosomal binding-site) and the dCas9 coding sequence[18]. **F.** Endogenous loci expressing long non-coding RNAs (lncRNAs) are engineered for monitoring lncRNA transcription. This is achieved by fusing lncRNAs with an tRNA scaffolds and sgRNA expression cassettes[19]. *Figure was re-used from a literature review accepted for publication in The CRISPR Journal.

were reported[9]. After initial proof-of-concept experiments showing that CRISPR-Cas9 genome editing works *in vivo*[9], several other studies developed strategies for correcting disease-causing mutations in somatic cells[7, 8, 10, 20].

Several therapeutic strategies were developed to date including *ex vivo* versus *in vivo* gene editing. *Ex vivo* strategies are mainly applied to editing blood cells. Cells are collected from patients, edited *ex vivo*, and edited cells are then transplanted back into the patient[10].This strategy could be used for treatment of hemoglobinopathies, including sickle-cell anaemia and β-thalassemia[21, 22] as well as for engineering CAR-T (chimeric antigen receptors T-cell) cells with superior performance[22, 23]. In contrast to *ex vivo* strategies, *in vivo* gene editing involves systemic or localised administration of CRISPR-Cas9[10]. Examples of *in vivo* therapeutical strategies involve editing of eye[24] or liver cells[25].

For example, Leber congenital amaurosis type 10 is a retinal dystrophy caused by mutations in the CEP290 gene. EDIT-101 is a candidate genome editing strategy that restored vision loss associated with these mutations in primates[24] and it is currently being tested in clinical trials[10]. As part of this strategy, Cas9 expression is controlled by tissue-specific promoters and packaged together with sgRNAs into adeno-associated virus 5 (AAV5). AAV5 vectors were delivered by sub-retinal injections to further ensure localised CRISPR delivery[24].

The liver is another excellent candidate for *in vivo* CRISPR therapeutics due to its discontinuous capillary structures facilitating accumulation of systemically administered nanoparticles[26]. Transthyretin Amyloidosis is a severe disease characterised by the accumulation of misfolded transthyretin (TTR) proteins in tissues such as nerves and heart. As most misfolded proteins are produced by the liver, knocking out TTR in the liver was proposed as a therapeutic strategy. NTLA-2001 was proposed as an *in vivo* genome editing formulation comprising of Cas9 messenger RNA (mRNA) and chemically modified sgRNAs targeting TTR,

loaded into liver-targeting lipid nanoparticles (LNPs). As part of a recent clinical trial, administration of NTLA-2001 to a small group of patients resulted in mild adverse events as well as reduction of TTR in the serum[25].

1.2 Importance of controlling CRISPR activity

Despite the recent advances in the clinic and great therapeutic potential, CRISPR-Cas9 off-target effects remain a safety concern. Common off-target effects of CRISPR-Cas9 involve directing Cas9 to non-desired genomic DNA regions and induction of unwanted DSBs[27]. Furthermore, DSBs induced by Cas9 were shown to promote chromosomal rearrangements and deletions[28] as well as to induce a p53-mediated DNA damage response[29].

In edited retinal pigment epithelial cells, selection against edited cells with functional p53 was reported. As p53 is an important tumour-supressing factor, selection against cells with functional p53 could promote accumulation of tumour-promoting mutations in edited cells[29]. Such findings were also supported by another study carried out in human pluripotent stem cells (hPSCs). In hPSCs, gene editing efficiencies were substantially lower than in other cell types. This effect was shown to be caused by toxic effects associated with the p53 response and death of edited cells. Such findings warned scientists to proceed with caution when developing cell replacement therapies based on hPSCs edited with CRISPR[30]. Apart from toxicity associated with DSBs, DNA repair pathways active in different tissues could further complicate the landscape of *in vivo* genome editing[31].

One of the classical ways of reducing CRISPR off-target effects involves modulating system expression through appropriate choice of delivery vectors. Different CRISPR delivery strategies components use viral vectors encoding Cas9 and sgRNA, delivery of Cas9 mRNA and sgRNA as well as delivery of Cas9 recombinant protein complexed with sgRNA. For example, delivery of Cas9 mRNA or Cas9 protein/sgRNA complexes could reduce off-target effects by limiting the duration of Cas9 expression

inside cells[32].

Alternative Synthetic Biology approaches were also suggested for controlling CRISPR activity and reducing off-target effects. Cas9 versions with decreased OFF-target effects were evolved through directed evolution methods in *Escherichia coli* (*E. coli*)[33]. Use of chemically modified sgRNAs that incorporate high proportion of modified DNA nucleotides could also reduce Cas9 binding to off-target sites[34]. Furthermore, sgRNA modifications that enable controlling CRISPR activity following exposure to light[35] or small molecules[36–40] were also developed. Methods involving extensive sgRNA engineering will be discussed in more detail later on.

Due to toxicity associated with DSBs, avoiding DSB production was explored as a potential strategy. HNH and RuvC are two functional domains within the Cas9 nuclease lobe, and each domain is responsible for cleaving one of the two strands of the DNA duplex. Inactivating point mutations in these domains enabled development of Cas9 enzymes that created DNA single-strand breaks (SSBs) rather than DSBs. These mutant Cas9 versions are known as Cas9 nickases[3]. For example, introduction of DSBs at unintended sites could be prevented by employing two Cas9 nickases to introduce two SSBs in close proximity to one another. Increased specificity is ensured as SSBs are repaired with high fidelity inside cells, while DSBs are formed when two single-strand breaks are formed in close proximity to one another[41].

Nicking Cas9 can serve as a landing pad for recruiting other DNA-modifying enzymes to target DNA sequences. Such interactions enabled development of second-generation CRISPR editors such as base editing[42, 43] and prime editing technologies[44, 45]. Base editors enable correction of DNA point mutations[42, 43], while prime editors enable creation of precise insertions or small deletions at desired genomic locations[44, 45]. Coupling prime editing with the use of serine integrases also enabled insertion of large fragments (up to 5kb) as well as

creation of large insertions, inversions and deletions at desired genomic locations[46].

Inactivation of both HNH and RuvC Cas9 domains results in a catalytically dead Cas9 (dCas9) version that has retained DNA binding activity, but is unable to cut DNA[47]. dCas9 also serves as a platform for recruiting transcriptional activation modules[48–50], transcriptional repressors[47] as well as transposases[51] to desired genomic locations.

1.3 Restricting CRISPR activity to tissues of interest *in vivo*

Restricting delivery of CRISPR components to disease-affected tissues is another way of controlling CRISPR activity *in vivo*. In this way, off-target effects are prevented in tissues that are not affected by the disease[52]. *Ex vivo* gene editing is a good way of overcoming editing in unintended tissues[10]. In terms of *in vivo* editing, localised injection of CRISPR components[24] appears to be a good strategy for editing disease-associated mutations in the eye and liver[25], as previously discussed. Nevertheless, systemic administration of CRISPR components to other tissues apart from the liver is more challenging[10, 26].

Standard methods for delivering CRISPR components include adeno-associated vectors (AAVs) and lipid nanoparticles (LNPs)[10]. AAVs appear to be the viral vectors of choice for CRISPR applications as their lifecycle does not involve genomic integration[53]. Natural AAV serotypes with specific tissue tropism were identified. For example, AAV1 serotypes are used for transgene expression in the brain, AAV8 in liver and muscle, AAV2 and AAV5 in retina and AAV6 in muscles. Further control of tissue-specific Cas9 expression could be achieved by including tissue-specific promoters within AAV vectors[54]. Apart from naturally-occurring AAV serotypes, additional engineering efforts enabled reducing liver targeting and increasing AAV tropism for desired tissues. This was achieved by genetically engineering the AAV capsid[55, 56].

Despite being promising vectors, AAVs have some limitations. Constructs packaged in AAV vectors should be below 4.2 kb in size[53]. Integration of AAV genome or AAV genome fragments at Cas9-induced DSBs was also observed[57], while prolonged Cas9 expression could lead to undesired off-target effects. Furthermore, patients might have pre-existing neutralising antibodies for AAVs and the costs of AAV production are very high[53]. Pre-existing immunity was also reported for Cas9[58]. Given the immunogenic potential of Cas9, an ideal *in vivo* based therapeutic would involve a single administration of a highly-efficient CRISPR formulation[52]. Achieving transient Cas9 expression might also be beneficial in this regard[32] by delivering mRNA or Cas9-sgRNA RNPs (ribonucleoproteins) through LNPs.

Recent advances were made towards engineering LNPs to target other organs apart from liver[59–61]. Standard LNP formulations include ionizable cationic lipids, cholesterol, polyethylene glycol as well as amphipathic phospholipids[62]. Incorporation of antibodies within lipid formulations represents a way of targeting LNPs to organs of interest. For example, anti-EGFR (epidermal growth factor receptor) antibodies enabled selective delivery of CRISPR components to ovarian tumours expressing high levels of EGFR. Docking interactions between antibody and receptor favoured LNP uptake in cells overexpressing EGFR[60].

Another way of modulating tissue-targeting properties of LNPs involves addition of various amounts of permanent cationic lipids. In this way, LNP formulations that preferentially target muscles, brain, lung[61] as well as epithelial cells, endothelial cells, B cells and T cells[59] were developed. While delivered systemically, serum proteins bind to LNPs and facilitate their tissue-specific uptake.

For example, the liver targeting LNPs used in clinical trials for the treatment of

Transthyretin Amyloidosis are recognised by the serum protein ApoE (Apolipoprotein E). ApoE coats the surface of LNPs and helps liver-specific uptake[25]. Mechanistic studies suggested that introduction of permanent cationic lipids within LNPs changes their tissue-targeting preferences by favouring LNP interaction with different serum proteins. Subsequent interaction between interacting serum proteins and cellular receptors overexpressed on the surfaces of different tissues, seems to play important roles in LNP uptake[62].

Current developments in viral and LNP delivery vectors improved transgene delivery to target tissues, but are not yet perfect[55, 59]. So far, targeted delivery strategies appear to heavily rely on the availability of cell-surface protein biomarkers[60, 62] such as membrane receptors which act as a proxy for cell identity. Molecular interactions between viral capsid proteins[63] or proteins associated with LNPs[62] favour uptake of delivery agents to different target tissues.

1.4 The unexplored potential of RNA biomarkers for cell identity

For cell types or diseased cells where finding precise surface biomarkers or tissue-specific promoters could be challenging, using RNA biomarkers to activate CRISPR activity might represent an excellent alternative solution[64]. The pool of membrane protein receptors is relatively small when compared to the pool of different species of RNAs expressed in different cell types. Membrane proteins were estimated to represent about 27% of the total proteins expressed by human cells[65]. Nevertheless, the ENCODE (Encyclopaedia of DNA Elements) project estimated that protein-coding exons represent only about 1.22% of the total human genome[66]. Therefore, while looking for cell-specific membrane proteins, one would be restricted to the information encoded in roughly only 0.33% of the human genome.

ENCODE also provided evidence that the human genome is extensively transcribed[67]. Approximately 80% of the human genome was predicted to be transcribed into various RNA species[66]. Therefore, identification of tissue-specific RNA biomarkers provides a much larger sample size. Different types of RNA are found within eukaryotic cells, including messenger RNA (mRNA), ribosomal RNA (rRNA) and transfer RNA (tRNA)[68].

Apart from these well-known RNA species, several types of non-coding transcripts were also identified[68]. This includes short non-coding RNAs as well as long non-coding RNAs (lncRNAs). Short non-coding RNAs are RNA species shorter than 200 nucleotides such as: microRNAs[69], small-interfering RNAs[69], small-nuclear RNAs[70], small nucleolar RNAs[71] as well as PIWI-interacting RNAs[72]. LncRNAs are non-protein coding transcripts that are larger than 200 nucleotides[73]. It is estimated that roughly 30% of the total lncRNAs have their origin in transposable elements[74].

Studies such as the Genotype-Tissue Expression (GTEx)[75] and ENCODE[76] performed extensive RNA sequencing experiments in different tissues of the human body. Furthermore, projects such as the Tabula Sapiens carried single-cell RNA sequences experiments for further understanding variations in gene expression that occur in different cell types[77]. Starting from such datasets, downstream studies suggested that approximately half of the annotated human genes displayed tissue-specific expression profiles[78].

Other sequencing projects such as GENCODE v7 aimed to characterise differential expression of lncRNAs across human tissues[79]. A large fraction of human lncRNAs display tissue-specific expression. Furthermore, these RNA species were often restricted to single cell lineages[80]. For example, large fractions of lncRNAs were specifically expressed in the brain[79] and catalogues of heart-specific lncRNAs are also available[81]. lncRNAs transcribed from super-enhancer regions could

also be good indicators of cell identity[82], as well as lncRNAs transcribed from transposable elements[83].

Smaller RNA species such as microRNAs were also shown to display tissue-specific expression. Expression of almost 2000 microRNAs was characterised in over 61 biopsies collected from different human tissues[84]. Some well-known examples of tissue-specific microRNAs involve: miR-122a (liver), mir-129 (brain), mir-200 (epithelial tissues) as well as miR-192 (gastrointestinal organs)[85]. Furthermore, tissue-specific variations in tRNA gene expression were also described[86].

1.5 Progress made towards development of CRISPR-based RNA sensors

Despite the availability of a large pool of tissue-specific RNA biomarkers, linking CRISPR activity with the detection of endogenous RNA species is limited by currently available RNA-detection technologies. Recent studies demonstrated that CRISPR activity could be controlled by sensing endogenous RNA molecules [15–17, 19, 87–94]. This was achieved by either engineering expression cassettes encoding CRISPR components or by designing RNA-responsive sgRNAs. In this section, current progress made towards the development of CRISPR-based RNA sensors is discussed.

1.5.1 RNA sensing through engineered expression cassettes

When delivered as plasmids, sgRNAs are typically expressed from polymerase III promoters, but with specific modifications in the construct design, expression using polymerase II promoters is also possible[95]. When transcribed by polymerase II, sgRNAs are capped and polyadenylated. Although the cap structure sterically blocks Cas9 activity, endogenous factors such as Drosha can be deployed to remove the cap.

For example, sgRNAs can be expressed as part of polycistronic transcripts that also encode short-hairpin RNAs (shRNAs). Subsequently, endogenous Drosha

processing of shRNAs results in the release of sgRNAs and the removal of the cap[96]. Alternative ways of releasing sgRNAs from polymerase II transcripts include RNA interference (RNAi), self-cleaving ribozymes, site-specific nucleases or tRNA scaffolds[95]. As tRNA scaffolds could also act as polymerase III promoters, minimal tRNA scaffolds without transcriptional activity can also be used[97].

The sensing system using engineered polymerase II cassettes was previously employed to detect endogenous microRNAs in mammalian cells. This was achieved by flanking sgRNAs with microRNA-recognition elements (MREs, Figure 1.1.C.). In cells that express microRNAs of interest, polymerase II transcripts are processed, resulting in the formation of active Cas9-sgRNA complexes[15]. Moreover, the stability of Cas9-encoding transcripts was also regulated by MREs. In this way, the presence of target microRNAs could silence the expression of Cas9[90].

Alternative ways of controlling CRISPR-Cas9 activity involves deploying or controlling the expression of anti-CRISPR proteins. Anti-CRISPR proteins inhibit the activity of Cas9-sgRNA complexes through different mechanisms, including stoichiometric and enzymatic inhibition and covalent Cas9 modifications or sgRNA cleavage[88]. To control the expression of one such protein that blocks Cas9 activity, AcrIIA4, specific MREs were included within the AcrIIA4 transcripts. As a result, AcrIIA4 is degraded in cells expressing microRNAs of interest, resulting in a specific tune-up of CRISPR activity[16].

A similar approach was used for restricting CRISPR activity to tissues of interest in mice. *Neisseria meningitidis* Cas9 (NmeCas9) and AcrIIC3 anti-CRISPR protein were included in two different adeno-associated viruses (AAVs). AcrIIC3 mRNA contained MREs for a liver-specific microRNA, miR-122 (Figure 1.1.D.). After systemic co-delivery of AAV vectors, heart and liver tissues were collected, and editing was shown to occur preferentially in the liver[17].

mRNA sensing was also demonstrated in (*E.coli*) by the Ribozyme-ENabled Detection of RNA (RENDR) technology. A dCas9 expression cassette was split into two transcripts. First transcript contained a ribosome-binding site (RBS), while the second transcript consisted of the dCas9 coding DNA sequence (CDS). Split ribozyme sequences and guide sequences complementary with the mRNA to be sensed were also included within the two mRNA sequences. The two engineered transcripts bind to their target mRNA in close proximity, leading to the formation of a functional ribozyme. A trans-splicing reaction enabled formation of a functional dCas9 expression cassette, consisting of both RBS and CDS (Figure 1.1.E.)[18].

Genome engineering was also employed for sensing lncRNAs. To this end, sgRNAs flanked by tRNA scaffolds were inserted within the 3' ends of the lncRNAs of interest (Figure 1.1.F.). sgRNAs were transcribed together with lncRNAs and processed by endogenous RNase P and RNase Z before interacting with Cas9[19].

1.5.2 RNA sensing through engineered sgRNAs

Several engineering efforts were made for designing CRISPR-based RNA sensors based on sgRNA engineering [87, 89, 91, 94, 98–103]. These sensors enabled efficient RNA detection *in vitro* and in bacterial system. Limited evidence is also provided that engineered sgRNAs could detect RNA triggers in mammalian cells. In order to facilitate a better understanding of these sensors, an introduction will be provided to existing sgRNA engineering approaches. Studies that identified positions at which sgRNAs could be engineered will be introduced, together with development of other conditional sgRNA designs. After that, RNA detectors based on sgRNA engineering will be presented.

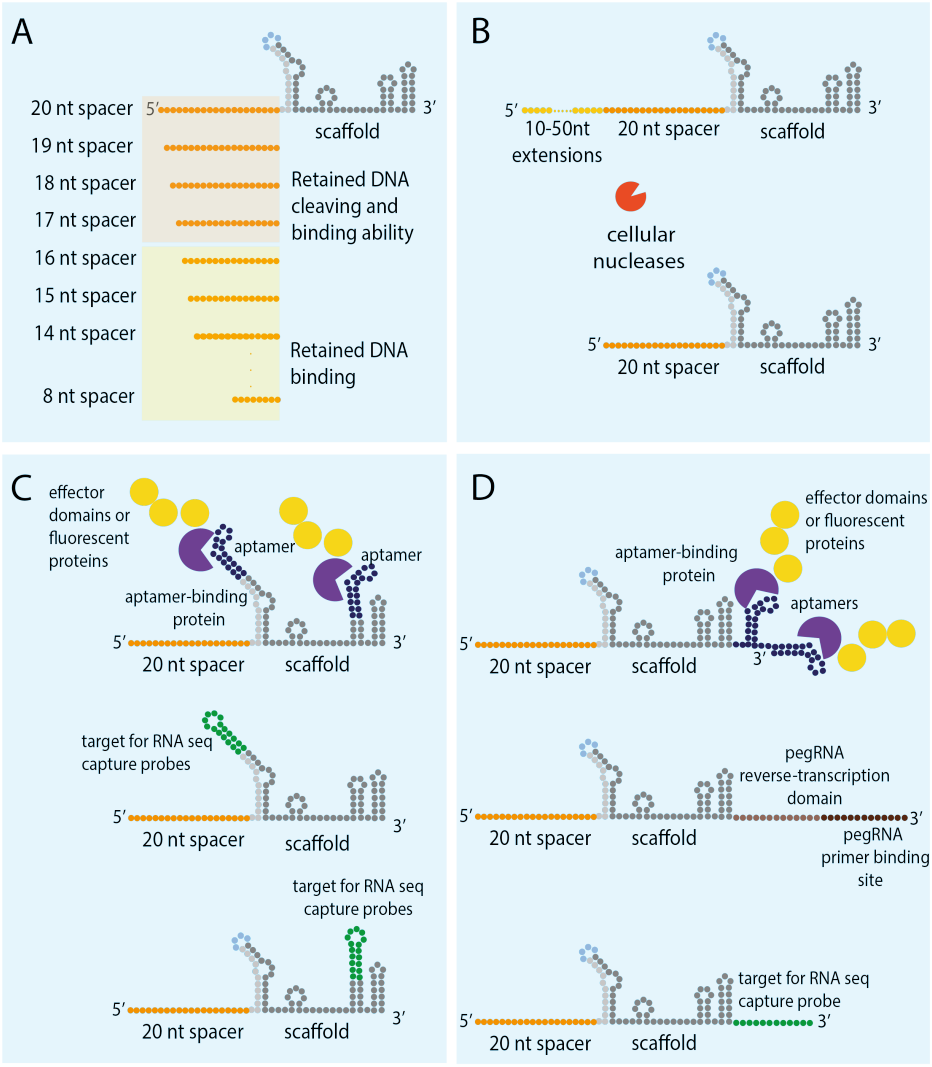


Figure 1.2: sgRNA engineering strategies

In this figure, sgRNA spacer sequences are represented in orange, the "GAAA" tetraloop in cyan and the sgRNA scaffold sequences in grey. **A.** Cas9-sgRNA complexes containing truncated spacers sequences (down to 17nt) maintain their ability to cut DNA [104], while Cas9-sgRNA complexes with shorter spacer truncations (down to 8nt) still maintain their ability to bind target DNA[105, 106]. **B.** Extended spacer sequences that do not form complex secondary structures are truncated back to canonical 20nt spacer sequences. Truncation is performed by an uncharacterised mechanism, likely involving cellular RNases[41, 107]. **C.** Examples of sgRNA scaffold engineering involve fusions with aptamer domains (dark purple) [50, 105, 108–110] or of sequences complementary with RNA sequencing capture probes (green)[111, 112]. **D.** Examples of sgRNA 3' end engineering involve fusion of sgRNAs with aptamer sequences (dark purple) [110], prime-editing reverse-transcription (light brown) and primer-binding sites (dark brown)[44] as well as fusion of sgRNAs with target sequences complementary with RNA sequencing capture probes (green) [111]. *Figure was re-used from a literature review accepted for publication in The CRISPR Journal.

Engineering of several sgRNA domains is possible

The sgRNA sequence is composed of 6 functional modules: spacer, lower stem, upper stem, nexus as well as two hairpins[14]. The spacer sequence has its origin in the crRNA, while the nexus and hairpins have their roots in the tracrRNA. The lower and upper stems are formed by base-pairing interactions between the crRNA and tracrRNA molecules fused using a minimal "GAAA" tetraloop (Figure 1.1.B.)[14]. The position of these modules with regard to Cas9 was determined in structural studies[113–115]. A crystal structure of the Cas9-sgRNA-DNA complex revealed that the tetraloop, the loop of the second hairpin, as well as the 5' and 3' sgRNA ends, point outside of the Cas9 structure[115].

Both structural[113–115] and functional experiments[3, 14, 104] suggested that engineering the 5' sgRNA end, tetraloop, loop 2, and the 3' terminus would be possible without massively disrupting sgRNA function. Whilst Nowak and colleagues[116] provided a comprehensive summary of the sgRNA engineering efforts before 2016, here I am also introducing more novel approaches.

Spacer modifications were attempted by truncating (Figure 1.2.A.) or extending (Figure 1.2.B.) the original 20 nucleotide (nt) sequence. Removing nucleotides from the 5' end or replacing ribonucleotides with deoxyribonucleotides at this end reduces Cas9 off-target effects[104, 117]. 17-19nt spacers were not associated with a significant reduction in cutting activity compared to original 20nt spacers[104]. Reducing the spacer length below 14-15nt inhibited Cas9 nuclease activity but did not prevent binding to its DNA target. Truncated spacers efficiently triggered transcriptional activation[105]. The average dwell time (duration of binding to DNA targets) of dCas9 is affected by the length of the spacer sequence. In mammalian cells, 11nt spacers resulted in an average dCas9 dwell time of 206 min, whereas 8nt spacers resulted in only 25 min of average binding time[118]. Mismatches within the last 5nt of the spacer[3] or substitution of RNA nucleotides with DNA nucleotides[117] can also be tolerated. Most of the nucleotides in the entire sgRNA

sequence could be replaced with DNA nucleotides[34, 119].

In a study attempting to improve Cas9 ON-target specificity, 30nt spacers tested were trimmed to 20nt by an uncharacterised cellular mechanism[41]. A similar phenomenon was observed for self-targeting sgRNAs with spacer sequences of up to 70nt[107]. Processing of 20nt spacer sequences was also observed when co-expressing a crRNA array with a tracrRNA in mammalian cells. Although without definitive evidence, crRNAs are believed to be processed by type III RNases in mammalian cells[5], reminiscing the crRNA maturation processes observed in *S. pyogenes*[13]. Extended spacer sequences that form hairpins appear more stable, as these secondary structures may protect sgRNAs from being recognised by endogenous nucleases[37, 40, 120, 121]. This includes hairpins that only partially cover the spacer sequences explored to reduce Cas9 off-target binding[120, 121].

Scaffold modifications (Figure 1.2.C.) were employed for recruiting engineered proteins to exposed sgRNA domains. MS2, PP7 and boxB are well-characterised aptamers bound by the MCP, PCP and N22 proteins, respectively. Such aptamers were embedded into the sgRNA tetraloop and loop 2[50, 105, 108–110]. By fusing aptamer-binding proteins with transcription factors[50, 110] or fluorescent proteins[108, 109, 122], modulation of gene expression and imaging of repetitive DNA loci was achieved. For instance, sgRNA scaffolds were modified to improve sgRNA folding or expression yields from polymerase III promoters[123]. sgRNA loops and the 3' scaffold ends were also engineered to include A/G-rich structures that could be captured by oligo dT RNA-sequencing capture probes[111] as well as other target-specific probes[112]. Scaffold modifications also enabled designing of sgRNAs capable of targeting their own DNA expression cassette. This was achieved by replacing the first 3 "GUU" scaffold nucleotides with a "GGG" PAM (palindromic associated motif) site. These designs enabled continuous barcode evolution for deep lineage tracing applications[107, 124].

3'end sgRNA modifications (Figure 1.2.D.) could also recruit engineered MS2, PP7 and Com proteins to Cas9 binding sites[110]. Additional modifications include appending Cas12a guide RNAs to the 3'end of the Cas9 sgRNA to improve co-expression of the two gRNAs at a single-cell level. These chimeric gRNAs are recognised by Cas12a that processes its own guide RNA, while releasing a functional Cas9 sgRNA[125]. CRISPR-Display technology suggests that fusing the 3' sgRNA end with lncRNA sequences is also possible[126]. Furthermore, the prime editing technology also relies on modifying the 3'end of the sgRNA. In the prime-editing gRNAs (pegRNA), a template for reverse transcription and the binding site for reverse transcriptase are both fused to the 3' end of the scaffold[44]. Similar to extended spacer sequences, unprotected 3' end sgRNA extensions also appear to be degraded. In a subsequent optimisation of pegRNA designs, structured RNA motifs were appended to the 3' end of the engineered pegRNAs (epegRNAs) for promoting their stability[45].

Engineering strategies for designing conditional sgRNAs

In contrast to non-disruptive sgRNA engineering, conditional sgRNA designs employ engineering strategies that interrupt sgRNA activity. A study was carried out to identify secondary structure domains that are essential for sgRNA activity (Figure 1.3.A.). Blocking the functional sgRNA domains using single-stranded DNA probes demonstrated that interfering with spacer and lower stem, nexus as well as hairpin 1 considerably inhibits sgRNA activity while blocking hairpin 2 does not cause a substantial inhibition[127].

Different methods were developed to precisely control sgRNA activity [35–40, 127]. Chemical sgRNA modifications enabled conditional sgRNA activation following exposure to light [35] (Figure 1.3.B.). Alternative sgRNA activation triggers included small molecules such as theophylline[36, 38–40] (Figure 1.3.C.), tetracycline or guanine[40]. Small molecule sensing was achieved by embedding specific aptamers within the sgRNA structure. Site-specific endoribonucleases such as Csy4 were also shown

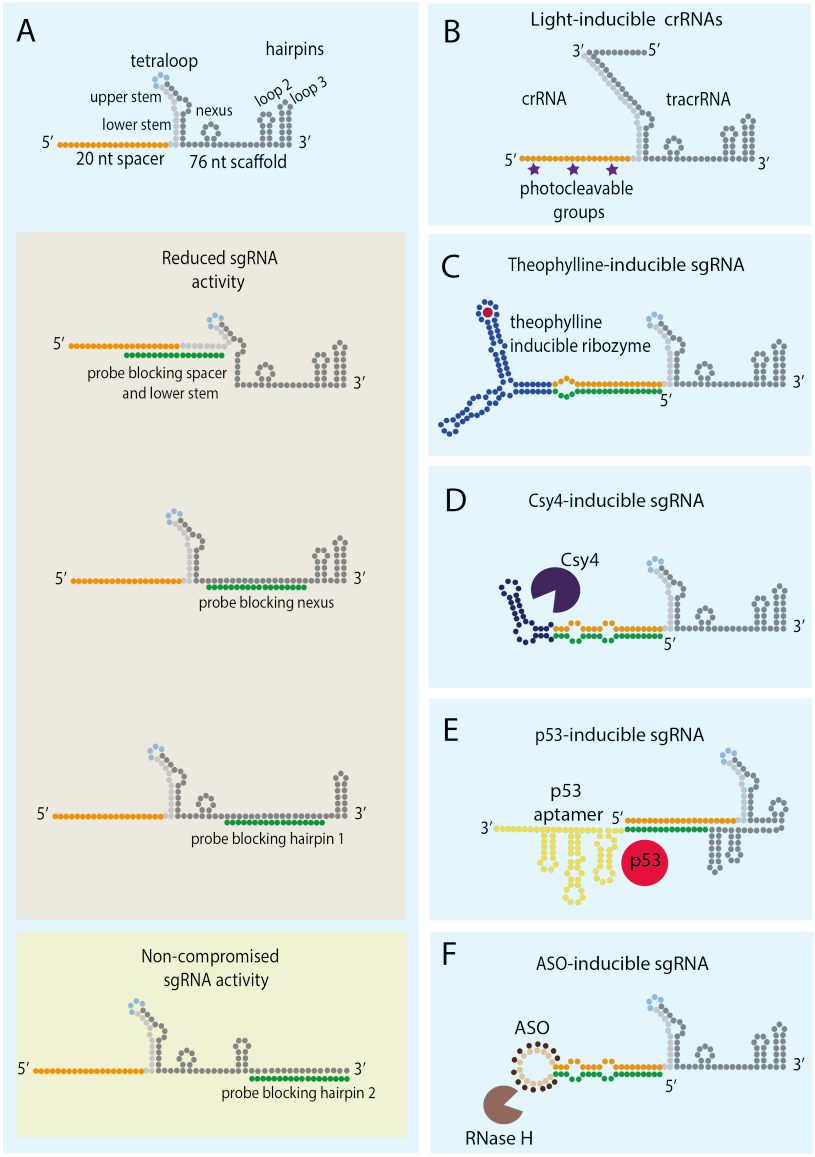


Figure 1.3: Conditional sgRNA designs

sgRNA spacer sequences are represented in orange, the "GAAA" tetraloop in cyan and the sgRNA scaffold sequences in grey. **A.** sgRNA function can be abolished by short nucleic acid probes complementary with important functional domains (green)[127]. **B.** Light-inducible crRNAs incorporate photocleavable groups (purple). Exposure to light results in the removal of caging groups and gRNA activation[35]. **C.** Theophylline-inducible sgRNAs form a complex secondary structure. sgRNA spacer is fused with a theophylline-inducible ribozyme (blue) and a spacer-blocking sequence (green). Theophylline (red) binding triggers ribozyme-mediated removal of the spacer-blocking sequences[40]. **D.** Csy4-inducible sgRNAs contain a Csy4-recognition sequence (dark blue) as well as a spacer-blocking sequence (green). Csy4 nuclease (purple) cleaves its recognition sequence, resulting in sgRNA activation[37]. **E.** p53-inducible sgRNA become activated following p53 (red) binding to its aptamer (yellow). This configuration change displaces the spacer-blocking sequence (green)[39]. **F.** Antisense oligos (ASO)-inducible sgRNA designs contain a loop sequence (light orange) and a spacer-blocking sequence (green). ASOs (dark brown) base-pair with the sgRNA loop, while and RNase H (light brown) cleaves RNA-DNA hybrids[37]. *Figure was re-used from a literature review accepted for publication in The CRISPR Journal.

to activate modified sgRNAs in mammalian cells [37] (Figure 1.3.D.) . However, all these approaches for conditional sgRNA engineering require additional exogenous inputs. CRISPR-based sensors capable of harvesting endogenous information to control Cas9 activity would massively benefit the development of therapeutic and research applications.

Early studies demonstrated that sgRNA activity can be controlled by endogenous proteins upregulated in cancer cells. Examples included proteins such as p53, NF-$\kappa\beta$, AFP, β-catenin and HSF1, for which well characterised high-affinity aptamers are available. These sensors were generated by extending the 3' sgRNA ends with sequences complementary to the spacer. Specific aptamers were also included within 3' extensions (Figure 1.3.E.). In cells without high expression of target proteins, 3' end extensions blocked the spacer, resulting in a non-functional sgRNA. However, in cells overexpressing target proteins, aptamer-protein interactions created a conformational change that exposed the spacer sequence and activated the sgRNA[39].

As high-affinity aptamers are not available for all endogenous factors, a different approach could involve controlling sgRNA activity by endogenous RNAs. Several studies suggested that single-stranded DNA oligos (ssDNA) could activate sgRNAs whose activity was inhibited by secondary structures[37, 101, 127]. Activation of such engineered sgRNAs called iSBH-sgRNAs (inducible spacer-blocking hairpin sgRNAs) was thought to depend on RNaseH hydrolysis of RNA molecules when hybridised to DNA [37](Figure 1.3.F.). Nevertheless, apart from enabling sgRNA activation in the presence of DNA antisense oligos (ASOs), detection of ssDNA would not have other relevant applications *in vivo*. Apart from ssDNA, single-stranded RNA (ssRNA) could also activate engineered sgRNAs. Interestingly, ssRNAs were reported to lead to better activation outcomes than ssDNA sequences *in vitro*. This was believed to be because double-stranded RNA (dsRNA) duplexes are more thermally stable than DNA-RNA hybrids[101].

RNA-responsive sgRNA designs

RNA strand displacement is a molecular mechanism in which a ssRNA invader attacks a preformed RNA duplex. A conformational change occurs as the ssRNA hybridises to one of the RNA duplex strands, leading to the displacement of the second duplex strand[128]. Starting from this principle, several OFF-ON (Figure 1.4) or ON-OFF (Figure 1.5) sgRNA switches were engineered for sensing RNA triggers. In most cases, this was achieved by engineering sgRNAs to fold into complex secondary structures. Both synthetic or endogenous ssRNA triggers can act as invaders and disrupt secondary structures within engineered sgRNAs. Several proof-of-concept experiments showed that such systems could be applicable for sensing RNA *in vitro*[101–103, 129]. Sensing synthetic triggers [87, 98–100, 129], RNA interference (RNAi) effectors [92, 94, 130, 131] and mRNAs [87, 89, 92, 94, 99] were demonstrated in live cells including bacteria and mammalian cells.

OFF-ON sgRNA switches are generated by engineering diverse components of the sgRNAs, to generate secondary structures incompatible with sgRNA function. This can be achieved, for instance, by blocking the spacer sequences [87, 89, 91, 94, 98–102](Figures 1.4.A-E,H.), or preventing the formation of essential scaffold domains [101, 103](Figures 1.4.F,G,I). Cognate RNA triggers delivered exogenously or sensed endogenously initiate strand-displacement reactions that result in the formation of active sgRNAs.

ON-OFF sgRNA switches are functional in the absence of a cognate RNA trigger. The introduction of cognate triggers results in the formation of secondary structures that are incompatible with the sgRNA function. Such designs are based on the formation of double-stranded RNA sequences upstream of the spacer sequence[102] (Figure 1.5.A.) or insertion of trigger-complementary sequences within the tetraloop and loop 2 [101](Figures 1.5.B,C).

In vitro sensing of RNA triggers was demonstrated in several publications[93,

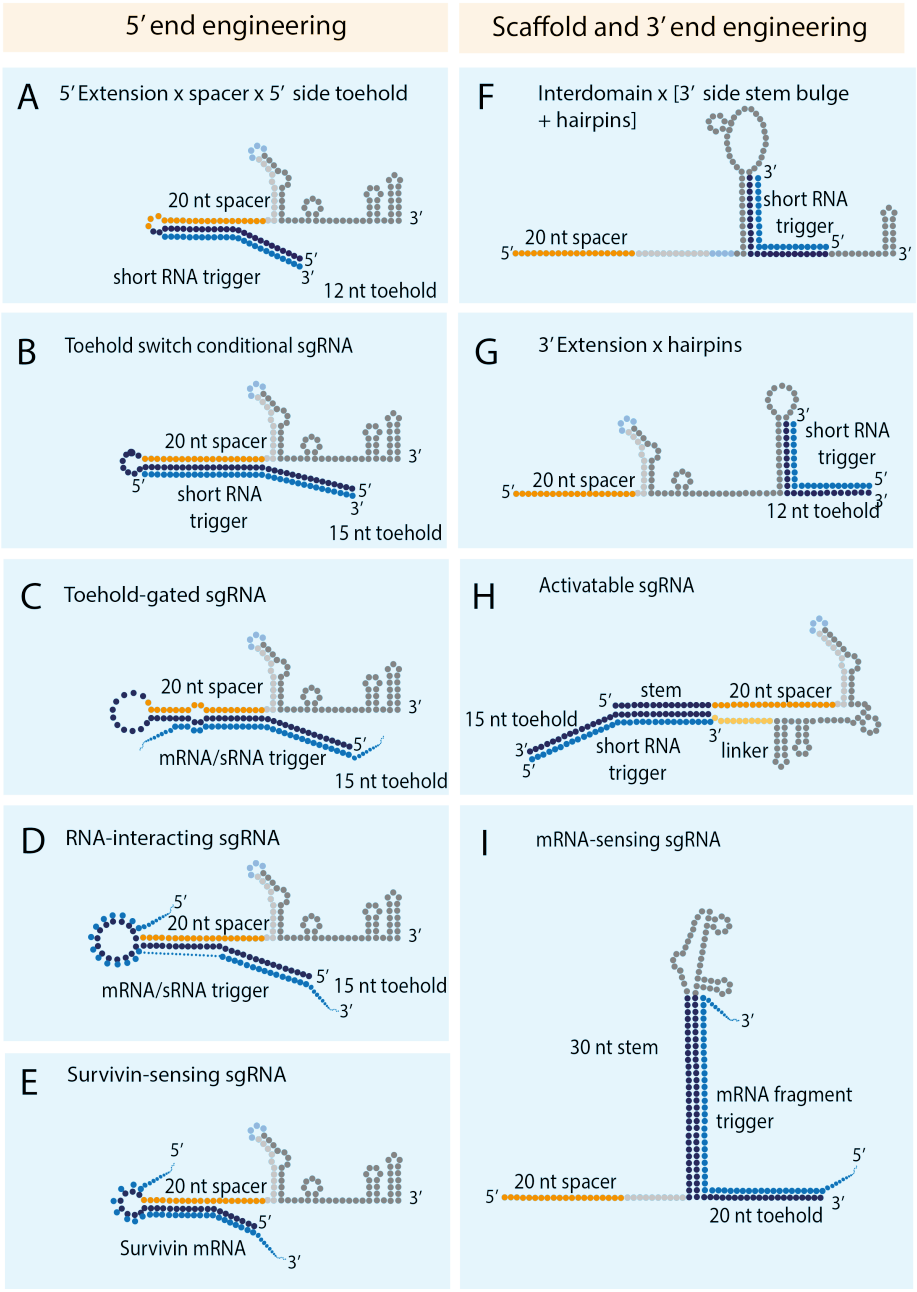


Figure 1.4: OFF-ON sgRNA switches

sgRNA switches are designed to fold into complex secondary structures that, in the ground state, inhibit sgRNA activity. Upon recognising complementary RNAs, engineered sgRNAs become activated. Figures A-I represent published engineering strategies for designing OFF-ON sgRNA switches. sgRNA modifications are represented in dark blue, while complementary RNA triggers are represented in light blue. Spacer sequences are represented in orange, the "GAAA" tetraloop in cyan and the sgRNA scaffold sequences in grey. **A.** 5' Extension x spacer x 5' side toehold design [101]. **B.** Toehold switch conditional sgRNA[132].**C.** Toehold-gated sgRNA[94].**D.** RNA-interacting sgRNA[87]. **E.** Survivin-sensing sgRNA[89]. **F.** Interdomain x [3' side stem bulge + hairpins] design[101]. **G.** 3' Extension x hairpins design[101]. **H.** Activatable sgRNA[102]. **I.** mRNA-sensing sgRNA[103]. *Figure was re-used from a literature review accepted for publication in The CRISPR Journal.

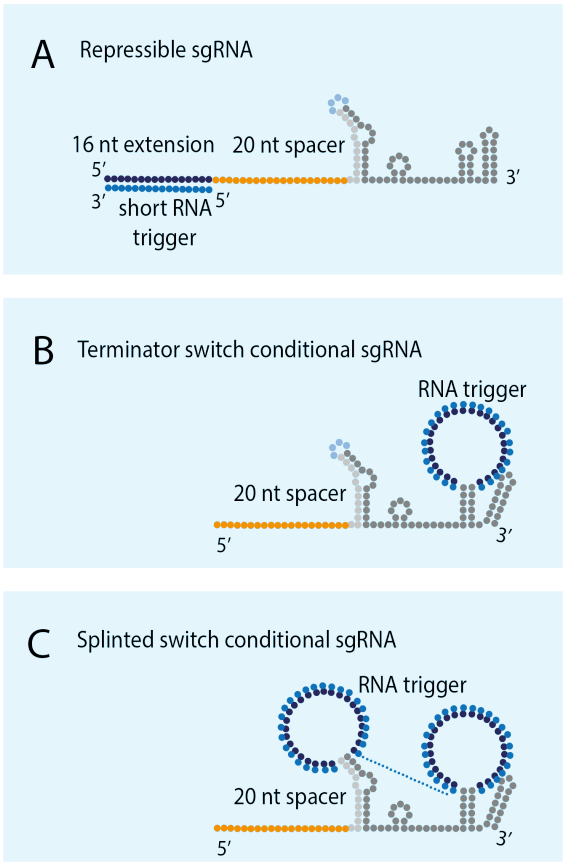


Figure 1.5: ON-OFF sgRNA switches

sgRNA switches are designed to fold into a secondary structure that is compatible with sgRNA activity and CRISPR function. Upon recognising complementary RNA triggers, engineered sgRNAs become inactive, repressing CRISPR function. Figures A-C represent published engineering strategies for designing ON-OFF sgRNA switches. sgRNA modifications are represented in dark blue, while complementary RNA triggers are represented in light blue. Spacer sequences are represented in orange, the "GAAA" tetraloop in cyan and the sgRNA scaffold sequences in grey. **A.** Repressible sgRNA[102]. **B.** Terminator switch conditional sgRNA[98]. **C.** Splinted switch conditional sgRNA[98]. *Figure was re-used from a literature review accepted for publication in The CRISPR Journal.

101–103, 129, 133]. Jakimo and colleagues proposed a series of designs (Figures [101]1.4.A,F,G) that sensed short synthetic RNA triggers. Double-stranded RNA structures formed immediately upstream of the spacer sequence inhibit Cas9 activity. From this observation, both activatable sgRNA (Figures 1.4.H.) and repressible sgRNA (Figure 1.5.A.) designs were generated. Activatable sgRNAs have complementarity between 5' and 3' sgRNA ends, while repressible sgRNAs are sgRNAs with extended spacers. When the supernumerary nucleotides in repressible sgRNAs hybridise with RNA triggers, Cas9 activity is repressed[102]. Even though the repressible sgRNA design works *in vitro*, it will be essential to test it in living cells, given that previous studies have suggested that extended spacer sequences get truncated by RNases[41, 107].

mRNA sensing sgRNA (Figure 1.4.I.) designs involve forming a secondary structure between the sgRNA tetraloop and the 3' end of sgRNA, including an unpaired single-stranded toehold that promotes hybridisation with incoming triggers and facilitates strand-displacement. The mRNA sensing sgRNA secondary structure is incompatible with the sgRNA function in an OFF state, and strand displacement activates the sgRNA. This study hypothesised that the longer the mRNA, the more secondary structures it forms, thus preventing sgRNA strand displacement[103]. Similar findings were reported for engineered Cas12a gRNAs[129].

tracrRNA modifications enable the designing of other types of OFF-ON switches. Recent studies have shown that SpCas9[93, 133] and *Campylobacter jejuni* Cas9 (CjeCas9) [133] tracrRNAs could be extensively modified without compromising Cas9 activity. For SpCas9, tracrRNA positions affecting interactions between tracrRNA and crRNA were identified and sequences complementary to mRNA triggers were introduced within tracrRNA at these sites. When modified tracrRNAs hybridised with mRNA triggers, crRNAs were processed from the trigger sequences, resulting in the formation of functional crRNA-tracrRNA complexes[93] (Figure

1.6.A.). A similar approach was also used for CjeCas9, and both systems were successfully employed for *in vitro* detection of SARS-CoV-2 RNA fragments[93, 133].

Sensing RNA triggers in cellular systems

Conditional sgRNAs with OFF-ON and ON-OFF logics were tested in *E. coli* and mammalian HEK293T cells[98, 103]. Terminator (Figure 1.5.B.) and splinted (Figure 1.5.C.) designs contain RNA-sensing loops at different positions within the sgRNA, and these loops do not impact normal Cas9 function (ON-state). Binding of a trigger RNA leads to the formation of a distorted sgRNA structure incompatible with sgRNA function, leading to an OFF-state. The terminator design has a single loop that hybridises with RNA triggers, while the splinted design has two complementary loops. These two designs sensed short RNA triggers expressed at high levels in *E. coli*. The terminator design was also able to sense short, nuclear RNA triggers expressed from U6 promoters in HEK293T cells[98].

OFF-ON switches relying on blocked spacer sequences (Figure 1.4.B.) displayed an undesired background activity in the OFF-state. Therefore, a new OFF-ON switch was proposed by splitting the sgRNA structure. In this design, hairpin 1 was truncated, and the hybridisation of the RNA trigger enabled the formation of a functional gRNA (Figure 1.6.B.). Further optimisations of conditional gRNA switches were enabled by flanking the synthetic RNA triggers with protective hairpin structures[134]. Although helpful with sensing synthetic RNA triggers, such protective structures cannot be added to endogenous RNA species to facilitate their detection.

RNA species used as endogenous sensed triggers include RNA interference (RNAi) effectors (bacterial small RNA and mammalian microRNAs) and messenger RNA (mRNA). In *E. coli*, toehold-gated sgRNAs (Figure 1.4.C.) detected plasmid-encoded or endogenous small RNAs such as micF, ryhB and oxyS[94]. Lin and colleagues proposed another mechanism for detecting small RNAs or mammalian microRNAs

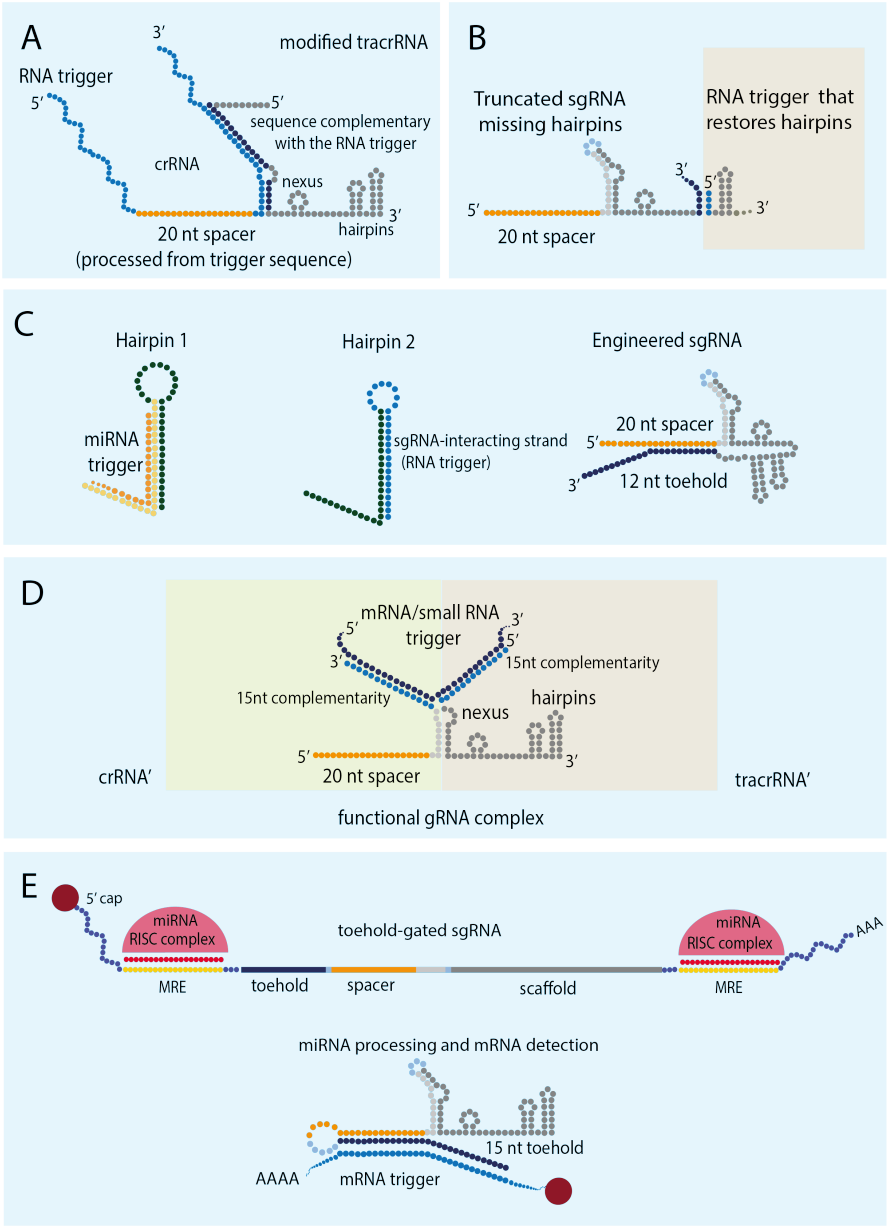


Figure 1.6: Other OFF-ON gRNA switches

A. tracrRNAs were modified to include sequences complementary with mRNA triggers. Upon hybridisation with modified tracrRNAs, mRNA triggers are processed into crRNAs. This results in the formation of functional crRNA-tracrRNA complexes[93]. **B.** sgRNA was inactivated by removing terminal hairpins. Recognition of RNA triggers restored hairpin 1, leading to sgRNA activation[132]. **C.** Sensing microRNAs using two RNA hairpins and modified sgRNAs. Target microRNA (orange) binds to a complementary sequence within hairpin 1 (yellow). Displaced hairpin 1 strand (green) hybridises with a sequence within hairpin 2 (green), leading to displacement of the RNA trigger (light blue). RNA trigger hybridises with the spacer-blocking region of the modified sgRNA (dark blue), resulting in sgRNA activation[91]. **D.** Endogenous mRNAs or small RNAs bridge connection between engineered crRNA' and tracrRNA' forming a functional gRNA complex[92]. **E.** Toehold-gated sgRNAs are expressed from polymerase II promoters. In the cytoplasm, microRNAs process toehold-gated sgRNAs. Complementary mRNA sequences promote sgRNA activation[99]. *Figure was re-used from a literature review accepted for publication in The CRISPR Journal.

that uses RNA circuits containing two hairpins and a modified sgRNA. sgRNA is engineered to adopt an OFF state by appending a 3' extension that hybridises with the spacer and a toehold that facilitates strand displacement. Sensing the microRNA target opens the first hairpin, leading to the opening of the second hairpin, which triggers a strand displacement reaction that activates the sgRNA (Figure 1.6.C.). Such a tri-component complex circuit design avoids sequence dependencies between the sgRNA and RNA trigger. This system enabled micF and ryhB small RNAs sensing in *E. coli* and mir-17, mir-16 and let-7 miRNAs in HEK293T cells[91].

mCherry mRNA was also sensed in *E. coli* using toehold-gated sgRNAs (Figure 1.4.C.). Three different toehold-gated sgRNAs were designed for sensing specific subsequences of the mCherry transcript. sgRNA activation efficiencies were measured using CRISPR-interference assays, and the best design resulted in up to 60% repression of a fluorescent reporter gene[94]. RNA-interacting sgRNA designs (Figure 1.4.D.) can also be used to sense mRNA. When sensing short transcripts in *E. coli*, this system yielded up to 26-fold induction/upregulation of fluorescence reporter. However, when full mRNA sequences were used as triggers, activation dropped to only 2.5 fold for mKate fluorophore and 10 fold for the HIV infectivity factor[87]. RFP mRNA and genes from the arsenic operon were also sensed in bacteria using modified tracrRNAs (Figure 1.6.A.). Engineered tracrRNAs hybridised to the mRNAs to be sensed, leading to production of crRNAs from bacterial mRNA species[93].

An alternative design proposed in *E.coli* involved truncated crRNAs (crRNA') and tracrRNAs (tracrRNA') that lost their ability to recognise each other through base-pairing interactions. 3' end of the crRNA' and 5' end of the tracrRNA' were fused to sequences complementary with target RNAs to be detected. Target RNAs bridged interactions between crRNA' and tracrRNA', resulting in the formation of a functional gRNA complex (Figure 1.6.D.). This approach was successfully used for sensing bacterial endogenous mRNA and small RNA species[92].

In mammalian cells, RNA detection is dependent on the RNA polymerase transcribing RNA triggers. Triggers expressed from polymerase II promoters were detected 3-fold less efficiently then triggers expressed using polymerase III promoters[99, 100]. Nevertheless, several studies reported detection of polymerase II-expressed mRNA[89, 99].

Expression of survivin mRNA was monitored in different cell lines, including MCF-7, HeLa and MB-435s. In this study, the sgRNA engineering and strand-displacement mediated mRNA sensing was coupled with other engineering strategies for signal amplification. The sgRNA fold adopted in this study involved blocking the spacer sequence by a 12nt dsRNA-blocking sequence and including a toehold gate to promote hybridisation with incoming triggers (Figure 1.4.E.). mRNA sensing was monitored by dCas9-VPR transcriptional activators and a reporter system including eight CRISPR-binding sites upstream of a fluorescence reporter comprising five fluorescent aptamer repeats. The survivin mRNA is highly expressed and is relatively short (429bp). Modulators of survivin expression were used to tune mRNA copy number. Calibration curves observed suggested that 120 mRNA copies/cell was the minimum detection limit for this system[89]. However, data presented in this study relied on a small number of microscopy images. And it is, therefore, not clear in which percentage of cells this technology might work and what is the background activity of the reporter system.

As mRNA is exported in the cytoplasm, another study reasoned that translocating engineered sgRNAs to the cytoplasm could improve mRNA detection. This was achieved by expressing toehold-gated sgRNAs under the control of polymerase II promoters and by using MREs to release these sgRNAs into the cytoplasm (Figure 1.6.E.). Activation of such designs requires presence of microRNA targets as well as presence of the mCherry mRNA. For miRNA-processed toehold-gated sgRNAs, presence of mCherry caused a roughly 1.7-fold change in CRISPRa activity as

measured using fluorescence reporters[99]. Nevertheless, certain applications might require higher dynamic ranges for mRNA detection and extra work should be done for improving activation in the presence of endogenous transcripts in eukaryotic cells.

To sum up, proof-of-concept studies suggest that engineered sgRNAs can detect miRNA[91, 99] or mRNA[89, 99] in eukaryotic cells. However, further work is required for optimising dynamic ranges of RNA detection. Furthermore, evidence that engineered sgRNAs are functional *in vivo* is currently lacking and there is no clear evidence that the engineering approaches adopted so far could be generalisable for the detection of other endogenous RNA sequences. Computational tools enabling users to design engineered sgRNAs for their desired RNA detection applications are also lacking as well as studies attempting to characterise molecular mechanisms underlying RNA detection by engineered sgRNAs.

*I would like to acknowledge that Section 1.5 was adapted from a literature review accepted for publication in The CRISPR Journal: ***Pelea*** *O, Fulga T, Sauka-Spengler T, RNA-responsive gRNAs for controlling CRISPR activity: current advances, future directions, and potential applications.* Figures1.1-1.6 were also reused from this publication.

1.6 Purpose of this study

The long-term aim of this project is providing proof-of-concept evidence that CRISPR activity can be restricted to tissues of interest *in vivo* (Figure 1.7). Due to their transparency and relative ease of manipulation, zebrafish embryos are a great model. In zebrafish, a standard way of restricting CRISPR activity to tissues of interest involves the use of tissue-specific promoters or enhancers that drive Cas9 expression[135] (Figure 1.7.A.). In target tissues, CRISPR activity is switched ON due to the presence of tissue-specific transcription factors that interact with their cognate promoters or enhancers.

The *in vivo* part of this project aims to test an alternative way of restricting CRISPR activity to tissues of interest in zebrafish embryos, through detection of endogenous RNA biomarkers (Figure 1.7.B.). Engineered RNA sensors would be delivered to embryonic cells in an inactive configuration. When reaching target organs, detection of endogenous RNA biomarkers would switch ON CRISPR activity. Due to the availability of tissue-specific whole RNA transcriptome databases (coding and non-coding)[136], this project is aiming to restrict CRISPR activity to the zebrafish neural crest cells.

As part of a bigger goal, in my thesis work, I aim to develop intermediary technologies required for restricting CRISPR activity to any target tissue based on detecting specific RNA species. Current examples of CRISPR-based RNA sensors involve engineered expression cassettes encoding CRISPR components or engineered sgRNAs. Due to their simplicity and compatibility with all CRISPR delivery methods[137], this work focused on engineered sgRNAs.

Engineered iSBH-sgRNAs (inducible spacer-blocking hairpin sgRNAs) proposed by Ferry et. al[37] were chosen as a starting point for this project. As these designs were originally developed for detecting Csy4 nucleases (Figure 1.3.D.) and ASOs (Figure 1.3.F.), the first step involved repurposing these designs towards RNA detection.

In Chapter 2, I test whether iSBH-sgRNA could detect RNA, while Chapter 3 brings further improvements to the original iSBH-sgRNA designs. The goal of Chapter 4 is to build a computational pipeline to enable future users of this technology to design iSBH-sgRNAs for their custom RNA detection applications. As most technology development projects massively benefit from mechanistic insights, Chapter 5 explores the mechanism of iSBH-sgRNA activation in the presence of RNA triggers. After developing and characterising this RNA-sensing technology

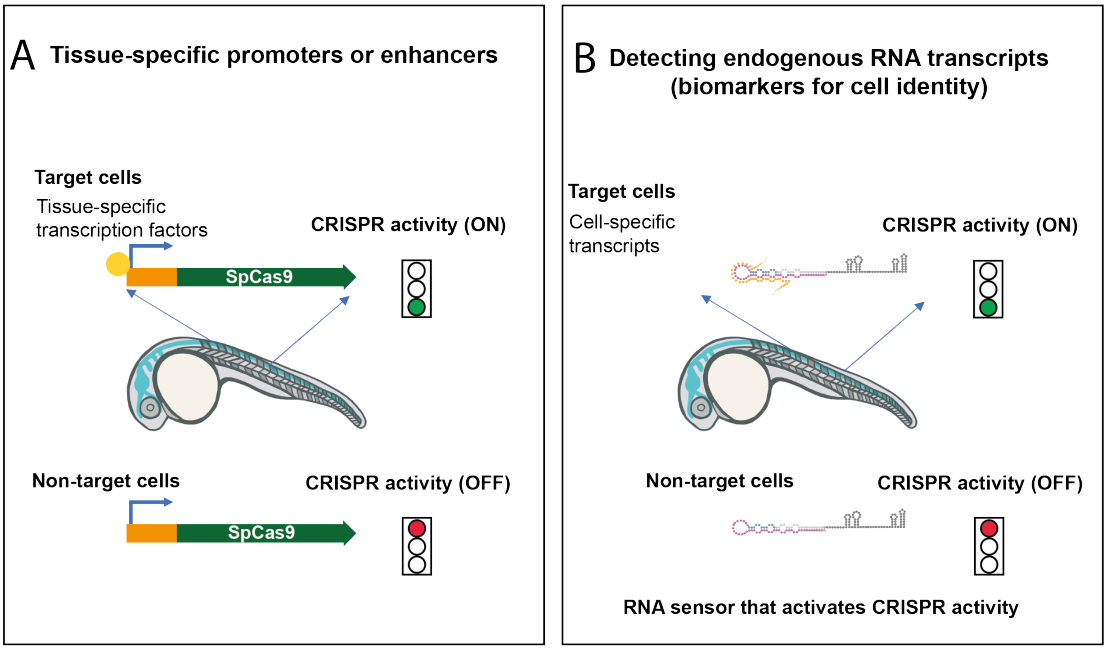


Figure 1.7: Restricting CRISPR activity to target cells through RNA detection *in vivo*

A. Tissue-specific promoters or enhancers are a standard way of achieving tissue-specific CRISPR activity in zebrafish embryos[135]. Presence of tissue-specific transcription factors bind to these cis-regulatory elements and drive Cas9 expression. **B.** Detecting endogenous RNA transcripts might be an alternative way of restricting CRISPR activity to cells of interest. CRISPR-based RNA sensors could be delivered to all embryonic cells in an inactive configuration. When reaching target cells, endogenous RNA biomarkers would trigger activation of CRISPR sensors. However, in order to accomplish this, additional technology development efforts are needed. My thesis aims, therefore, to further develop CRISPR-based RNA sensors based on sgRNA engineering and to adapt such technologies for *in vivo* applications.

in HEK293T cells (Chapters 2-5), in Chapter 6, I perform the tests of the iSBH-sgRNA technology in zebrafish embryos. Finally, Chapter 7 aims to provide insights into future directions required for efficiently implementing the iSBH-sgRNA technology *in vivo* and Chapter 8 describes the methodology and protocols I developed, employed and used in my thesis.

2

Testing if iSBH-sgRNAs could be repurposed for RNA detection

2.1 Introduction to first-generation iSBH-sgRNA designs

Native sgRNAs have two functional components: the spacer and scaffold sequences. The spacer sequence, complementary to the CRISPR-targeting sequence (CTS), determines sgRNA specificity, while the scaffold sequence stabilises the interactions between the sgRNA and Cas9 [1, 3] (Figure 2.1). Ferry et al., 2017[37] proposed an engineering approach to design inducible sgRNA, where a spacer blocking hairpin was appended at the 5' end of the spacer. Therefore, apart from the spacer and scaffold structures, the design contained a 14nt loop and a spacer* sequence partially complementary with the spacer sequence. Of a particular interest were the ASO-inducible designs developed by Ferry et al. (Figure 1.3.F., Figure 2.1). I will refer to these designs as first-generation iSBH-sgRNAs (inducible spacer-blocking hairpin sgRNAs)[37].

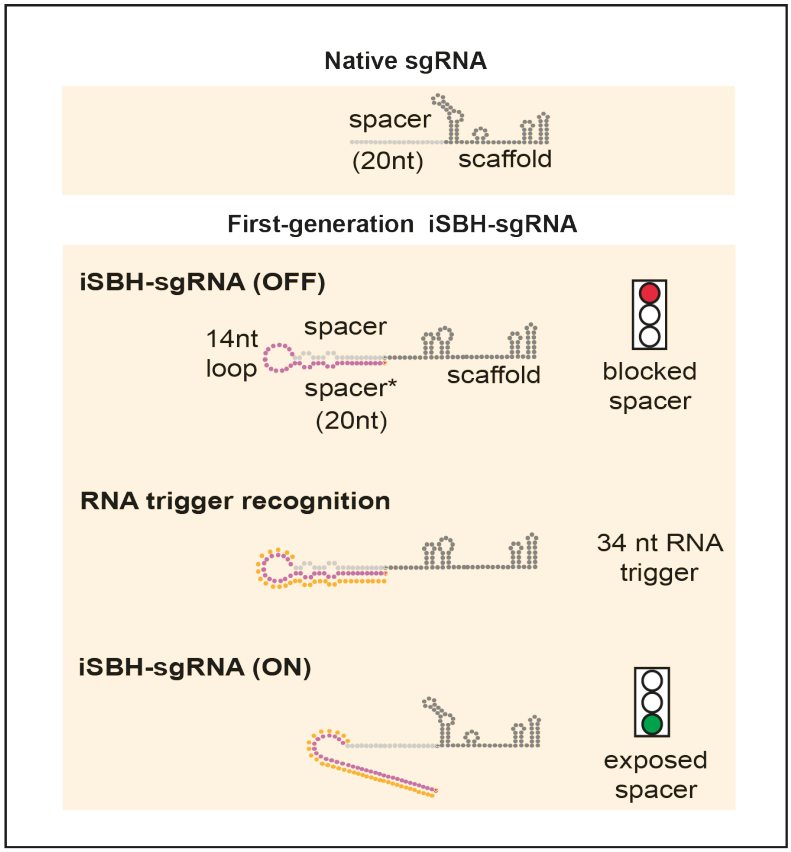


Figure 2.1: Native sgRNA vs RNA-sensing first-generation iSBH-sgRNA

Native sgRNA sequences are composed of spacer and scaffold sequences[3]. iSBH-sgRNAs fold into complex secondary structures that interfere with the Cas9 ability to recognise target DNA sequences (OFF state) [37]. ISBH-sgRNAs were designed by extending the 5' end of the spacer sequence with a 14nt loop and a spacer* sequence partially complementary with the spacer. Bulges were also introduced within the iSBH-sgRNA sequence in order to ensure that the interaction between the spacer* and RNA trigger is more energetically favourable. In the ON state, iSBH-sgRNAs recognise complementary RNA triggers and become activated, enabling Cas9 to perform its function. Short RNA triggers are complementary with the iSBH-sgRNA loop and spacer* sequence.

2.2 Introduction to the experimental pipeline for testing iSBH-sgRNAs in HEK293T cells

The aim of this chapter is to test if first-generation iSBH-sgRNAs[37] could be activated by short RNA triggers (Figure 2.1). To address this question, five novel iSBH-sgRNAs were designed, each featuring a distinct sgRNA spacer sequence. For each iSBH-sgRNA, RNA triggers complementary with the loop and the spacer*

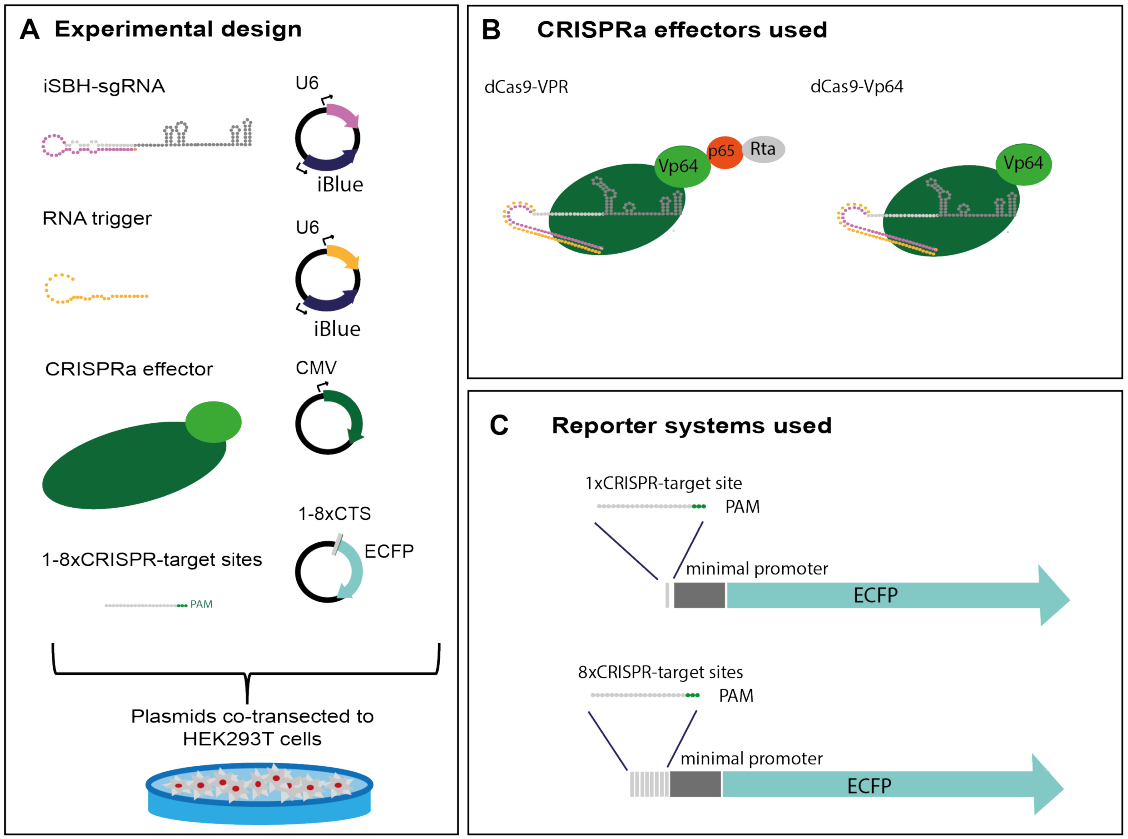


Figure 2.2: Experimental pipeline for testing iSBH-sgRNAs in HEK293T cells **A.** Experimental design. iSBH-sgRNAs and RNA triggers were cloned in mammalian U6 expression vectors. 1 or 8 CRISPR-target sequences (CTSs) were also cloned upstream of an ECFP fluorescent reporter. I co-transfected the 3 cloned plasmids as well as CRISPRa effectors into HEK293T cells. Fluorescence outputs were measured by Flow Cytometry. **B.** CRISPRa effector used. Effectors used include dCas9-VPR[48] and dCas9-Vp64[49]. In contrast with dCas9-Vp64, dCas9-VPR is a stronger transcriptional activator due to the presence of the additional p65 and Rta domains. **C.** Reporter systems used. 1xCTS (CRISPR-target sequence) reporters contain a single Cas9-binding region upstream from the ECFP reporter. 8xCTS reporters[138] contain 8 repeats of the Cas9-binding region upstream from the ECFP reporter. For promoting expression of the ECFP reporters, minimal adenovirus major late promoters were also introduced within these cassettes.

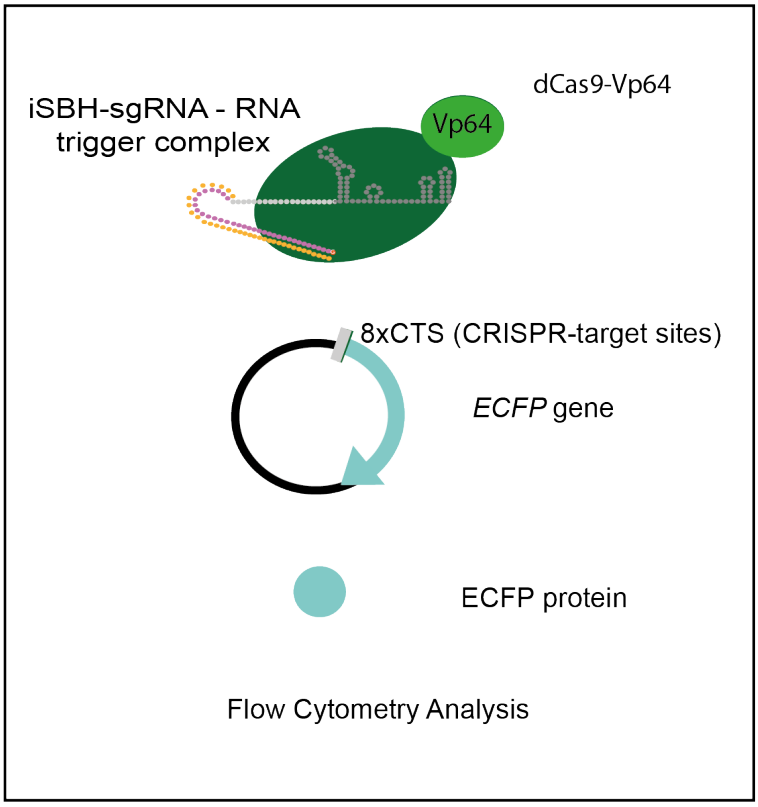


Figure 2.3: Expected iSBH-sgRNA system behaviour inside HEK293T cells

Inside cells, RNA triggers are expected to bind to complementary iSBH-sgRNAs, inducing iSBH-sgRNA activation. Activated iSBH-sgRNAs are recognised by CRISPRa effectors and drive ECFP production from a fluorescent reporter. In this particular example, activated iSBH-sgRNAs interact with dCas9-Vp64[49] and drive ECFP production from an 8xCTS-ECFP reporter[138]. Following reporter induction, ECFP production could be monitored by Flow Cytometry.

sequences were also designed (Figure 2.1). To prevent their degradation by cellular nucleases, RNA triggers were flanked by 5' and 3' hairpins. For simplicity reasons, hairpins are not displayed in the figures, but more details could be found in the Supplementary Materials (Chapter 9).

iSBH-sgRNA and corresponding RNA triggers were cloned in mammalian plasmids, under the control of U6 promoters. Plasmids were co-transfected into HEK293T cells together with a CRISPR activator (CRISPRa) and a fluorescent reporter cassette designed to monitor CRISPRa activity (Figure 2.2.A). CRISPR activators used included dCas9-VPR[48] and dCas9-Vp64[49] (Figure 2.2.B). CRISPRa re-

porters containing either a single CRISPR-target sequence (1xCTS) or multiple (8x CTS)[138] sequences were employed for driving expression of the ECFP (Enhanced Cyan Fluorescent Protein) reporter (Figure 2.2.C).

Inside the cells, RNA triggers are expected to bind to complementary iSBH-sgRNA sequences and thus expose spacer sequences. As RNA triggers are complementary to the loop as well as the backfold, the interaction between iSBH-sgRNA trigger is expected to be more energetically favourable than the interaction between the spacer and spacer* components. Cas9 transcriptional activators bind to activated iSBH-sgRNAs, resulting in production of ECFP from the CRISPRa reporter (Figure 2.3).

2.3 First-generation iSBH-sgRNA designs detect short RNA triggers

The first test was performed using dCas9-VPR as well as 1xCTS-ECFP reporters (Figure 2.4.A.). Data showed that iSBH-sgRNA could be activated by RNA triggers, and observed ON state activation was comparable to activation seen in cells transfected with native sgRNAs (spacer only, no iSBH fold). In comparison, for all five iSBH-sgRNA sequences, the presence of a closed hairpin (OFF state, absence of RNA trigger) significantly reduced sgRNA's ability to drive reporter expression. Nevertheless, iSBH-sgRNAs still displayed background activation levels in the OFF state.

2.4 Improving iSBH-sgRNA OFF states by changing CRISPRa reporter systems

Next, I tested the system using a weaker CRISPRa effector (dCas9-Vp64 [49]) in complex with the 8xCTS-ECFP [138] reporter cassette (Figure 2.4.B.). Changing these components not only reduced the OFF-state activation but also increased the intensity of the ECFP signals detected by Flow Cytometry in the ON-state (Figure 2.5). This suggested that some iSBH-sgRNAs might not form perfect OFF

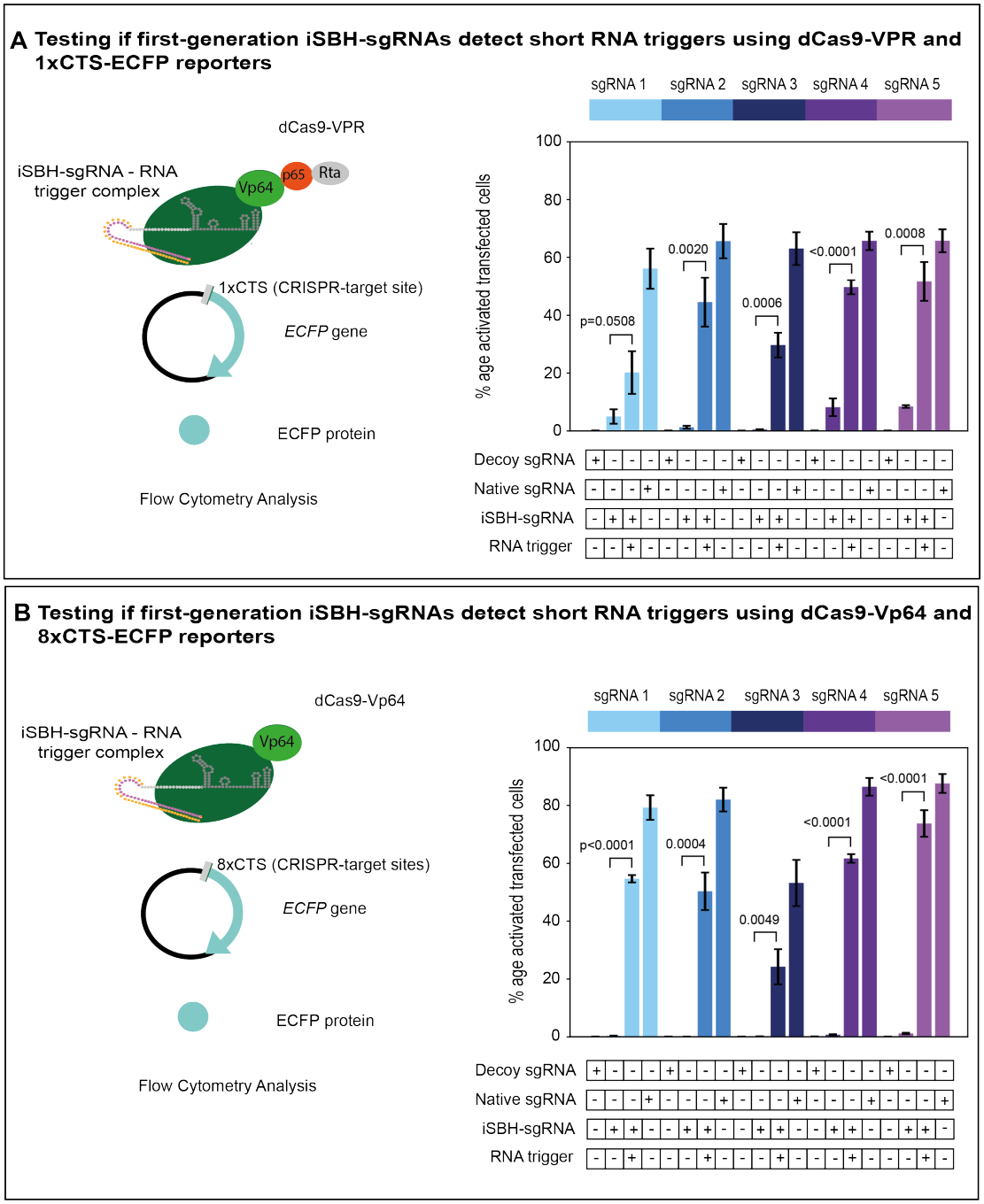


Figure 2.4: Testing if first-generation iSBH-sgRNAs detect short RNA triggers in HEK293T cells

A. Testing if first-generation iSBH-sgRNAs detect short RNA triggers using dCas9-VPR and 1xCTS-ECFP reporters. Starting from five different sgRNA spacer sequences, I designed 5 different iSBH-sgRNA sequences. For each iSBH-sgRNA, corresponding RNA triggers and 1xCTS-ECFP reporters were also designed. Ability of first-generation iSBH-sgRNA designs to drive expression of the ECFP reporter was assessed in the absence or presence of complementary RNA triggers. **B.** Testing if first-generation iSBH-sgRNAs detect short RNA triggers using dCas9-Vp64 and 8xCTS-ECFP reporters. The same 5 iSBH-sgRNAs and corresponding RNA triggers were tested using dCas9-Vp64 and 8xCTS-ECFP reporters. * Figure shows mean +/- standard deviation values measured for 3 biological replicates. P-values were determined through unpaired t-tests.

switches at a molecular level. This could be due to the fact that a percentage of molecules from the cellular pool of iSBH-sgRNAs might not adopt desired secondary structures. Stronger activators such as dCas9-VPR propagate this background noise. However, weaker activators such as dCas9-Vp64, which require concomitant binding of several effectors at the promotor to drive downstream gene expression, mask this noise. Unless specified otherwise, all following figures were generated using dCas9-Vp64 and 8xCTS reporters.

2.5 Activation of first-generation iSBH-sgRNA designs is trigger-specific

To develop the iSBH-sgRNA technology for detecting endogenous RNAs, it is imperative to ensure that complementary RNA triggers exert a very high level of specificity of binding to iSBH-sgRNAs. To ensure that iSBH-sgRNA activation is specific, I tested activation for all 25 iSBH-sgRNAs/RNA trigger combinations (Figure 2.6). Data shows that iSBH-sgRNAs only become activated in the presence of their corresponding triggers, suggesting exquisite activation specificity of the system. Furthermore, orthogonality between different iSBH-sgRNA and trigger pairs suggested that multiple iSBH-sgRNA trigger pairs could be incorporated withing genetic circuits and used in parallel for performing different tasks.

2.6 Closing remarks

Data presented in this chapter shows that iSBH-sgRNA technology could be repurposed towards RNA detection. Furthermore, engineered iSBH-sgRNA designs display high levels of specificity towards their complementary RNA triggers. Thus, changing CRISPRa reporter systems has had a great influences on the measured iSBH-sgRNA OFF-state and ON-state activation. This suggests that a fraction

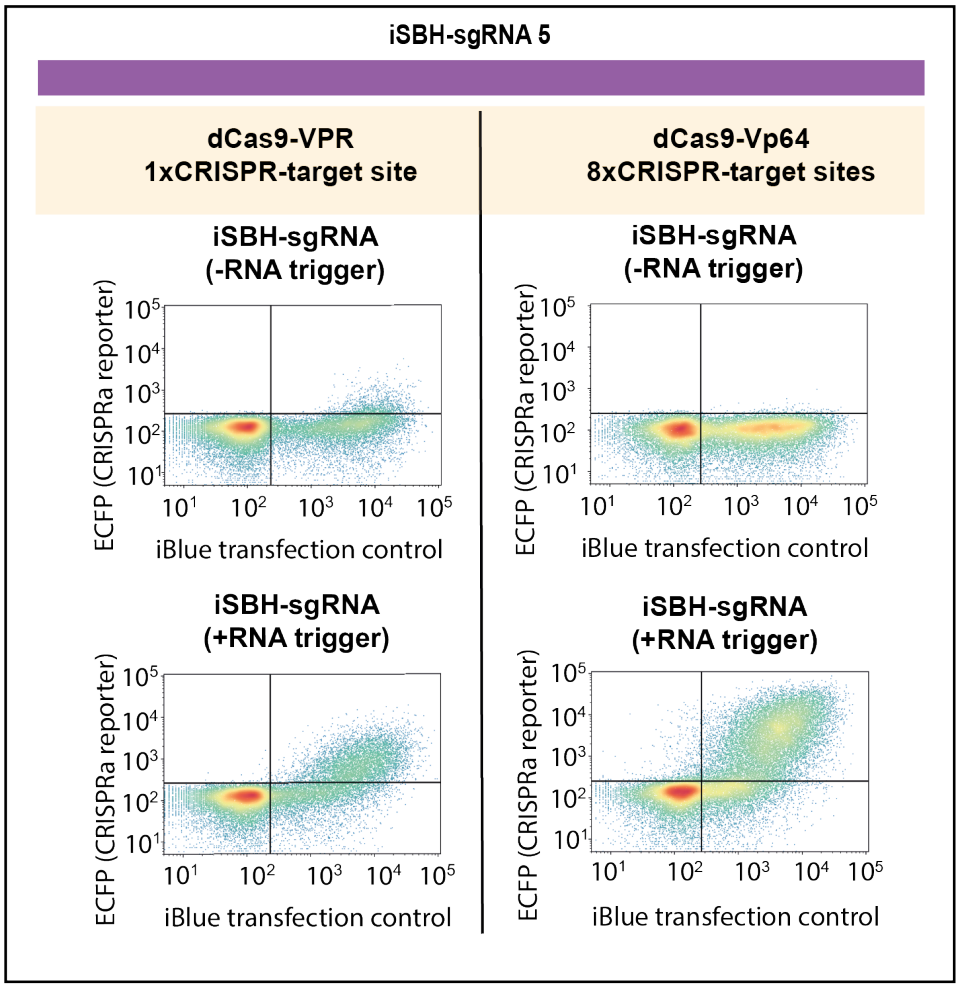


Figure 2.5: Representative Flow Cytometry plots

Plots compare ECFP fluorescence observed for iSBH-sgRNA 5 activation. In the left, OFF state and ON state activity is displayed using CRISPRa assays consisting of the 1xCTS reporters and dCas9-VPR. In the right, OFF state and ON state activity is displayed using CRISPRa assays consisting of the 8xCTS reporters and dCas9-Vp64. As iBlue fluorescent gene is encoded in plasmids encoding both iSBH-sgRNAs and RNA triggers, this marker serves as a transfection control.

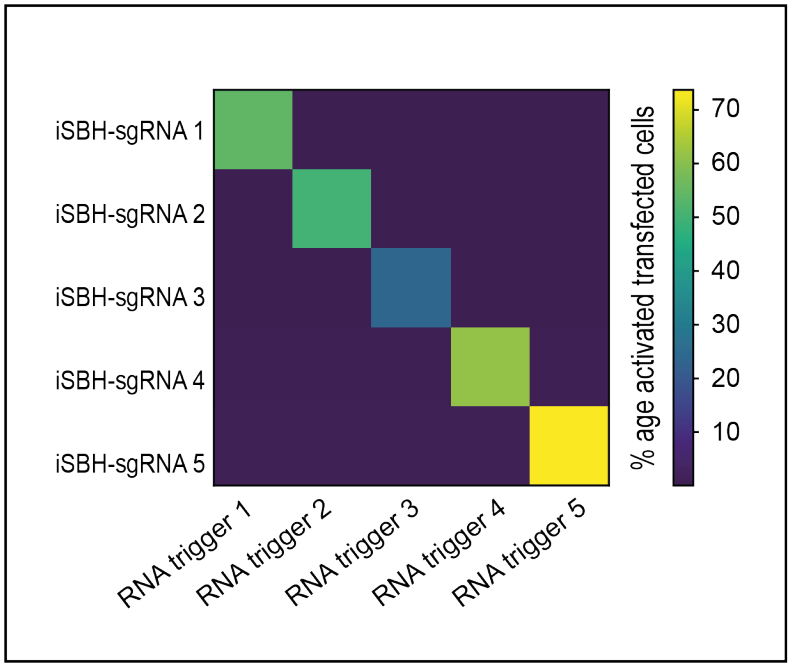


Figure 2.6: iSBH-sgRNA activation depends on complementary RNA triggers

An orthogonality test was performed, in which the the 5 iSBH-sgRNA designs were tested against all 5 RNA triggers. Activation is only deteced in the presence of matching iSBH-sgRNA and RNA trigger pairs.

of the iSBH-sgRNAs expressed by the cells do not adopt the desired secondary structure. Depending on downstream applications of the iSBH-sgRNA technology, extra work may be required to further reduce the iSBH-sgRNA OFF-state activation.

3

Further iSBH-sgRNA design optimisation and characterisation

3.1 Introduction to second-generation iSBH-sgRNA designs

The main aim of this thesis is to optimise the iSBH-sgRNA technology towards detecting endogenous RNA. Therefore, the ability to detect longer RNA triggers is essential, as the majority of endogenous RNA species are longer than the 34nt RNA triggers detected in initial proof-of-concept experiments (Figure 2.4).

Previous studies suggested that extending the length of hybridisation probes improves RNA targeting [139, 140]. Starting from this principle, second-generation iSBH-sgRNAs were designed by extending the length of the iSBH-sgRNA backfold complementary with the RNA trigger. A 10nt extension was introduced between the spacer and loop sequences, resulting in a 30nt backfold. Increasing the size of iSBH-sgRNAs enabled increasing the size of complementary RNA triggers to 44 rather than 34nt (Figure 3.1).

In this chapter I aim to test whether second-generation iSBH-sgRNA designs have improved features towards detection of longer RNA triggers. Furthermore,

different RNA trigger features that could influence iSBH-sgRNA activation are also explored.

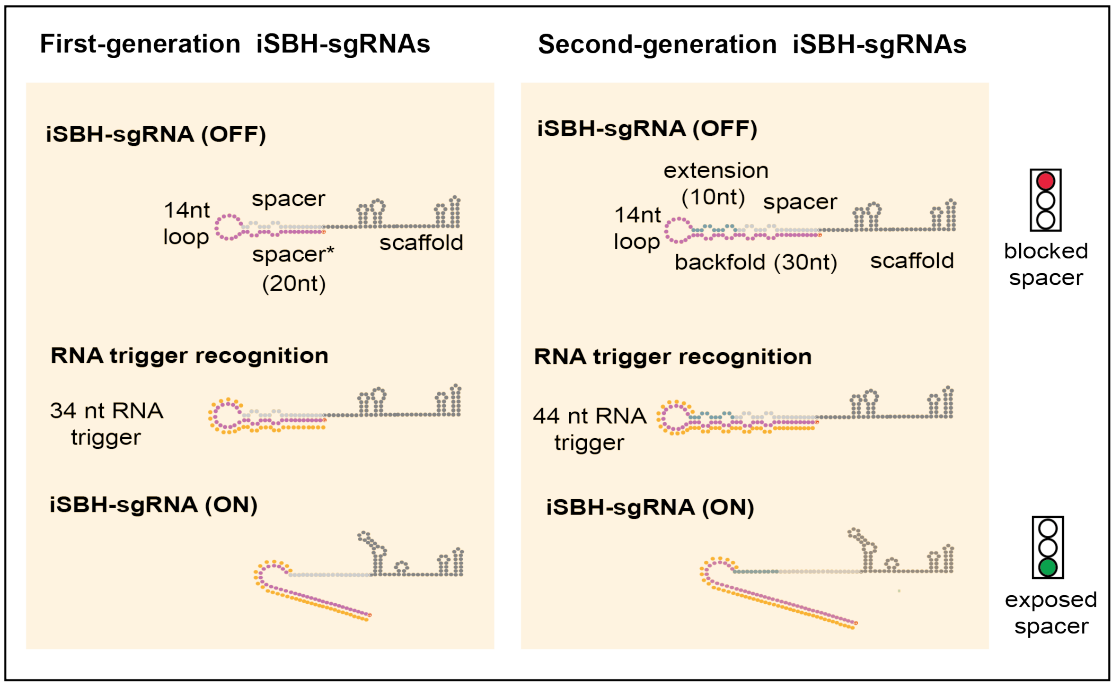


Figure 3.1: Comparing first-generation and second-generation iSBH-sgRNA designs

In contrast with first-generation iSBH-sgRNAs, second-generation designs contain a longer hairpin structure. A 10nt extension region was inserted between the spacer and loop sequences. This enabled increasing the size of the backfold sequence to 30nt. Short RNA triggers complementary with second-generation iSBH-sgRNAs have 44nt and are complementary with loop and backfold sequences.

3.2 Testing if second-generation iSBH-sgRNA designs detect RNA triggers

In the initial experiment, I sought to test if second-generation iSBH-sgRNA designs are silent in an OFF-state and still able to detect short RNA triggers. Performance of second-generation iSBH-sgRNA designs was tested using 6 different combinations of CRISPRa reporter systems (Figure 3.2). Similarly with the assessment performed for first-generation designs, I used 1xCTS-ECFP and 8xCTS-ECFP reporters. In

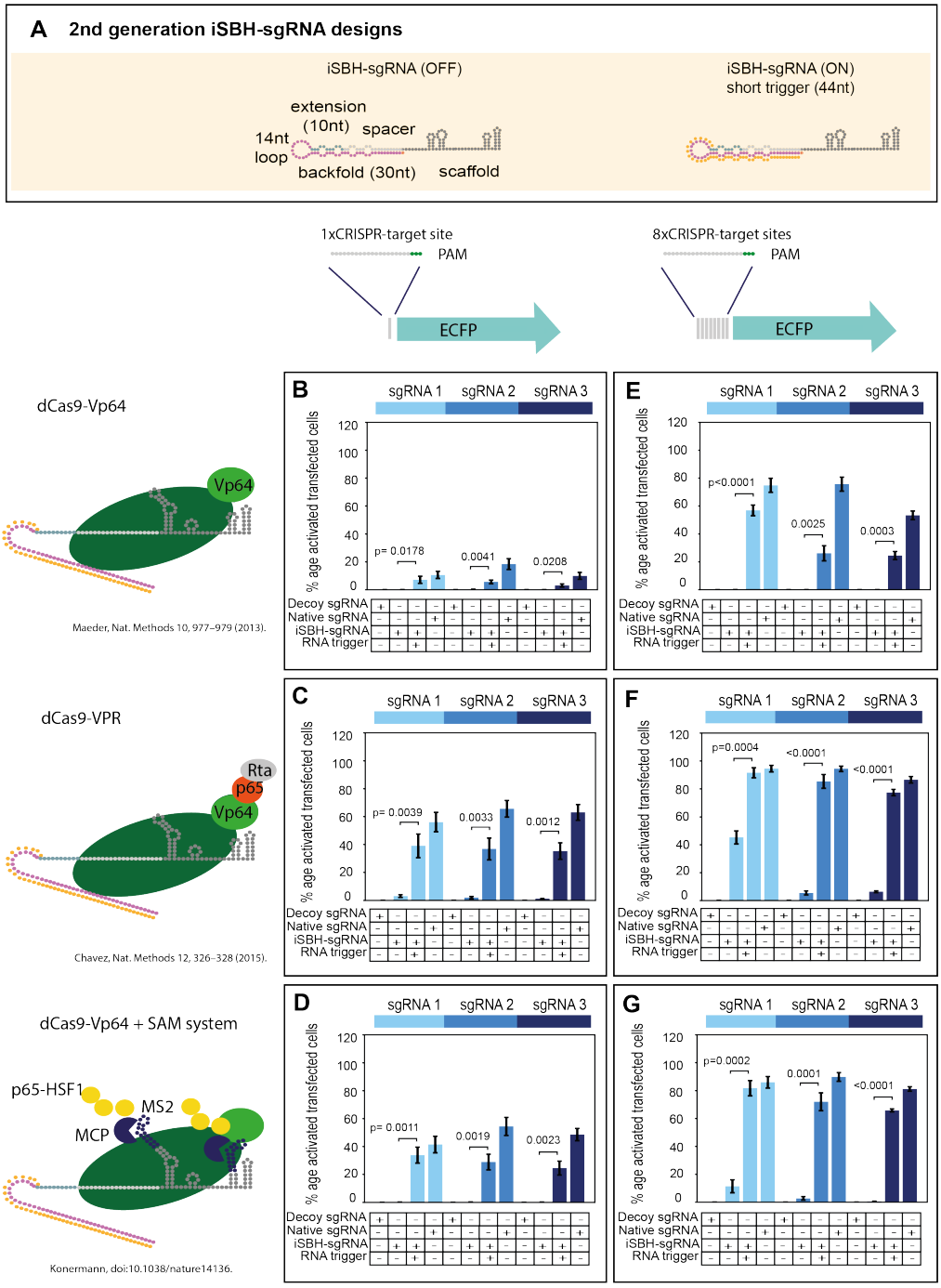


Figure 3.2: CRISPRa reporters of choice influence ON/OFF ratios of second generation iSBH-sgRNA designs while detecting short RNA triggers

A.Second-generation iSBH-sgRNA designs and corresponding short RNA triggers. **B.**Testing second-generation iSBH-sgRNAs using dCas9-VP64 and 1xCTS-ECFP reporters. **C.**Testing second-generation iSBH-sgRNAs using dCas9-VPR and 1xCTS-ECFP reporters. **D.**Testing second-generation iSBH-sgRNAs using dCas9-VP64, the SAM booster system and 1xCTS-ECFP reporters. **E.**Testing second-generation iSBH-sgRNAs using dCas9-VP64 and 8xCTS-ECFP reporters. **F.**Testing second-generation iSBH-sgRNAs using dCas9-VPR and 8xCTS-ECFP reporters. **G.**Testing second-generation iSBH-sgRNAs using dCas9-VP64, the SAM booster system and 8xCTS-ECFP reporters. * Figure shows mean +/- standard deviation values measured for 3 biological replicates. P-values were determined through unpaired t-tests.

terms of CRISPRa effectors, dCas9-Vp64 was included as a weak activator and dCas9-VPR as a stronger activator. Furthermore, dCas9-Vp64 was also tested in conjunction with the SAM amplification system[141].

In concordance with the observations made for first-generation iSBH-sgRNA designs, dCas9-Vp64 and 8xCTS-ECFP have the cleanest OFF-state (Figure 3.2.E.). Nevertheless, combinations between the 8xCTS-ECFP reporter and dCas9-VPR (Figure 3.2.F.) or dCas9-Vp64 and the SAM system (Figure 3.2.G.) substantially improved the ON-state activation, while compromising on a clean OFF-state. This finding further supports the idea that a certain percentage of iSBH-sgRNA molecules might not adopt desired secondary structures and that CRISPRa reporters could amplify or mask this noise.

3.3 Second-generation iSBH-sgRNA display superior performances while detecting longer RNA triggers

In the next round of experiments, I compared the ability of the first- and second-generation iSBH-sgRNAs to detect longer RNA triggers (Figure 3.3). In a first instance, I designed longer RNA triggers by appending a 100nt flank to the 3' end of short RNA triggers (100nt 3' flanks). For first-generation designs, activation in the presence of the trigger with 100nt 3' flank was efficient only for iSBH-sgRNA 1 (Figure 3.3). For second generation iSBH-sgRNA designs, notable activation was observed for all 3 iSBH-sgRNAs tested and ECFP production was detected in 15 to 30% of the transfected cells. This result confirmed the hypothesis that second-generation iSBH-sgRNA designs have superior abilities of detecting longer RNA triggers.

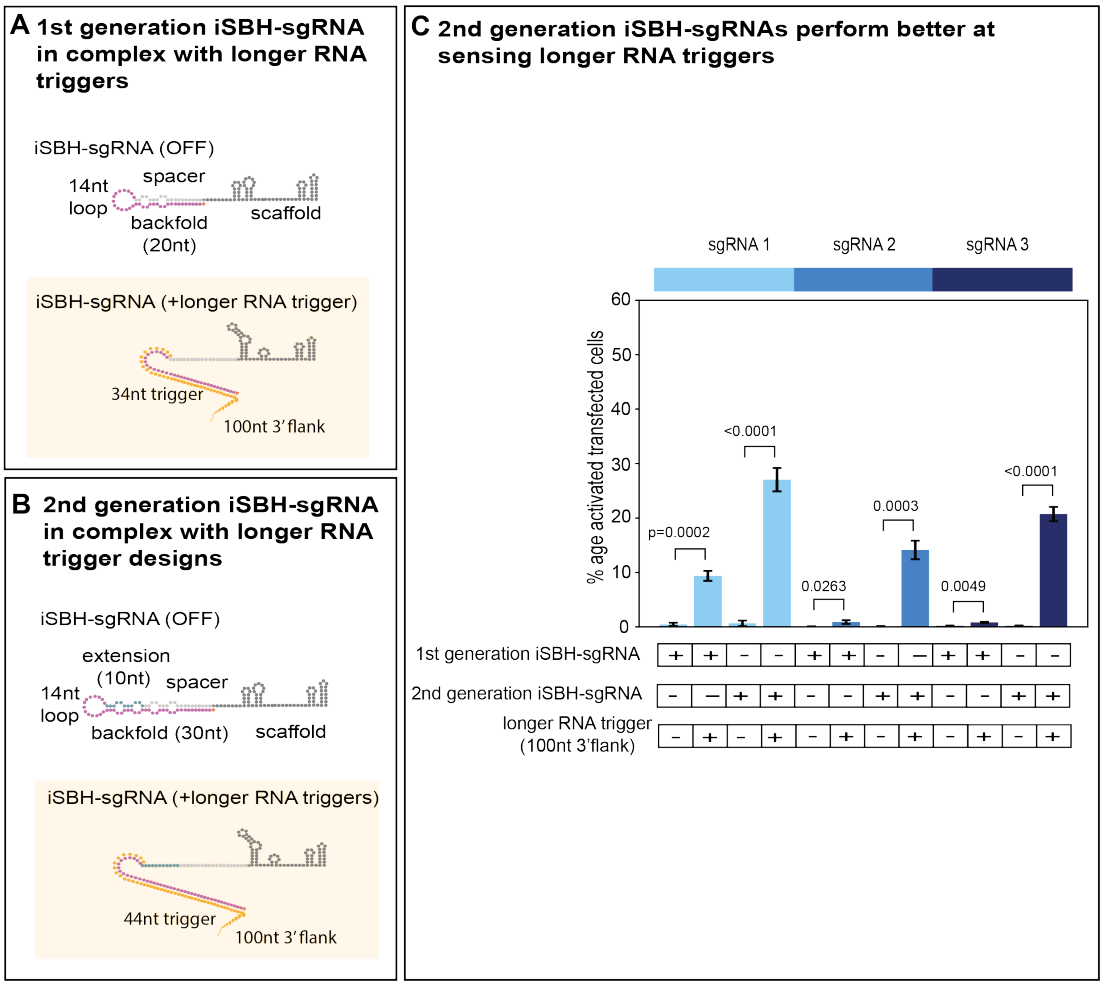


Figure 3.3: Second generation iSBH-sgRNAs detect longer RNA triggers in HEK293T cells

A. First-generation iSBH-sgRNA in complex with longer RNA triggers. Longer RNA triggers complementary with first-generation iSBH-sgRNAs have a 34nt sequence complementary with the loop and spacer* iSBH-sgRNA sequences. Triggers also have a 100nt flanking sequence immediately downstream from the iSBH-sgRNA complementary region. **B.** Second-generation iSBH-sgRNA in complex with longer RNA triggers. Second-generation designs contain a longer hairpin structure. A 10nt extension region was inserted between the spacer and loop sequences. This enabled increasing the size of the backfold sequence to 30nt. Longer RNA triggers complementary with 2nd generation iSBH-sgRNAs were designed, including triggers with 100nt 3' flanks, 100nt 5' flanks as well as 100nt 5' and 3' flanks. All trigger designs contain 44nt sequences complementary with the loops and backfold of the second-generation iSBH-sgRNAs. **C.** Second-generation iSBH-sgRNAs perform better at sensing longer RNA triggers. Ability of first-generation and second-generation iSBH-sgRNAs to sense 100nt 3' flank triggers was assessed. * Figure shows mean +/- standard deviation values measured for 3 biological replicates. P-values were determined through unpaired t-tests.

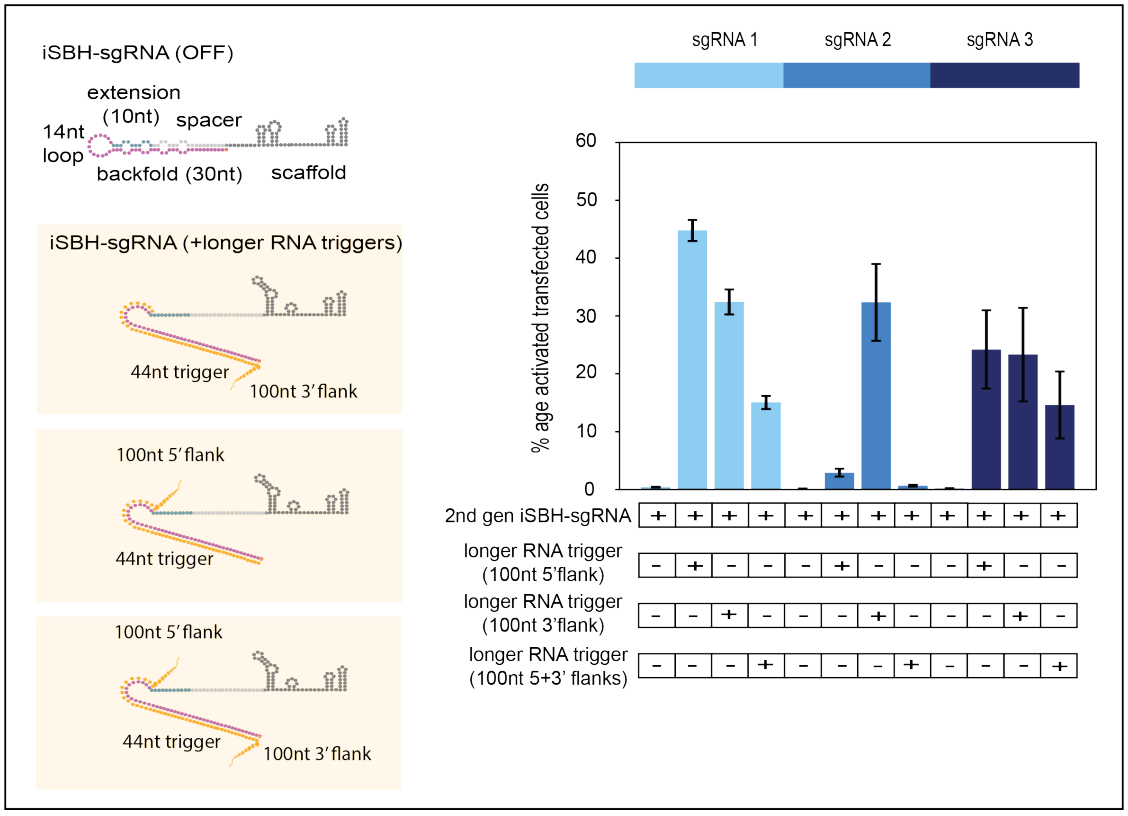


Figure 3.4: Assessing if RNA trigger length influences activation of second-generation iSBH-sgRNAs

Ability of second-generation iSBH-sgRNAs to detect different triggers with 100nt 5' flanks, 100nt 3' flanks, and 100nt 5' and 3' flanks was assessed. *Figure shows mean +/- standard deviation values measured for 3 biological replicates.

3.4 Trigger features that influence second-generation iSBH-sgRNA activation

Next, I investigated how other trigger features might influence iSBH-sgRNA activation. In particular, I was aiming to understand how even longer RNA triggers and triggers forming complex secondary structures could affect activation. Additionally, I attempted to characterise iSBH-sgRNA activation while expressing RNA triggers from different mammalian promoters as well as the impact of cytoplasmic trigger RNA export on iSBH-sgRNA activation.

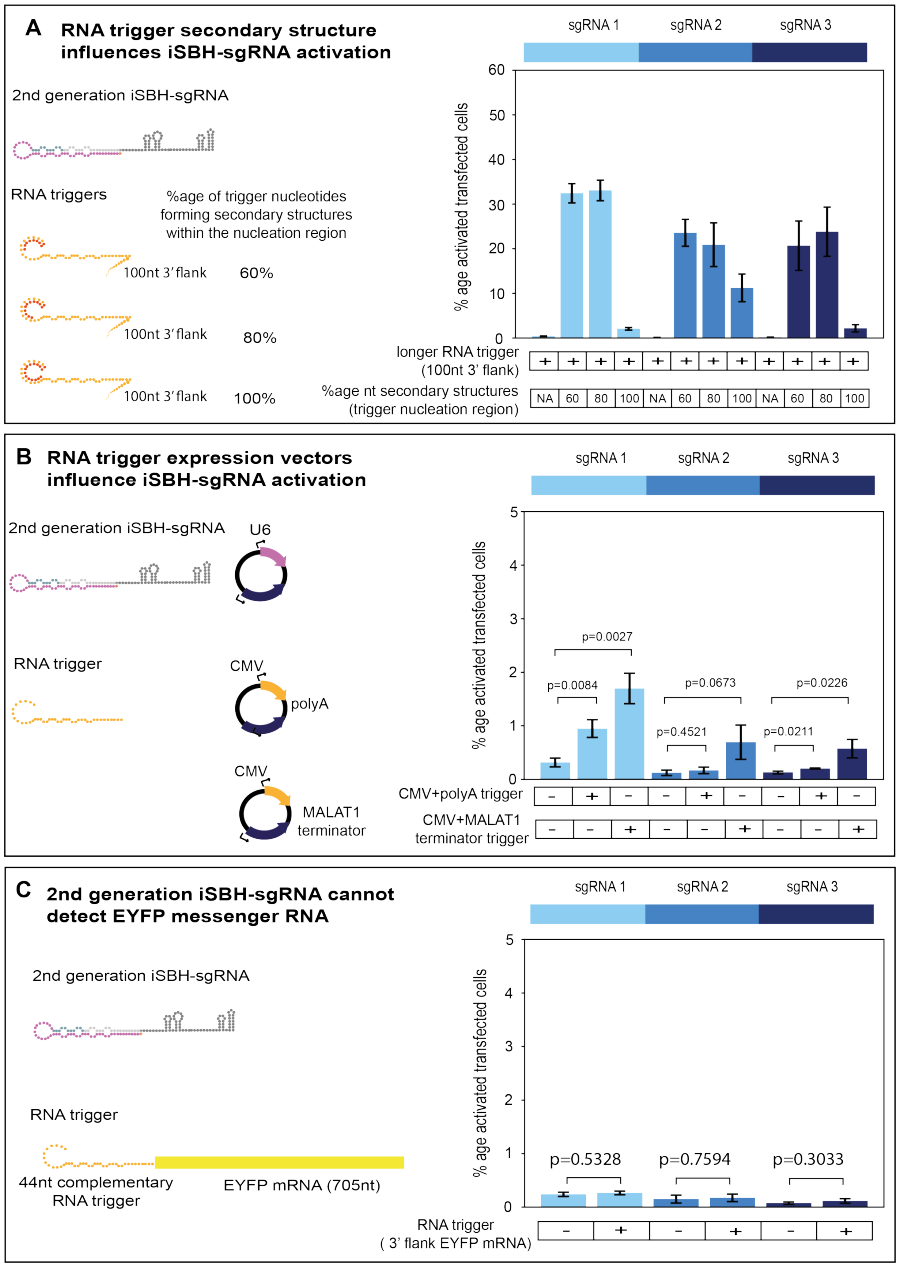


Figure 3.5: Trigger features that influence activation of second generation iSBH-sgRNAs

A. RNA trigger secondary structures influence iSBH-sgRNA activation. Trigger sub-sequences interacting with the iSBH-sgRNA loops were included within increasingly complex secondary structures. In this way, 60%, 80% or 100% of the nucleotides found within this trigger sub-sequence formed base-pairing interactions with other nucleotides within the RNA trigger. **B.** RNA trigger expression vectors influence iSBH-sgRNA activation. RNA triggers are expressed from polymerase II promoters (CMV, cytomegalovirus). Polyadenylation (poly(A)) signals or MALAT1 terminators were added to the 3' end of the expression cassettes. Triggers that have a poly(A) tail are expected to be exported in the cytoplasm [142], while triggers with the MALAT1 terminator are expected to be retained within the nucleus [97]. **C.** Second-generation iSBH-sgRNAs cannot detect EYFP messenger RNA. The 44nt RNA trigger was cloned downstream of the EYFP mRNA before the poly(A) signal. * Figure shows mean +/- standard deviation values measured for 3 biological replicates. P-values were determined through unpaired t-tests.

3.4.1 Trigger length and secondary structure

To test whether second-generation iSBH-sgRNAs could also detect other longer RNA trigger designs (Figure 3.4), I appended 100nt flanks at the 5' end (100nt 5' flank), 3' end (100nt 3' flank) or at both ends (100nt 5'+3' flanks) of the original 44nt trigger sequence.

My data suggested that shorter triggers, having 100nt extensions at either 5' or 3' end, were detected more efficiently than triggers flanked at both ends (Figure 3.4). This is consistent with previous studies suggesting that efficiency of detecting RNA using engineered gRNA designs drops with the increase in the length of RNAs triggers[87, 103, 129]. A possible explanation is that longer the RNAs form multiple secondary structures, preventing RNA triggers from interacting with engineered gRNAs[103]. Indeed, increasing formation of secondary structures within the RNA triggers could lead to decrease iSBH-sgRNA activation (Figure 3.5.A.).

3.4.2 Mammalian expression vectors used to express RNA triggers

In previous experiments short RNA triggers were expressed from polymerase III promoters (U6). To get a grasp of the feasibility of sensing different endogenous RNA species, I carried out a series of exploratory experiments where I expressed synthetic RNA triggers from different mammalian promoters. A previous study suggested that triggers expressed from polymerase II promoters are detected less efficiently than triggers expressed from polymerase III promoters. This is believed to be caused by nuclear export of polymerase II transcripts [100]. Data presented here also shows that placing short triggers under the control of mammalian CMV (cytomegalovirus) promoters resulted in a maximum 1% of transfected cells turning ECFP positive in the presence of polyadenylation signals (Figure 3.5.B.). Furthermore, no significant activation was detected when attempting to detect RNA triggers embedded in long messenger RNA sequences (Figure 3.5.C.).

I also tested RNA triggers expressed from polymerase II promoters that could not be exported from the nucleus. This was achieved by replacing polyadenylation signals (poly(A)) with MALAT1 terminators, known to promote nuclear retention[97]. Nuclear retention of RNA triggers appeared to improve iSBH-sgRNA activation; nevertheless the percentage activated cells only reached a maximum of 2% (Figure 3.5.B.). This suggested that nuclear export by itself may not be the only cause for a lower activation in the presence of polymerase II-expressed RNA triggers. Other factors such as concentration, additional increase in size due to addition of polyadenylation signals or MALAT1 terminators and the biology behind RNA trigger production may play significant roles in iSBH-sgRNA activation.

3.5 Closing remarks

In summary, data presented in this chapter suggests that second-generation iSBH-sgRNAs are superior to first-generation designs when detecting longer RNA triggers. In terms of trigger RNA features, second-generation designs could efficiently detect RNAs of up to about 300 nucleotides, when expressed from U6 promoters. Expressing RNA triggers from CMV promoters massively decreased iSBH-sgRNA activation. Despite not fully characterising the range of RNA species that could act as triggers, data presented in this chapter suggests that choosing smaller RNA triggers that are relatively free from secondary structures would be a good choice. Furthermore, choosing RNA triggers with nuclear localisation and preferably expressed from polymerase III promoters could provide additional advantages when designing an iSBH-sgRNA sensor with a higher activation rate.

4

Designing iSBH-sgRNAs for custom CRISPR experiments

4.1 Limitations of second-generation iSBH-sgRNAs and need for modular iSBH-sgRNA designs

In the second-generation iSBH-sgRNAs, the requirement for complementarity means that the spacer sequence restricts RNA trigger sequence choices. As such, these designs do not allow users to detect RNA triggers of choice, while targeting Cas9 to any desired CTS. Endogenous RNA detection would be facilitated by developing a platform for designing modular iSBH-sgRNAs where spacer and trigger-sensing components can be chosen independently from one another (Figure 4.1).

A way of achieving this involves reducing complementarity between the RNA trigger and iSBH-sgRNAs (Figure 4.2.A.) as well as reducing the extent of complementarity between spacer sequences and CTSs (Figure 4.2.B.). In a modular iSBH-sgRNA prototype, RNA triggers are only complementary with the loop and the first 15nt of the iSBH-sgRNA backfold, and iSBH-sgRNAs have only 17nt spacer sequences complementary with CTSs. To mitigate the impact of truncating RNA trigger sequences on the affinity between iSBH-sgRNAs and RNA triggers, I increased the size of iSBH-sgRNA loops (Figure 4.1).

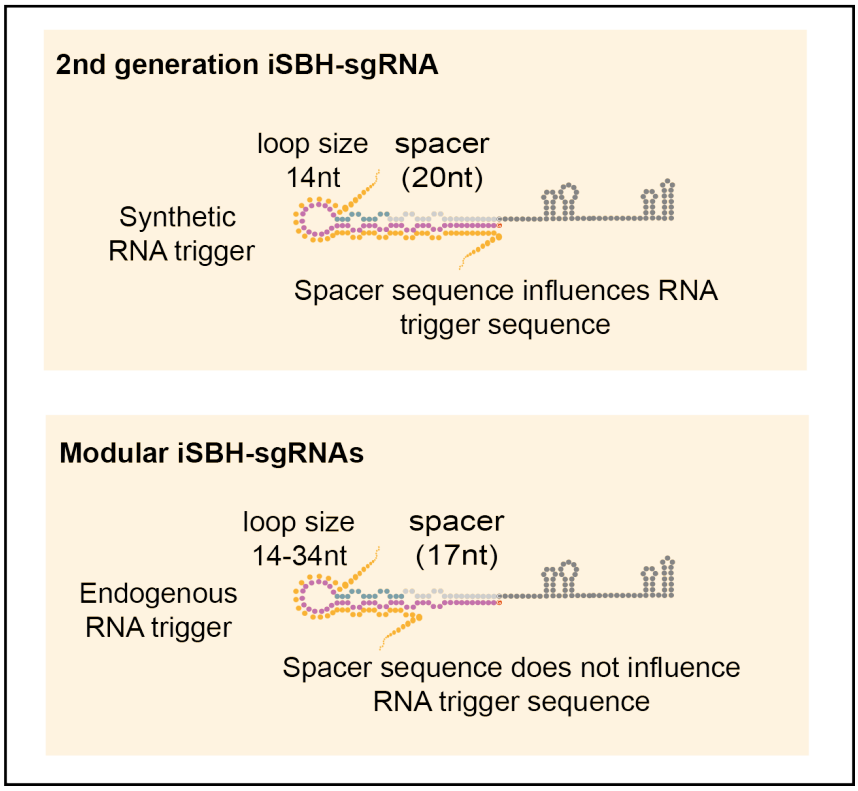


Figure 4.1: Modular iSBH-sgRNA designs enable spatial separation of spacer and trigger-sensing sequences

RNA triggers for second-generation iSBH-sgRNA designs are fully complementary with the iSBH-sgRNA loop and backfold sequences, meaning that the sgRNA spacers influence RNA trigger sequences. In order to facilitate design of iSBH-sgRNAs capable of sensing input RNA triggers of choice, while targeting Cas9 to any CTS of choice, modular iSBH-sgRNA designs are proposed. For modular iSBH-sgRNAs, triggers are only complementary with the iSBH-sgRNA loop and with the first 15nt of the backfold. In order to increase affinity between iSBH-sgRNAs and RNA triggers, the size of the loop sequence could be further increased. Efficient separation between trigger-sensing and spacer sequence is also achieved by reducing the complementarity between the spacer sequence and CTS from 20 to 17nt.

4.2 MODesign computational pipeline

Starting from the modular iSBH-sgRNA prototype, I developed the MODesign computational pipeline that enables iSBH-sgRNA design starting from input RNA triggers, sgRNA spacers, and desired loop sizes (Figure 4.3). MODesign calculates the size of the iSBH-sgRNA trigger-sensing component and creates a list of all potential trigger sub-sequences having that particular size. For each sub-sequence,

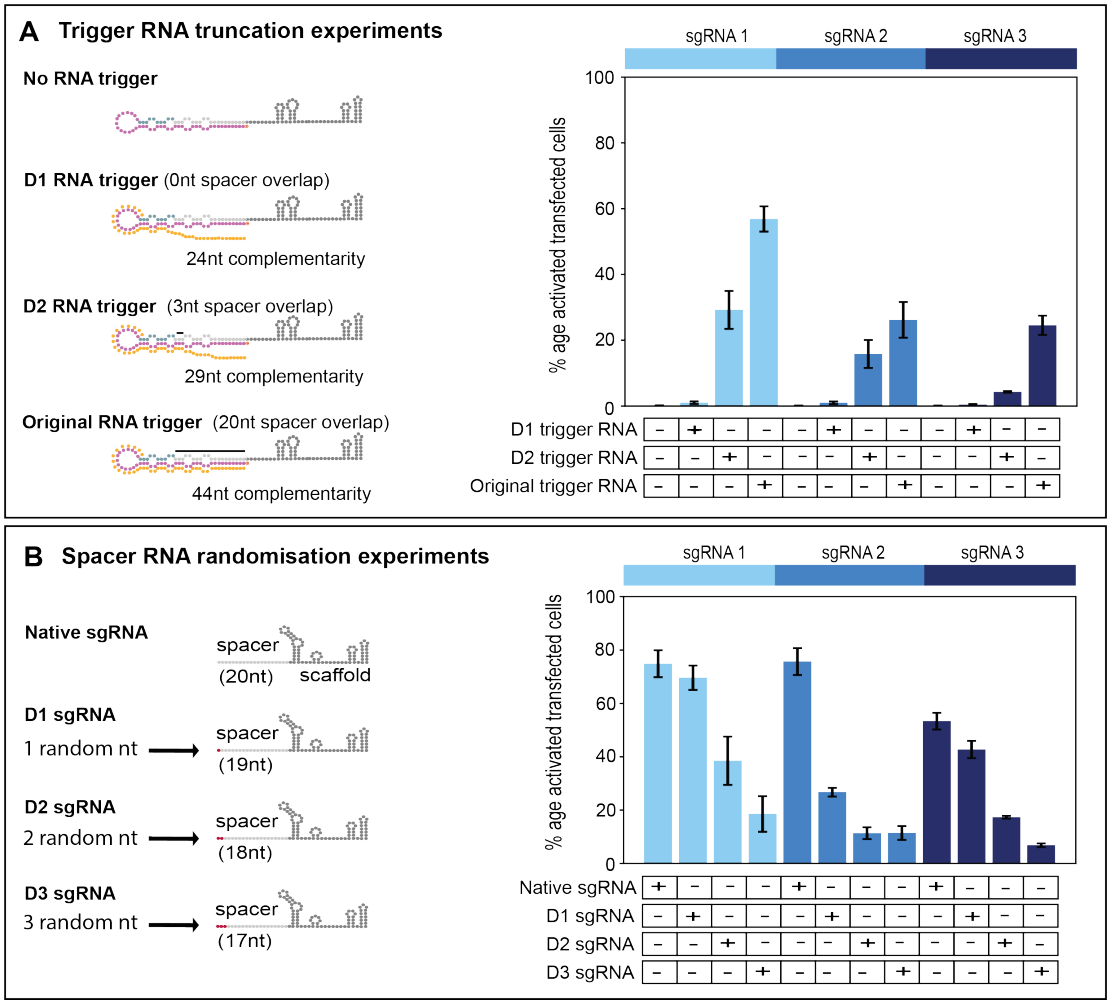


Figure 4.2: Exploring design space for modular iSBH-sgRNA designs

A. Trigger RNA truncation experiments. In the original RNA trigger design, the trigger sequence was influenced by the 20nt sgRNA spacer sequence. A way of avoiding this involves to truncating the trigger sequence, so that there is little to no overlap with the spacer sequence. In D1 design, the 20nt spacer sgRNA sequence does not influence at all the sequence of the RNA trigger. In D2 design, the first 3nt of the spacer are still influencing the sequence of the RNA trigger. **B.** Spacer RNA randomisation experiments. Here, I reduced the complementary between the sgRNA spacer sequence and the CTS. This experiment was carried out to accommodate D2 RNA trigger designs and to completely abolish the dependency between the sgRNA spacer and trigger sequences. Native sgRNAs that matched 19, 18 and 17nt of the CTS were designed and tested. * Figure shows mean +/- standard deviation values measured for 3 biological replicates.

it determines the reverse complementary region, and it inserts it between extension and spacer* sequences, while filling in the extension sequence. Next, MODesign checks if resulting modular iSBH-sgRNA sequences fold into desired RNA structures using NuPACK[143] and outputs all designs that adopt a correct fold.

Depending on input parameters, MODesign simulations could produce a list of multiple modular iSBH-sgRNA outputs. For proof-of-concept experiments, I decided to make educated guesses in deciding which iSBH-sgRNAs to select for experimental validation (Figure 4.3). A first criteria was to prioritise sequences predicted by NuPACK to adopt desired structures with high probabilities. I believe that a higher probability of adopting desired secondary structures would reduce the number of iSBH-sgRNAs that do not adopt perfect OFF switches within the cellular pool. Based on previous experimental evidence showing that trigger secondary structures could influence activation (Figure 3.5.A.), priority was also given to sequences that hybridise to RNA trigger sub-sequences lacking complex secondary structures.

In the modular iSBH-sgRNA design, triggers are designed to base-pair only with the iSBH-sgRNA loop and the first 15nt of the backfold, while iSBH-sgRNA spacers have only 17nt complementarity with the CTS (Figure 4.1). Nevertheless, by random chance, depending on the trigger sensing sequence, modular iSBH-sgRNAs could have extra complementarity between RNA triggers and the last 15 nucleotides of the backfold or more than 17nt complementarity with the CTS. As these features are beneficial for iSBH-sgRNA activation and for detecting CRISPRa activity (Figure 4.2), extra priority was also given to sequences displaying these features.

4.3 Experimentally validating MODesign outputs

In a first validation experiment, I run 3 MODesign simulations for designing iSBH-sgRNAs capable of sensing a U6 driven RNA trigger whose sequence corresponded to a 146nt mouse α-globin enhancer RNA (trigger D). Initial modular iSBH-sgRNAs had different sgRNA spacer sequences and 14nt loops. Outputs resulting from

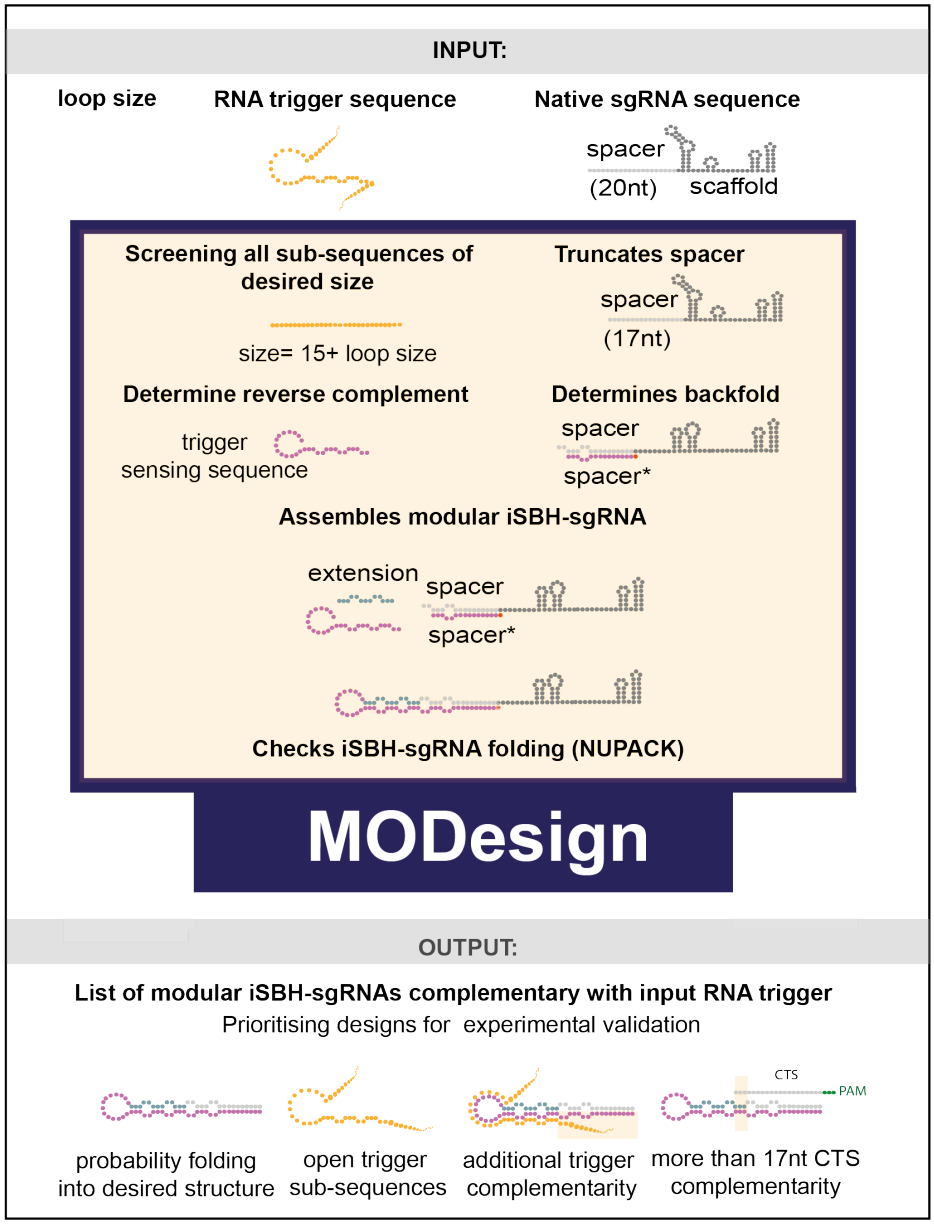


Figure 4.3: Computational pipeline for modular iSBH-sgRNA design

MODesign enables users to design modular iSBH-sgRNAs starting from desired RNA trigger, sgRNA spacer sequence and loop sizes. MODesign calculates the size of trigger-sensing sequences and it creates a list of all trigger sub-sequences having that size. Script determines the reverse complement of these sequences that could act as trigger-sensing sequences. iSBH-sgRNAs are assembled through adding spacer*, trigger-sensing sequences, extension, spacer and scaffold sequences. Extension sequences are engineered to be partially complementary with trigger-sensing sequences. When inserted within the iSBH-sgRNA structure, trigger-sensing sequences should form a loop. Before producing a list of output sequences, iSBH-sgRNA folding is checked using NuPACK[143]. Each simulation could result in multiple modular iSBH-sgRNA designs. Designs chosen for further testing were selected based on the probability of folding into the iSBH-sgRNA structure and lack of trigger secondary structures in the iSBH-sgRNA complementary region. Priority is also given to iSBH-sgRNAs that (by chance) had more than 17nt complementary with the CTS and have more longer stretches of complementary with the input RNA trigger sequence.

MODesign were cloned and co-transfected to HEK293T cells together with plasmids expressing trigger D from U6 promoters. My results suggest that the modular designs generated by the MODesign algorithm had good OFF-state activity and were activated by RNA trigger D (Figure 4.4). These tests were carried out using dCas9-Vp64 and 8xCTS-ECFP reporters.

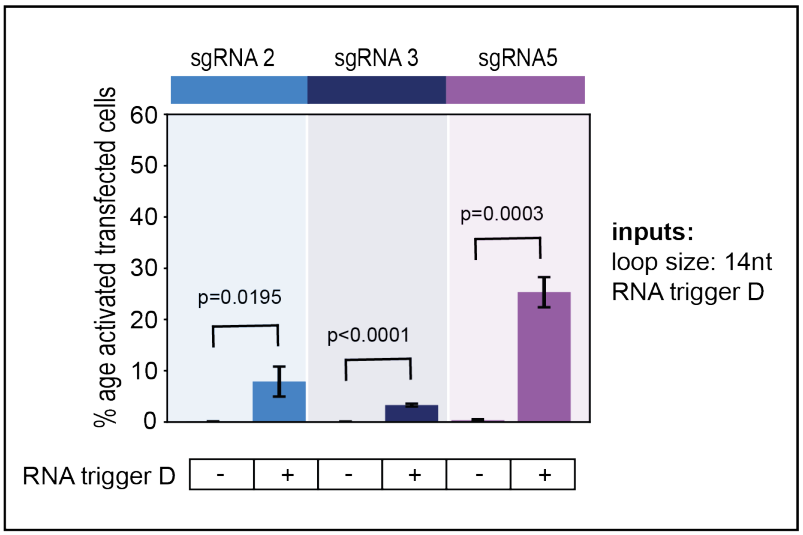


Figure 4.4: Designing modular iSBH-sgRNAs starting from one RNA trigger and 3 sgRNAs

3 MODesign simulations were carried out. The aim of these simulations was to design iSBH-sgRNAs capable of sensing trigger RNA D (146nt eRNA sequence). In each simulation, a different sgRNA sequence was used and desired loop size (14nt) was kept constant between simulations. Selected designs were transfected to HEK293T cells together with the RNA trigger D sequence (expressed from a U6 promoter). Tests were carried out using dCas9-Vp64 and 8xCTS-ECFP reporters. Figure shows mean +/- standard deviation values measured for 3 biological replicates. P-values were determined through unpaired t-tests.

After seeing that modular iSBH-sgRNAs can be designed starting from different sgRNA spacer sequences, a second test involved testing if these designs could detect different input triggers. Trigger RNA A (146nt), trigger RNA B (267nt), trigger RNA C (268nt) and trigger RNA D (146nt) were chosen for this purpose. Triggers A, B and C involved a mix of zebrafish enhancer RNAs and repetitive elements specifically upregulated in the neural crest[136]. For mammalian cell tests, these

triggers were expressed from the control of U6 promoters (Figure 4.5). For each trigger, I ran 11 MODesign simulations incrementally extending the loop size, while keeping sgRNA 2 spacer input constant. HEK293T validation experiments showed that choosing modular iSBH-sgRNAs that detect the 4 U6-expressed triggers is possible (Figure 4.5.A.) and best results are displayed in the (Figure 4.5.B.) . As expected, triggers A and D are detected more efficiently due to their smaller size. Furthermore, for these two triggers, addition of the SAM effector system[141] boosts ON-state activation without compromising too much on the OFF-state. Orthogonality tests suggest that activation of modular iSBH-sgRNA designs is specifically conditioned by complementary RNA triggers (Figure 4.5.C.).

4.4 Closing remarks

This chapter aimed to develop the MODesign algorithm. MODesign enables future users of the iSBH-sgRNA technology to design their own modular iSBH-sgRNAs starting from desired RNA trigger and sgRNA sequences.

Strategies for tuning modular iSBH-sgRNA designs involves increasing loop sizes, but a fine balance should also be considered between increasing loop size and maintaining a high proportion of iSBH-sgRNAs that adopt the correct fold in the OFF-state. As inspired from other studies that designed antisense oligos, that used locked nucleic acid (LNA) modifications for increasing RNA affinity, such modifications could perhaps also be included in the iSBH-sgRNA designs when delivered as synthetic, chemically synthesised RNAs[144].

sgRNA spacer sequence, GC content, folding probability, loop size or sequence motifs might also play important roles in iSBH-sgRNA performance. Nevertheless, the exact contribution of these factors is not yet understood. Machine learning algorithms and large datasets would be required for developing accurate predictions. In the absence of good prediction tools, multiple modular iSBH-sgRNAs should

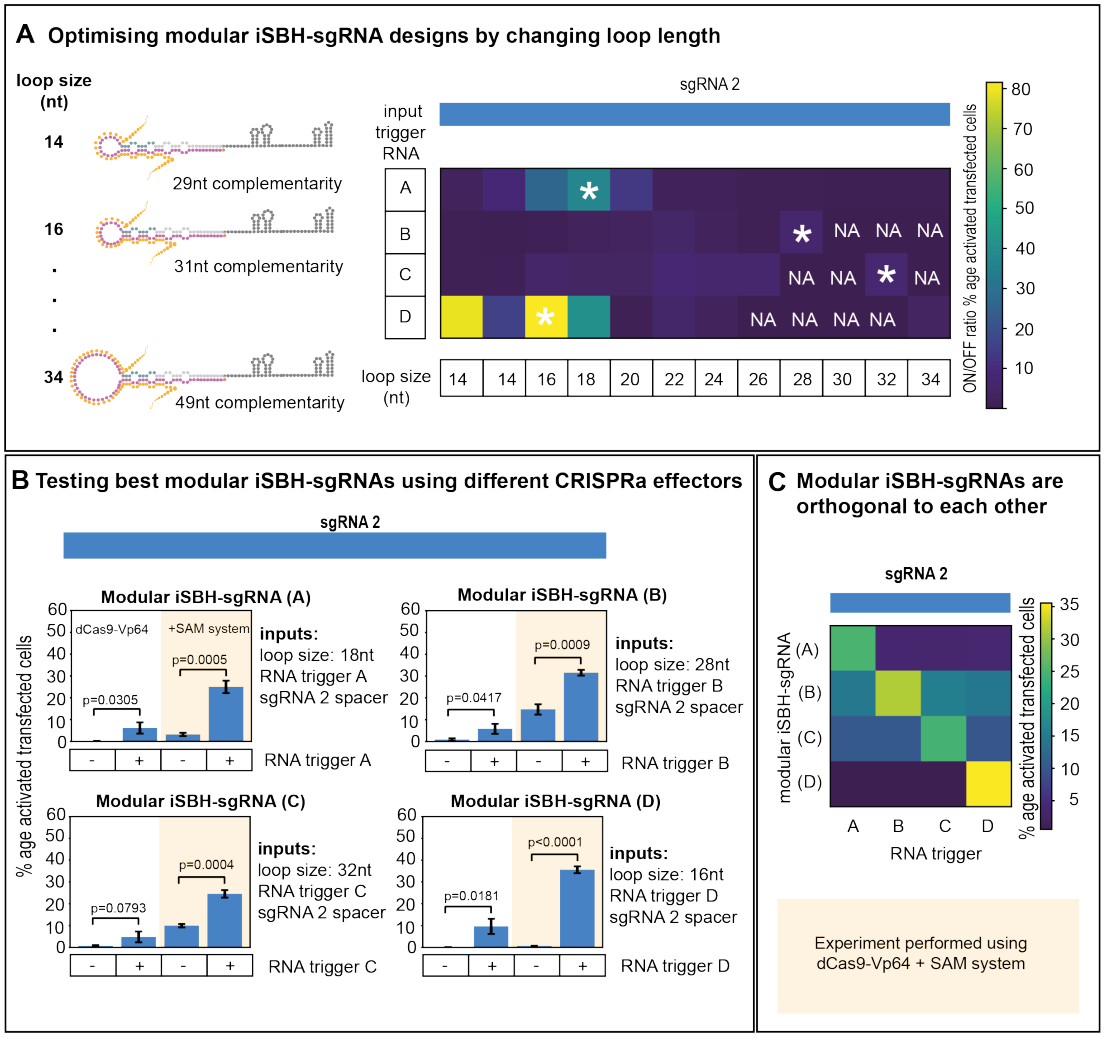


Figure 4.5: Designing modular iSBH-sgRNAs starting from 4 RNA triggers and one sgRNA sequence

A. Optimising modular iSBH-sgRNA designs by changing loop length. Multiple MODesign simulations were runfor designing iSBH-sgRNAs capable of sensing trigger RNA A (146nt repetitive RNA sequence), trigger RNA B (267nt repetitive RNA sequence), trigger RNA C (268nt repetitive RNA sequence) and trigger RNA D (146nt eRNA sequence). Each simulation contained different loop size specifications and sgRNA 2 sequence was kept constant between simulations. Selected iSBH-sgRNA designs were transfected to HEK293T cells together with corresponding RNA trigger sequences expressed from U6 promoters. For a loop size of 14nt, 2 different modular iSBH-sgRNAs were tested for each trigger. These iSBH-sgRNAs hybridised with different trigger sub-sequences. *NA* represents conditions where the RNA trigger sequence did not allow designing of iSBH-sgRNAs with a particular loop size. iSBH-sgRNA designs with best ON/OFF ratios observed for different RNA triggers are marked with *. **B.** Testing best modular iSBH-sgRNAs using different CRISPRa effectors. Graphs show mean +/- standard deviation values measured for 3 biological replicates. P-values were determined through unpaired t-tests. **C.** Modular iSBH-sgRNAs are orthogonal to each other. 4 modular iSBH-sgRNAs (A,B,C and D) were co-transfected to HEK293T cells and all iSBH-sgRNA: RNA trigger combinations were tested.

be designed for individual RNA detection applications, while best designs should be validated experimentally.

5

Insights into the mechanism of iSBH-sgRNA activation

5.1 Formation of double-stranded RNA structures and its implications

Next, I sought to investigate the mechanism of iSBH-sgRNA activation to further benefit iSBH-sgRNA technology development. It is known that in eukaryotic cells, double-stranded RNAs are recognised and cleaved by endogenous RNA processing pathways such as RNA interference (RNAi, Figure 5.1) [145, 146].

Interaction between iSBH-sgRNAs and RNA triggers leads to the formation of long double-stranded RNA structures and similar structures were reported to act as a non-canonical substrate for Dicer[146, 149]. Therefore, I tested whether double-stranded RNA processing mechanisms occur during delivery and subsequent activation of iSBH-sgRNAs.

I proposed two scenarios for the iSBH-sgRNA activation (Figure 5.2). In a first scenario, activation would happen due to RNA strand-displacement and resulting RNA duplexes would not be processed. If this mechanism is true, sizes of the iSBH-sgRNA and RNA triggers are expected to remain constant after activation.

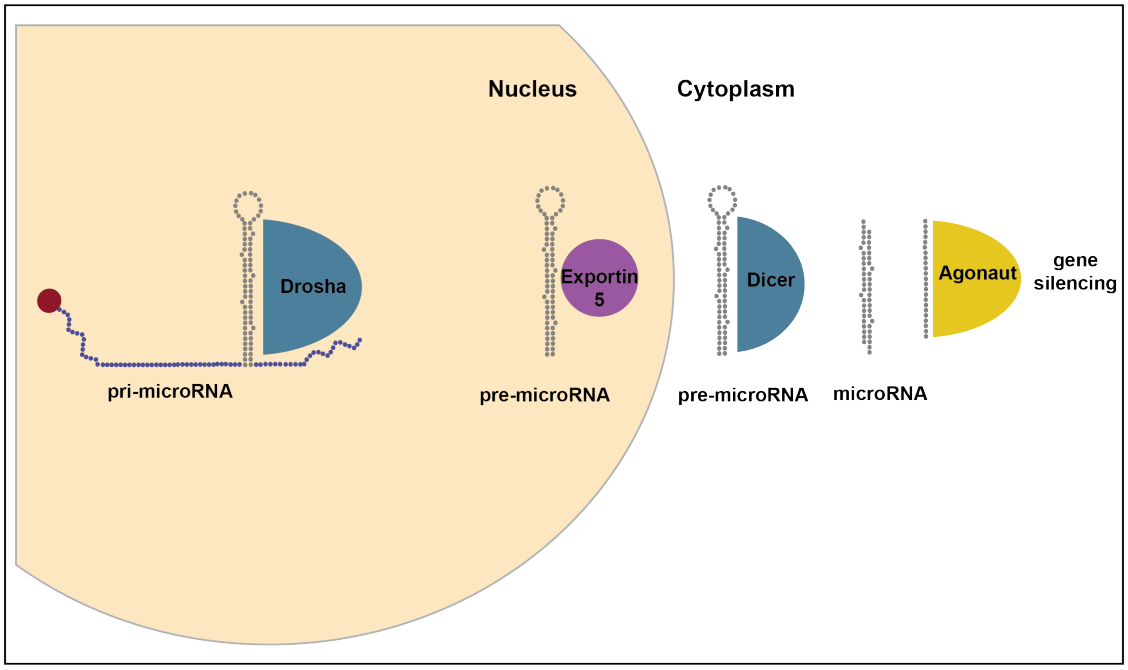


Figure 5.1: RNA interference pathway

RNAi is a conserved double-stranded RNA processing mechanism through which gene expression is silenced at a post-transcriptional level[147]. MicroRNAs are a class of highly characterised RNAi effectors in eukaryotic cells[69]. They are encoded as part of longer transcript precursors known as pri-microRNAs. In the nucleus, pri-microRNAs are recognised by Drosha and processed into short-hairpin RNAs known as pre-microRNAs. Pre-microRNAs are exported into the cytoplasm by Exportin 5, where they are further processed by Dicer. Dicer cleaves the hairpin structure, resulting in the formation of short 19-23 nucleotides double-stranded RNAs that contain 3' overhangs. From this double-stranded RNA duplex, the strand with less thermostable 5'ends tend to associate with Argonaut proteins. microRNA-Argonaut complexes drive gene silencing and the most common mechanism involves endonucleolytic cleavage of target RNAs[148].

A second scenario involves double-stranded RNA processing. As a consequence, iSBH-sgRNAs and RNA triggers would be truncated following activation. To test these scenarios, I carried out RNA circularisation assays[97] to measure the size of iSBH-sgRNAs (Figure 5.3.A.) and RNA triggers (Figure 5.4.A.) following activation.

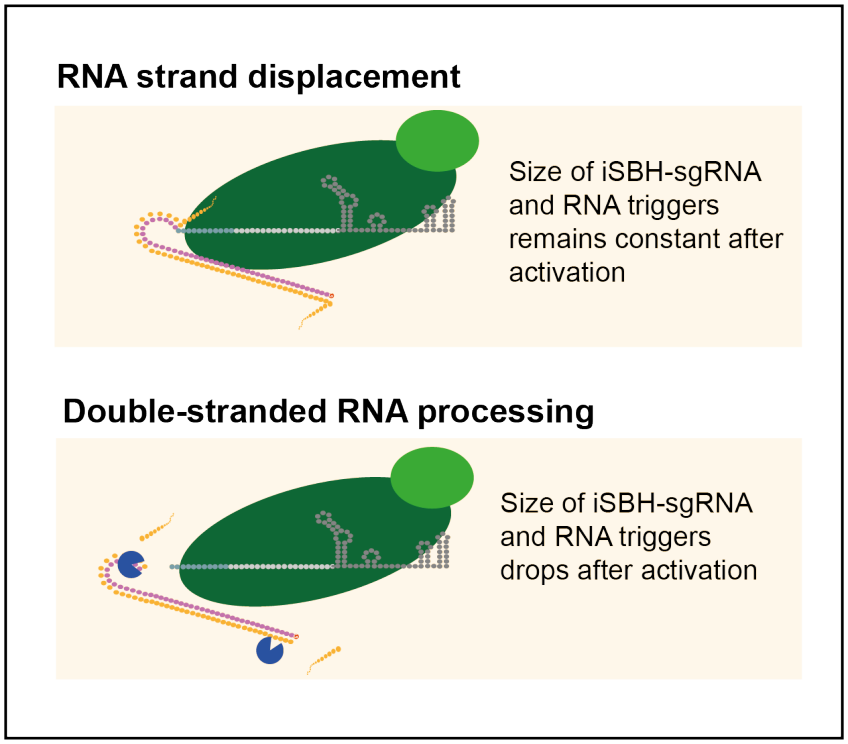


Figure 5.2: Potential scenarios for the mechanism of iSBH-sgRNA activation

Interaction between iSBH-sgRNAs and RNA trigger leads to the formation of long double-stranded RNA structures. A potential activation mechanism might involve RNA strand displacement and formation of stable molecular complexes between the iSBH-sgRNA and the RNA trigger sequence. Supposing this scenario is correct, the size of the iSBH-sgRNA and RNA triggers are expected to remain constant after activation. A second scenario involves double-stranded RNA processing. If this is correct, iSBH-sgRNAs and RNA trigger sequences are expected to be truncated.

5.2 Measuring the size of iSBH-sgRNAs following activation

In the first instance, I assessed the size of second-generation iSBH-sgRNA designs in the presence or absence of short, complementary RNA triggers and dCas9-Vp64 (Figure 5.3.B.). In the absence of RNA triggers, recovered iSBH-sgRNA RT-PCR bands matched the expected 137bp size. In the presence of dCas9-Vp64 and complementary RNA triggers, the bands decreased to 81 bp. This was similar in size to bands recovered from non-engineered native sgRNA control, suggesting that engineered sgRNA components (extension, loop, backfold) were removed from

iSBH-sgRNAs during activation. This observation was confirmed by sequencing PCR products recovered from truncated samples (Figure 5.3.C.).

In the absence of dCas9-Vp64, iSBH-sgRNA truncation was noticeable but very weak for 2 out of 3 iSBH-sgRNAs tested (Figure 5.3.B.). This indicates that dCas9 might be required for efficient iSBH-sgRNA processing. The exception was iSBH-sgRNA 1, however, the size of the truncated product was larger than the size of the truncated products recovered in the presence of dCas9. Taken together, these results suggests that formation of sgRNAs with 20nt spacers might be dCas9-dependent. In the OFF-state, iSBH-sgRNAs were unable to bind to dCas9. However, when RNA triggers are present, the conformational change enables iSBH-sgRNAs to bind to dCas9, leading to iSBH-sgRNA truncation. These results are also supported by previous studies reporting that engineered sgRNAs that have extended spacer sequences are processed to original 20nt spacer sequences[41, 107].

5.3 Measuring the size of RNA triggers following activation

I then explored the fate of RNA triggers during activation and measured the sizes of longer RNA triggers with 100nt 3' flanks in various experimental conditions (Figure 5.4.A.). The results demonstrate that RNA triggers are truncated specifically in the presence of matching iSBH-sgRNAs. When co-transfected together with iSBH-sgRNAs with incompatible hairpins, RNA triggers remain intact. Interestingly, unlike iSBH-sgRNAs, truncation is not dCas9-dependent (Figure 5.4.B.). Sequencing results confirmed deletions in the truncated RNA triggers that overlap with the 44nt sequence complementary with iSBH-sgRNAs (Figure 5.4.C.).

5.4 Closing remarks

In summary, the results presented in this chapter suggests the involvement of endogenous double-stranded RNA processing mechanisms when iSBH-sgRNAs are

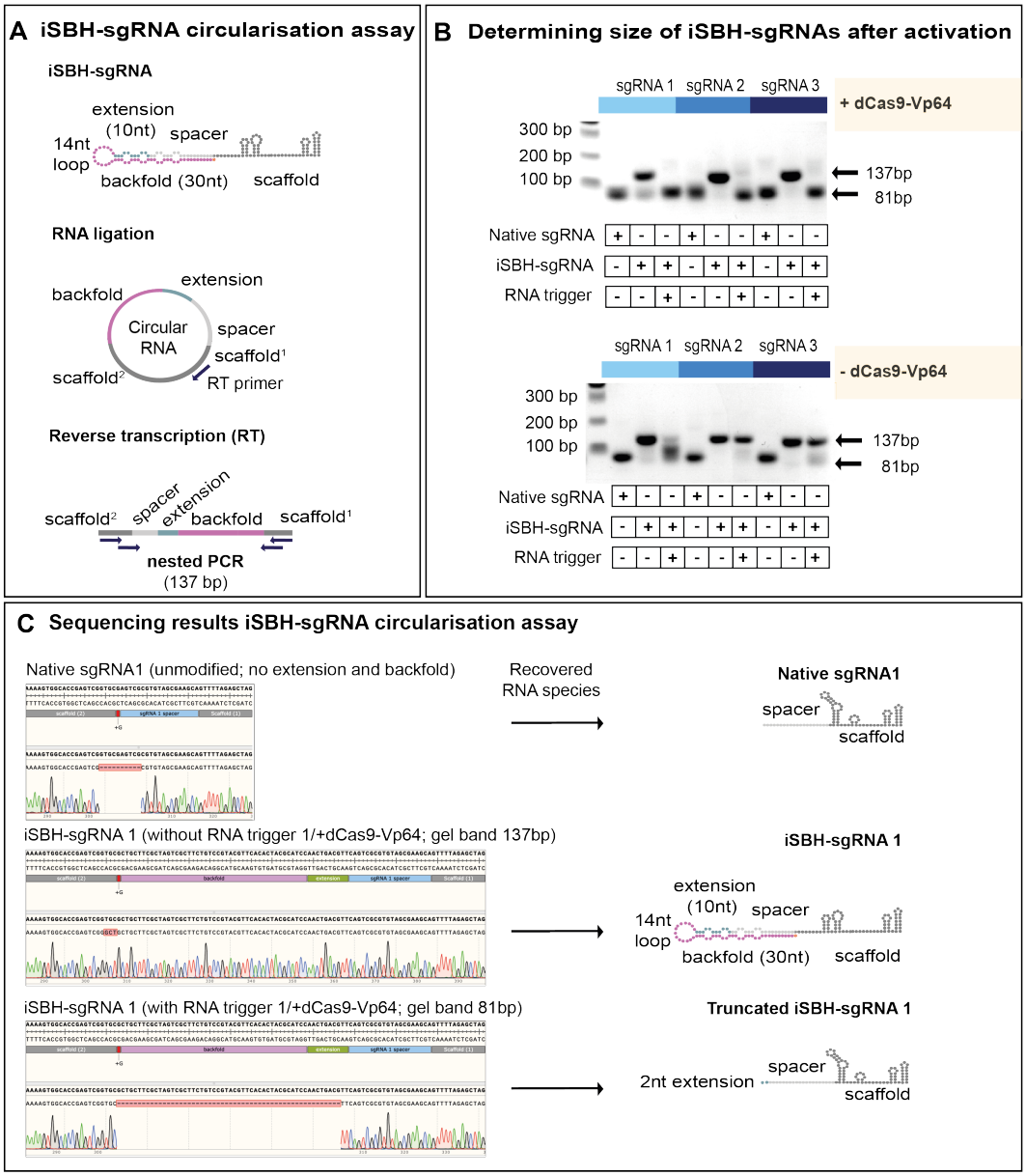


Figure 5.3: Measuring the size of iSBH-sgRNA following activation

A.iSBH-sgRNA circularisation assay. Cells were transfected with system components, followed by RNA extraction and ligation. Reverse transcription (RT) was performed on circular RNAs by using RT primers complementary with the sgRNA scaffold. The size of the RT products was determined by two sequential PCR reactions. PCR primers annealed with the scaffold^2 and scaffold^1 sequences, which are the scaffold sequences found downstream and upstream from the RT primer. For a full-length iSBH-sgRNA sequence, a second PCR product of 137bp is expected, while for a non-engineered native sgRNA, an 81nt product is expected. **B.** Determining the size of the iSBH-sgRNA after activation. Assays were performed in the presence or absence of complementary 44nt, short RNA triggers and dCas9-Vp64. Non-engineered, native sgRNA controls were also included. **C.** Sequencing results. PCR products were submitted for Sanger Sequencing. Data displays sequencing results obtained for iSBH-sgRNA 1 in the presence of dCas9-Vp64. Top sequencing trace was recovered from cells transfected with non-engineered, native sgRNA. Middle and lower sequencing traces resulted from cells transfected with iSBH-sgRNAs in the absence or presence of complementary RNA triggers.

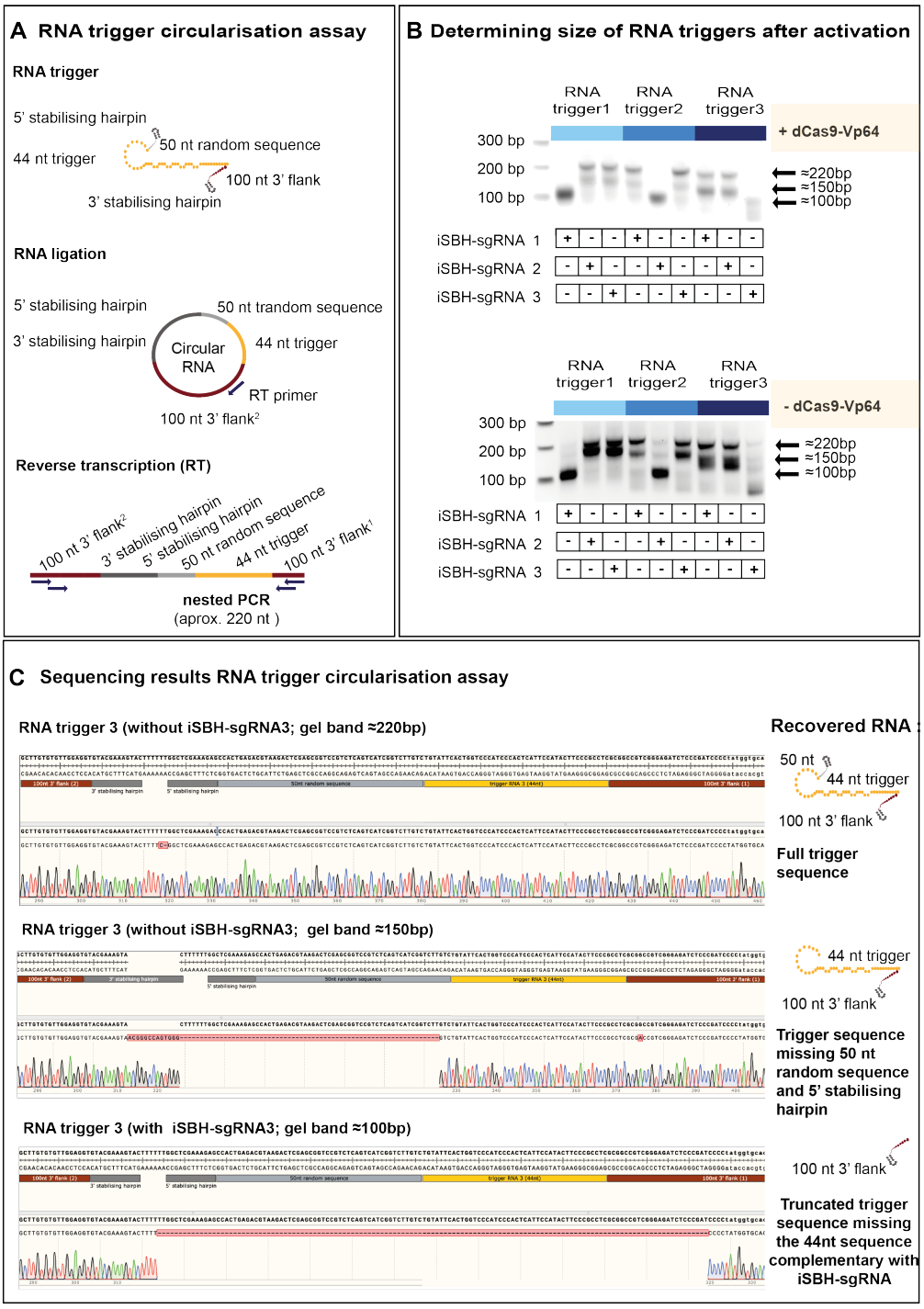


Figure 5.4: Measuring the size of RNA triggers following activation

A. RNA trigger circularisation assay. After transfection, RNA extraction and RT, RNA trigger size was determined by nested PCR. PCR primers annealed with the 100nt 3' flank^2 and 100nt 3' flank^1 sequences, which are the flank sequences downstream and upstream from the RT primer. For full-length RNA triggers, 220bp PCR bands are expected. **B.** Determining the size of the RNA triggers after activation. Assays were performed in the presence or absence of a complementary iSBH-sgRNAs and dCas9-Vp64. **C.** Sequencing results. PCR products obtained for RNA trigger 3 were submitted for Sanger Sequencing. Top and middle traces represent RNA trigger sequences recovered in the absence of complementary iSBH-sgRNA (220 and 150bp bands), while bottom trace was recovered in the presence of complementary iSBH-sgRNAs.

deployed. Nevertheless, additional mechanism(s) of activation yet to be characterised could also be involved.

RNA circularisation assays suggest that iSBH-sgRNAs are truncated to native sgRNA sequences in the presence of complementary RNA triggers as well as dCas9. This suggests that, in an OFF-state, iSBH-sgRNAs are unable to interact with dCas9. Following binding to complementary RNA triggers, iSBH-sgRNAs are able to interact with dCas9, resulting in sgRNA processing (Figure 5.5.A.). iSBH-sgRNA truncation to native sgRNA sequences is consistent with previous observations reported in studies that attempted to extend spacer sequences[41, 107] and might be important for efficient Cas9 function.

Circularisation assays also suggest that RNA triggers are truncated in the presence of complementary iSBH-sgRNA sequences. Nevertheless, the presence or absence of dCas9-Vp64 did not influence trigger truncation, suggesting that RNA-RNA interactions could be responsible for trigger truncation. A potential explanation is the presence of the iSBH-sgRNA hairpin, which have similar shapes as pri-microRNAs (Figure 5.5.B.). Thus, a subset of iSBH-sgRNAs from the cellular pool could act as a substrate for Drosha[150, 151] (Figure 5.5.C.). Drosha processing could, in turn, produce microRNA-like molecules called siRNAs (small-interfering RNAs) that subsequently lead to truncation of RNA trigger sequences. This also results in a net decrease of the concentration of RNA triggers that could activate iSBH-sgRNAs (Figure 5.5.D.).

Future engineering strategies will be required to ensure that iSBH-sgRNAs remain intact inside the cellular environment. Efforts should also be made to prevent RNA trigger truncation. Perhaps, next engineering efforts should aim to design iSBH-sgRNA sequences that resemble natural RNAs as little as possible. Chemical modifications could, therefore be the way forward in future iSBH-sgRNA designs and potential engineering strategies will be discussed in Chapter 7.

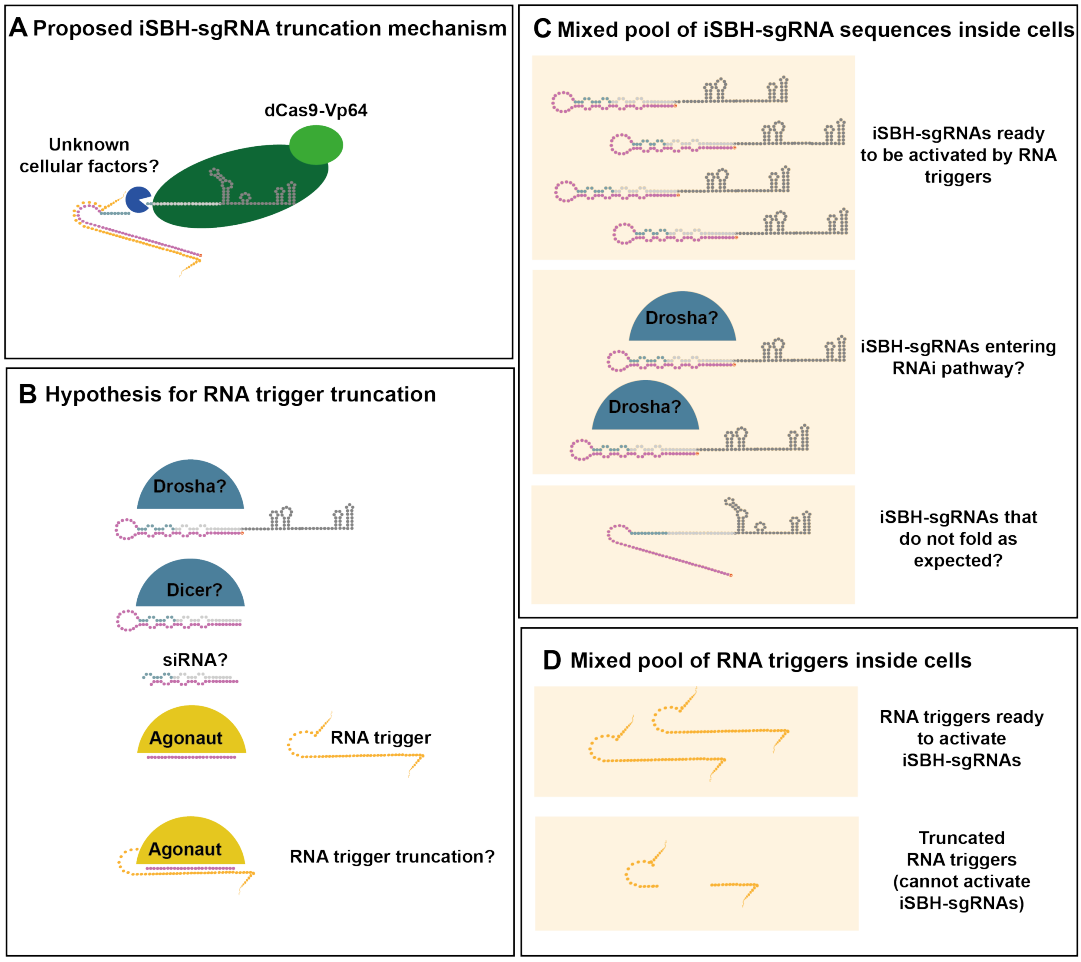


Figure 5.5: Potential mechanisms that cause iSBH-sgRNA and RNA trigger truncation

A. Proposed iSBH-sgRNA truncation mechanism. Circularisation assays suggest that iSBH-sgRNA truncation happens in the presence of complementary RNA triggers and dCas9-Vp64. Nevertheless, other cellular factors might also be involved in this process. **B.** Hypothesis for RNA trigger truncation. iSBH-sgRNAs resemble pri-microRNA precursors. A fraction of cellular iSBH-sgRNAs might therefore be processed through the RNAi pathway. This could, in turn, result in the production of siRNAs and trigger RNA truncation. **C.** Mixed pool of iSBH-sgRNA sequences inside cells. In the absence of RNA triggers, iSBH-sgRNAs could adapt 3 different conformations. A first iSBH-sgRNA fraction might adopt the right fold and be ready for activation. Another fraction might enter RNAi pathway. The third fraction might not adopt desired iSBH-sgRNA shapes and cause the the observed OFF-state background activation. **D.** Mixed pool of RNA triggers inside cells. A fraction of cellular RNA triggers might be lost due to RNA truncation. Truncated triggers are not expected to promote iSBH-sgRNA activation.

6

Testing performance of iSBH-sgRNAs *in vivo*

After showing that iSBH-sgRNAs detect RNA triggers in HEK293T cells, I next endeavoured to develop this technology towards *in vivo* RNA detection. Zebrafish is an excellent model due to their relative ease of manipulation, transparency, small size, and key roles in developmental studies. Furthermore, available RNA sequencing databases, transgenic lines as well as ease of generating new transgenics were a great resource for further development of my system[136]. Of particular interest were RNA sequencing databases that characterised neural crest-specific transcripts localised in either nuclear or cytoplasmic compartments[136]. These data sets would enable selection of tissue-specific RNA biomarkers with different properties. Starting from these biomarkers, modular iSBH-sgRNAs could be designed and tested for their ability to activate CRISPR in a tissue-specific fashion *in vivo*.

6.1 Generating transgenic zebrafish lines

6.1.1 Transgenic line design

Available transgenic lines, such as Sox10:BirA-mCherry line with fluorescently labelled neural crest cells[136] (Figure 6.1.A.) acted as a starting point for further development of transgenics reporting on the tissues where CRISPRa is active.

For this purpose, two new lines were generated, one encoding for dCas9-Vp64 transcriptional activator (Figure 6.1.B.), and the other, encoding an 8xCTS-ECFP reporter (Figure 6.1.C.). Both new lines I developed were generated in the Sox10:BirA-mCherry transgenic background to allow fluorescent tracking of neural crest cells in which putative endogenous RNA triggers were enriched.

Hence, the transgenic zebrafish lines were compatible with further proof-of-concept experiments testing if iSBH-sgRNAs could detect transcripts specifically upregulated in the neural crest. In future experiments, the two lines would be crossed and at least 25% of resulting embryos would carry both 8xCTS-ECFP and dCas9-VP64 alleles. These embryos will be injected with iSBH-sgRNAs. The MCP-p65-HSF1 SAM effector component[50] (Figure 3.2.D,G.) could also be delivered by injection, in order to boost CRISPRa activity. ECFP signals would enable monitoring CRISPRa activity across different tissues (Figure 6.2.A.). Therefore, if iSBH-sgRNAs detect neural crest-specific transcripts as endogenous triggers, zebrafish embryos would co-express mCherry and ECFP in the neural crest cells (Figure 6.2.B.).

6.1.2 Establishment of transgenic lines

Embryonic injections and growing injected embryos to adulthood; Transgenic lines were generated by injecting plasmid vectors together with Tol2 mRNA into incrossed Sox10:BirA-mCherry embryos (Figure 6.1.B.). Embryos resulting from these injections constituted the founding generation (F0) of transgenics, characterised by mosaic transgene integration across tissues. F0 fish were grown up to 4 months until they reached reproduction age.

Establishing founders; For the dCas9-Vp64 line, 12 F0 fish reached adulthood (Figure 6.3.A.). In order to identify which of these fish have successful germline transmission of the Vp64 transgene, a series of crossing experiments were performed, followed by genotyping of resulting embryos (Figure 6.3.B-D.)

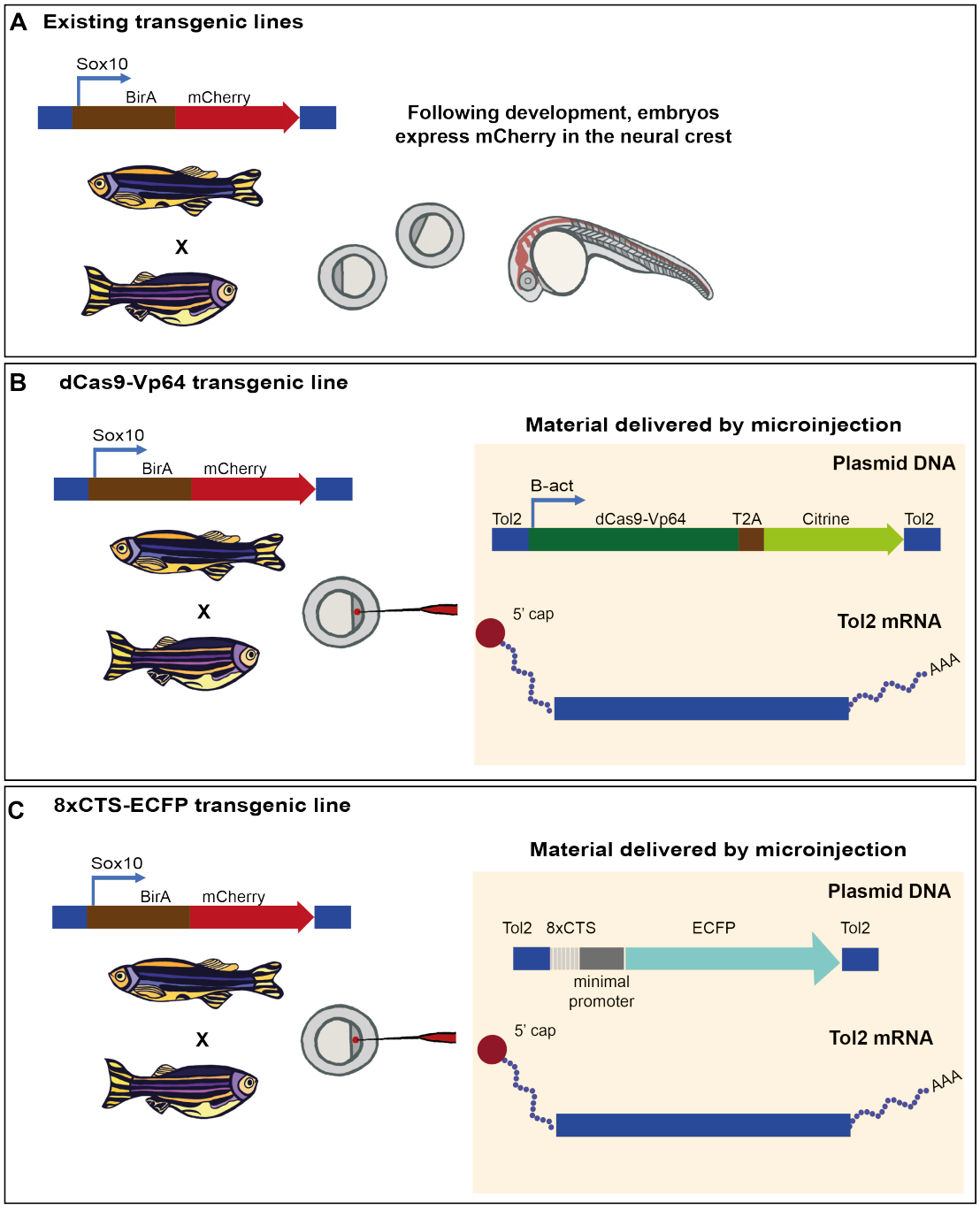


Figure 6.1: Generating zebrafish transgenic line design

A. Existing transgenic lines. Sox10:BirA-mCherry line[136] enables fluorescent labelling of neural crest cells. **B.** dCas9-Vp64 transgenic line. dCas9-Vp64 expression cassette was placed under the control of a ubiquitous, β-actin promoter. Construct was flanked by Tol2 arms, enabling for transposon-mediated integration into the zebrafish genome[152]. **C.** 8xCTS-ECFP transgenic line. 8xCTS-ECFP construct contained 8 CRISPR-target sequences compatible with one of the best performing sgRNAs tested in HEK293T cells. Construct was flanked by Tol2 arms, enabling for transposon-mediated integration into the zebrafish genome[152]. *I would like to thank Dr. Martyna Lukoseviciute and Dr. Filipa Simoes for performing the initial injections for the creation of dCas9-Vp64 and 8xCTS-ECFP transgenic lines.

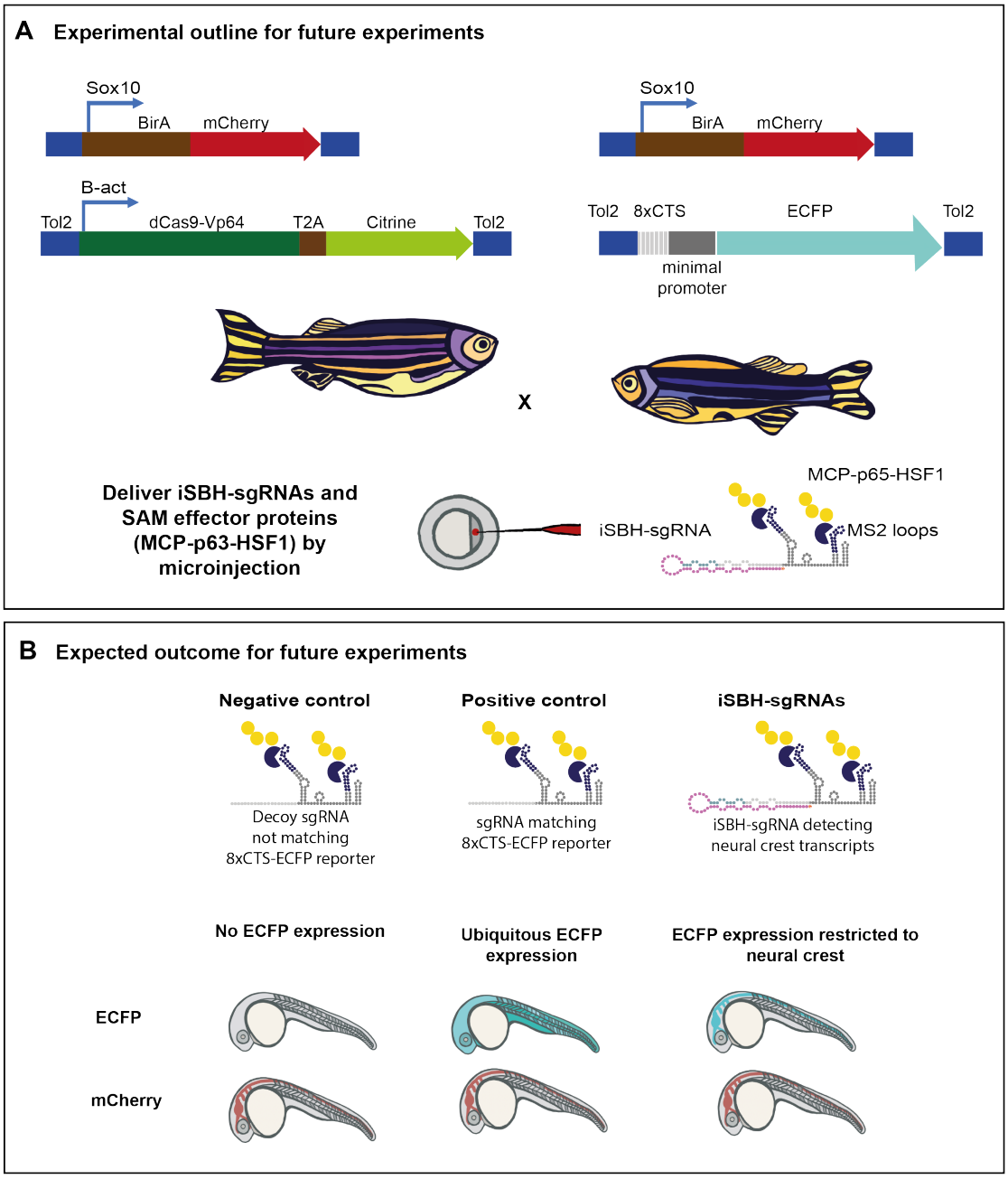


Figure 6.2: Vision for future *in vivo* RNA detection experiments

A. Experimental outline for future experiments. Future experiments would involve crossing dCas9-Vp64 and 8xCTS-ECFP transgenic lines. iSBH-sgRNAs and SAM effector molecules (MCP-p65-HSF1 [50]) would be delivered to resulting embryos through microinjection. **B.** Expected outcome for future experiments. For embryos injected with sgRNAs not-matching 8xCTS-ECFP reporter, no ECFP expression is expected. Embryos injected with matching sgRNAs (without the iSBH-sgRNA fold) are expected to express ECFP across all tissues. In an example experimental scenario, iSBH-sgRNAs will be designed starting from neural crest-specific RNA biomarkers. Embryos injected with these iSBH-sgRNAs are expected to express ECFP only in the neural crest. Thus, neural crest cells would then be expected to co-express mCherry and ECFP.

After identifying a founder, the next step was to **establish the first (F1) generation** of fish encoding dCas9-Vp64 in all their cells - both somatic and in the germline. F1 generation was established by crossing founders with Sox10:BirA-mCherry fish. Progeny resulting from these crosses were screened for the presence Sox10:BirA-mCherry as well as dCas9-Vp64 transgene (Figure 6.4).

Several rounds of crosses and genotyping experiments were carried out until 45 embryos double-positive for Sox10:BirA-mCherry and dCas9-Vp64 transgenes were identified. These embryos were grown for 4 more months until they reached adulthood and used in subsequent experiments (Figure 6.4.D.). A similar strategy was also adopted for generating the 8xCTS-ECFP transgenic line.

6.1.3 Validating transgenic lines

When F1 generation reached adulthood, the next step was to functionally validate the generated transgenic lines. This was carried out by crossing dCas9-Vp64 and 8xCTS-ECFP lines and injecting the resulting embryos with different sgRNA sequences (Figure 6.5.A.). In a negative control, embryos were injected with decoy sgRNAs, non-complementary with the 8xCTS-ECFP reporter. In a positive control, native sgRNAs (without the iSBH fold) complementary with the 8xCTS-ECFP reporter were injected. Negative control embryos were expected to display no ECFP fluorescence, while positive control embryos were expected to express ECFP (Figure 6.5.B.). In the first validation experiment, sgRNAs were cloned into zebrafish mini-vectors compatible with genomic integration using the Ac/Ds-transposition system[135]. Citrine expression from one of the Ac/Ds components was used as a a readout for injection quality (Figure 6.5.C.).

In the negative control, 9 injected fish were selected based on Citrine expression and imaged by fluorescence microscopy. None of these embryos expressed ECFP. In

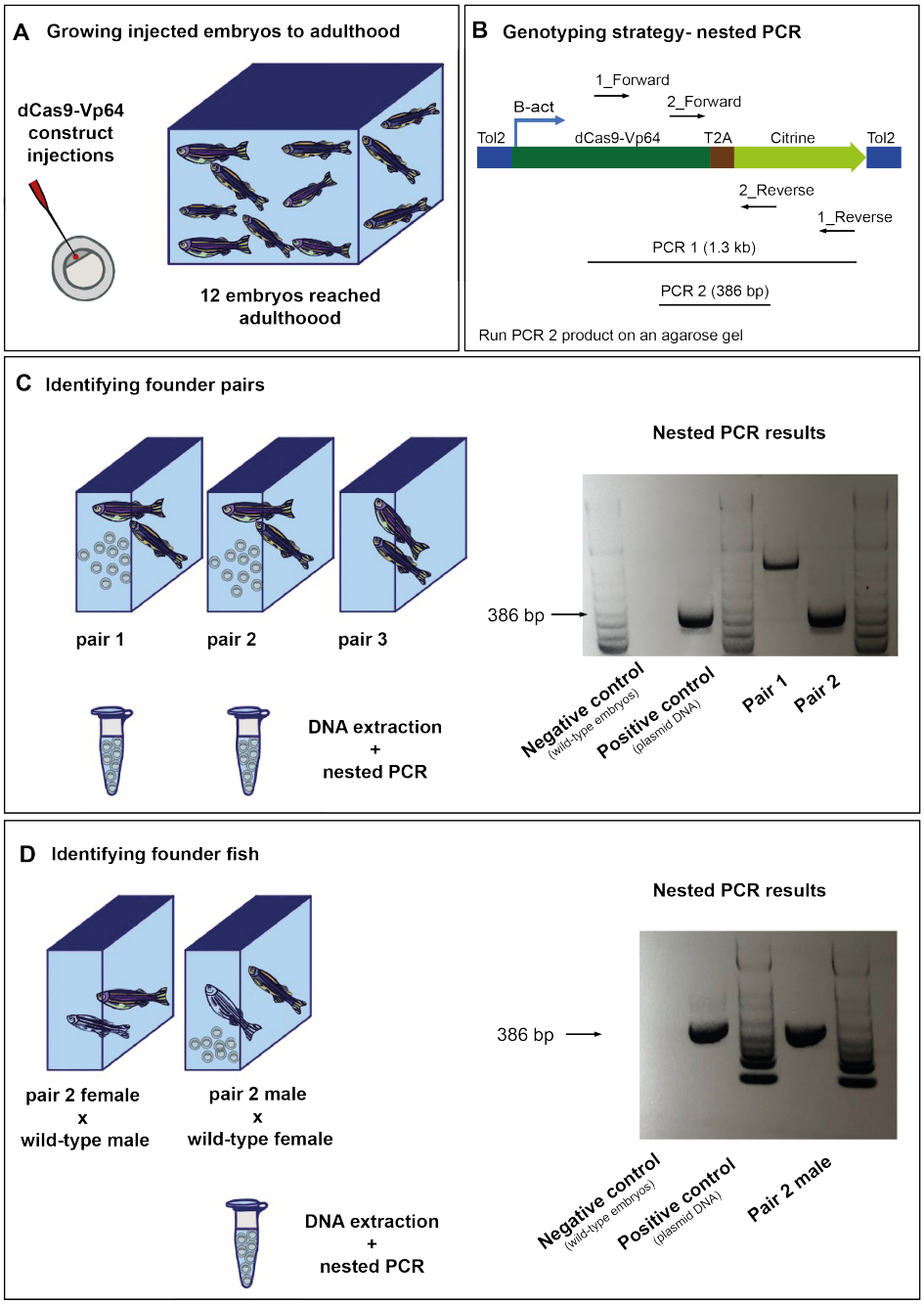


Figure 6.3: Identifying founders for the dCas9-Vp64 transgenic line

A. Growing injected embryos to adulthood. For dCas9-Vp64 initial injections (Figure 6.1.B.), 12 fish reached adulthood. **B.** Genotyping strategy- nested PCR. Transgene presence was confirmed through two PCR rounds. Products resulting from PCR1 were used as a template for PCR2. **C.** Identifying founder pairs. In a first cross, 3 F0 females were crossed with 3 F0 males for narrowing down potential founders. From these 3 pairs, only pairs 1 and 2 successfully spawned. Each cross resulted in more than 50 viable embryos that were collected and subjected to DNA extraction. PCR results indicated that a fish from the second pair is a founder. For pair 1, PCR products did not have the right size. **D.** Identifying founder fish. To identify whether the male or the female from pair 2 is a founder, the two fish were outcrossed with wild-type lines. PCR results indicated that pair 2 male as the founder for dCas9-Vp64 line. *I would like to thank Sarah Mayes for setting up all fish crosses displayed in this figure.

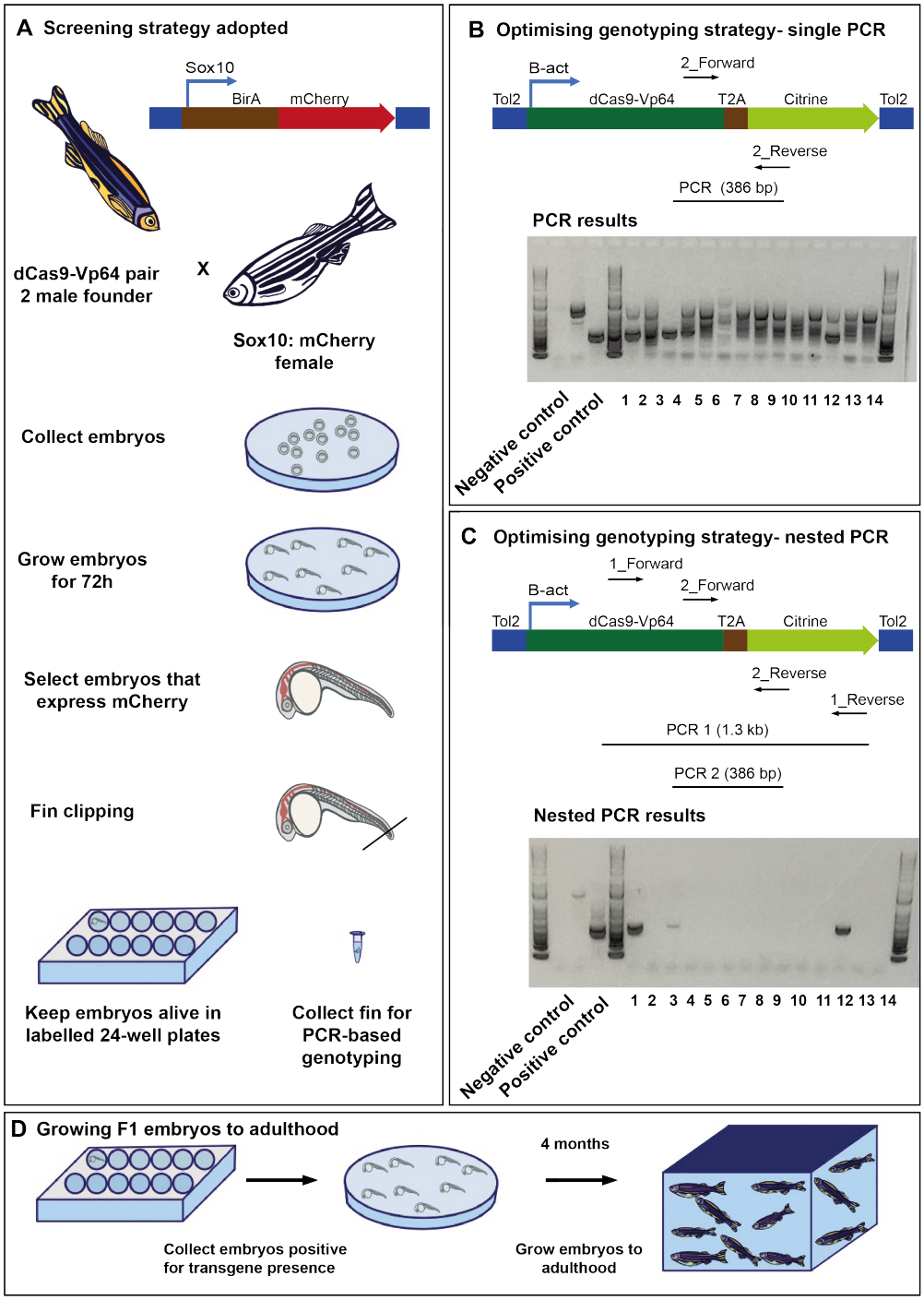


Figure 6.4: Establishing the F1 generation of transgenics

A. Screening strategy adopted. F1 generation was established by crossing dCas9-Vp64 founders with Sox10:BirA-mCherry fish. Embryos were grown up to 3 days post-fertilisation, followed by selection of mCherry positive fish by microscopy. DNA samples were collected from individual embryos using fin clipping. Presence of the dCas9-Vp64 transgene in individual embryos was confirmed by PCR. Initial PCR optimisation was required, while nested PCR-based genotyping resulted in cleaner results (Figure 6.4.B., Figure 6.4.C.). **B.** Optimising genotyping strategy- single PCR. **C.** Optimising genotyping strategy- nested PCR. **D.** Growing F1 embryos to adulthood. Embryos positive for both Sox10: BirA-mCherry and dCas9-Vp64 transgene were grown to adulthood and used in subsequent experiments. *I would like to thank Sarah Mayes for setting up all fish crosses displayed in this figure.

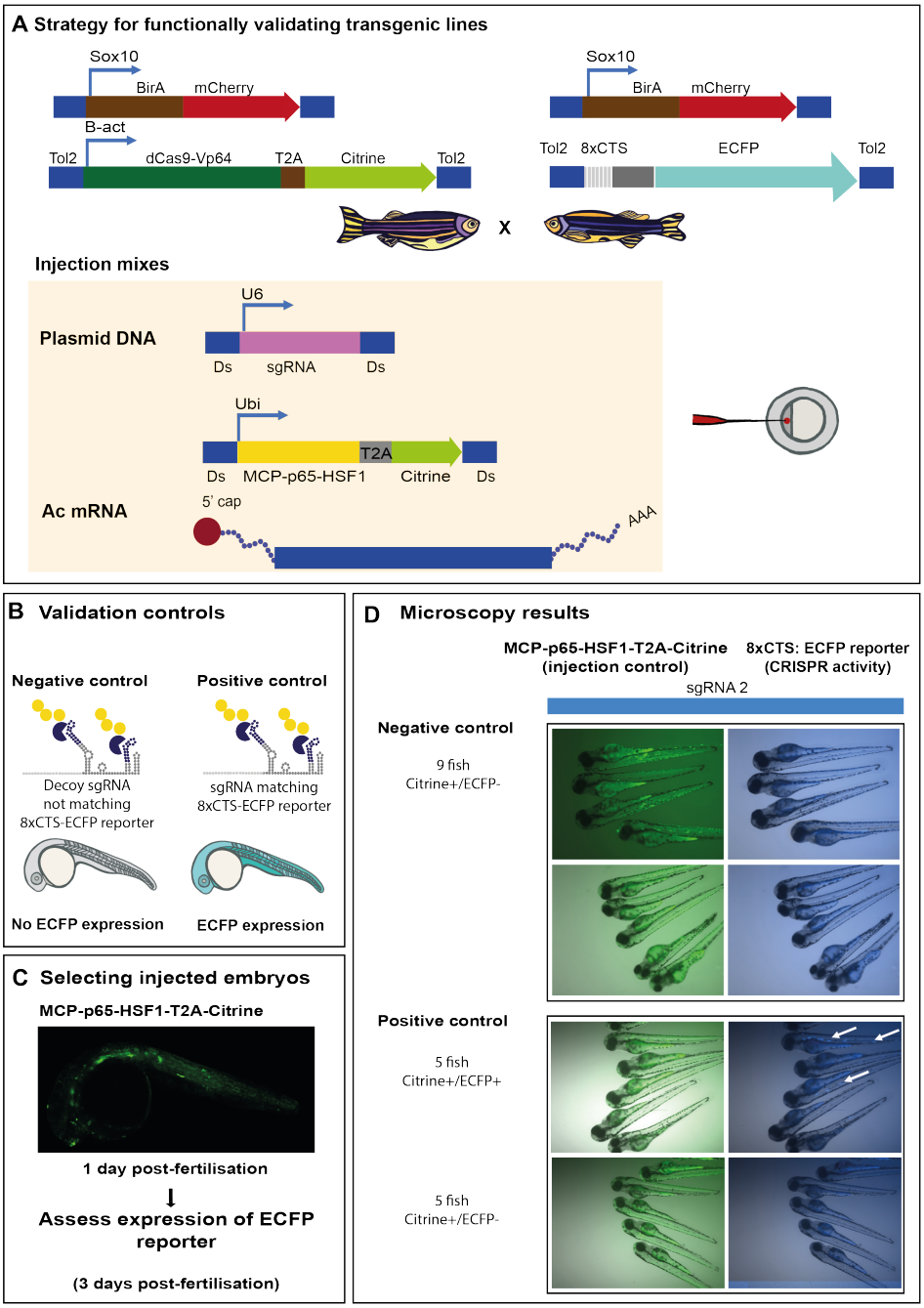


Figure 6.5: Validating transgenic lines

A. Strategy for functionality validating transgenic lines. Embryos resulted from crossing dCas9-Vp64 and 8xCTS-ECFP lines were injected with control sgRNA sequences. sgRNAs were expressed under the control of a fish U6 promoter and contained the modified sgRNA scaffold compatible with the SAM amplification system. SAM effector proteins (MCP-p63-HSF1)[141] were encoded into a second Ac/Ds vector, together with a Citrine fluorescent protein. This construct was expressed under the control of the ubiquitin promoter. The two vectors were co-injected with Ac transposase mRNA[135]. **B.** Validation controls. Embryos were injected with sgRNAs matching or not matching the 8xCTS-ECFP reporter. **C.** Selecting injected embryos. At 1-day post-fertilisation, embryos that were successfully injected displayed mosaic Citrine expression. **D.** Microscopy results. Representative examples for the two injections are displayed in this figure. Fish were imaged at 3 days post-fertilisation and a few examples of ECFP-positive cells are marked using arrows.

the positive control, 10 injected fish were selected and 5 fish (50%) expressed ECFP in a limited number of cells. This reflects the segregation of both dCas9-Vp64 and 8xCTS reporter alleles into approximately 50% offspring produced from the crosses, as one parent contained a single dCas9-Vp64 allele integration and the other parent contained multiple integrations of 8xCTS-ECFP allele (Figure 6.5.D.).

6.2 Testing if iSBH-sgRNAs detect synthetic RNA triggers *in vivo*

After showing that transgenic lines express ECFP in the presence of non-engineered, native sgRNAs, I proceeded with testing the iSBH-sgRNA system (Figure 6.6). To assess if factors required for iSBH-sgRNA activation are also present in fish, I performed injections with second-generation iSBH-sgRNA 2, one of the designs previously validated in HEK293T cells.

iSBH-sgRNAs and their RNA triggers as well as SAM effectors (MCP-p63-HSF1) were injected into zebrafish embryos using the Ac/Ds integration system as described in the previous section (Figure 6.6.A). In the absence of RNA triggers, embryos are expected to be ECFP negative, while in the presence of complementary RNA triggers, embryos should express ECFP (Figure 6.6.B).

Indeed, injected embryos observed using fluorescence microscopy demonstrated the expected ECFP readouts, suggesting that iSBH-sgRNA system is functional *in vivo* (Figure 6.7). However, ECFP expression was noticeable only in a very small number of embryos. In contrast with the native sgRNA (without iSBH fold) positive control for which ECFP expression was present in 50% of the fish, ECFP expression was only present in about 10% fish for iSBH-sgRNA ON condition. Furthermore, ECFP expression was restricted to a very small number of cells.

Data presented here suggests that the iSBH-sgRNA system cannot be utilised

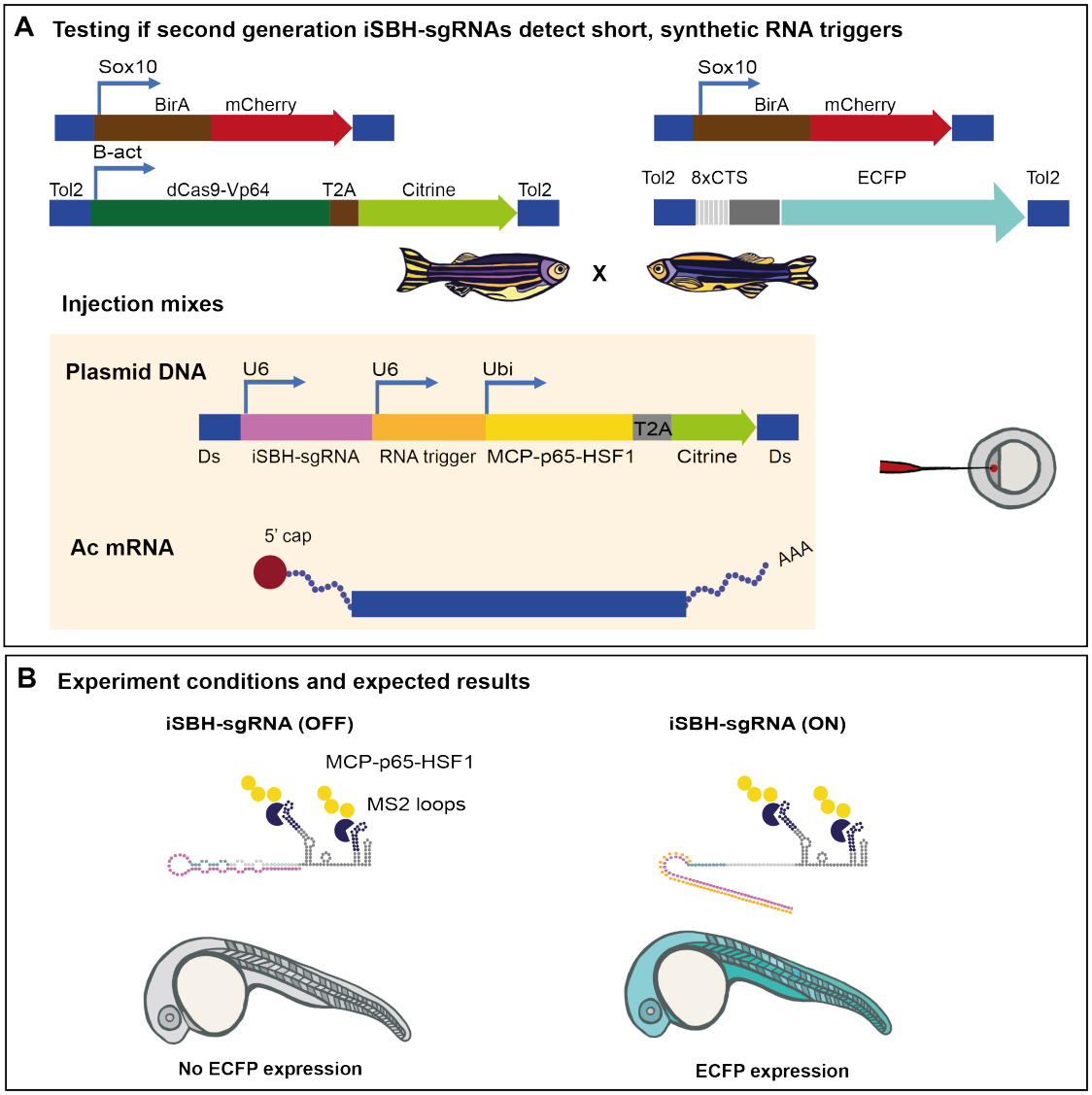


Figure 6.6: Strategy for testing if iSBH-sgRNAs detect synthetic RNA triggers *in vivo*

A.Testing if second-generation iSBH-sgRNAs detect short, synthetic RNA triggers. iSBH-sgRNAs, RNA triggers as well as SAM effectors (MCP-p63-HSF1) were injected into zebrafish embryos using Ac/Ds integration system[141]. Single-injection vectors ensured delivery of all system componentes to individual cells. Citrine labelled tissues where SAM effectors are expressed, enabling selection of successfully injected embryos. **B.** Experiment conditions and expected results. iSBH-sgRNAs were injected in the absence or presence of complementary RNA triggers. Presence of RNA triggers is expected to turn on ECFP production.

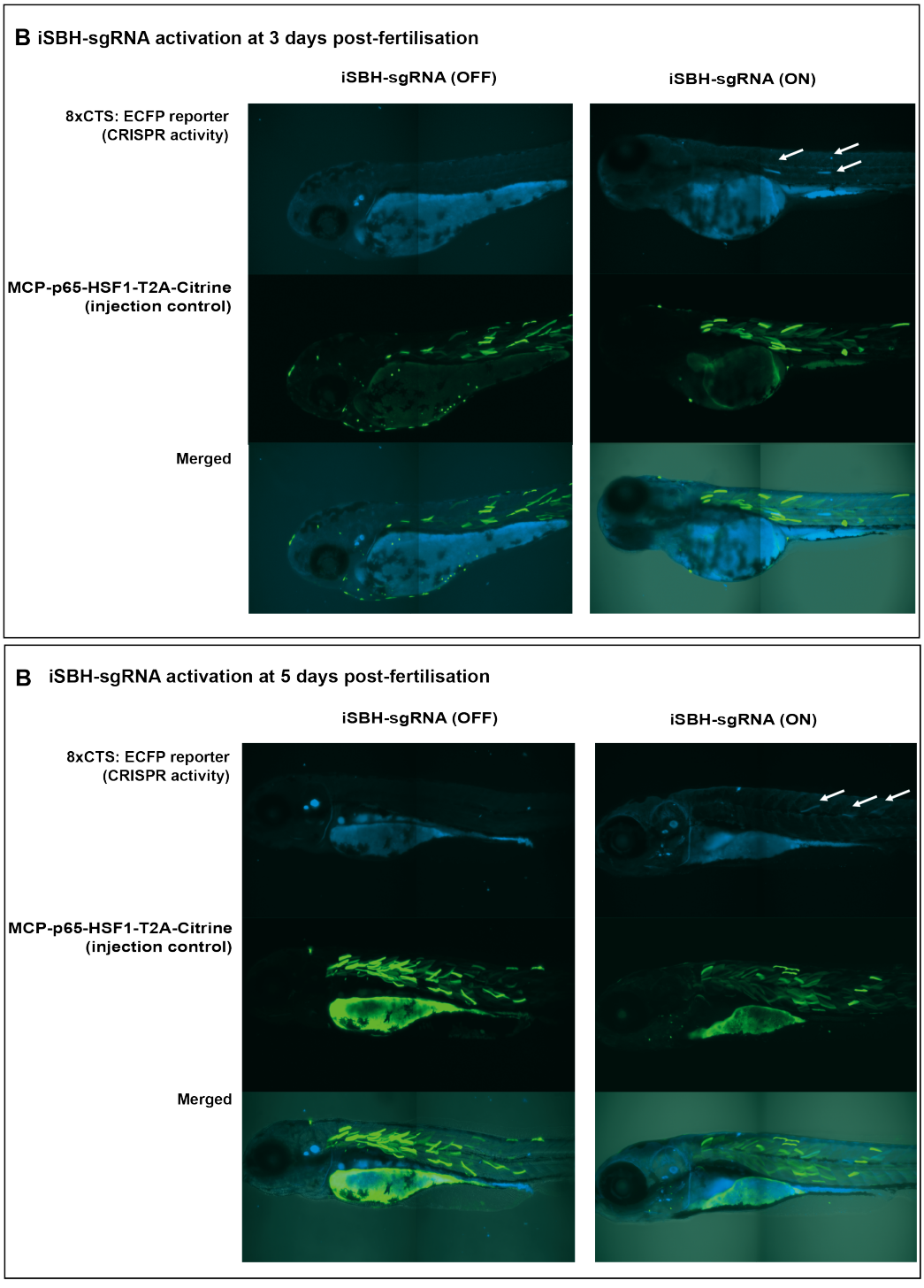


Figure 6.7: Testing if iSBH-sgRNAs detect synthetic RNA triggers *in vivo*

A. iSBH-sgRNA activation at 3 days post-fertilisation. Images display embryos injected with iSBH-sgRNAs in the absence and presence of complementary RNA triggers. Fish imaged for the iSBH-sgRNA (OFF) condition were genotyped to confirm dCas9-Vp64 and 8xCTS-ECFP presence. Arrows indicate ECFP positive cells. **B.** iSBH-sgRNA activation at 5 days post-fertilisation. Images display embryos injected with iSBH-sgRNAs in the absence and presence of complementary RNA triggers. Fish imaged for the iSBH-sgRNA (OFF) condition were genotyped to confirm dCas9-Vp64 and 8xCTS-ECFP presence. Arrows indicate ECFP positive cells.*I would like to thank Prof. Tatjana Sauka-Spengler for helping me image the iSBH-sgRNA (ON) condition at 5 days post-fertilisation.

under current conditions for endogenous RNA detection. An intermediary step would involve optimising sgRNA delivery in fish and ensuring the CRISPRa reporters are evenly activated across tissues.

6.3 Troubleshooting sgRNA delivery

In order to improve CRISPRa activity at a single-cell level, I focused on optimising delivery of positive control native sgRNA (without iSBH fold). Closer examination of embryos injected with a single Ac/Ds vector encoding for MCP-p63-HSF1 (with Citrine fluorophore) as well as the positive control sgRNA (Figure 6.8.A.) revealed sporadic expression of construct components in a non-homogeneous fashion.

This phenomenon could result from a combination of issues, for example, mosaic Ac/Ds-mediated construct integration as well as reduced integration efficiency due to large construct size. Additional causes might involve integration in heterochromatic regions or poor promoter activity. Silencing of transposable elements during development is also a well-documented phenomenon[153, 154] and that might also interfere with efficient sgRNA expression.

Furthermore, within the zebrafish community, DNA based delivery has a well-known limitation of uneven transgene expression. RNA based delivery leads to more homogeneous expression across tissues [155] and is favoured where experimental conditions allow. Given this, I decided to explore a new strategy and deliver chemically-synthesised sgRNAs to embryos.

In the absence of Cas9, sgRNAs are highly unstable, having half-lives of approximately 15 minutes[156]. To be able to detect RNA triggers expressed at later developmental stages, more stable sgRNAs with longer half-lives would be essential. Therefore, I tested IDT's proprietary sgRNA-XT chemical modification, reported to have increased resistance against nucleases (Figure 6.9.A.)[60]. Injecting these

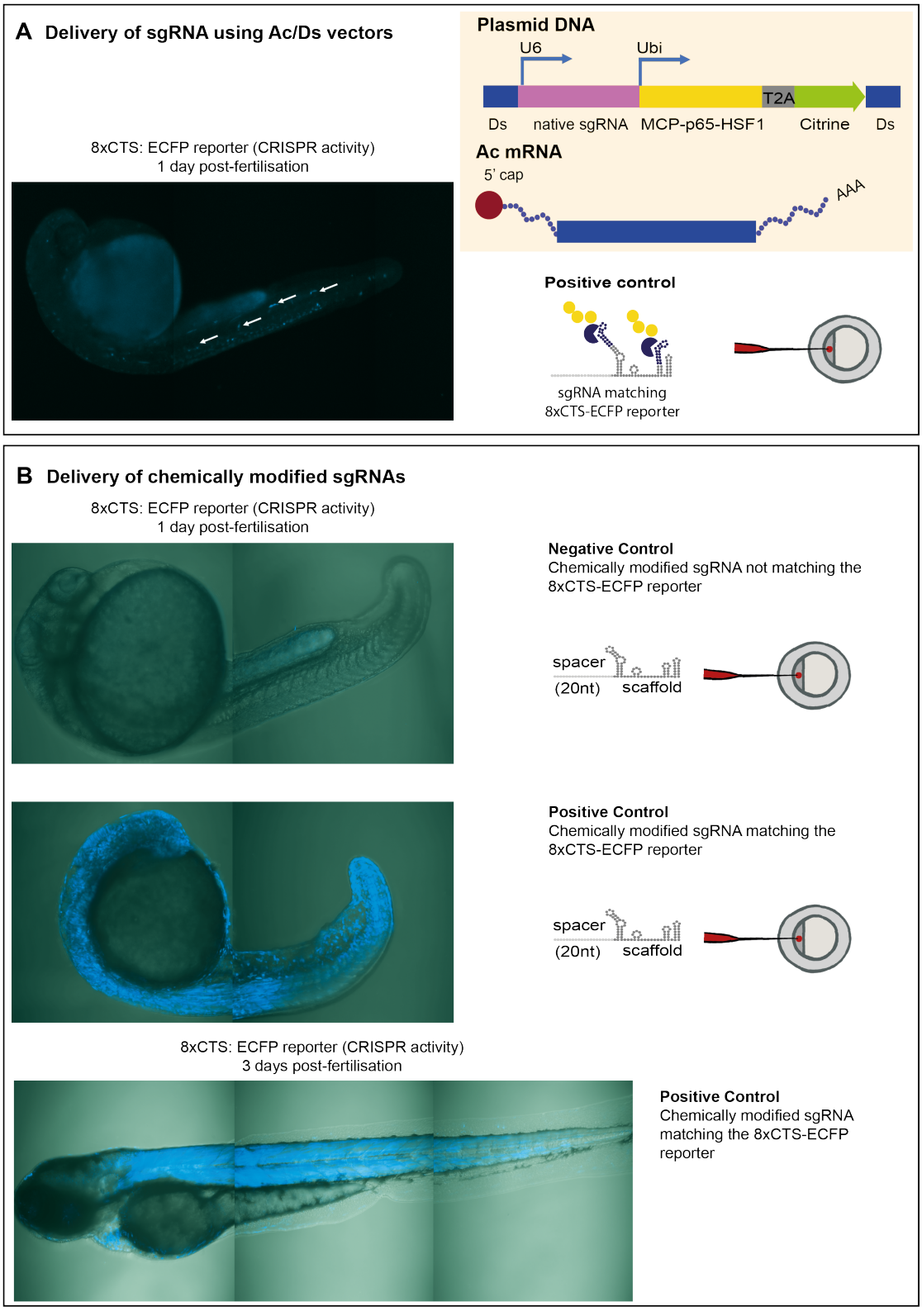


Figure 6.8: Troubleshooting sgRNA expression across tissues

A. Delivery of sgRNA using Ac/Ds vector. Positive controls injected with native sgRNA without the iSBH fold are displayed here. Arrows point towards a few ECFP positive cells. **B.** Delivery of chemically modified sgRNAs. Embryos were injected with negative and positive control sgRNAs. For positive control, images are taken at 1 day as well as 3 days post-fertilisation. *I would like to thank Sarah Mayes for performing some of these injections and imaging experiments together with me.

chemically modified sgRNAs into embryos resulted in significantly more robust activation of the 8xCTS-ECFP reporter (Figure 6.8.B) with persistent ECFP expression lasting several days post-injection.

6.4 Testing a first version of iSBH-sgRNAs with chemical modifications *in vivo*

Due to improved positive control sgRNA activity in the presence of chemical modifications, I next tested chemically modified iSBH-sgRNAs. A first iteration of chemically-modified iSBH-sgRNAs were designed (Figure 6.9.B.) with chemical modifications employed towards protecting the 5'end of the iSBH-sgRNA as well as the sgRNA scaffold. However, when co-injected with chemically synthesised RNA triggers, these modifications did not lead to activation of the 8xCTS-ECFP reporter (Figure 6.9.E., Figure 6.9.F.). Further work is therefore required to investigate positions at which chemical modifications should be placed in order to promote iSBH-sgRNA activity *in vivo*.

Available mechanistic data suggest that iSBH-sgRNAs get truncated during activation (Figure 5.5). In a first iteration of chemically modified iSBH-sgRNAs, truncation would lead to production of sgRNAs with unprotected spacer sequences. In this scenario, 5'-3' sgRNA degradation is highly likely to occur. Alternatively, iSBH-sgRNA hairpins might be recognised and processed by RNA interference factors such as Drosha[157]. Potential differences in Drosha levels or affinity for iSBH-sgRNAs in HEK293T cells vs zebrafish embryos, might also lead to the inefficient iSBH-sgRNA activity observed in embryos.

Therefore, future iterations of chemical modification strategies will be employed to optimise function of iSBH-sgRNA systems *in vivo*. These modifications would involve protecting the spacer sequence from degradation, reducing interaction with cellular nucleases and increasing iSBH-sgRNA binding affinities for their RNA

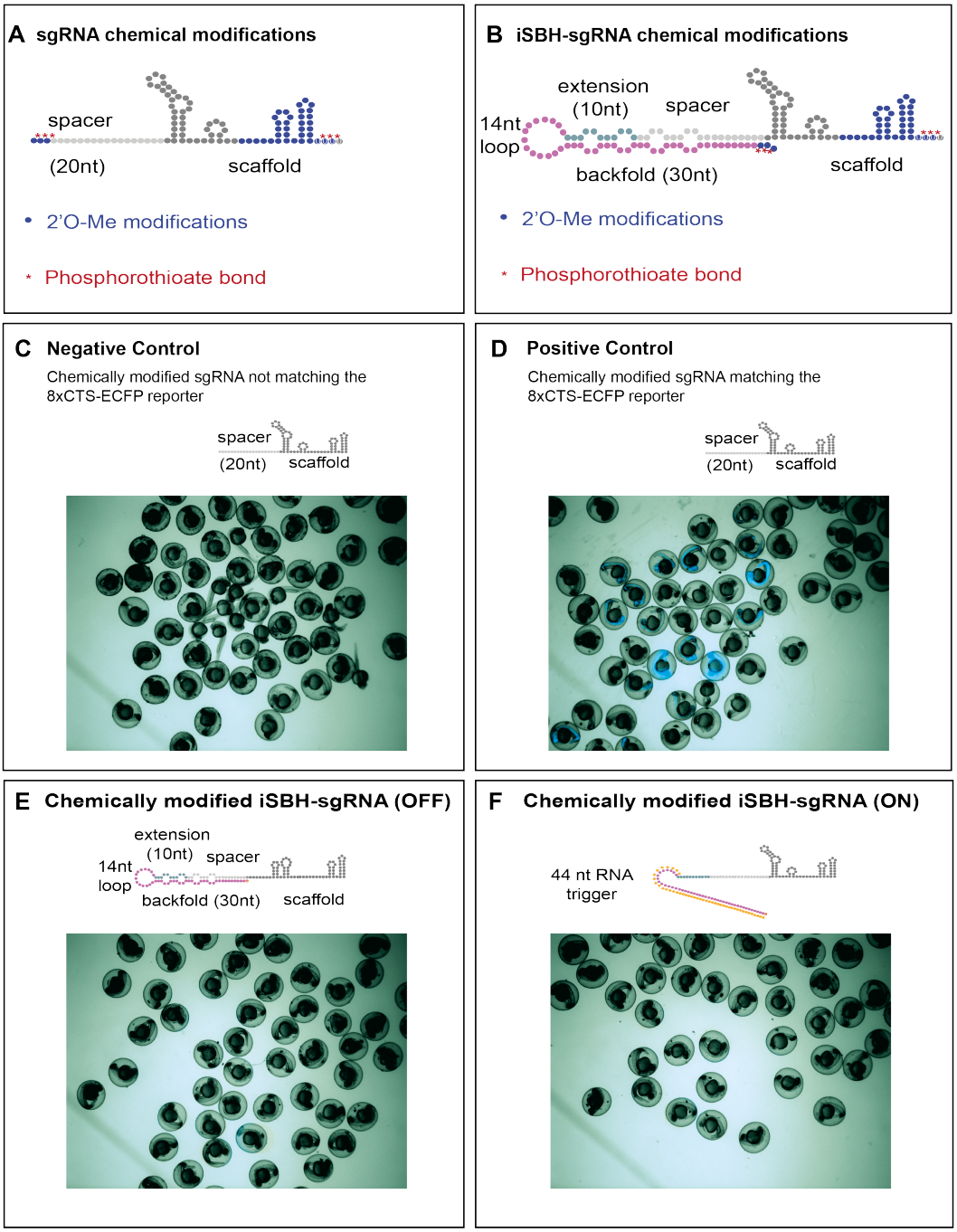


Figure 6.9: Testing a first version of iSBH-sgRNA with chemical modifications

A. sgRNA chemical modifications. Chemical modifications for the IDT proprietary sgRNA XT involve 2'O-methyl (Me) modifications as well as phosphorothioate (PS) bonds [60]. **B.** iSBH-sgRNA chemical modifications. A first attempt to insert chemical modifications within the iSBH-sgRNA structure involved protecting the 5' end of the backfold as well as the scaffold sequence. **C.** Negative control. 1 day post-fertilisation embryos injected with native sgRNAs that do not match the CRISPRa reporter. **D.** Positive control. Embryos injected with native sgRNAs matching the CRISPRa reporter. **E.** Chemically modified iSBH-sgRNAs (OFF). Embryos injected with iSBH-sgRNAs in the absence of complementary RNA triggers. **F.** Chemically modified iSBH-sgRNAs (ON). Embryos injected with iSBH-sgRNAs in the presence of complementary RNA triggers. *I would like to thank Sarah Mayes for performing these injections together with me. Her help enabled injection of multiple sgRNA conditions in the same day.

triggers. A full description of potential chemical modifications that would be explored is discussed in Chapter 7.

6.5 Closing remarks

Results presented in this chapter suggest that iSBH-sgRNA systems could, in theory, be repurposed towards RNA detection *in vivo*. Expression of iSBH-sgRNAs from DNA expression cassettes is limited by inefficient iSBH-sgRNA activation in the presence of complementary RNA triggers as evidenced by detectable CRISPRa activity in a small number of embryos and a limited number of activated cells. Embryos injected with positive control sgRNAs (without the iSBH-sgRNA fold) also displayed mosaic CRISPRa activation.

Delivery issues were resolved by injecting chemically modified native sgRNAs with improved resistance against nucleases[60]. Although initial *in vivo* trials utilising iSBH-sgRNAs with first version of chemical modification were unsuccessful, injection of chemically modified iSBH-sgRNAs remains the most viable option subject to future work establishing improved versions of modification. After optimising detection of synthetic RNA triggers across tissues, future screening experiments could be carried out for identifying the classes of endogenous RNAs that can act as triggers *in vivo*.

7

Discussion and Further Directions

7.1 Results summary

As a follow up to the original study carried out by Ferry et al.[37], this study shows that iSBH-sgRNAs can be repurposed towards RNA detection in mammalian cells. Limited evidence is also provided that iSBH-sgRNAs could be adopted for *in vivo* RNA detection, in transgenic zebrafish embryos.

In HEK293T cells, first-generation iSBH-sgRNAs successfully detect short RNA triggers expressed from U6 promoters. Nevertheless, these designs are not efficient at detecting longer RNA triggers. In contrast with first-generation designs, second-generation iSBH-sgRNAs present an improved ability to detect longer RNA triggers. This was achieved by extending the length of the iSBH-sgRNA backfold complementary with the RNA trigger. A subsequent design improvement represented modular iSBH-sgRNA designs. Modular designs enable spatial-separation of two iSBH-sgRNA modules: sequence required for RNA trigger detection as well as the sequence required for targeting Cas9 to desired DNA target sequences. Modular designs also incorporate larger loop sizes that substantially improve iSBH-sgRNA ability to detect longer RNA triggers.

Experiments carried out *in vivo* also suggest that iSBH-sgRNA designs could detect RNA triggers. Two zebrafish transgenic lines encoding for dCas9-Vp64 as well as 8xCTS-ECFP reporters were generated for monitoring CRISPRa activity using fluorescence. When injecting iSBH-sgRNAs as well as RNA triggers into transgenic embryos, fluorescence is observed in a small number of cells. This suggests that the endogenous machinery required for iSBH-sgRNA activation might also be present in fish. Nevertheless, before proceeding with understanding the classes of endogenous RNAs that could act as good triggers *in vivo*, further work is required towards optimising iSBH-sgRNA delivery and activation at a single-cell level. Injecting positive control sgRNAs without the iSBH fold suggested that DNA-based delivery of sgRNA does not lead to even CRISPRa activity across tissues. Injection of highly modified sgRNAs[158] leads to improved outcomes. Transgenic embryos injected with chemically modified sgRNAs display CRISPRa activity across most tissues. This shows that chemical modifications are a feasible path forward for optimising iSBH-sgRNA activity *in vivo*. Nevertheless, further work is necessary to decipher positions at which chemical modifications are tolerated.

This study also presents findings and developments in characterising iSBH-sgRNA activation mechanism, understanding how CRISPR activators and reporters affect activation, building the MODesign pipeline as well as preliminary results for understanding the range of RNA species that could act as triggers in cellular systems.

iSBH-sgRNA activation mechanism was characterised through RNA circularisation assays[97]. Data suggests that iSBH-sgRNAs are truncated in the presence of RNA triggers and dCas9-Vp64. A potential explanation is that, in the OFF-state, iSBH-sgRNAs are not able to bind to dCas9-Vp64. When RNA triggers are present, a conformational change enables iSBH-sgRNAs to bind to dCas9-Vp64, leading to iSBH-sgRNA truncation. This hypothesis is also supported by previous studies reporting that engineered sgRNAs with extended spacer sequences are processed

to original 20nt spacer sequences[41, 107]. RNA triggers are also truncated in the presence of complementary iSBH-sgRNA sequences. I hypothesise that factors from the RNA interference[146, 159] pathway might be responsible for this observation. A small subset of iSBH-sgRNA sequences might be recognised by Drosha and Dicer and processed into siRNAs. In turn, siRNAs interact with Argonaut proteins and cleave complementary RNA triggers.

MODesign is a computational algorithm I developed to enable users to design iSBH-sgRNAs for their custom RNA detection and genome editing applications. Users need to specify the sequence of the RNA trigger they would like to detect, the DNA sequence where they would like Cas9 to bind as well as their desired length for iSBH-sgRNA loops. MODesign will suggest a panel of iSBH-sgRNAs sequences that could be used for that particular application.

CRISPR activators of choice as well as the number of Cas9 binding sites in the fluorescent reporters used could significantly impact ON/OFF ratios behind iSBH-sgRNA activation. In the cellular pool of iSBH-sgRNAs, molecules that do not adopt desired secondary structures contribute to baseline OFF -state leakiness. Using stronger activators such as dCas9-VPR propagates this noise. Nevertheless, weaker activators such as dCas9-Vp64, which require concomitant binding of several effectors at the promotor to drive downstream gene expression, mask this noise.

Different varieties of RNA triggers with varying sizes and properties were cloned in mammalian expression vectors. Vectors were co-transfected with complementary iSBH-sgRNA in HEK293T cells in equimolar amounts. RNA triggers of up to 300 nucleotides expressed from U6 promoters were efficiently detected. Detection efficiency decreased with increasing RNA trigger length. Biology behind RNA trigger expression and cellular localisation impacts iSBH-sgRNA activation. Expressing RNA triggers from CMV polymerase II promoters significantly reduces iSBH-sgRNA

activation. Furthermore, data suggests that, in their current developments, iSBH-sgRNAs cannot detect messenger RNA. This suggests that additional increase in RNA trigger size, changes in RNA trigger expression levels as well as the biology behind RNA trigger production may play significant roles in iSBH-sgRNA activation.

Future work is needed to develop iSBH-sgRNA technology towards *in vivo* applications and existing RNA targeting technologies could provide useful insights. Furthermore, understanding the range of endogenous RNAs that could act as triggers is essential, as well as understanding iSBH-sgRNA design rules. The following sections explore lessons that could be learnt from other similar technologies as well as the unmet RNA-detection niches where iSBH-sgRNAs could significantly contribute. Further strategies for developing iSBH-sgRNAs towards therapeutically relevant applications are also discussed, together with other potential applications of this technology.

7.2 Lessons to be learnt from similar technologies

7.2.1 Chemical modification improve probe design

Extensive research into the usage of antisense oligonucleotides (ASOs) and short interfering RNAs (siRNAs) to target RNA discovered that chemical modifications prevent degradation by nucleases, reduce immune stimulation as well improve RNA targeting properties. Three different design iterations provided extra layers of improvements in drug-like properties of nucleic acid therapeutics[160].

First generation chemical modifications involved the use of nucleic acids with phosphorothioate (PS) bonds (Figure 7.1.A.). In a standard RNA chain, ribose groups are linked through phosphodiester bonds. For PS bonds, the non-bridging oxygen (O) within the phosphate group is replaced by sulphur (S). Such modifications conferred resistance against nucleases. Nevertheless, such modifications were not very useful on their own and reduce ASO affinities for their RNA targets[161].

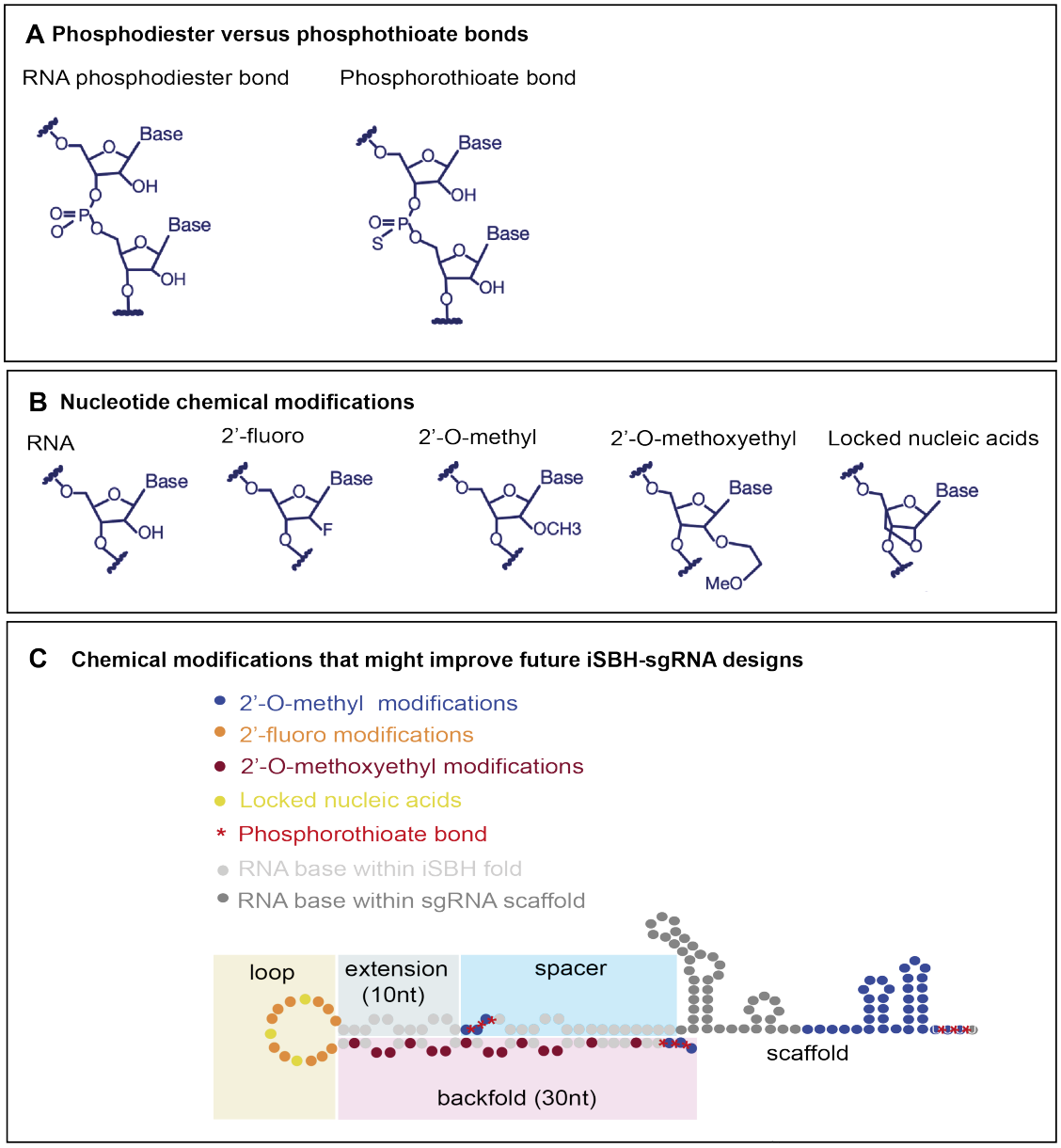


Figure 7.1: Chemical modifications that could improve iSBH-sgRNA designs

A. Phosphodiester versus phosphothioate bonds. RNA nucleotides are naturally linked through phosphodiester bonds. In modified phosphothioate bonds, one of the oxygen (O) atoms in the phosphate group is replaced by sulphur (S). This figure was inspired from Khvorova et al. [160]. **B.** Nucleotide chemical modifications. Example of nucleotide modifications that could improve iSBH-sgRNA designs include: 2'-fluoro, 2'-O-methyl, 2'-O-methoxyethyl and locked nucleic acids (LNA). This figure was inspired from Khvorova et al. [160]. **C.** Chemical modifications that might improve future iSBH-sgRNA designs. Scaffold modifications as well as modifications at beginning of the backfold structure could prevent exonucleolytic degradation. As iSBH-sgRNAs get truncated following trigger recognition, chemical modifications at the 5' end of the spacer would also prevent further exonucleolytic degradation. Chemical modifications in the loop structure as well as the backfold structure could improve iSBH-sgRNA affinity for RNA triggers. Chunky modifications in the backfold might prevent recognition by endogenous endonucleases, but also improve iSBH-sgRNA duplex stability.

Second-generation chemical modification involved modifications of the 2' oxygen (O) in the ribose sugar (Figure 7.1.B.). Example of ribose modifications include 2'-O-methyl, 2'-O- methoxyethyl as well as 2'-fluoro[160]. 2'-O-methyl modifications are also found in nature and they served as a starting point for development of even more potent modifications such as 2'-O- methoxyethyl. 2'-O-methoxyethyl provides even more protection against nucleases and, each modified nucleotide could promote an increase in melting temperature between 0.9 and 1.7°C[160]. siRNAs with 2'-fluoro modifications display resistance against nucleases, reduced immune-system activation as well as present improved RNA affinity[162]. 2'-fluoro modifications could increase probe temperature with up to 2.5 °C per modified nucelotide[160]. Furthermore, such modifications were shown to have minimal effects on RNA folding[162].

Third-generation improvements involved the use of locked nucleic acid modifications (LNA). LNA also contain modifications in the ribose group. These chemical modifications involve formation of a methylene bridge between the ribose 2'-O and 4'-C (carbon)[160, 161]. LNA modifications are the most efficient chemical modifications introduced in strand-displacement nucleic acid probes. LNA modifications reduce the flexibility of the RNA backbone and decrease the entropy loss upon hybridisation, while favouring more efficient nucleobase interactions[163]. Such modification enable an increase in melting temperature between 4 and 8°C per modified nucleotide[161].

Chemical modifications are also used in improving sgRNA stability for CRISPR applications[164]. In the absence of Cas9, sgRNAs were shown to be highly unstable. When not bound to Cas9, sgRNAs could be detected in the cells for about 15 min. Cas9 binding increases sgRNA stability for up to 4h[165]. Extra sgRNA stability could be achieved through preventing sgRNA recognition by nucleases through chemical modifications. PS bonds, 2'-O-methyl groups[165], LNA[166] as well as by DNA nucleotides are tolerated at certain sgRNA positions[34]. Due to their small

size, 2'-O-fluoro modifications appear to be the most-widely tolerated modification across different positions, while chunkier modifications such as 2'-O- methoxyethyl are less tolerated[167].

Chemically modified sgRNAs that I tested in zebrafish embryos contain PS bonds between the first 4 and last 4 sgRNA nucleotides [60] (Figure 6.9.A.). These sgRNAs display extensive 2'-O-methyl modifications in several bases from the scaffold as well as in the first 3 spacer nucleotides. Apart from these modifications, iSBH-sgRNA designs could further benefit from modified nucleotides. Such modifications would help improving iSBH-sgRNA folding, increase affinity for RNA targets, reduce immune stimulation as well as increase overall iSBH-sgRNA resistance against nucleases. Current ASO and siRNA improvements have mostly looked into ways of improving affinity of linear hybridisation probes for RNA[160, 161]. Nevertheless, as iSBH-sgRNAs form a hairpin structure, studies understanding the effect on chemical modifications on RNA folding and hairpin stability[162, 168–170] might be more informative.

Modifications such as 2'-O-methyl were well tolerated within loop structures[168, 169], while 2'-fluoro modifications also had minimal interference with RNA folding[162]. In contrast, LNA increased rigidity of the backbone and could impact folding. For hairpins with "GAAA" tetraloops, LNA modifications in the loop destabilised hairpin formation. However, hairpin destabilisation was way more eminent for DNA rather than RNA hairpins[171]. For molecular beacons (nucleic acid probes that form a hairpin structure) with larger loops, LNA modifications appeared to maintain hairpin stability. LNAs were tolerated when beacons contained alternating modified and non-modified nucleotides along both stem and loop structures[170]. Furthermore, addition of only 3 LNA nucleotides within an 11 nucleotides loop, appeared to massively increase beacon affinity for their RNA targets[168].

Starting from these findings, as well as previous insights into iSBH-sgRNA activation mechanism, strategies for incorporating chemical modifications into iSBH-sgRNA designs could also be devised (Figure 7.1.C.). As iSBH-sgRNAs might be recognised by endogenous RNAi factors (Figure 5.5.C.), a strategy for preventing this could involve the use of large chemical modifications. When inserted at certain key positions, such modifications could prevent interaction between iSBH-sgRNAs and cellular enzymes by introducing steric clashes. Furthermore, RNA circularisation assay suggests that iSBH-sgRNAs get truncated to original sgRNA sequences following activation. This suggest that modifications at the 5' end of the spacer might be beneficial for preventing exonucleolytical degradation of truncated sgRNA. As iSBH-sgRNA truncation might be important for sgRNA activity, chemical modifications should not inhibit this process. dCas9 might be involved in this process and it is very likely that truncation might happen within the extension sequence. Chemical modifications within extension sequences might therefore be avoided.

Other modifications could be employed for avoiding iSBH-sgRNA background activity in an OFF state as well as improving iSBH-sgRNA affinity for endogenous RNA triggers. For iSBH-sgRNAs that display undesired background activity in the OFF state (Figure 5.5.C.), chemical modifications could be included in the stem structure for promoting duplex stability. Affinity for RNA triggers could be improved through modifying loop structure as well as iSBH-sgRNA bulges. The loop structure could be extensively modified with 2'-fluoro bases[160, 162] as well as a few LNA modifications for improving iSBH-sgRNA affinity for RNA triggers. Extra modifications in the backfold bulges might further increase iSBH-sgRNA affinity for complementary triggers, while preventing recognition by endogenous nucleases.

7.2.2 Probe design rules

Extensive studies were carried out for understanding how probe design could improve RNA targeting efficiency. Apart from chemical modifications, aspects such as probe-

binding sites, probe size as well as other sequence specific parameters appear to be very important.

Probe binding site; The position at which hybridisation occurs is essential, as RNA secondary structures and RNA-binding proteins likely impede probe binding. Several studies sought to identify probes that hybridise with open RNA structures[172–178]. *In vitro* experiments shown that 20nt oligos hybridise with different efficiencies to different sub-sequences within a target mRNA[176]. Similar results were reported for Cas13 tiling experiments, where well-performing gRNAs preferentially hybridised with specific mRNA sub-sequences[179].

Alternative approaches for identifying structurally open RNA sub-sequences include using *in-silico* prediction tools such as NuPACK[180] or ViennaRNA[181]. Although not perfect, *in silico* predictions are a good start for choosing accessible targets. Sequences predicted by NuPACK as being inaccessible also appeared inaccessible in experimental assays. However, sequences predicted by NuPACK as accessible had variable accessibilities when tested experimentally[176]. Sequencing protocols developed for mapping RNA secondary structures[182–184] may also offer a significant benefit in choosing appropriate trigger sub-sequences while considering the structure of a given RNA within a cellular context.

Probe size; In studies aiming to target mRNAs using sequences with extended complementarity, no screening experiments were carried out for identifying open RNA structures. In such studies, increasing the length of the RNA probe resulted in improved targeting[139, 140]. Synthetic ADAR guide RNAs could recruit endogenous ADAR enzymes to mRNAs of interest, resulting in RNA editing[139, 185, 186]. Progressively longer ADAR gRNAs ranging from 40nt to 151nt were tested, and longer gRNAs were shown to cause better editing outcomes[139]. Similar results were reported for RADARS (Reprogrammable ADAR Sensors)[187] and trans-splicing RNAs[140], suggesting that longer hybridisation probes might outcompete

RNA secondary structures within endogenous RNAs. Therefore, increasing the length of the RNA-detection regions might also improve the affinity of the iSBH-sgRNAs for RNA triggers. Such findings are also supported by different iterations of the iSBH-sgRNA designs where increasing size of the stem as well as the loop improved detection of longer RNA triggers. Nevertheless, for iSBH-sgRNAs, caution should be displayed when choosing the loop size, as longer loop extensions might destabilise the iSBH fold.

Sequence-specific parameters; A study attempted to identify parameters that characterise effective splicing-modulating ASOs. Parameters such as melting temperature (Tm) and GC contents as well as the amount of energy released upon ASO binding were important for ASO activity. For efficient ASOs, Tm and GC contents were higher. Also, for effective ASOs, higher amounts of energy were predicted to be released after binding[174].

iSBH-sgRNA design rules could also be understood through similar high throughput experiments. Depending on input parameters, MODesign simulations could produce a list of multiple modular iSBH-sgRNA outputs. For proof-of-concept experiments, educated guesses were made for selecting iSBH-sgRNAs for further experimental validation. A first criteria was to prioritise sequences predicted by NuPACK[180] to adopt desired structures with high probabilities. Based on previous experimental evidence, priority was also given to sequences that hybridise to RNA trigger sub-sequences lacking complex secondary structures. Extra priority was also given to modular iSBH-sgRNAs that have more than 17nt complementary with CRISPR target sites as well as extended complementary with RNA triggers. Nevertheless, other criteria such as sgRNA spacer sequence, GC content, loop size or sequence motifs might also play important roles in iSBH-sgRNA performance. The exact contribution of these factors is not yet understood. In order to fully understand

design rules, large screening experiments will be required where multiple iSBH-sgRNAs will be designed for different RNA triggers. Machine learning algorithms might also assist with the development of accurate prediction scores.

7.2.3 Avoiding off-target effects

A common off-target effect of sgRNAs is directing Cas9 to non-desired genomic DNA regions[27]. In iSBH-sgRNAs, this issue could be achieved by further engineering pre-validated sgRNA sequences with minimal off-target effects. Additionally, iSBH-sgRNAs might have additional off-target effects due to non-specific iSBH-sgRNA activation. Furthermore, formation of long RNA duplexes might be a concern in eukaryotic cells.

Non-specific iSBH-sgRNA activation could occur due to recognition of endogenous RNA triggers with similar sequences. Algorithms developed for designing highly-specific ASOs or siRNAs could also assist with choosing iSBH-sgRNA trigger sub-sequences[188]. Poor iSBH-sgRNA designs that do not fold into perfect switches might also be a source of off-target effects. In the future, extra care should be taken in order to ensure that observed Cas9 activity is linked to detecting endogenous RNA triggers of interest.

As observed from RNAi effectors[189], formation of long RNA duplexes could be a concern due to potential activation of type I interferon (IFN) response. IFN response is an innate immune mechanism involved in defence against RNA viruses. Double-stranded RNAs are recognised by RNA sensors such as RIG-I, MDA5 and TLR3[190]. After dsRNA detection, a signalling cascade is triggered, leading to non-specific degradation of cellular RNAs and translation arrest. In this way, cells attempt to block viral replication and synthesis of viral components[191]. IFN response was also associated with cellular toxicity following delivery of *in vitro* transcribed sgRNAs with 5' triphosphate groups[192].

Testing whether dsRNA duplexes resulting from iSBH-sgRNA interaction with endogenous RNAs activate IFN responses would be necessary while optimizing this technology. If type I interferon gets activated, chemical modifications such as 2'-O-methyl and 2'-fluoro could avoid TLR activation[193].

7.3 iSBH-sgRNA technology in the existing landscape of CRISPR-based RNA sensors

Eukaryotic cells express many different classes of transcripts that have different properties in terms of their size, expression levels and subcellular localization. One could imagine that sensing different types of RNA would pose different challenges. In this context, engineering CRISPR-based sensors towards detecting all meaningful classes of RNA would be essential. Examples of published CRISPR sensors involve engineered expression cassettes as well as engineered sgRNAs. Published examples of engineered cassettes have clear applications in microRNA detection. Sensing of longer and more complex RNA species such as mRNA might be more efficient using technologies such as RADARS[187] or RENDR (Figure 1.1.E.)[18]. In these technologies, hybridisation probes up to 200nt enable mRNA targeting and sensing. On the other side, classes of transcripts that could be detected by engineered sgRNA are not yet characterized. Nevertheless, in theory, these technologies are scalable to the detection of multiple transcript classes.

7.3.1 iSBH-sgRNAs vs engineered CRISPR expression cassettes

Delivery of Cas9-sgRNA ribonucleoprotein (RNP) complexes or Cas9 mRNA is often preferred for therapeutic applications as rapid vector degradation reduces off-target effects associated with prolonged Cas9 expression[32]. While iSBH-sgRNAs are, in theory compatible with both delivery formats, expression cassettes are only compatible with RNA-based delivery. For example, expression cassettes based on microRNA response elements (MREs) facilitate both CRISPR activation[15–17] or in inactivation[90] in target organs.

In terms of **CRISPR activation in target tissues**, published examples involve controling anti-CRISPR proteins[16, 17] or processing sgRNAs[15] in the presence of microRNAs. Addition of an extra anti-CRISPR protein into CRISPR delivery systems require extra adjustments for co-loading two different mRNAs as well as sgRNAs into LNPs (lipid nanoparticles). Furthermore, similar with Cas9, anti-CRISPR proteins[88] also have bacterial origin and pose an extra risk for immunogeneicity. Flanking sgRNAs by microRNA-responsive elements and incorporating them into mRNA-like sequences could be an elegant approach for restricting CRISPR-activity to organs of interest[15]. Nevertheless, when delivered *in vivo*, sgRNAs are more efficient in the presence of chemical modifications[164]. Therefore, further chemical modifications of sgRNA-expressing cassettes is a logical step forward.

Technologies such as RADARS[187] or RENDR [18] have potential to be developed toward CRISPR activation in the presence of mRNA. Proof-of-concept RENDR designs (Figure 1.1.E.) have shown that dCas9 expression could be induced in bacterial cells in the presence of target mRNA[18]. Other technologies based on mRNA trans-splicing were also implemented in mammalian cells[140], suggesting that RENDR is scalable to eukaryotic systems. RADARS repurpose endogenous RNA molecules to act as ADAR guide RNAs and edit a synthetic transgene encoding therapeutic payloads. STOP codons introduced at the beginning of therapeutic payloads prevent transgene expression. ADAR-mediated RNA editing enable removal of STOP codons, leading to payload expression. RADARS were initially developed and tested in mammalian cells, where induction of IL6 (interleukin 6) was demonstrated in the presence of target mRNAs[187]. Replacing IL6 with Cas9 in RADAR expression cassettes might also enable conditional CRISPR activation in target cells.

A big advantage of iSBH-sgRNA technology over expression cassettes that enable

CRISPR activation in target cells relies on its simplicity. iSBH-sgRNAs only involve minimal modifications in the sgRNA sequence and, in theory, will be compatible with existing CRISPR delivery platforms.

Technologies that enable **CRISPR inactivation in target tissues** are more elegant. Introduction of MREs in the Cas9-encoding transcript[90] is a minimal change made to the Cas9 mRNA and, in theory, such cassettes will be compatible with existing RNA-only CRISPR delivery formulations. Furthermore, the level of microRNA-mediated construct derepresion could be tuned by introducing mismatches in the MREs or by increasing the number of MREs within a coding transcript. For example, addition of 4 MREs complementary with miR-17 enabled almost complete transgene silencing in mammalian cells[194]. As many targeted delivery systems leads to high concentrations in the liver, despite also reaching their intended organs[53], a potential way of preventing CRISPR activity in the liver could involve Cas9 silencing by microRNAs. Incorporating of MREs specific to liver microRNAs such as miR-122a[195], could enable degradation of Cas9 mRNA in the liver, but not in on-target tissues. In contrast with this method, iSBH-sgRNAs would enable conditional CRISPR activation in on-target tissues.

7.3.2 iSBH-sgRNAs vs other engineered sgRNAs

Several sgRNA designs[87, 89, 92, 94, 98–103, 129] were proposed for RNA sensing. Nevertheless, all these designs were tested in different conditions and against different RNA triggers. A head-to-head comparison between the performance of these designs is currently lacking, and it is hard to pinpoint a design that performs better than others. Specific designs have a clear advantage, as they enable users to detect any given RNA trigger and direct Cas9 to any DNA sequence of choice [87, 91, 92, 98, 102, 103, 129, 134] and modular iSBH-sgRNA designs also have this feature. This was made possible by decoupling RNA trigger sensing component from the spacer modules.

The fact that iSBH-sgRNAs were extensively optimised in eukaryotic systems represents a clear advantage. Available data suggests that extending iSBH-sgRNA stem and loop sequences could improve design performance while detecting longer RNA triggers. Dynamic ranges could be optimised through either sgRNA design or through appropriate choices of CRISPRa effectors and reporters. iSBH-sgRNA systems were designed and tested starting from a panel of sgRNA spacers and RNA triggers, suggesting the scalability of the system towards other RNA sequences. Furthermore, orthogonality tests suggest the iSBH-sgRNAs display high levels of specificity towards their complementary RNA triggers. Nevertheless, a deep understanding of the rules governing folding and the importance of several thermodynamic parameters in iSBH-sgRNA designs are yet to be understood. High throughput studies combined with machine learning approaches will be required for identifying optimal design rules.

To date, iSBH-sgRNA was the first CRISPR-based RNA sensor based on sgRNA engineering to be tested *in vivo*. Available mechanistic data would enable designing chemically modified sgRNAs with superior performances. So far, iSBH-sgRNA designs are the only RNA sensors based on sgRNA engineering for which chemical modifications were considered. Furthermore, in contrast with other available platforms, iSBH-sgRNAs would be easy to test and implement in other labs due to the availability of computational tools for automatic iSBH-sgRNA design.

7.4 Characterising endogenous RNAs that could act as triggers *in vivo*

In order to develop the iSBH-sgRNA technology towards meaningful applications, it is imperative to understand the range of endogenous RNAs that could be detected efficiently. Current technologies already enable efficient mRNA[187] or microRNA[15–17, 90, 91] detection in mammalian cells. Data presented in this thesis suggests that iSBH-sgRNAs cannot detect mRNA. Due to increased length and complexity of secondary structure, technologies relying on long hybridization

probes (up to 200nt)[139, 140] might be a better approach for detecting mRNA or complex long non-coding RNA structures. Despite existing microRNA-detection technologies, iSBH-sgRNAs is an additional tool well-suited to microRNA triggers. In particular, microRNAs would be interesting targets as they tend to be recycled back to the cellular pool after cleaving their RNA targets[196]. Introduction of MREs in the iSBH-sgRNA loop would enable microRNA-mediated cleavage of the iSBH-sgRNA backfold. Furthermore, due to the cellular recycling process, a single microRNA might enable activation of multiple iSBH-sgRNA molecules.

In light of the data presented in this thesis, a competition-free niche for future innovation has opened up for the development of CRISPR-based RNA sensors to efficiently detect small RNA species other than microRNAs and up to 300 nucleotides in size. Apart from choosing endogenous RNA triggers up to 300 nucleotides, relative concentrations of iSBH-sgRNAs and RNA triggers would also be important. The maximum level of functional Cas9-sgRNA complexes that could be formed depends on the endogenous level of RNA triggers. Therefore, choosing triggers that have comparable concentrations with iSBH-sgRNA would be essential. Other important factors that should be considered when choosing endogenous RNA triggers include RNA accessibility in terms of secondary structures, subcellular localisation as well as aspects associated with co-transcriptional RNA processing. Triggers that are relatively free-from secondary structures would be ideal. In terms of subcellular localisation, triggers found in the nucleus might be preferred in order to ensure their co-localisation with iSBH-sgRNAs and Cas9. As iSBH-sgRNA technology was optimised for detecting polymerase III transcripts, endogenous RNAs transcribed by this polymerase would also be a great bonus. On top of that, choosing RNAs that have important roles in determining cellular identity would be essential.

After improving iSBH-sgRNA activity in zebrafish embryos through chemical modifications, further screening experiments would enable characterising classes of endogenous RNAs that can act as triggers. Transgenic zebrafish lines generated

would be a good starting point for these experiments. Existing RNA sequencing data-sets characterising nuclear transcripts upregulated in neural crest cells[136] could be a great resource for identifying transcripts that match ideal-trigger criteria. For each trigger candidate, it would be important to design several iSBH-sgRNAs hybridising with different trigger sub sequences. This would minimise the effect of iSBH-sgRNA design variability in understanding classes of transcripts that could be detected.

MIC-Drop platform[197] might also enable designing of efficient screening strategies. This would involve packaging individual iSBH-sgRNAs designs into lipid droplets. Individual droplets containing a single iSBH-sgRNA design could be mixed together in order to create a library of droplet encapsulating different iSBH-sgRNAs. A single droplet would be injected per embryo, enabling testing of multiple designs in a time-efficient manner. At one day post-fertilisation, embryos that selectively express ECFP in the neural crest could be selected and the identity of well-performing iSBH-sgRNAs could be identified through sequencing experiments. After identifying gene expression profiles associated with good RNA triggers, this technology could be further scaled up towards other therapeutical and research applications.

7.5 Future applications of this technology

Once optimised for efficient detection of endogenous RNAs, iSBH-sgRNAs could be used in detecting disease-related RNA biomarkers and restricting CRISPR activity to disease-affected cells. Nevertheless, iSBH-sgRNA technology might also enable further applications in basic research as well as *in vitro* diagnostics.

7.5.1 Potential iSBH-sgRNA applications within the CRISPR therapeutic space

CRISPR therapeutic strategies might benefit from further iSBH-sgRNA developments. Several therapeutic approaches were proposed to date, involving both *ex vivo* and *in vivo* genome editing[52, 198]. For cell types or diseased cells where finding

precise surface biomarkers or tissue-specific promoters could be challenging, using RNA biomarkers to activate Cas9 activity might represent an excellent alternative solution[17]. Novel therapeutic strategies could combine existing targeted or non-targeted delivery methods with iSBH-based RNA sensors for increasing the specificity of such therapeutics. For example, iSBH-sgRNAs could be delivered to cells in an OFF state. Once inside a cell, these sensors could sense the presence of transcript biomarkers and then only activate their therapeutic function.

In particular, iSBH-sgRNA technology could be used in conjunction with other *in vivo* therapeutic formulations aiming to correct mutations outside of the liver. Despite the fact that delivery formulations might accumulate in the liver, iSBH-sgRNAs would reduce unintended editing events in this organ. Once inside target cells, CRISPR therapeutic activity would be activated by detecting small endogenous RNA biomarkers that act as indicators for the cellular identity. Examples of disorders for which this could be useful include CRISPR-based strategies aiming to correct sickle-cell disease by delivering LNPs *in vivo*[199] as well as strategies aiming to correct Duchenne Muscular Dystrophy[7].

7.5.2 Potential iSBH-sgRNA applications in basic research

By analysing cellular transcripts and their levels, information could be extracted about the cell identity and the environmental challenges the particular cells face[200, 201]. Using RNA as a biomarker for cell identity might enable restricting CRISPR activity to different tissues or developmental stages[98]. This could enable knocking out important developmental genes in limited subsets of cells without compromising viability of a model organisms. Expression of RNA biomarkers might also help identify cells in which specific pathways are turned on. Examples could consist of cells with active DNA-damage repair mechanisms, cells exposed to different environmental stimuli, cells at different cell-cycle stages, cells that upregulate specific metabolic pathways, etc.

Coupling RNA sensing with CRISPR transcriptional activators might enable the fluorescent labelling of cells expressing transcripts of interest. Such approaches could offer an alternative to existing RNA detection methods[202]. Methods such as hybridisation chain reaction[202–214], fluorescent aptamers[202, 215–220] or fluorescent RNA-binding proteins[179, 202, 221–229], enable visualisation of RNA foci as well as their subcellular localisation. Nevertheless, such methods do not allow the linking of RNA detection to any downstream applications.

In contrast with existing RNA detection methods, iSBH-sgRNAs would enable coupling of RNA detection with functions such as: DNA cleavage[3], base editing[42, 43], prime editing[44, 45], CRISPR transposases[51] or transcriptional modulators[48–50, 230]. For example, iSBH-sgRNAs might enable production of lineage tracing barcodes[231] in cells that express given RNA biomarkers. Alternatively, CRISPRa effectors could be employed for driving expression of ubiquitin ligases[99, 100] or biotin ligases. Ubiquitin ligases conjugated with antibodies could promote target protein degradation[99, 100] in cells of interest, while biotin ligases could enable labelling of RNAs or proteins[136, 232] for downstream purification and analysis.

7.5.3 Potential iSBH-sgRNA applications for *in vitro* diagnostics

CRISPR-based diagnostics are currently gaining popularity and a variety of diagnostic tools are available for detecting disease-associated RNAs through Cas13 or Cas12 collateral cleavage activity[233–238]. Despite the fact that existing Cas13 or Cas12 diagnostics[233–238] are promising, having different platforms that rely on different components would be beneficial in a pandemic. Developing diagnostic tests that differ as much as possible from each other would avoid supply chain bottlenecks.

Different RNA-sensing sgRNA designs were already adopted in diagnostics[93, 133] and methods for amplifying signals for Cas9 *in vitro* cleavage assays were also proposed[93]. Furthermore, implementation of RNA logic gates is also possible

using strand-displacement based circuits[102, 103, 129]. Logic gates could increase the confidence of diagnostics by simultaneously analysing two or more relevant RNA biomarkers. Starting from these examples, iSBH-sgRNA-based diagnostics could also be designed. In theory, detection of disease associated RNAs, including viral RNA is possible. By engineering the iSBH-sgRNAs to be highly specific for their probes of interest, different tests could also be built for sensing different alleles or viral variants.

*I would like to acknowledge that Sections 7.2.2, 7.2.3 and 7.3-7.5 were adapted from a literature review accepted for publication in The CRISPR Journal: ***Pelea*** *O, Fulga T, Sauka-Spengler T, RNA-responsive gRNAs for controlling CRISPR activity: current advances, future directions, and potential applications.* I would also like to acknowledge that the idea of using MREs to degrade Cas9 expression in the liver (Section 7.3.1) emerged from a brainstorming session taking place at the NCCR RNA & Disease Summer School (Saas-Fee, Switzerland, 2022).

8 Materials and Methods

8.1 *In silico* design of iSBH-sgRNAs

First generation iSBH-sgRNA designs were designed by inputting different spacer sequences into a computational pipeline generated by Ferry et al.: **http://apps.molbiol.ox.ac.uk/iSBHfold/**[37]. Main features of iSBH-sgRNA designs include the spacer (20nt), loop (14nt), spacer* (20nt) and scaffold sequences. Spacer* sequence is the reverse complement of the spacer sequence, and it was modified to contain mismatches at positions: 11-12 and 16-17. The NuPACK[180] fold corresponding to these sequences (excluding scaffold and extra GC sequences*****) is: ((((((((((..(((..(((.....)))..)))..))))))))))

Second generation iSBH-sgRNA designs were generated by introducing 10nt extensions between loop and spacer sequences of the first-generation design. The random 10nt extension sequences had a GC content ranging between 40 and 60%. Extension* sequence is generated by determining the reverse complement of the extension sequence, with mismatches integrated at positions 1-2 and 6-7. NuPACK[180] fold corresponding to this design (excluding scaffold and extra GC sequences*****) is: ((((((((((..(((..(((..(((..(((.....)))..)))..)))..)))..)))..))))))))))

Modular iSBH-sgRNA designs were generated using the MODesign computational pipeline. The overall structure of this computational pipeline can be seen in the Figure 4.3. This pipeline takes 3 input sequences: iSBH-sgRNA loop size, the RNA trigger to be sensed an input sgRNA sequence that targets Cas9 to desired CTSs (CRISPR target sequences). **MODesign** calculates the size of the iSBH-sgRNA trigger-sensing component and creates a list of all potential trigger sub-sequences having that particular size. For each sub-sequence, it determines the reverse complementary region, and it inserts it between extension and spacer* sequences, while filling in the extension sequence. MODesign checks if resulting modular iSBH-sgRNA sequences fold into desired RNA structures using NuPACK[180] and outputs all designs that adopt a correct fold. NuPACK[180] fold corresponding to this design is (excluding scaffold and extra GC sequences*****): ((((((((((..(((..(((..(((..(((.....)))..)))..)))..)))..)))..))))))))))

A MODesign scoring algorithm was also implemented, based on the iSBH-sgRNA probability of adopting desired secondary structures (N), percentage of nucleotides that are free-from secondary structures in the trigger sub-sequence (M), percentage of trigger nucleotides complementary with the iSBH-sgRNA backfold (P) and percentage of spacer nucleotides complementary with the CTS (Q). Scoring formula used was N*N*N*M*P*Q. For different simulations, modular iSBH-sgRNAs with higher scores were selected for experimental validation.

For first generation designs, RNA triggers complementary with the iSBH-sgRNA loop and spacer* sequences were designed. For second generation designs, RNA triggers are complementary with the loop, extension* and spacer* sequences. CTSs were designed by adding PAM (palindromic associated motifs) sites downstream from the spacer sequence. All sgRNA spacers, iSBH-sgRNAs, triggers and CTS reporter sequences could be found in the Supplementary Material.

***** At the beginning of each iSBH-sgRNA, an extra GC sequence was added. The

extra G was inserted in order to promote transcription from U6 polymerase III promoter[6]. The extra C was added in order to increase base-pairing complementary with the first G in the sgRNA scaffold.

8.2 Molecular cloning

All mammalian plasmids expressing Cas9 transcriptional activators were acquired from Addgene: dCas9-Vp64 (47107), dCas9-VPR (63798) as well as Cas9_pX458 (48138).

Native sgRNAs, iSBH-sgRNAs, RNA triggers and 1xCTS repeats were cloned in mammalian-expression vectors by inserting annealed single-stranded DNA oligos (IDT) into appropriate plasmid backbones. iSBH sequences, as well as sgRNA spacer sequences were cloned between BbsI restriction sites in the pcDNA3.1-U6_sgRNA_6xT-SV40_iBue_PA plasmid, generated by Ferry et al., 2017 . 1xCTS sequences were cloned in p035_pause-HBG-CFP-pA (XbaI, AscI restriction sites; plasmid received as a gift from Dr. David Knapp). Short trigger sequences were cloned within U6-TM2Emp-6T_iBlue plasmid (gift from Dr. Quentin Ferry) between BbsI restriction sites. U6-TM2Emp-6T_iBlue plasmid also encoded for two hairpin structures aiming to protect 34 and 44nt short triggers from degradation.

Forward and reverse oligos were treated with polynucleotide kinase (PNK) and incubated at 37°C for 30 minutes. Oligos were denatured and re-annealed by heating up samples to 95°C and decreasing the temperature to 25°C (at a rate of 2°C/minute). Backbone plasmids were digested according to the NEB protocols, followed by gel extraction. Ligation was carried out using 100ng backbone and 0.5μl annealed oligos in a total volume of 10μl. 3 μl ligation product were transformed into DH5α *E. coli* (Invitrogen, 18265017) cells.

8xCTS sequences were cloned in the P2-ECFP-pA (Addgene 26280) plasmid generated by Nissim et al[138]. The original 8xCTS sequence was removed (NheI

restriction sites) and replaced with a cloning landing pad containing numerous restriction sites. Resulting vector was named Landing_Pad_8xCTS-ECFP-pA. ssDNA oligos encoding for 2xCTS repeats were ordered from IDT and amplified by PCR. PCR products were separated into two restriction reactions- SacI/Esp3I and Esp3I/SacI respectively. The two digestion products were cloned into the Landing_Pad_8xCTS-ECFP-pA, leading to a 4xCTS-ECFP reporter. In a subsequent round of cloning, 4xCTSs were amplified by PCR, digested with NheI/ApaI and cloned back into the 4xCTS-ECFP reporter, leading to the final 8xCTS-ECFP constructs. Primer sequences and IDT oligos used for cloning 8xCTS reporters could be found in the Supplementary Material.

100nt flank U6 triggers were cloned in the U6-TM2Emp-6T_iBlue_ModFlanks plasmid. In a first round of cloning, backbone was digested using BbsI and short trigger sequences were cloned by annealing and ligation. Further rounds of cloning involved amplification of 100nt flanks by PCR and standard restriction-ligation cloning. 100nt 5' flanks were digested with BsmBI and 100nt 3' flanks were digested with NotI/NheI. Final size of triggers with 100nt 3' flank extension was 191nt for first-generation iSBH-sgRNA designs and 200nt for second-generation iSBH-sgRNA designs. Size of triggers with 100nt 5' flanks was 175nt, while size of triggers with 100nt 5'+3' flanks was 270nt. Sizes exclude two hairpin sequences that protect 5' and 3' trigger ends from degradation.

CMV+poly(A) and CMV+MALAT1 triggers were cloned using a two-step protocol. Initially, the pcDNA3.1 cloning cassette was inserted in the U6-TM2Emp-6T_iBlue using MfeI/NheI; cloning replaced U6 cloning cassette with a polII cloning cassette. Then, short triggers were amplified from U6 plasmids and cloned in the polII cassette using NotI/AflII. **mRNA triggers** were cloned into the pcDNA3.1_EYFP plasmid (gift from Dr. Toni Baeumler). Triggers were amplified from U6 expressing 100nt 5' 100nt flank triggers and cloned downstream from the EYFP using XbaI/ApaI.

Repetitive RNA trigger sequences for testing modular iSBH-sgRNA designs were cloned in the U6-TM2Emp-6T_iBlue plasmid between BbsI restriction sites using type II S restriction-cloning. Trigger sequences were split into smaller sub-sequences and ordered from IDT as single-stranded oligos. Primers annealed to oligo sequences and contained long flaps including extra trigger sub-sequences and BbsI restriction sites. Flaps enabled assembly of longer trigger sub-sequences by PCR. PCR products were loaded in a 1% agarose gel, followed by gel extraction, BbsI digestion and PCR clean-up. Depending on the trigger size, 2 or 3 digested PCR products were ligated to 100ng backbone using a 3:1 molar ratio for inserts to backbone. Example of IDT oligos and ordered primers for cloning triggers A and B are available in the Supplementary Material.

Tol2(B-act:dCas9-Vp64-T2A-Citrine) vector was generated by Gibson Assembly. Vector backbone was amplified by PCR from pMTB2-NLS-Cas9-2a-Citrine plasmid (gift from Dr. Vanessa Chong-Morrison), while the dCas9-Vp64 cassette was amplified from the Sox10:BAC_Cas9m4-Vp64-2a-Citrine bacterial artificial chromosome (gift from Dr. Vanessa Chong-Morrison). **Tol2(8xCTS-ECFP) reporter** was generated by standard restriction-digestion cloning. pMTB2-NLS-Cas9-2a-Citrine was digested using MfeI and XhoI restriction enzymes (NEB). Insert was amplified by PCR and digested with BbsI (NEB) to produce sticky ends compatible with plasmid ligation.

iSBH-sgRNAs and triggers for zebrafish expression were cloned in AcDs_Mini Vector_ U6a-sgRNA-MS2 (gift from Dr. Vanessa Chong-Morrison). SAM effector plasmids for zebrafish expression were also modified from AcDs_MiniVector_ubb-MCP-p65-HSF1-2a-Cerulean (gift from Dr. Vanessa Chong-Morrison) by replacing Cerulean with Citrine using standard restriction-digestion cloning. AcDs vectors containing ubb-MCP-p65-HSF1-2a-Cerulean, iSBH-sgRNAs as well as RNA triggers were cloned using the Gibson Assembly.

All PCRs were carried out using Phusion High-Fidelity PCR Master Mix with GC Buffer (NEB), gel extractions using the QIAquick Gel Extraction Kit (QIAGEN) and PCR clean-up reactions using the MinElute PCR Purification Kit (QIAGEN). DNA concentrations were estimated by NanoDrop, standard ligations were performed using T4 DNA ligase (NEB), while Gibson assembly reactions using the NEBuilder(R) HiFi DNA Assembly Cloning Kit (NEB). Transformed bacteria were grown for 24h at 32°C, while single-colonies were grown for 20h at 37°C in ampicillin-containing LB media. Plasmid DNA was extracted using the QIAprep Spin Miniprep Kit (QIAGEN) and constructs were validated by Sanger sequencing (Eurofins Genomics) prior to being transfected into HEK293T cells or injected into zebrafish embryos.

8.3 Maintaining HEK293T cell lines

HEK293T cells were grown in in full media (Thermo-Fisher Dulbecco's modified Eagle's medium supplemented with 10% foetal-bovine serum). Cells were grown at 37°C and 5% CO2 and passaged every 2 days. Cells have been *Mycoplasma* tested using the Eurofins Genomics Mycoplasmacheck services.

8.4 Transfecting HEK293T cell lines

Transfection was performed using Sigma-Aldrich polyethyleneimine (PEI). For each well, transfection DNA cocktails were mixed with 50μl GIBKO Opti-MEM. 1.5μl PEI/ µg DNA ratio was maintained for all experiments. To each well, 250ng Cas9 effector was transfected with 250ng sgRNA-backbone plasmid (pcDNA3.1_SV40-iBlue-pA), 250ng trigger-encoding plasmids and 125ng 1xCTS or 8xCTS-ECFP reporter plasmids.

8.5 Flow Cytometry sample preparation

48h after transfection, cells were washed with PBS and incubated for 10min with 50μl trypsin-EDTA. Trypsin was inactivated using 250μl FACS buffer (PBS with 10% foetal bovine serum), followed by cell filtration (Falcon 70μm White Cell Strainer) and transfer to Flow Cytometry tubes. Cells were stored on ice prior analysis using the BD LSR Fortessa Analyzer (BD Biosciences).

8.6 Flow Cytometry data analysis

For each condition, 100,000 events were acquired, and data analysis was performed using a Python script developed starting from the FlowCal package[239]. HEK293T cells were identified by plotting SSC-A (side-scattering area) and FSC-A (forward scattering area) values. Single-cells were identified by plotting SSC-H (side-scattering height) against SSC-A values. Then, script plots the iBlue (640-670nm) transfection control reporter against ECFP (405-450nm) reporter for monitoring CRISPR activity.

Gates for iBlue level were set up in the untransfected control, so that only 0.1% of untransfected cells are iBlue+. Gates for the ECFP control were set up in the reporter control, where the sgRNA transfected is not complementary with the nxCTS-ECFP reporter. ECFP gate was set up in such a way that around 0.1% of the cells in the reporter condition were ECFP positive. Displayed percentage activated transfected cells was calculated using the following formula: count(ECFP+/iBlue+ cells)/ [count(ECFP+/iBlue+ cells)+ count(ECFP-/iBlue+ cells)].

All bar-graphs presented in the percentage of activated transfected cells measured for 3 different transfections carried out in 3 different days. The error bars represent the +/- standard deviations for these 3 biological replicates. Displayed p-values were calculated using a non-paired t-test.

8.7 RNA circularisation assays

RNA circularisation protocol was adapted from Knapp et al.[97]. 48h after transfection, cells were washed with PBS, followed by incubation with 50μl trypsin. Trypsin was inactivated using 500μl FACS buffer. Cells were transferred to a 1.5ml Eppendorf tube, followed by centrifugation at 300xg for 5 min. Supernatant was removed and cells were resuspended in 500ml PBS followed by another centrifugation step at 300xg for 5 min. Supernatant was removed followed by snap-freezing of cellular pellets. Cells were stored at -80°C prior to RNA extraction using the Charge Switch Total RNA cell Kit (ThermoFisher). All indications specified in the kit were followed, except from using 1/5 of suggested volume of buffers. The optional DNase treatment step was also included, and the RNA was eluted in 20μl elution buffer.

The ligation reaction was set up by mixing 10μl RNA, 2μl T4 RNA ligase buffer (NEB), 1.9μl H2O, 4μl 50% PEG 8000 (NEB), 0.1μl 10mM ATP, 1μl T4 RNA ligase (NEB), 1μl SUPERase in RNAase inhibitor (ThermoFisher). Reaction was incubated for 4h at room temperature, followed by another round of RNA extraction using Charge Switch Total RNA cell Kit. In this second round of RNA extraction, 1/10 off specified buffer volumes were used, the optional DNase treatment was not performed, and RNA was eluted in 12.5μl elution buffer.

10μl circular RNAs were subjected to reverse transcription (RT) using the Quanti-Tect Rev. Transcription Kit (QIAGEN). Manual specifications were followed, but, provided RT primers were replaced with custom made primers that specifically bind to regions of interest. According to user specifications, primers were diluted to 0.7 µM and RT reactions were incubated for 30min at 42°C.

Desired sequences were amplified using two subsequent PCR reactions using the Phusion High-Fidelity PCR Master Mix with GC Buffer (NEB). All extension steps were carried out in a 15s timeframe and the first PCR round comprised of 10 cycles. 2μl RT product were used as a template in the first reaction. The product

of first reaction was diluted 1/10 and 1μl of this dilution served as a template in the second round of PCR (25 cycles). Second PCR products were mixed with Gel Loading Dye, purple (NEB) followed by loading of 2μl mixture into a 2% agarose gel. For optimal results, gels were run for approximately 90min. Second PCR primers contained NotI/NheI restriction sites that enabled fragment cloning into the pcDNA3.1 plasmid followed by Sanger sequencing.

Circularisation assay primer sequences could be found in the Supplementary Material.

8.8 Zebrafish husbandry

Zebrafish experiments were carried out according to regulated procedures authorised by the UK Home Office within the framework of the Animals (Scientific Procedures) Act 1986. Embryos used were derived from AB zebrafish strains.

8.9 Synthesis of mRNA for embryonic injections

Templates for IVT were linearised by either restriction digestion or PCR, followed by purification using the QIAquick PCR Purification Kit (QUIAGEN). mRNA was synthesised using the mMESSAGE mMACHINETM SP6 Transcription Kit (ThermoFisher) according to producer specifications. Following treatment with TURBO DNase, RNA was purified using the MonarchR RNA Cleanup Kit (NEB). Integrity of the purified RNA was determined by running RNA in a 1% agarose gel, while RNA concentration was measured using the Qubit™ RNA Broad Range assay (ThermoFisher).

8.10 Generation of zebrafish transgenics

dCas9-Vp64 and 8xCTS-ECFP transgenic lines were generated by manipulating the existing TgBAC(Sox10:BirA-mCherry) transgenic lines generated by Trinh et. al.[136]. Single-cell embryos generated by incrossing TgBAC(Sox10:BirA-Cherry)

fish were injected with plasmid DNA containing Tol2 recombination arms as well as Tol2 mRNA (Figure 6.1)[152]. Injection mixtures contained 2μl of plasmid DNA (200ng/μl), 1.5μl Tol2 mRNA (160 ng/μl) as well as 0.5μl Phenol Red. 2nL mixture was injected using a PICOSPRITZER III injector.

Following injections, embryos were kept in a 28°C incubator. 6h post-injection, fertilised embryos were selected and transferred to E3 media. At 1 day post-fertilisation, dead embryos were removed. Surviving embryos were grown for 4 months following subsequent genotyping.

A strategy consisting of two rounds of crosses was employed for identifying founders (Figure 6.3). In a first round, potential male and female founders were incrossed for identifying founder pairs. Once a founder pair was identified, the next step was determining whether the male or the female from that cross contains the transgene. This was achieved by outcrossing males and females with wild-type fish. For each cross, DNA was extracted from a pool of embryos at 1 day post-fertilisation. Extraction was carried out using the PureLink Genomic DNA Mini kit (ThermoFischer), while opting for a 1h lysis step. Extracted DNA was measured by NanoDrop and 100ng DNA was added to a first nested PCR reaction (10 cycles). Products of the first PCR reaction were diluted 1:10 and 2μl products were transferred to a second PCR reaction (29 cycles). PCRs were carried out using the Phusion High-Fidelity PCR Master Mix with GC Buffer (NEB) in a total volume of 20ul, while opting for an extension time on 1min/kb. Results were assessed by running PCR products in a 2% agarose gel.

After founder identification, the next steps involved generation of the first generation of fish (F1) encoding the transgene in all cells of their body (Figure 6.4). Founders were outcrossed with TgBAC(Sox10:BirA-Cherry) adult fish. At 1-3 days post-fertilisation, embryos expressing the Sox10:BirA-mCherry transgene were selected by assessing the mCherry expression using an Olympus MVX microscope. At 3 days

post-fertilisation, embryos were fin clipped. Fins were transferred to 50μl lysis buffer (25mM NaOH, 0.2mM EDTA). Samples were boiled at 95°C for 45min, following by cooling at 4°C for 5min. 50μl neutralisation buffer (40mM Tris-HCl) were added to each sample and nested PCR was carried out for determining transgene presence. PCR reactions were set up in a similar way as for founder identification, except that 5μl DNA were added to the first PCR reaction. F1 embryos were subsequently grown for another 4 months. Subsequent experiments were carried out in embryos resulted from incrossing F1 adult fish.

All genotyping primer sequences could be found in the Supplementary Material.

8.11 Microinjection of zebrafish embryos using the Ac/Ds system

First generation (F1) fish encoding dCas9-Vp64 and 8xCTS-ECFP CRISPR reporter were incrossed. Resulting embryos were injected with Ac mRNA as well as plasmid DNA containing Ds transposase-recognition sequences[135]. Reaction mixtures contained 3μl plasmid DNA (266ng/μl), 0.5μl Ac mRNA (150ng/μl) and 0.5μl Phenol Red. 2nL mixture were injected into single-cell embryos.

Following injections, embryos were kept in a 28°C incubator. 6h post-injection, fertilised embryos were selected and transferred to E3 media. At 1-day post-fertilisation, dead embryos were removed and surviving embryos were screened for the expression of construct DNA. This was achieved by assessing Citrine expression under the Olympus MVX microscope.

8.12 Microinjection of zebrafish embryos with chemically modified sgRNAs

F1 fish encoding dCas9-Vp64 and 8xCTS-ECFP CRISPR reporter were incrossed. Resulting embryos were injected with IDT sgRNA XT[60]. Reaction mixtures

contained 1μl sgRNA (1μg/μl), 2.5μl H2O and 0.5μl Phenol Red. 2nL mixture were injected into single-cell embryos.

Following injections, embryos were kept in a 28°C incubator. 6h post-injection, fertilised embryos were selected and transferred to E3 media. At 1-day post-fertilisation, dead embryos were removed. Surviving embryos were screened for ECFP production using an MVX microscope. Representative embryos were also imaged by confocal microscopy.

8.13 Microinjection of zebrafish embryos with chemically modified iSBH-sgRNAs

F1 fish encoding dCas9-Vp64 and 8xCTS-ECFP CRISPR reporter were incrossed. Resulting embryos were injected with chemically modified iSBH-sgRNAs as well as chemically synthesised RNA triggers. iSBH-sgRNAs were co-injected together with complementary and non-complementary triggers. Reaction mixtures contained 1μl iSBH-sgRNA (1μg/ul), 1μl RNA trigger (1μg/ul), 1.5μl H2O and 0.5μl Phenol Red. 2nL mixture were injected into single-cell embryos. 6h post-injection, fertilised embryos were selected and transferred to E3 media. At 1-day post-fertilisation, dead embryos were removed. Surviving embryos were screened for ECFP production using an MVX microscope.

8.14 Confocal microscopy

Embryos were anesthetised in MS222 and mounted in 1% low-melting point agarose (Invitrogen) dissolved in E3 media. Embryos were imaged on a Zeiss780 LSM upright confocal microscope using a 10x objective. For DNA-based injections, Citrine expression labelled tissues where construct was expressed, while ECFP expression labelled tissues where CRISPR systems were active. The same microscope settings were maintained in between imaging different samples injected with DNA constructs. Embryos injected with iSBH-sgRNAs, but no triggers were genotyped following

injection, for confirming presence of dCas9-Vp64 and the 8xCTS-ECFP reporters. Due to the strength of the signal, laser power had to be decreased for samples injected with chemically modified sgRNAs.

8.15 Code and plasmid availability

In terms of code availability, several scripts will be made available on GitHub together with my next publication: **Pelea O, Mayes S, Ferry Q, Fulga T, Sauka-Spengler T, Development of RNA sensors for controlling CRISPR activity *in vivo*, Research Article (manuscript in progress)**. This will include: The MODesign pipeline, MODesign scoring algorithm, Scripts for automating the iSBH-sgRNA oligo design, Flow Cytometry analysis script.

Plasmids mentioned in this section will be made available via AddGene together with my next publication: **Pelea O, Mayes S, Ferry Q, Fulga T, Sauka-Spengler T, Development of RNA sensors for controlling CRISPR activity *in vivo*, Research Article (manuscript in progress)**. This will include: pcDNA3.1_SV40-iBlue-pA,U6-TM2Emp-6T_iBlue,U6-TM2Emp-6T_iBlue-ModFlanks,p035_pause-HBG-CFP-pA,Landing_Pad_8xCTS-ECFP-pA, sgRNA1_8xCTS-ECFP-pA, sgRNA2_8xCTS-ECFP-pA,sgRNA3_8xCTS-ECFP-pA,gRNA4_8xCTS-ECFP-pA, sgRNA5_8xCTS-ECFP-pA, Tol2(B-act:dCas9-Vp64-T2A-Citrine), Tol2(8xCTS-ECFP).

9
Supplementary Materials

Native sgRNA sequences

+G Spacer scaffold

sgRNA 1:
GCAGTCGCGTGTAGCGAAGCAGTTTTAGAGCTAGAAATAGCAAGTTAAAATAAG
GCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC

sgRNA 2:
GTAAGTCGGAGTACTGTCCTGTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGG
CTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC

sgRNA 3:
GCATGACATTATTCCCGCCTCGTTTTAGAGCTAGAAATAGCAAGTTAAAATAAG
GCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC

sgRNA 4:
GCCAGGTAATGACTAGCAGGAGTTTTAGAGCTAGAAATAGCAAGTTAAAATAAG
GCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC

sgRNA 5:
GCACTACTTCAGTAAGAGTCGTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGG
CTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC

*For sgRNAs with spacer sequences starting with an A, C or T, an extra G was added for facilitating expression from U6 vectors

Native sgRNAs with randomised spacer sequences

+G randomised nt Spacer scaffold

sgRNA 1 (D1):
GAGTCGCGTGTAGCGAAGCAGTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGG
CTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC

sgRNA 1 (D2):
GTGGTCGCGTGTAGCGAAGCAGTTTTAGAGCTAGAAATAGCAAGTTAAAATAAG
GCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC

sgRNA 1 (D3):
GTGATCGCGTGTAGCGAAGCAGTTTTAGAGCTAGAAATAGCAAGTTAAAATAAG
GCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC

sgRNA 2 (D1):
GCTAAGTCGGAGTACTGTCCTGTTTTAGAGCTAGAAATAGCAAGTTAAAATAAG
GCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC

sgRNA 2 (D2):
GCAAAGTCGGAGTACTGTCCTGTTTTAGAGCTAGAAATAGCAAGTTAAAATAAG
GCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC

sgRNA 2 (D3):

GCATAGTCGGAGTACTGTCCTGTTTTAGAGCTAGAAATAGCAAGTTAAAATAAG
GCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC

sgRNA 3 (D1):
GATGACATTATTCCCGCCTCGTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGG
CTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC

sgRNA 3 (D2):
GGTGACATTATTCCCGCCTCGTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGG
CTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC

sgRNA 3 (D3):
GGGGACATTATTCCCGCCTCGTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGG
CTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC

First-generation iSBH-sgRNA sequences

+GC spacer* loop spacer scaffold

*An additional GC sequence was added to all iSBH-sgRNA sequences. G helps expression from U6 promoter, while C base-pairs with the first C from the scaffold

First-generation iSBH-sgRNA 1:
GCTGCTTCGCTAGTCGCTTCTGCACACTACGCATCCCAGTCGCGTGTAGCGAA
GCAGTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACT
TGAAAAAGTGGCACCGAGTCGGTGC

First-generation iSBH-sgRNA 2:
GCAGGACAGTACAGCGAGATACGAACATCCCTAACAGTAAGTCGGAGTACTGT
CCTGTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACT
TGAAAAAGTGGCACCGAGTCGGTGC

First-generation iSBH-sgRNA 3:
GCGAGGCGGGAAGTATGGAATGGACCAGTGAATACACATGACATTATTCCCGC
CTCGTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACT
TGAAAAAGTGGCACCGAGTCGGTGC

First-generation iSBH-sgRNA 4:
GCTCCTGCTAGTACTTAAATGGGCTCCCCACATCAACCAGGTAATGACTAGCAG
GAGTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTG
AAAAAGTGGCACCGAGTCGGTGC

First-generation iSBH-sgRNA 5:
GCGACTCTTACTCCAGTTCTGCATCTTGTTCGTCCAGCACTACTTCAGTAAGAG
TCGTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTG
AAAAAGTGGCACCGAGTCGGTGC

Second-generation iSBH-sgRNA sequences

+GC spacer* extension*loop extensionspacer scaffold
Backfold=spacer*+extension+loop

*An additional GC sequence was added to all iSBH-sgRNA sequences. G helps expression from U6 promoter, while C base-pairs with the first C from the scaffold

Second-generation iSBH-sgRNA 1:
GCTGCTTCGCTAGTCGCTTCTGTCCGTACGTTCACACTACGCATCCAACTGACG
TTCAGTCGCGTGTAGCGAAGCAGTTTTAGAGCTAGAAATAGCAAGTTAAAATAA
GGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC

Second-generation iSBH-sgRNA 2:
GCAGGACAGTACAGCGAGATACCAGCTTAGTCGAACATCCCTAACAGACATAG
CCCGTAAGTCGGAGTACTGTCCTGTTTTAGAGCTAGAAATAGCAAGTTAAAATA
AGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC

Second-generation iSBH-sgRNA 3:
GCGAGGCGGGAAGTATGGAATGAGTGGGATGGGACCAGTGAATACACCAAAC
CAGACATGACATTATTCCCGCCTCGTTTTAGAGCTAGAAATAGCAAGTTAAAAT
AAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC

Modular iSBH-sgRNA sequences

+GC spacer* trigger-sensing region extension spacer scaffold

*An additional GC sequence was added to all iSBH-sgRNA sequences. G helps expression from U6 promoter, while C base-pairs with the first C from the scaffold

Modular iSBH-sgRNA (trigger D/sgRNA2/loop=14nt):
GCAGGACAGTACAGCGAGACAGTGTTCTCTGCACAGATAAGGACAAGCAATGA
AATCTGAGTCGGAGTACTGTCCTGTTTTAGAGCTAGAAATAGCAAGTTAAAATAA
GGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC

Modular iSBH-sgRNA (trigger D/sgRNA3/loop=14nt):
GCGAGGCGGGAAGTATGGACAGTGTTCTCTGCACAGATAAGGACAAGCATTGA
AATCTGGACATTATTCCCGCCTCGTTTTAGAGCTAGAAATAGCAAGTTAAAATAA
GGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC

Modular iSBH-sgRNA (trigger D/sgRNA5/loop=14nt):
GCGACTCTTACTCCAGTCAGTGTTCTCTGCACAGATAAGGACAAACGTGATGAG
TTCACCTACTTCAGTAAGAGTCGTTTTAGAGCTAGAAATAGCAAGTTAAAATAAG
GCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC

Modular iSBH-sgRNA (trigger A/sgRNA2/loop=14nt (1)):
GCAGGACAGTACAACGAAAACTGATCATCACCACAACACAACTAAAGGTCTTGA
GAAGTAGTCGGAGTACTGTCCTGTTTTAGAGCTAGAAATAGCAAGTTAAAATAA
GGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC

Modular iSBH-sgRNA (trigger A/sgRNA2/loop=14nt (2)):
GCAGGACAGTACAACGAACGACTAAACTGATCATCACCACAACACGGATATGTT
GCGTCAGTCGGAGTACTGTCCTGTTTTAGAGCTAGAAATAGCAAGTTAAAATAA
GGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC

Modular iSBH-sgRNA (trigger A/sgRNA2/loop=16nt):
GCAGGACAGTACAGCGAACGACTAAACTGATCATCACCACAACACAACGATATG
TTGCGTCAGTCGGAGTACTGTCCTGTTTTAGAGCTAGAAATAGCAAGTTAAAAT

AAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC

Modular iSBH-sgRNA (trigger A/sgRNA2/loop=18nt):
GCAGGACAGTACAGCGAACGACTAAACTGATCATCACCACAACACAACTAGATA
TGTTGCGTCAGTCGGAGTACTGTCCTGTTTTAGAGCTAGAAATAGCAAGTTAAA
ATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC

Modular iSBH-sgRNA (trigger A/sgRNA2/loop=20nt):
GCAGGACAGTACAACGAACGACTAAACTGATCATCACCACAACACAACTAAAGA
TATGTTGCGTCAGTCGGAGTACTGTCCTGTTTTAGAGCTAGAAATAGCAAGTTA
AAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC

Modular iSBH-sgRNA (trigger A/sgRNA2/loop=22nt):
GCAGGACAGTACAGCGAAACACAACTAAACTGATCATCACCACAACACAACTAA
CAGCGTAGGAGTGAGTCGGAGTACTGTCCTGTTTTAGAGCTAGAAATAGCAAGT
TAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC

Modular iSBH-sgRNA (trigger A/sgRNA2/loop=24nt):
GCAGGACAGTACAGCGAAACACAACTAAACTGATCATCACCACAACACAACTAA
ACCAGCGTAGGAGTGAGTCGGAGTACTGTCCTGTTTTAGAGCTAGAAATAGCAA
GTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC

Modular iSBH-sgRNA (trigger A/sgRNA2/loop=26nt):
GCAGGACAGTACAGCGAAACACAACTAAACTGATCATCACCACAACACAACTAA
ACTGCAGCGTAGGAGTGAGTCGGAGTACTGTCCTGTTTTAGAGCTAGAAATAG
CAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGG
TGC

Modular iSBH-sgRNA (trigger A/sgRNA2/loop=28nt):
GCAGGACAGTACAGCGAAAACTGATCATCACCACAACACAACTAAACTGATCAT
CACCACGGTCTTGAGAAGTAGTCGGAGTACTGTCCTGTTTTAGAGCTAGAAATA
GCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCG
GTGC

Modular iSBH-sgRNA (trigger A/sgRNA2/loop=30nt):
GCAGGACAGTACAGCGAAAACTGATCATCACCACAACACAACTAAACTGATCAT
CACCACAAGGTCTTGAGAAGTAGTCGGAGTACTGTCCTGTTTTAGAGCTAGAAA
TAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGT
CGGTGC

Modular iSBH-sgRNA (trigger A/sgRNA2/loop=32nt):
GCAGGACAGTACAGCGAAAACTGATCATCACCACAACACAACTAAACTGATCAT
CACCACAACAGGTCTTGAGAAGTAGTCGGAGTACTGTCCTGTTTTAGAGCTAGA
AATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGA
GTCGGTGC

Modular iSBH-sgRNA (trigger A/sgRNA2/loop=34nt):
GCAGGACAGTACAGCGAAAACTGATCATCACCACAACACAACTAAACTGATCAT
CACCACAACACAGGTCTTGAGAAGTAGTCGGAGTACTGTCCTGTTTTAGAGCTA
GAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACC
GAGTCGGTGC

Modular iSBH-sgRNA (trigger B/sgRNA2/loop=14nt (1)):
GCAGGACAGTACGACGAGATGAAGTGAAAAACAGAGGAAGGAAGGAGTTGGTC

AAGTCAAGTCGGAGTACTGTCCTGTTTTAGAGCTAGAAATAGCAAGTTAAAATAA
GGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC

Modular iSBH-sgRNA (trigger B/sgRNA2/loop=14nt (2)):
GCAGGACAGTACAACGAACAACGACCAAATGTCTCTAGTCATCCCCACAGGTG
GAAGTTAGTCGGAGTACTGTCCTGTTTTAGAGCTAGAAATAGCAAGTTAAAATAA
GGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC

Modular iSBH-sgRNA (trigger B/sgRNA2/loop=16nt):
GCAGGACAGTACAGCGAGGGGCAGATGAAGTGAAAAACAGAGGAAGGACACG
GCATAGGCCAGTCGGAGTACTGTCCTGTTTTAGAGCTAGAAATAGCAAGTTAAA
ATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC

Modular iSBH-sgRNA (trigger B/sgRNA2/loop=18nt):
GCAGGACAGTACAGCGAGGGCAGATGAAGTGAAAAACAGAGGAAGGAAGGTC
AAGTCAGATGCAGTCGGAGTACTGTCCTGTTTTAGAGCTAGAAATAGCAAGTTA
AAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC

Modular iSBH-sgRNA (trigger B/sgRNA2/loop=20nt):
GCAGGACAGTACAGCGAGGGGCAGATGAAGTGAAAAACAGAGGAAGGAAGGA
CACGGCATAGGCCAGTCGGAGTACTGTCCTGTTTTAGAGCTAGAAATAGCAAGT
TAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC

Modular iSBH-sgRNA (trigger B/sgRNA2/loop=22nt):
GCAGGACAGTACAGCGAGATGAAGTGAAAAACAGAGGAAGGAAGGATTGTGG
GAGTTGGTCAAGTCAAGTCGGAGTACTGTCCTGTTTTAGAGCTAGAAATAGCAA
GTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC

Modular iSBH-sgRNA (trigger B/sgRNA2/loop=24nt):
GCAGGACAGTACAGCGAGATGGGGGCAGATGAAGTGAAAAACAGAGGAAGGA
AGGATCAGATGCAACCAAGTCGGAGTACTGTCCTGTTTTAGAGCTAGAAATAGC
AAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC

Modular iSBH-sgRNA (trigger B/sgRNA2/loop=26nt):
GCAGGACAGTACAGCGAGGCGATCACAACGACCAAATGTCTCTAGTCATCCCC
TGCTTGTCAGTGTATTCGAGTCGGAGTACTGTCCTGTTTTAGAGCTAGAAATAG
CAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC

Modular iSBH-sgRNA (trigger B/sgRNA2/loop=28nt):
GCAGGACAGTACAGCGAGATGAAGTGAAAAACAGAGGAAGGAAGGATTGTGG
GATTATGAGTTGGTCAAGTCAAGTCGGAGTACTGTCCTGTTTTAGAGCTAGAAA
TAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTG
C

Modular iSBH-sgRNA (trigger C/sgRNA2/loop=14nt (1)):
GCAGGACAGTACGACGAAGACAGACTGTGAAAGGAGAGGCGAGGGATTTATCA
GGATGTAGTCGGAGTACTGTCCTGTTTTAGAGCTAGAAATAGCAAGTTAAAATA
AGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC

Modular iSBH-sgRNA (trigger C/sgRNA2/loop=14nt (2)):
GCAGGACAGTACGACGAAGTGCGCAGCGCGCCTCTCTCTCTCGCTCGGCAAG
CTAAGCAAGTCGGAGTACTGTCCTGTTTTAGAGCTAGAAATAGCAAGTTAAAAT
AAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC

Modular iSBH-sgRNA (trigger C/sgRNA2/loop=16nt):

GCAGGACAGTACAGCGAACAGACAGACTGTGAAAGGAGAGGCGAGGGATCAA
TGTCAATCTAGTCGGAGTACTGTCCTGTTTTAGAGCTAGAAATAGCAAGTTAAAA
TAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC

Modular iSBH-sgRNA (trigger C/sgRNA2/loop=18nt):
GCAGGACAGTACAGCGAACAGACAGACTGTGAAAGGAGAGGCGAGGGAGGTC
AATGTCAATCTAGTCGGAGTACTGTCCTGTTTTAGAGCTAGAAATAGCAAGTTAA
AATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC

Modular iSBH-sgRNA (trigger C/sgRNA2/loop=20nt):
GCAGGACAGTACAGCGAACAGACAGACTGTGAAAGGAGAGGCGAGGGAGGAG
TCAATGTCGTTCTAGTCGGAGTACTGTCCTGTTTTAGAGCTAGAAATAGCAAGTT
AAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC

Modular iSBH-sgRNA (trigger C/sgRNA2/loop=22nt):
GCAGGACAGTACAGCGAACAGACAGACTGTGAAAGGAGAGGCGAGGGAGGAG
GGTCAATGTCAATCTAGTCGGAGTACTGTCCTGTTTTAGAGCTAGAAATAGCAA
GTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC

Modular iSBH-sgRNA (trigger C/sgRNA2/loop=24nt):
GCAGGACAGTACAGCGAAGACAGACTGTGAAAGGAGAGGCGAGGGAGGAGGG
ATGATTTATCAGGATGTAGTCGGAGTACTGTCCTGTTTTAGAGCTAGAAATAGC
AAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGT
GC

Modular iSBH-sgRNA (trigger C/sgRNA2/loop=26nt):
GCAGGACAGTACAGCGAACAGACAGACTGTGAAAGGAGAGGCGAGGGAGGAG
GGATGATCAATGTCGTTCTAGTCGGAGTACTGTCCTGTTTTAGAGCTAGAAATA
GCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCG
GTGC

Modular iSBH-sgRNA (trigger C/sgRNA2/loop=32nt):
GCAGGACAGTACAGCGAGACTTTGACAGACAGACTGTGAAAGGAGAGGCGAG
GGAGGAGGGATGCTGGATGTATAAGAGTCGGAGTACTGTCCTGTTTTAGAGCT
AGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCAC
CGAGTCGGTGC

Modular iSBH-sgRNA (trigger D/sgRNA2/loop=14nt (1)):
GCAGGACAGTACAGCGAGACAGTGTTCTCTGCACAGATAAGGACAAGCAATGA
AATCTGAGTCGGAGTACTGTCCTGTTTTAGAGCTAGAAATAGCAAGTTAAAATAA
GGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC

Modular iSBH-sgRNA (trigger D/sgRNA2/loop=14nt (2)):
GCAGGACAGTACAGCGAAGTGTTCTCTGCACAGATAAGGACAAACATGTTAAGA
CTACAAGTCGGAGTACTGTCCTGTTTTAGAGCTAGAAATAGCAAGTTAAAATAA
GGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC

Modular iSBH-sgRNA (trigger D/sgRNA2/loop=16nt):
GCAGGACAGTACAGCGAGACAGTGTTCTCTGCACAGATAAGGACAAACGCAAT
GAAATCTGAGTCGGAGTACTGTCCTGTTTTAGAGCTAGAAATAGCAAGTTAAAA
TAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC

Modular iSBH-sgRNA (trigger D/sgRNA2/loop=18nt):
GCAGGACAGTACAGCGAAGACAGTGTTCTCTGCACAGATAAGGACAAACACAG
TTAACTATGTAGTCGGAGTACTGTCCTGTTTTAGAGCTAGAAATAGCAAGTTAAA

ATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC

Modular iSBH-sgRNA (trigger D/sgRNA2/loop=20nt):
GCAGGACAGTACAGCGAAGCAAATGTGTTGCCAAAAAGGATGCTTTAGAGACG
GCTTCACTGTTGAGTCGGAGTACTGTCCTGTTTTAGAGCTAGAAATAGCAAGTT
AAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC

Modular iSBH-sgRNA (trigger D/sgRNA2/loop=22nt):
GCAGGACAGTACAGCGAGAGACAGTGTTCTCTGCACAGATAAGGACAAACATT
AAGATTACAAGGTCAGTCGGAGTACTGTCCTGTTTTAGAGCTAGAAATAGCAAG
TTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC

Modular iSBH-sgRNA (trigger D/sgRNA2/loop=24nt):
GCAGGACAGTACAGCGAAGAGACAGTGTTCTCTGCACAGATAAGGACAAACAT
TATGAGTTCACGATCTAGTCGGAGTACTGTCCTGTTTTAGAGCTAGAAATAGCA
AGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTG
C

Modular iSBH-sgRNA (trigger D/sgRNA2/loop=34nt):
GCAGGACAGTACAGCGAGACAGTGTTCTCTGCACAGATAAGGACAAACATTATT
CAGAGGGAGTACGCATTGAAATCTGAGTCGGAGTACTGTCCTGTTTTAGAGCTA
GAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACC
GAGTCGGTGC

*An additional GC sequence was added to all iSBH-sgRNA sequences. G helps expression from U6 promoter, while C base-pairs with the first C from the scaffold

RNA triggers for first-generation iSBH-sgRNAs

hairpin linker 34nt trigger 100nt 3' flank linker hairpin
random sequence

Non-complementary RNA trigger (transfected in all OFF states):
GGCTCGAAAGAGCCACGGGTCTTCGAGCATTGGAAGTCTAGATACCTGCTTTA
CGAAGACCTATGCTAGTACGAAAGTACTAGC

First-generation iSBH-sgRNA 1 trigger:
GGCTCGAAAGAGCCACGGATGCGTAGTGTGCAGAAGCGACTAGCGAAGCAAT
GCTAGTACGAAAGTACTAGC

First-generation iSBH-sgRNA 2 trigger:
GGCTCGAAAGAGCCACTGTTAGGGATGTTCGTATCTCGCTGTACTGTCCTGCAT
GCTAGTACGAAAGTACTAGC
*GC complementary with extra GC sequences on the iSBH-sgRNA 1

First-generation iSBH-sgRNA 3 trigger:
GGCTCGAAAGAGCCACTGTATTCACTGGTCCATTCCATACTTCCCGCCTCATGC
TAGTACGAAAGTACTAGC

First-generation iSBH-sgRNA 4 trigger:
GGCTCGAAAGAGCCACTTGATGTGGGGAGCCCATTTAAGTACTAGCAGGAATG
CTAGTACGAAAGTACTAGC

First-generation iSBH-sgRNA 5 trigger:
GGCTCGAAAGAGCCACTGGACGAACAAGATGCAGAACTGGAGTAAGAGTCATG
CTAGTACGAAAGTACTAGC

First-generation iSBH-sgRNA 1 trigger 100nt 3' flank:
GGCTCGAAAGAGCCACTGAGACGTAAGACTCGAGCGGTCCGTCTCAGTCAGTA
TGGTCTTGGGATGCGTAGTGTGCAGAAGCGACTAGCGAAGCAGGCGGCCGTA
TGAGGGACAATTGGAGaagtgaattatataaatataaagtagtaaaaattgaaccattaggagtagcacc
caccaagGCAAAGAGAAGAGTGGTGCAGGTACGAAAGTAC

First-generation iSBH-sgRNA 2 trigger 100nt 3' flank:
GGCTCGAAAGAGCCACTGAGACGTAAGACTCGAGCGGTCCGTCTCAGTCATAG
TATATGTGTGTTAGGGATGTTCGTATCTCGCTGTACTGTCCTGGCGGCCGCATA
GTTAAGCCAGTATCTGCtccctgcttgtgtgttggaggtcgctgagtagtgcgcgagcaaaatttaagctac
aacaaGGCAAGGCTTGACCGACAAGTACGAAAGTAC

First-generation iSBH-sgRNA 3 trigger 100nt 3' flank:
GGCTCGAAAGAGCCACTGAGACGTAAGACTCGAGCGGTCCGTCTCAGTCACTA
TCTATATCTGTATTCACTGGTCCATTCCATACTTCCCGCCTCGGCGGCCGTCGG
GAGATCTCCCGATCCCCtatggtgcactctcagtacaatctgctctgatgccgcatagttaagccagtatct
GCTCCCTGCTTGTGTGTTGGAGGTGTACGAAAGTAC

RNA triggers for second-generation iSBH-sgRNAs

hairpin linker 100nt 5' flank 44nt trigger 100nt 3' flank linker hairpin

Second-generation iSBH-sgRNA 1 trigger:
GGCTCGAAAGAGCCACGGATGCGTAGTGTGAACGTACGGACAGAAGCGACTA
GCGAAGCAATGCTAGTACGAAAGTACTAGC

Second-generation iSBH-sgRNA 2 trigger:
GGCTCGAAAGAGCCACTGTTAGGGATGTTCGACTAAGCTGGTATCTCGCTGTA
CTGTCCTATGCTAGTACGAAAGTACTAGC

Second-generation iSBH-sgRNA 3 trigger:
GGCTCGAAAGAGCCACTGTATTCACTGGTCCCATCCCACTCATTCCATACTTCC
CGCCTCATGCTAGTACGAAAGTACTAGC

Second-generation iSBH-sgRNA 1 trigger 100nt 3' flank:
GGCTCGAAAGAGCCACTGAGACGTAAGACTCGAGCGGTCCGTCTCAGTCAGTA
TGGTCTTGGGATGCGTAGTGTGAACGTACGGACAGAAGCGACTAGCGAAGCAG
CGGCCGTATGAGGGACAATTGGAGaagtgaattatataaatataaagtagtaaaaattgaaccatta
ggagtagcacccaccaagGCAAAGAGAAGAGTGGTGCAGGTACGAAAGTAC

Second-generation iSBH-sgRNA 2 trigger 100nt 3' flank:
GGCTCGAAAGAGCCACTGAGACGTAAGACTCGAGCGGTCCGTCTCAGTCATAG
TATATGTGTGTTAGGGATGTTCGACTAAGCTGGTATCTCGCTGTACTGTCCTGC
GGCCGCATAGTTAAGCCAGTATCTGCtccctgcttgtgtgttggaggtcgctgagtagtgcgcgagc
aaaatttaagctacaacaaGGCAAGGCTTGACCGACAAGTACGAAAGTAC

Second-generation iSBH-sgRNA 3 trigger 100nt 3' flank:

GGCTCGAAAGAGCCACTGAGACGTAAGACTCGAGCGGTCCGTCTCAGTCATCG
GTCTTGTCTGTATTCACTGGTCCCATCCCACTCATTCCATACTTCCCGCCTCGC
GGCCGTCGGGAGATCTCCCGATCCCCtatggtgcactctcagtacaatctgctctgatgccgcatag
ttaagccagtatctGCTCCCTGCTTGTGTGTTGGAGGTGTACGAAAGTAC

Second-generation iSBH-sgRNA 1 trigger 100nt 5' flank:
GGCTCGAAAGAGCCACCCTATGGTGCACTCTCAGTACaatctgctctgatgccgcatagtta
agccagtatctgctccctgcttgtgtgttggaGGTCGCTGAGTAGTGCGCGAGCGTCAGTATGG
TCTTGGGATGCGTAGTGTGAACGTACGGACAGAAGCGACTAGCGAAGCAGCG
GCCGCATGCTAGTACGAAAGTACTAGC

Second-generation iSBH-sgRNA 2 trigger 100nt 5' flank:
GGCTCGAAAGAGCCACGCAAAATTTAAGCTACAACAAGGCaaggcttgaccgacaattg
catgaagaatctgcttagggttaggcgttttgcgCTGCTTCGCGATGTACGGGCCAGGTCATAGT
ATATGTGTGTTAGGGATGTTCGACTAAGCTGGTATCTCGCTGTACTGTCCTGCG
GCCGCATGCTAGTACGAAAGTACTAGC

Second-generation iSBH-sgRNA 3 trigger 100nt 5' flank:
GGCTCGAAAGAGCCACGTCGACGGATCGGGAGATCTCCCGATCCCctatggtgcact
ctcagtacaatctgctctgatgccgcatagttaagcCAGTATCTGCTCCCTGCTTGTGTGTCATCG
GTCTTGTCTGTATTCACTGGTCCCATCCCACTCATTCCATACTTCCCGCCTCGC
GGCCGCATGCTAGTACGAAAGTACTAGC

Second-generation iSBH-sgRNA 1 trigger 100nt 5'+3' flank:
GGCTCGAAAGAGCCACCCTATGGTGCACTCTCAGTACAATCTGCTCTGATGCC
GCATAGTTAAGCCAGTATCTGCTCCCTGCTTGTGTGTTGGAGGTCGCTGAGTAG
TGCGCGAGCGTCAGTATGGTCTTGGGATGCGTAGTGTGAACGTACGGACAGAA
GCGACTAGCGAAGCAGCGGCCGTATGAGGGACAATTGGAGAAGTGAATTATAT
AAATATAAAGTAGTAAAAATTGAACCATTAGGAGTAGCACCCACCAAGGCAAAG
AGAAGAGTGGTGCAGGTACGAAAGTAC

Second-generation iSBH-sgRNA 2 trigger 100nt 5'+3' flank:
GGCTCGAAAGAGCCACGCAAAATTTAAGCTACAACAAGGCAAGGCTTGACCGA
CAATTGCATGAAGAATCTGCTTAGGGTTAGGCGTTTTGCGCTGCTTCGCGATGT
ACGGGCCAGGTCATAGTATATGTGTGTTAGGGATGTTCGACTAAGCTGGTATCT
CGCTGTACTGTCCTGCGGCCGCATAGTTAAGCCAGTATCTGCTCCCTGCTTGT
GTGTTGGAGGTCGCTGAGTAGTGCGCGAGCAAAATTTAAGCTACAACAAGGCA
AGGCTTGACCGACAAGTACGAAAGTAC

Second-generation iSBH-sgRNA 3 trigger 100nt 5'+3' flank:
GGCTCGAAAGAGCCACGTCGACGGATCGGGAGATCTCCCGATCCCCTATGGT
GCACTCTCAGTACAATCTGCTCTGATGCCGCATAGTTAAGCCAGTATCTGCTCC
CTGCTTGTGTGTCATCGGTCTTGTCTGTATTCACTGGTCCCATCCCACTCATTC
CATACTTCCCGCCTCGCGGCCGTCGGGAGATCTCCCGATCCCCTATGGTGCAC
TCTCAGTACAATCTGCTCTGATGCCGCATAGTTAAGCCAGTATCTGCTCCCTGC
TTGTGTGTTGGAGGTGTACGAAAGTAC

Longer RNA triggers with engineered secondary structures for second-generation iSBH-sgRNAs

hairpin linker that base-pairs with 44nt trigger 44nt trigger 100nt 3' flank linker hairpin

Second-generation iSBH-sgRNA 1 trigger 100nt 3’ flank (60%):
GGCTCGAAAGAGCCACTGAGACGTAAGACTCGAGCGGTCCGTCTCAGTCAGTA
TGGTCTTCGGATGCGTAGTGTGAACGTACGGACAGAAGCGACTAGCGAAGCAG
CGGCCGTATGAGGGACAATTGGAGaagtgaattatataaatataaagtagtaaaaattgaaccatta
ggagtagcacccaccaagGCAAAGAGAAGAGTGGTGCAGGTACGAAAGTAC

Second-generation iSBH-sgRNA 2 trigger 100nt 3’ flank (60%):
GGCTCGAAAGAGCCACTGATCCCTAAGACTCGAGCGGTCCGTCTCAGTCATAG
TATATGTGTGTTAGGGATGTTCGACTAAGCTGGTATCTCGCTGTACTGTCCTGC
GGCCGCATAGTTAAGCCAGTATCTGCTCCCTGCTTGTGTGTTGGAGGTCGCTG
AGTAGTGCGCGAGCAAAATTTAAGCTACAACAAGGCAAGGCTTGACCGACAAG
TACGAAAGTAC

Second-generation iSBH-sgRNA 3 trigger 100nt 3’ flank (60%):
GGCTCGAAAGAGCCACTGAGTGAATAGACTCGAGCGGTCCGTCTCAGTCATCG
GTCTTGTCTGTATTCACTGGTCCCATCCCACTCATTCCATACTTCCCGCCTCGC
GGCCGTCGGGAGATCTCCCGATCCCCTATGGTGCACTCTCAGTACAATCTGCT
CTGATGCCGCATAGTTAAGCCAGTATCTGCTCCCTGCTTGTGTGTTGGAGGTGT
ACGAAAGTAC

Second-generation iSBH-sgRNA 1 trigger 100nt 3’ flank (80%):
GGCTCGAAAGAGCCACTGAGACACTACGCTCGAGCGGTCCGTCTCAGTCAGTA
TGGTCTTCGGATGCGTAGTGTGAACGTACGGACAGAAGCGACTAGCGAAGCAG
CGGCCGTATGAGGGACAATTGGAGAAGTGAATTATATAAATATAAAGTAGTAAA
AATTGAACCATTAGGAGTAGCACCCACCAAGGCAAAGAGAAGAGTGGTGCAGG
TACGAAAGTAC

Second-generation iSBH-sgRNA 2 trigger 100nt 3’ flank (80%):
GGCTCGAAAGAGCCACTGAGAACATCCCCTCGAGCGGTCCGTCTCAGTCATAG
TATATGTGTGTTAGGGATGTTCGACTAAGCTGGTATCTCGCTGTACTGTCCTGC
GGCCGCATAGTTAAGCCAGTATCTGCTCCCTGCTTGTGTGTTGGAGGTCGCTG
AGTAGTGCGCGAGCAAAATTTAAGCTACAACAAGGCAAGGCTTGACCGACAAG
TACGAAAGTAC

Second-generation iSBH-sgRNA 3 trigger 100nt 3’ flank (80%):
GGCTCGAAAGAGCCACTGAGACCAGTGACTCGAGCGGTCCGTCTCAGTCATCG
GTCTTGTCTGTATTCACTGGTCCCATCCCACTCATTCCATACTTCCCGCCTCGC
GGCCGTCGGGAGATCTCCCGATCCCCTATGGTGCACTCTCAGTACAATCTGCT
CTGATGCCGCATAGTTAAGCCAGTATCTGCTCCCTGCTTGTGTGTTGGAGGTGT
ACGAAAGTAC

Second-generation iSBH-sgRNA 1 trigger 100nt 3’ flank (100%):
GGCTCGAAAGAGCCACCACACTACGCATCCCGAGCGGTCCGTCTCAGTCAGTA
TGGTCTTGGGATGCGTAGTGTGAACGTACGGACAGAAGCGACTAGCGAAGCAG
CGGCCGTATGAGGGACAATTGGAGAAGTGAATTATATAAATATAAAGTAGTAAA
AATTGAACCATTAGGAGTAGCACCCACCAAGGCAAAGAGAAGAGTGGTGCAGG
TACGAAAGTAC

Second-generation iSBH-sgRNA 2 trigger 100nt 3’ flank (100%):
GGCTCGAAAGAGCCACGAACATCCCTAACACGAGCGGTCCGTCTCAGTCATAG
TATATGTGTGTTAGGGATGTTCGACTAAGCTGGTATCTCGCTGTACTGTCCTGC
GGCCGCATAGTTAAGCCAGTATCTGCTCCCTGCTTGTGTGTTGGAGGTCGCTG
AGTAGTGCGCGAGCAAAATTTAAGCTACAACAAGGCAAGGCTTGACCGACAAG
TACGAAAGTAC

Second-generation iSBH-sgRNA 3 trigger 100nt 3' flank (100%):
GGCTCGAAAGAGCCACGACCAGTGAATACACGAGCGGTCCGTCTCAGTCATCG
GTCTTGTCTGTATTCACTGGTCCCATCCCACTCATTCCATACTTCCCGCCTCGC
GGCCGTCGGGAGATCTCCCGATCCCCTATGGTGCACTCTCAGTACAATCTGCT
CTGATGCCGCATAGTTAAGCCAGTATCTGCTCCCTGCTTGTGTGTTGGAGGTGT
ACGAAAGTAC

RNA triggers for second-generation iSBH-sgRNA expressed from CMV promoters

CMV enhancer CMV promoter linker hairpin linker trigger linker hairpin linker bGH poly(A) signal MALAT1 terminator

Second-generation iSBH-sgRNA 1 (CMV+pA) trigger:

GACATTGATTATTGACTAGTTATTAATAGTAATCAATTACGGGGTCATTAGTTCAT
AGCCCATATATGGAGTTCCGCGTTACATAACTTACGGTAAATGGCCCGCCTGGC
TGACCGCCCAACGACCCCCGCCCATTGACGTCAATAATGACGTATGTTCCCATA
GTAACGCCAATAGGGACTTTCCATTGACGTCAATGGGTGGAGTATTTACGGTAA
ACTGCCCACTTGGCAGTACATCAAGTGTATCATATGCCAAGTACGCCCCCTATT
GACGTCAATGACGGTAAATGGCCCGCCTGGCATTATGCCCAGTACATGACCTT
ATGGGACTTTCCTACTTGGCAGTACATCTACGTATTAGTCATCGCTATTACCATG
GTGATGCGGTTTTGGCAGTACATCAATGGGCGTGGATAGCGGTTTGACTCACG
GGGATTTCCAAGTCTCCACCCCATTGACGTCAATGGGAGTTTGTTTTGGCACCA
AAATCAACGGGACTTTCCAAAATGTCGTAACAACTCCGCCCCATTGACGCAAAT
GGGCGGTAGGCGTGTACGGTGGGAGGTCTATATAAGCAGAGCTCTCTGGCTAA
CTAGAGAACCCACTGCTTACTGGCTTATCGAAATTAATACGACTCACTATAGGG
AGACCCAAGCTGGCTAGCGTTTAAACTTAAGGCTCGAAAGAGCCACGGATGCG
TAGTGTGAACGTACGGACAGAAGCGACTAGCGAAGCAATGCTAGTACGAAAGT
ACTAGCTTTTTTGCTAGCAGGCATGCTGGGGGCCGCTCGAGTCTAGAGGGCCC
GTTTAAACCCGCTGATCAGCCTCGACTGTGCCTTCTAGTTGCCAGCCATCTGTT
GTTTGCCCCTCCCCCGTGCCTTCCTTGACCCTGGAAGGTGCCACTCCCACTGT
CCTTTCCTAATAAAATGAGGAAATTGCATCGCATTGTCTGAGTAGGTGTCATTCT
ATTCTGGGGGGTGGGGTGGGGCAGGACAGCAAGGGGGAGGATTGGGAAGAC
AATAGCAGGCATGCTGGGGATGCGGTGGGCTCTATGG

Second-generation iSBH-sgRNA 1 (CMV+MALAT1) trigger:
GACATTGATTATTGACTAGTTATTAATAGTAATCAATTACGGGGTCATTAGTTCAT
AGCCCATATATGGAGTTCCGCGTTACATAACTTACGGTAAATGGCCCGCCTGGC
TGACCGCCCAACGACCCCCGCCCATTGACGTCAATAATGACGTATGTTCCCATA
GTAACGCCAATAGGGACTTTCCATTGACGTCAATGGGTGGAGTATTTACGGTAA
ACTGCCCACTTGGCAGTACATCAAGTGTATCATATGCCAAGTACGCCCCCTATT
GACGTCAATGACGGTAAATGGCCCGCCTGGCATTATGCCCAGTACATGACCTT
ATGGGACTTTCCTACTTGGCAGTACATCTACGTATTAGTCATCGCTATTACCATG
GTGATGCGGTTTTGGCAGTACATCAATGGGCGTGGATAGCGGTTTGACTCACG
GGGATTTCCAAGTCTCCACCCCATTGACGTCAATGGGAGTTTGTTTTGGCACCA
AAATCAACGGGACTTTCCAAAATGTCGTAACAACTCCGCCCCATTGACGCAAAT
GGGCGGTAGGCGTGTACGGTGGGAGGTCTATATAAGCAGAGCTCTCTGGCTAA
CTAGAGAACCCACTGCTTACTGGCTTATCGAAATTAATACGACTCACTATAGGG
AGACCCAAGCTGGCTAGCGTTTAAACTTAAGGCTCGAAAGAGCCACGGATGCG
TAGTGTGAACGTACGGACAGAAGCGACTAGCGAAGCAATGCTAGTACGAAAGT
ACTAGCTTTTTTGCTAGCAGGCATGCTGGGGGCCGCGATTCGTCAGTAGGGTT
GTAAAGGTTTTTCTTTTCCTGAGAAAACAACCTTTTGTTTTCTCAGGTTTTGCTTT

TTGGCCTTTCCCTAGCTTTAAAAAAAAAAAAGCAAAAGACGCTGGTGGCTGGCA
CTCCTGGTTTCCAGGACGGGGTTCAAGTCCCTGCGGTGTCTTTGCTT

Second-generation iSBH-sgRNA 2 (CMV+pA) trigger:
GACATTGATTATTGACTAGTTATTAATAGTAATCAATTACGGGGTCATTAGTTCAT
AGCCCATATATGGAGTTCCGCGTTACATAACTTACGGTAAATGGCCCGCCTGGC
TGACCGCCCAACGACCCCCGCCCATTGACGTCAATAATGACGTATGTTCCCATA
GTAACGCCAATAGGGACTTTCCATTGACGTCAATGGGTGGAGTATTTACGGTAA
ACTGCCCACTTGGCAGTACATCAAGTGTATCATATGCCAAGTACGCCCCCTATT
GACGTCAATGACGGTAAATGGCCCGCCTGGCATTATGCCCAGTACATGACCTT
ATGGGACTTTCCTACTTGGCAGTACATCTACGTATTAGTCATCGCTATTACCATG
GTGATGCGGTTTTGGCAGTACATCAATGGGCGTGGATAGCGGTTTGACTCACG
GGGATTTCCAAGTCTCCACCCCATTGACGTCAATGGGAGTTTGTTTTGGCACCA
AAATCAACGGGACTTTCCAAAATGTCGTAACAACTCCGCCCCATTGACGCAAAT
GGGCGGTAGGCGTGTACGGTGGGAGGTCTATATAAGCAGAGCTCTCTGGCTAA
CTAGAGAACCCACTGCTTACTGGCTTATCGAAATTAATACGACTCACTATAGGG
AGACCCAAGCTGGCTAGCGTTTAAACTTAAGGCTCGAAAGAGCCACTGTTAGG
GATGTTCGACTAAGCTGGTATCTCGCTGTACTGTCCTATGCTAGTACGAAAGTA
CTAGCTTTTTTGCTAGCAGGCATGCTGGGGGCCGCTCGAGTCTAGAGGGCCCG
TTTAAACCCGCTGATCAGCCTCGACTGTGCCTTCTAGTTGCCAGCCATCTGTTG
TTTGCCCCTCCCCCGTGCCTTCCTTGACCCTGGAAGGTGCCACTCCCACTGTC
CTTTCCTAATAAAATGAGGAAATTGCATCGCATTGTCTGAGTAGGTGTCATTCTA
TTCTGGGGGGTGGGGTGGGGCAGGACAGCAAGGGGGAGGATTGGGAAGACA
ATAGCAGGCATGCTGGGGATGCGGTGGGCTCTATGG

Second-generation iSBH-sgRNA 2 (CMV+MALAT1) trigger:
GACATTGATTATTGACTAGTTATTAATAGTAATCAATTACGGGGTCATTAGTTCAT
AGCCCATATATGGAGTTCCGCGTTACATAACTTACGGTAAATGGCCCGCCTGGC
TGACCGCCCAACGACCCCCGCCCATTGACGTCAATAATGACGTATGTTCCCATA
GTAACGCCAATAGGGACTTTCCATTGACGTCAATGGGTGGAGTATTTACGGTAA
ACTGCCCACTTGGCAGTACATCAAGTGTATCATATGCCAAGTACGCCCCCTATT
GACGTCAATGACGGTAAATGGCCCGCCTGGCATTATGCCCAGTACATGACCTT
ATGGGACTTTCCTACTTGGCAGTACATCTACGTATTAGTCATCGCTATTACCATG
GTGATGCGGTTTTGGCAGTACATCAATGGGCGTGGATAGCGGTTTGACTCACG
GGGATTTCCAAGTCTCCACCCCATTGACGTCAATGGGAGTTTGTTTTGGCACCA
AAATCAACGGGACTTTCCAAAATGTCGTAACAACTCCGCCCCATTGACGCAAAT
GGGCGGTAGGCGTGTACGGTGGGAGGTCTATATAAGCAGAGCTCTCTGGCTAA
CTAGAGAACCCACTGCTTACTGGCTTATCGAAATTAATACGACTCACTATAGGG
AGACCCAAGCTGGCTAGCGTTTAAACTTAAGGCTCGAAAGAGCCACTGTTAGG
GATGTTCGACTAAGCTGGTATCTCGCTGTACTGTCCTATGCTAGTACGAAAGTA
CTAGCTTTTTTGCTAGCAGGCATGCTGGGGGCCGCGATTCGTCAGTAGGGTTG
TAAAGGTTTTTCTTTTCCTGAGAAAACAACCTTTTGTTTTCTCAGGTTTTGCTTTT
TGGCCTTTCCCTAGCTTTAAAAAAAAAAAAGCAAAAGACGCTGGTGGCTGGCAC
TCCTGGTTTCCAGGACGGGGTTCAAGTCCCTGCGGTGTCTTTGCTT

Second-generation iSBH-sgRNA 3 (CMV+pA) trigger:
GACATTGATTATTGACTAGTTATTAATAGTAATCAATTACGGGGTCATTAGTTCAT
AGCCCATATATGGAGTTCCGCGTTACATAACTTACGGTAAATGGCCCGCCTGGC
TGACCGCCCAACGACCCCCGCCCATTGACGTCAATAATGACGTATGTTCCCATA
GTAACGCCAATAGGGACTTTCCATTGACGTCAATGGGTGGAGTATTTACGGTAA
ACTGCCCACTTGGCAGTACATCAAGTGTATCATATGCCAAGTACGCCCCCTATT
GACGTCAATGACGGTAAATGGCCCGCCTGGCATTATGCCCAGTACATGACCTT
ATGGGACTTTCCTACTTGGCAGTACATCTACGTATTAGTCATCGCTATTACCATG
GTGATGCGGTTTTGGCAGTACATCAATGGGCGTGGATAGCGGTTTGACTCACG
GGGATTTCCAAGTCTCCACCCCATTGACGTCAATGGGAGTTTGTTTTGGCACCA

AAATCAACGGGACTTTCCAAAATGTCGTAACAACTCCGCCCCATTGACGCAAAT
GGGCGGTAGGCGTGTACGGTGGGAGGTCTATATAAGCAGAGCTCTCTGGCTAA
CTAGAGAACCCACTGCTTACTGGCTTATCGAAATTAATACGACTCACTATAGGG
AGACCCAAGCTGGCTAGCGTTTAAACTTAAGGCTCGAAAGAGCCACTGTATTCA
CTGGTCCCATCCCACTCATTCCATACTTCCCGCCTCATGCTAGTACGAAAGTAC
TAGCTTTTTTGCTAGCAGGCATGCTGGGGGCCGCTCGAGTCTAGAGGGCCCGT
TTAAACCCGCTGATCAGCCTCGACTGTGCCTTCTAGTTGCCAGCCATCTGTTGT
TTGCCCCTCCCCCGTGCCTTCCTTGACCCTGGAAGGTGCCACTCCCACTGTCC
TTTCCTAATAAAATGAGGAAATTGCATCGCATTGTCTGAGTAGGTGTCATTCTAT
TCTGGGGGGTGGGGTGGGGCAGGACAGCAAGGGGGAGGATTGGGAAGACAA
TAGCAGGCATGCTGGGGATGCGGTGGGCTCTATGG

Second-generation iSBH-sgRNA 3 (CMV+MALAT1) trigger:
GACATTGATTATTGACTAGTTATTAATAGTAATCAATTACGGGGTCATTAGTTCAT
AGCCCATATATGGAGTTCCGCGTTACATAACTTACGGTAAATGGCCCGCCTGGC
TGACCGCCCAACGACCCCCGCCCATTGACGTCAATAATGACGTATGTTCCCATA
GTAACGCCAATAGGGACTTTCCATTGACGTCAATGGGTGGAGTATTTACGGTAA
ACTGCCCACTTGGCAGTACATCAAGTGTATCATATGCCAAGTACGCCCCTATT
GACGTCAATGACGGTAAATGGCCCGCCTGGCATTATGCCCAGTACATGACCTT
ATGGGACTTTCCTACTTGGCAGTACATCTACGTATTAGTCATCGCTATTACCATG
GTGATGCGGTTTTGGCAGTACATCAATGGGCGTGGATAGCGGTTTGACTCACG
GGGATTTCCAAGTCTCCACCCCATTGACGTCAATGGGAGTTTGTTTTGGCACCA
AAATCAACGGGACTTTCCAAAATGTCGTAACAACTCCGCCCCATTGACGCAAAT
GGGCGGTAGGCGTGTACGGTGGGAGGTCTATATAAGCAGAGCTCTCTGGCTAA
CTAGAGAACCCACTGCTTACTGGCTTATCGAAATTAATACGACTCACTATAGGG
AGACCCAAGCTGGCTAGCGTTTAAACTTAAGGCTCGAAAGAGCCACTGTATTCA
CTGGTCCCATCCCACTCATTCCATACTTCCCGCCTCATGCTAGTACGAAAGTAC
TAGCTTTTTTGCTAGCAGGCATGCTGGGGGCCGCGATTCGTCAGTAGGGTTGT
AAAGGTTTTTCTTTTCCTGAGAAAACAACCTTTTGTTTTCTCAGGTTTTGCTTTTT
GGCCTTTCCCTAGCTTTAAAAAAAAAAAAGCAAAAGACGCTGGTGGCTGGCACT
CCTGGTTTCCAGGACGGGGTTCAAGTCCCTGCGGTGTCTTTGCTT

EYFP mRNA with 3’ UTR trigger sequences

EYFP mRNA linker hairpin linker 100nt 5’ flank linker44nt trigger linker hairpin

EYFP mRNA + second-generation iSBH-sgRNA 1 trigger:
ATGAGCAGCGGCGCCCTGCTGTTCCACGGCAAGATCCCCTACGTGGTGGAGAT
GGAGGGCGATGTGGATGGCCACACCTTCAGCATCCGCGGTAAGGGCTACGGC
GATGCCAGCGTGGGCAAGGTGGATGCCCAGTTCATCTGCACCACCGGCGATG
TGCCCGTGCCCTGGAGCACCCTGGTGACCACCCTGACCTACGGCGCCCAGTG
CTTCGCCAAGTACGGCCCCGAGCTGAAGGATTTCTACAAGAGCTGCATGCCCG
ATGGCTACGTGCAGGAGCGCACCATCACCTTCGAGGGCGATGGCAATTTCAAG
ACCCGCGCCGAGGTGACCTTCGAGAATGGCAGCGTGTACAATCGCGTGAAGCT
GAATGGCCAGGGCTTCAAGAAGGATGGCCACGTGCTGGGCAAGAATCTGGAG
TTCAATTTCACCCCCCACTGCCTGTACATCTGGGGCGATCAGGCCAATCACGG
CCTGAAGAGCGCCTTCAAGATCTGCCACGAGATCGCCGGCAGCAAGGGCGATT
TCATCGTGGCCGATCACACCCAGATGAATACCCCCATCGGCGGCGGCCCCGT
GCACGTGCCCGAGTACCACCACATGAGCTACCACGTGAAGCTGAGCAAGGATG
TGACCGATCACCGCGATAATATGAGCCTGACGGAGACCGTGCGCGCCGTGGAT
TGCCGCAAGACCTACCTGTAACTCGAGTCTAGACCGGCTCGAAAGAGCCACCC
TATGGTGCACTCTCAGTACAATCTGCTCTGATGCCGCATAGTTAAGCCAGTATC
TGCTCCCTGCTTGTGTGTTGGAGGTCGCTGAGTAGTGCGCGAGCGTCAGTATG
GTCTTGGGATGCGTAGTGTGAACGTACGGACAGAAGCGACTAGCGAAGCAGC

GGCCGCATGCTAGTACGAAAGTACT

EYFP mRNA + second-generation iSBH-sgRNA 2 trigger:
ATGAGCAGCGGCGCCCTGCTGTTCCACGGCAAGATCCCCTACGTGGTGGAGAT
GGAGGGCGATGTGGATGGCCACACCTTCAGCATCCGCGGTAAGGGCTACGGC
GATGCCAGCGTGGGCAAGGTGGATGCCCAGTTCATCTGCACCACCGGCGATG
TGCCCGTGCCCTGGAGCACCCTGGTGACCACCCTGACCTACGGCGCCCAGTG
CTTCGCCAAGTACGGCCCCGAGCTGAAGGATTTCTACAAGAGCTGCATGCCCG
ATGGCTACGTGCAGGAGCGCACCATCACCTTCGAGGGCGATGGCAATTTCAAG
ACCCGCGCCGAGGTGACCTTCGAGAATGGCAGCGTGTACAATCGCGTGAAGCT
GAATGGCCAGGGCTTCAAGAAGGATGGCCACGTGCTGGGCAAGAATCTGGAG
TTCAATTTCACCCCCCACTGCCTGTACATCTGGGGCGATCAGGCCAATCACGG
CCTGAAGAGCGCCTTCAAGATCTGCCACGAGATCGCCGGCAGCAAGGGCGATT
TCATCGTGGCCGATCACACCCAGATGAATACCCCCATCGGCGGCGGCCCCGT
GCACGTGCCCGAGTACCACCACATGAGCTACCACGTGAAGCTGAGCAAGGATG
TGACCGATCACCGCGATAATATGAGCCTGACGGAGACCGTGCGCGCCGTGGAT
TGCCGCAAGACCTACCTGTAACTCGAGTCTAGACCGGCTCGAAAGAGCCACGC
AAAATTTAAGCTACAACAAGGCAAGGCTTGACCGACAATTGCATGAAGAATCTG
CTTAGGGTTAGGCGTTTTGCGCTGCTTCGCGATGTACGGGCCAGGTCATAGTA
TATGTGTGTTAGGGATGTTCGACTAAGCTGGTATCTCGCTGTACTGTCCTGCGG
CCGCATGCTAGTACGAAAGTACT

EYFP mRNA + second-generation iSBH-sgRNA 3 trigger:
ATGAGCAGCGGCGCCCTGCTGTTCCACGGCAAGATCCCCTACGTGGTGGAGAT
GGAGGGCGATGTGGATGGCCACACCTTCAGCATCCGCGGTAAGGGCTACGGC
GATGCCAGCGTGGGCAAGGTGGATGCCCAGTTCATCTGCACCACCGGCGATG
TGCCCGTGCCCTGGAGCACCCTGGTGACCACCCTGACCTACGGCGCCCAGTG
CTTCGCCAAGTACGGCCCCGAGCTGAAGGATTTCTACAAGAGCTGCATGCCCG
ATGGCTACGTGCAGGAGCGCACCATCACCTTCGAGGGCGATGGCAATTTCAAG
ACCCGCGCCGAGGTGACCTTCGAGAATGGCAGCGTGTACAATCGCGTGAAGCT
GAATGGCCAGGGCTTCAAGAAGGATGGCCACGTGCTGGGCAAGAATCTGGAG
TTCAATTTCACCCCCCACTGCCTGTACATCTGGGGCGATCAGGCCAATCACGG
CCTGAAGAGCGCCTTCAAGATCTGCCACGAGATCGCCGGCAGCAAGGGCGATT
TCATCGTGGCCGATCACACCCAGATGAATACCCCCATCGGCGGCGGCCCCGT
GCACGTGCCCGAGTACCACCACATGAGCTACCACGTGAAGCTGAGCAAGGATG
TGACCGATCACCGCGATAATATGAGCCTGACGGAGACCGTGCGCGCCGTGGAT
TGCCGCAAGACCTACCTGTAACTCGAGTCTAGACCGGCTCGAAAGAGCCACGT
CGACGGATCGGGAGATCTCCCGATCCCCTATGGTGCACTCTCAGTACAATCTG
CTCTGATGCCGCATAGTTAAGCCAGTATCTGCTCCCTGCTTGTGTGTCATCGGT
CTTGTCTGTATTCACTGGTCCCATCCCACTCATTCCATACTTCCCGCCTCGCGG
CCGCATGCTAGTACGAAAGTACT

Truncated RNA triggers for second-generation iSBH-sgRNAs

hairpin linker trigger randomized nt linker hairpin

trigger + randomized nt= 44nt

Second-generation iSBH-sgRNA 1 trigger (D1):
GGCTCGAAAGAGCCACGGATGCGTAGTGTGAACGTACGGAATCCGTGATACTA
ACGCCGGATGCTAGTACGAAAGTACTAGC

Second-generation iSBH-sgRNA 1 trigger (D2):

GGCTCGAAAGAGCCACGGATGCGTAGTGTGAACGTACGGACAGAATGATACTA
ACGCCGGATGCTAGTACGAAAGTACTAGC

Second-generation iSBH-sgRNA 2 trigger (D2):
GGCTCGAAAGAGCCACTGTTAGGGATGTTCGACTAAGCTGACGCGAATTGCGC
TGCAATAATGCTAGTACGAAAGTACTAGC

Second-generation iSBH-sgRNA 2 trigger (D1):
GGCTCGAAAGAGCCACTGTTAGGGATGTTCGACTAAGCTGGTATCAATTGCGC
TGCAATAATGCTAGTACGAAAGTACTAGC

Second-generation iSBH-sgRNA 3 trigger (D1):
GGCTCGAAAGAGCCACTGTATTCACTGGTCCCATCCCACTATGAATCGCTGAGA
TAGGAAATGCTAGTACGAAAGTACTAGC

Second-generation iSBH-sgRNA 4 trigger (D2):
GGCTCGAAAGAGCCACTGTATTCACTGGTCCCATCCCACTCATTCTCGCTGAGA
TAGGAAATGCTAGTACGAAAGTACTAGC

Endogenous trigger sequences for modular iSBH-sgRNA designs

hairpin linker endogenous trigger sequence linker hairpin

RNA trigger A:
GGCTCGAAAGAGCCACCTGATGATCAGTTTAGTCGTGTTGTGGTGATGATCAGT
TTAGTTGTGTTGTGGTGATGATCAGTTTAGTTGTGTTGTGGTGATGATCAGTTTA
GTCGTGTTGTGGTGATGATCAGTTTAGTCGTGTTGTGGTGATGATCTGTTTAGA
TGCTAGTACGAAAGTACTAGC

RNA trigger B:
GGCTCGAAAGAGCCACCTCAACAAAAGGTTAAAAATGGTTTAGTCACATGTATG
CTAACTTTCAATCTACAAAACTACAACCATCATCATAATCCCACAATCCTTCCTTC
CTCTGTTTCACTTCATCTGCCCCCATCAAAGAAAGCAGGGGATGACTAGAGACA
TTTGGTCGTTGTGATCGCCCGCCGTGAAATTGGCCCTTGGGCACGGATGATCG
GACGTAAGTCTCAAACGCCGGTTGGTGAGCTGTGAGGTGTTGGATTTCTGCGT
TTATTACTGATAATGCTAGTACGAAAGTACTAGC

RNA trigger C:
GGCTCGAAAGAGCCACTCTTAAAGTGCCAGCTTTCCCCTTGAAGTTCTGCCTAA
GGTCATCCCTCCTCCCTCGCCTCTCCTTTCACAGTCTGTCTGTCAAAGTCCCGT
CAGTCAGCCCGAGCCCGCGCCCCCGCCCCTGCACGCGCCTCGGGCACTCGC
GTTCTGTCGCTCTGGAATGAGCGAGCGAGCGAGAGAGAGAGGCGCGCTGCGC
ACTCCCGACTGGACAACAAGACTGTTTGTGGTTCTGAGACGCGAAAGCGAATG
TGGGTCAAAATCTGACAAGATGCTAGTACGAAAGTACTAGC

RNA trigger D:
GGCTCGAAAGAGCCACCACTGGCTTAGGAGTCTCAGCTCTGGGTACTCCCTCT
GAATAATGTTTGTCCTTATCTGTGCAGAGAACACTGTCTCTAAAGCATCCTTTGG
CAACACATTTGCTCAATCAACTACTGAATTGGTTGTTAAAATTAATTTTCCTTTAT
GCTAGTACGAAAGTACTAGC

Chemically modified sgRNAs and iSBH-sgRNAs tested in zebrafish

GC spacer* extension*loop extensionspacer scaffoldextra Us

Backfold=spacer*+extension+loop

*An additional GC sequence was added to all iSBH-sgRNA sequences to maintain consistency.

sgRNA 2 with chemical modifications:
mG*mU*mA*rArGrUrCrGrGrArGrUrArCrUrGrUrCrCrUrGrUrUrUrUrArGrArGrCrUrAr
GrArArArUrArGrCrArArGrUrUrArArArArUrArArGrGrCrUrArGrUrCrCrGrUrUmAmUm
CmAmAmCmUmUmGmAmAmAmAmAmGmUmGmGmCmAmCmCmGmAmGmUm
CmGmGmUmGmCmU*mU*mU*rU

Second-generation iSBH-sgRNA 2 with chemical modifications:
mG*mC*mA*rGrGrArCrArGrUrArCrArGrCrGrArGrArUrArCrCrArGrCrUrUrArGrUrCrG
rArArCrArUrCrCrCrUrArArCrArGrArCrArUrArGrCrCrCrGrUrArArGrUrCrGrGrArGrUr
ArCrUrGrUrCrCrUrGrUrUrUrUrArGrArGrCrUrArGrArArArUrArGrCrArArGrUrUrArArA
rArUrArArGrGrCrUrArGrUrCrCrGrUrUmAmUmCmAmAmCmUmUmGmAmAmAmA
mAmGmUmGmGmCmAmCmCmGmAmGmUmCmGmGmUmGmCmU*mU*mU*rU

m= 2'OMe RNA base
r= ribobase
*= phosphorothioate bond

Chemically modified RNA triggers tested in zebrafish

hairpin linker 34nt trigger 100nt 3' flank linker hairpin
random sequence

Non-complementary RNA trigger with chemical modifications:
mG*mG*mC*rUrCrGrArArArGrArGrCrCrArCrGrGrGrUrCrUrUrCrGrArGrCrArUrUrGr
GrArArGrUrCrUrArGrArUrArCrCrUrGrCrUrUrUrArCrGrArArGrArCrCrUrArUrGrCrUrA
rGrUrArCrGrArArArGrUrArCrU*mA*mG*mC

Second-generation iSBH-sgRNA 2 trigger with chemical modifications:
mG*mG*mC*rUrCrGrArArArGrArGrCrCrArCrUrGrUrUrArGrGrGrArUrGrUrUrCrGrAr
CrUrArArGrCrUrGrGrUrArUrCrUrCrGrCrUrGrUrArCrUrGrUrCrCrUrArUrGrCrUrArGr
UrArCrGrArArArGrUrArCrU*mA*mG*mC

m= 2'OMe RNA base
r= ribobase
*= phosphorothioate bond

1xCTS sequences

sequence matching sgRNA PAM site

1xCTS sgRNA 1: CAGTCGCGTGTAGCGAAGCAAGG

1xCTS sgRNA 2: GTAAGTCGGAGTACTGTCCTAGG

1xCTS sgRNA 3: CATGACATTATTCCCGCCTCAGG

1xCTS sgRNA 4: CCAGGTAATGACTAGCAGGAAGG

1xCTS sgRNA 5: GCACTACTTCAGTAAGAGTCAGG

8xCTS sequences

Linker 8x(sequence matching sgRNA PAM site linker) linker

8xCTS sgRNA 1:
TCTAGAGAGCTCCAGTCGCGTGTAGCGAAGCACGGCTATGAAAATATTAACAGT
CGCGTGTAGCGAAGCAGGGACTAGTAGGGTCAGTAGGAGCTCCAGTCGCGTG
TAGCGAAGCACGGCTATGAAAATATTAACAGTCGCGTGTAGCGAAGCAGGGAC
TAGTGGGCCCCTAGAGAGCTCCAGTCGCGTGTAGCGAAGCACGGCTATGAAAA
TATTAACAGTCGCGTGTAGCGAAGCAGGGACTAGTAGGGTCAGTAGGAGCTCC
AGTCGCGTGTAGCGAAGCACGGCTATGAAAATATTAACAGTCGCGTGTAGCGA
AGCAGGGACTAGTGCTAGC

8xCTS sgRNA 2:
TCTAGAGAGCTCGTAAGTCGGAGTACTGTCCTCGGCTATGAAAATATTAAGTAA
GTCGGAGTACTGTCCTGGGACTAGTAGGGTCAGTAGGAGCTCGTAAGTCGGAG
TACTGTCCTCGGCTATGAAAATATTAAGTAAGTCGGAGTACTGTCCTGGGACTA
GTGGGCCCCTAGAGAGCTCGTAAGTCGGAGTACTGTCCTCGGCTATGAAAATA
TTAAGTAAGTCGGAGTACTGTCCTGGGACTAGTAGGGTCAGTAGGAGCTCGTA
AGTCGGAGTACTGTCCTCGGCTATGAAAATATTAAGTAAGTCGGAGTACTGTCC
TGGGACTAGTGCTAGC

8xCTS sgRNA 3:
TCTAGAGAGCTCCATGACATTATTCCCGCCTCCGGCTATGAAAATATTAACATG
ACATTATTCCCGCCTCGGGACTAGTAGGGTCAGTAGGAGCTCCATGACATTATT
CCCGCCTCCGGCTATGAAAATATTAACATGACATTATTCCCGCCTCGGGACTAG
TGGGCCCCTAGAGAGCTCCATGACATTATTCCCGCCTCCGGCTATGAAAATATT
AACATGACATTATTCCCGCCTCGGGACTAGTAGGGTCAGTAGGAGCTCCATGAC
ATTATTCCCGCCTCCGGCTATGAAAATATTAACATGACATTATTCCCGCCTCG
GGACTAGTGCTAG

8xCTS sgRNA 4:
TCTAGAGAGCTCCCAGGTAATGACTAGCAGGACGGCTATGAAAATATTAACCAG
GTAATGACTAGCAGGAGGGACTAGTAGGGTCAGTAGGAGCTCCCAGGTAATGA
CTAGCAGGACGGCTATGAAAATATTAACCAGGTAATGACTAGCAGGAGGGACT
AGTGGGCCCCTAGAGAGCTCCCAGGTAATGACTAGCAGGACGGCTATGAAAAT
ATTAACCAGGTAATGACTAGCAGGAGGGACTAGTAGGGTCAGTAGGAGCTCCC
AGGTAATGACTAGCAGGACGGCTATGAAAATATTAACCAGGTAATGACTAGCAG
GAGGGACTAGTGCTAG

8xCTS sgRNA 5:
TCTAGAGAGCTCGCACTACTTCAGTAAGAGTCCGGCTATGAAAATATTAAGCAC
TACTTCAGTAAGAGTCGGGACTAGTAGGGTCAGTAGGAGCTCGCACTACTTCA
GTAAGAGTCCGGCTATGAAAATATTAAGCACTACTTCAGTAAGAGTCGGGACTA
GTGGGCCCCTAGAGAGCTCGCACTACTTCAGTAAGAGTCCGGCTATGAAAATA
TTAAGCACTACTTCAGTAAGAGTCGGGACTAGTAGGGTCAGTAGGAGCTCGCA
CTACTTCAGTAAGAGTCCGGCTATGAAAATATTAAGCACTACTTCAGTAAGAGTC
GGGACTAGTGCTAG

*8xCTS sequences displayed here are in reverse orientation

Oligos and primer sequences ordered for cloning 8xCTS reporters

Oligo 2xCTS sgRNA 1:
GGACCCTACTAGTCCCTGCTTCGCTACACGCGACTGTTAATATTTTCATAGCCG
TGCTTCGCTACACGCGACTGGAGCTCCTACTGACCC

Oligo 2xCTS sgRNA 2:
GGACCCTACTAGTCCCAGGACAGTACTCCGACTTACTTAATATTTTCATAGCCG
AGGACAGTACTCCGACTTACGAGCTCCTACTGACCC

Oligo 2xCTS sgRNA 3:
GGACCCTACTAGTCCCGAGGCGGGAATAATGTCATGTTAATATTTTCATAGCCG
GAGGCGGGAATAATGTCATGGAGCTCCTACTGACCC

Oligo 2xCTS sgRNA 4:
GGACCCTACTAGTCCCTCCTGCTAGTCATTACCTGGTTAATATTTTCATAGCCGT
CCTGCTAGTCATTACCTGGGAGCTCCTACTGACCC

Oligo 2xCTS sgRNA 5:
GGACCCTACTAGTCCCGACTCTTACTGAAGTAGTGCTTAATATTTTCATAGCCG
GACTCTTACTGAAGTAGTGCGAGCTCCTACTGACCC

Primers used for amplifying 2xCTS oligos and cloning 4xCTS-ECFP reporters
OP_2xCTS_FW: ACTGCCGTCTCGGACCCTACTAGTCCC
OP_2xCTS_RW: ACCTACGTCTCGGGTCAGTAGGAGCTC

Primers used for amplifying 4xCTS plasmids and cloning 8xCTS-ECFP reporters
OP_4xCTS_NheI_FW: AGAGATACCGGGCGCTAGCACTAGTCCC
OP_4xCTS_ApaI_RW: GCCCCCCGCTAGTGGGCCCCTAGAGAGCTC

Oligos and primer sequences ordered for repetitive trigger sequences

Trigger A

OP_trig_A_oligo_1:
GAAGAGGAAGACCTCCACCTGATGATCAGTTTAGTCGTGTTGTGGTGATGATC
AGTTTAGTTGTGTTGCTGTCTTCGCACTACGC

OP_trig_A_oligo_1_FW_BbsI: GAAGAGGAAGACCTCCAC
OP_trig_A_oligo_1+2_RW_BbsI: GCGTAGTGCGAAGACAG

OP_trig_A_oligo_2:
CCACGACTGGAAGACTGGTTGTGGTGATGATCAGTTTAGTTGTGTTGTGGTGA
TGATCAGTTTAGTCCTGTCTTCGCACTACGC

OP_trig_A_oligo_2+3_FW_BbsI: CCACGACTGGAAGACTG

OP_trig_A_oligo_1+2_RW_BbsI: GCGTAGTGCGAAGACAG

OP_trig_A_oligo_3:
CCACGACTGGAAGACTGAGTCGTGTTGTGGTGATGATCAGTTTAGTCGTGTTG
TGGTGATGATCTGTTTAGATGCCGGTCTTCTCGTC

OP_trig_A_oligo_2+3_FW_BbsI: CCACGACTGGAAGACTG
OP_trig_A_oligo_3_RW_BbsI: GACGAGAAGACCGGCAT

Trigger B

OP_trig_B_oligo_1:
CATGTATGCTAACTTTCAATCTACAAAACTACAACCATCATCATAATCCCACAA TCCTTCCTTCCTCTGTTTTTCACTTCATCTGCCCCC

OP_trig_B_oligo_1_FW_BbsI:
AAGAGGAAGACCTCCACCTCAACAAAAGGTTAAAAATGGTTTAGTCACATGTAT GCTAACTTTCAATCTAC

OP_trig_B_oligo_1_RW_BbsI:
ATGCAGAAGACCACCTGCTTTCTTTGATGGGGGCAGATGAAGTGAAAAAC

OP_trig_B_oligo_2:
CATTTGGTCGTTGTGATCGCCCGCCGTGAAATTGGCCCTTGGGCACGGATGAT CGGACGTAAGTCTCAAACGCCGGTTGGTGAGCTGTGA

OP_trig_B_oligo_2_FW_BbsI:
ACTGGAAGACTGCAGGGGATGACTAGAGACATTTGGTCGTTGTGATCG

OP_trig_B_oligo_2_RW_BbsI:
TACTAGAAGACTGGCATTATCAGTAATAAACGCAGAAATCCAACACCTCACAGC TCACCAACC

Primer sequences for RNA circularisation assays

Second-generation iSBH-sgRNA circularisation assay

Second-generation iSBH-sgRNA 1 primers

OP_Ca_nes1_F: TTATCAACTTGAAAAAGTGGCAC
OP_Ca_nes1_R+RT: GACTAGCCTTATTTTAACTTGC
(primer also used for RT)
OP_CA_nes2_NheI_F: AAGCTGGCTAGCAAAAGTGGCACCGAGTC
OP_CA_nes2_sgRNA1_NotI_R: ACTCGAGCGGCCGCCTAGCTCTAAAACTGCTTCG

Second-generation iSBH-sgRNA 2 primers

OP_Ca_nes1_F: TTATCAACTTGAAAAAGTGGCAC
OP_Ca_nes1_R+RT: GACTAGCCTTATTTTAACTTGC
(primer also used for RT)
OP_CA_nes2_NheI_F: AAGCTGGCTAGCAAAAGTGGCACCGAGTC
OP_CA_ nes2_sgRNA2_NotI_R: ACTCGAGCGGCCGCCTAGCTCTAAAACAGGACAG

Second-generation iSBH-sgRNA 3 primers

OP_Ca_nes1_F: TTATCAACTTGAAAAAGTGGCAC
OP_Ca_nes1_R+RT: GACTAGCCTTATTTTAACTTGC
(primer also used for RT)
OP_CA_nes2_NheI_F: AAGCTGGCTAGCAAAAGTGGCACCGAGTC

OP_CA_nes2_sgRNA3_NotI_R: ACTCGAGCGGCCGCCTAGCTCTAAAACGAGGCG

RNA triggers circularisation assay

Second-generation iSBH-sgRNA 1 trigger 100nt 3' flank primers

OP_CA_nes1_trig1_F: CCACCAAGGCAAAGAGAAG
OP_CA_nes1_trig1_R+RT: CTACTCCTAATGGTTCAATTTTTACTAC
(primer also used for RT)
OP_CA_nes2_trig1_F_NheI: AAGCTGGCTAGCCAAAGAGAAGAGTGGTGCAG
OP_CA_nes2_trig1_R_NotI:
ACTCGAGCGGCCGCCAATTTTTACTACTTTATATTTATATAATTCACTTCTC

Second-generation iSBH-sgRNA 2 trigger 100nt 3' flank primers

OP_CA_nes1_trig2_F: TACAACAAGGCAAGGCTTG
OP_CA_nes1_trig2_R+RT: CTTAAATTTTGCTCGCGCAC
(primer also used for RT)
OP_CA_nes2_trig2_F_NheI: AAGCTGGCTAGCCAAGGCTTGACCGACAAG
OP_CA_nes2_trig2_R_BbsI: GAGCCGAAGACCCGGCCGCCGACCTCCAACACACAAG

Second-generation iSBH-sgRNA 3 trigger 100nt 3' flank primers

OP_CA_nes1_trig3_F: AGTATCTGCTCCCTGCTTG
OP_CA_nes1_trig3_R+RT: TTAACTATGCGGCATCAGAG
(primer also used for RT)
OP_CA_nes2_trig3_F_NheI: AAGCTGGCTAGCGCTTGTGTGTTGGAGGTG
OP_CA_nes2_trig3_R_NotI: ACTCGAGCGGCCGCGAGCAGATTGTACTGAGAGTG

Primer sequences- zebrafish genotyping

OP_dCas9_Vp64_nes1_F: AAATAAGCGAATTCTCCAAAAG
OP_dCas9_Vp64_nes1_R:CTACGTCTCGCCGCATGTTAGCAGACTTCCTCTGCC
CTCCGAGCCACCGCCCTTGTACAGCTCGTCCATG
OP_dCas9_Vp64_nes2_F: TGCTTGGTTCGGATGCCCTTGATGAC
OP_dCas9_Vp64_nes2_R: GGCGAAGCACATCAGGCCGTAGC
OP_8xCTS _nes1_F: GGTACGGGAGGTACTTGGAGCG
OP_8xCTS _nes1_R: CACCACCCCGGTGAACAGCTC
OP_8xCTS _nes2_F: CTTGGAGCGGCCGCAATAAATATC
OP_8xCTS _nes2_R: CTTGCTCACCATGGTCAACGC

References

[1] Thomas Gaj. "ZFN, TALEN and CRISPR/Cas based methods for genome engineering". In: *Trends Biotechnol.* 31.7 (2014), pp. 397–405.

[2] Philippe Horvath and Rodolphe Barrangou. "CRISPR/Cas, the immune system of Bacteria and Archaea". In: *Science* 327.5962 (2010), pp. 167–170.

[3] Martin Jinek et al. "A Programmable Dual-RNA-Guided DNA Endonuclease in Adaptative BActerial Immunity". In: *Science* 337 (2012), pp. 816–822.

[4] Martin Jinek et al. "RNA-programmed genome editing in human cells". In: *eLife* 2013.2 (2013), pp. 1–9.

[5] Cong Le et al. "Multiplex Genome Engineering Using CRISPR/Cas Systems". In: *Science* (2013), pp. 819–824.

[6] F Ann Ran et al. "Genome engineering using the CRISPR-Cas9 system". In: *Nature Protocols* 8.11 (2013), pp. 2281–2308.

[7] Leonela Amoasii et al. "Gene editing restores dystrophin expression in a canine model of Duchenne muscular dystrophy". In: *Science* 362.6410 (2018), pp. 86–91.

[8] Ergin Beyret et al. "Single-dose CRISPR–Cas9 therapy extends lifespan of mice with Hutchinson–Gilford progeria syndrome". In: *Nature Medicine* 25.3 (2019), pp. 419–422. URL: http://dx.doi.org/10.1038/s41591-019-0343-4.

[9] Chengzu Long et al. "Prevention of muscular dystrophy in mice by CRISPR/Cas9-mediated editing of germline DNA". In: *Science* 345.6201 (2014), pp. 1184–1188.

[10] Ashley E Modell et al. "CRISPR-based therapeutics: current challenges and future applications". In: *Trends in Pharmacological Sciences* 43.2 (2022), pp. 151–161. URL: https://doi.org/10.1016/j.tips.2021.10.012.

[11] Kira S Makarova and Eugene V Koonin. "Annotation and Classification of CRISPR-Cas Systems". In: *Methods in Molecular Biology* (2015), pp. 1–366.

[12] Kira S Makarova et al. "Evolutionary classification of CRISPR–Cas systems: a burst of class 2 and derived variants". In: *Nature Reviews Microbiology* 18.2 (2020), pp. 67–83.

[13] Elitza Deltcheva et al. "CRISPR RNA maturation by trans-encoded small RNA and host factor RNase III". In: *Nature* 471.7340 (2011), pp. 602–607.

[14] Alexandra E Briner et al. "Guide RNA functional modules direct Cas9 activity and orthogonality". In: *Molecular Cell* 56.2 (2014), pp. 333–339. URL: http://dx.doi.org/10.1016/j.molcel.2014.09.019.

[15] Xi-wen Wang et al. "A microRNA-inducible CRISPR–Cas9 platform serves as a microRNA sensor and cell-type-specific genome regulation tool". In: *Nature Cell Biology* (2019). URL: http://dx.doi.org/10.1038/s41556-019-0292-7.

[16] Moe Hirosawa, Yoshihiko Fujita, and Hirohide Saito. "Cell-Type-Specific CRISPR Activation with MicroRNA-Responsive AcrllA4 Switch". In: *ACS Synthetic Biology* 8.7 (2019), pp. 1575–1582.

[17] Jooyoung Lee et al. "Tissue-restricted genome editing in vivo specified by microRNA-repressible anti-CRISPR proteins". In: *Rna* (2019), Vol.25, No11.

[18] Lauren Gambill et al. "A split ribozyme that links detection of a native RNA to orthogonal protein outputs". In: *bioRxiv preprint* (2022), pp. 1–33.

[19] Ni Gao et al. "Endogenous promoter-driven sgRNA for monitoring the expression of low-abundance transcripts and lncRNAs". In: *Nature Cell Biology* 23.1 (2021), pp. 99–108.

[20] Gokul Kesavan. "Innovations in CRISPR-Based Therapies". In: *Molecular Biotechnology* 0123456789 (2021). URL: https://doi.org/10.1007/s12033-021-00411-x.

[21] Chiara Antoniani et al. "Induction of fetal hemoglobin synthesis by CRISPR/Cas9-mediated editing of the human b-globin locus". In: *Blood* 131.17 (2018), pp. 1960–1973. URL: http://dx.doi.org/10.1182/blood-2017-10-811505.

[22] Long Shi et al. "CRISPR knock out CTLA-4 enhances the anti-tumor activity of cytotoxic T lymphocytes". In: *Gene* 636.May (2017), pp. 36–41. URL: http://dx.doi.org/10.1016/j.gene.2017.09.010.

[23] You Lu et al. "Safety and feasibility of CRISPR-edited T cells in patients with refractory non-small-cell lung cancer". In: *Nature Medicine* 26.5 (2020), pp. 732–740.

[24] Morgan L Maeder et al. "Development of a gene-editing approach to restore vision loss in Leber congenital amaurosis type 10". In: *Nature Medicine* 25.2 (2019), pp. 229–233. URL: http://dx.doi.org/10.1038/s41591-018-0327-9.

[25] Julian D Gillmore et al. "CRISPR-Cas9 In Vivo Gene Editing for Transthyretin Amyloidosis". In: *New England Journal of Medicine* 385.6 (2021), pp. 493–502.

[26] Cornelia Lorenzer et al. "Going beyond the liver: Progress and challenges of targeted delivery of siRNA therapeutics". In: *Journal of Controlled Release* 203 (2015), pp. 1–15. URL: http://dx.doi.org/10.1016/j.jconrel.2015.02.003.

[27] Xuebing Wu et al. "Genome-wide binding of the CRISPR endonuclease Cas9 in mammalian cells". In: *Nature Biotechnology* 32.7 (2014), pp. 670–676.

[28] Michael Kosicki, Kärt Tomberg, and Allan Bradley. "Repair of double-strand breaks induced by CRISPR – Cas9 leads to large deletions and complex rearrangements". In: *Nature Biotechnology* (2018). URL: http://dx.doi.org/10.1038/nbt.4192.

[29] Emma Haapaniemi et al. "CRISPR – Cas9 genome editing induces a p53-mediated DNA damage response". In: *Nature Medicine* 24.July (2018). URL: http://dx.doi.org/10.1038/s41591-018-0049-z.

[30] Robert J Ihry et al. "P53 inhibits CRISPR-Cas9 engineering in human pluripotent stem cells". In: *Nature Medicine* 24.7 (2018), pp. 939–946. URL: http://dx.doi.org/10.1038/s41591-018-0050-6.

[31] Joana da Silva, Mathilde Meyenberg, and Joanna I Loizou. "Tissue specificity of DNA repair: the CRISPR compass". In: *Trends in Genetics* 37.11 (2021), pp. 958–962.

[32] Chang Liu et al. "Delivery strategies of the CRISPR-Cas9 gene-editing system for therapeutic applications". In: *Journal of Controlled Release* 266.June (2017), pp. 17–26. URL: https://doi.org/10.1016/j.jconrel.2017.09.012.

[33] Jungjoon K Lee et al. "Directed evolution of CRISPR-Cas9 to increase its specificity". In: *Nature Communications* 9.1 (2018). URL: http://dx.doi.org/10.1038/s41467-018-05477-x.

[34] Paul D Donohoue et al. "Article Conformational control of Cas9 by CRISPR hybrid RNA-DNA guides mitigates off-target activity in T cells ll Article Conformational control of Cas9 by CRISPR hybrid RNA-DNA guides mitigates off-target activity in T cells". In: *Molecular Cell* 81.17 (2021), pp. 3637–3649. URL: https://doi.org/10.1016/j.molcel.2021.07.035.

[35] Elena V Moroz-Omori et al. "Photoswitchable gRNAs for Spatiotemporally Controlled CRISPR-Cas-Based Genomic Regulation". In: *ACS Central Science* (2020).

[36] Hao Chen et al. "Aptazyme-mediated direct modulation of post-transcriptional sgRNA level for conditional genome editing and gene expression". In: *Journal of Biotechnology* 288 (2018), pp. 23–29. URL: https://doi.org/10.1016/j.jbiotec.2018.10.011.

[37] Quentin R V Ferry, Radostina Lyutova, and Tudor A Fulga. "Rational design of inducible CRISPR guide RNAs for de novo assembly of transcriptional programs". In: *Nature Communications* 8 (2017), pp. 1–10. URL: http://dx.doi.org/10.1038/ncomms14633.

[38] Kale Kundert et al. "Controlling CRISPR-Cas9 with ligand-activated and ligand-deactivated sgRNAs". In: *Nature Communications* (2019), pp. 1–11. URL: http://dx.doi.org/10.1038/s41467-019-09985-2.

[39] Yuchen Liu et al. "Directing cellular information flow via CRISPR signal conductors". In: *Nature Methods* 13.11 (2016), pp. 938–944.

[40] Weixin Tang, Johnny H Hu, and David R Liu. "Aptazyme-embedded guide RNAs enable ligand-responsive genome editing and transcriptional activation". In: *Nature Communications* 8 (2017), pp. 1–8. URL: http://dx.doi.org/10.1038/ncomms15939.

[41] F Ann Ran et al. "Double Nicking by RNA-Guided CRISPR Cas9 for Enhanced Genome Editing Specificity". In: *CELL* 154.6 (2013), pp. 1380–1389. URL: http://dx.doi.org/10.1016/j.cell.2013.08.021.

[42] Nicole M Gaudelli et al. "Programmable base editing of T to G C in genomic DNA without DNA cleavage". In: *Nature* 551.7681 (2017), pp. 464–471. URL: http://dx.doi.org/10.1038/nature24644.

[43] Alexis C Komor et al. "Programmable editing of a target base in genomic DNA without double-stranded DNA cleavage". In: *Nature* 533.7603 (2016), pp. 420–424. URL: http://dx.doi.org/10.1038/nature17946.

[44] Andrew V Anzalone et al. “Search-and-replace genome editing without double-strand breaks or donor DNA”. In: *Nature* 576.7785 (2019), pp. 149–157. URL: http://dx.doi.org/10.1038/s41586-019-1711-4.

[45] James W Nelson et al. “Engineered pegRNAs improve prime editing efficiency”. In: *Nature Biotechnology* 40.3 (2022), pp. 402–410.

[46] Andrew V Anzalone et al. “Programmable deletion, replacement, integration and inversion of large DNA sequences with twin prime editing”. In: *Nature Biotechnology* 40.5 (2022), pp. 731–740.

[47] Lei S Qi et al. “Repurposing CRISPR as an RNA-guided platform for sequence-specific control of gene expression”. In: *Cell* 152.5 (2013), pp. 1173–1183. URL: http://dx.doi.org/10.1016/j.cell.2013.02.022.

[48] Alejandro Chavez et al. “Highly efficient Cas9-mediated transcriptional programming”. In: *Nature Methods* 12.4 (2015), pp. 326–328.

[49] Morgan L Maeder et al. “CRISPR RNA-guided activation of endogenous human genes”. In: *Nature Methods* 10.10 (2013), pp. 977–979. URL: http://dx.doi.org/10.1038/nmeth.2598.

[50] Silvana Konermann et al. “Genome-scale transcriptional activation by an engineered CRISPR-Cas9 complex”. In: *Nature* 517.7536 (2015), pp. 583–588. URL: http://dx.doi.org/10.1038/nature14136.

[51] Mahere Rezazade Bazaz, Mohammad M Ghahramani Seno, and Hesam Dehghani. “Transposase-CRISPR mediated targeted integration (TransCRISTI) in the human genome”. In: *Scientific Reports* 12.1 (2022), pp. 1–18. URL: https://doi.org/10.1038/s41598-022-07158-8.

[52] Jennifer A Doudna. “The promise and challenge of therapeutic genome editing”. In: *Nature* 578 (2020), pp. 229–236. URL: http://dx.doi.org/10.1038/s41586-020-1978-5.

[53] Matthew Behr et al. “In vivo delivery of CRISPR-Cas9 therapeutics: Progress and challenges”. In: *Acta Pharmaceutica Sinica B* 11.8 (2021), pp. 2150–2171. URL: https://doi.org/10.1016/j.apsb.2021.05.020.

[54] Cia Hin Lau and Yousin Suh. “In vivo genome editing in animals using AAV-CRISPR system: Applications to translational research of human disease”. In: *F1000Research* 6.0 (2017), pp. 1–21.

[55] David Goertsen et al. “AAV capsid variants with brain-wide transgene expression and decreased liver targeting after intravenous delivery in mouse and marmoset”. In: *Nature Neuroscience* 25.1 (2022), pp. 106–115.

[56] Mohammadsharif Tabebordbar et al. “Directed evolution of a family of AAV capsid variants enabling potent muscle-directed gene delivery across species”. In: *Cell* 184.19 (2021), pp. 4919–4938. URL: https://doi.org/10.1016/j.cell.2021.08.028.

[57] Killian S Hanlon et al. “High levels of AAV vector integration into CRISPR-induced DNA breaks”. In: *Nature Communications* 10.1 (2019), pp. 1–11.

[58] Carsten T Charlesworth et al. “Identification of preexisting adaptive immunity to Cas9 proteins in humans”. In: *Nature Medicine* 25.2 (2019), pp. 249–254. URL: http://dx.doi.org/10.1038/s41591-018-0326-x.

[59] Qiang Cheng et al. “Selective organ targeting (SORT) nanoparticles for tissue-specific mRNA delivery and CRISPR–Cas gene editing”. In: *Nature Nanotechnology* 15.4 (2020), pp. 313–320. URL: http://dx.doi.org/10.1038/s41565-020-0669-6.

[60] Daniel Rosenblum et al. “CRISPR-Cas9 genome editing using targeted lipid nanoparticles for cancer therapy”. In: *Science Advances* 6.47 (2020).

[61] Tuo Wei et al. “Systemic nanoparticle delivery of CRISPR-Cas9 ribonucleoproteins for effective tissue specific genome editing”. In: *Nature Communications* (2020), pp. 1–12. URL: http://dx.doi.org/10.1038/s41467-020-17029-3.

[62] Sean A Dilliard, Qiang Cheng, and Daniel J Siegwart. “On the mechanism of tissue-specific mRNA delivery by selective organ targeting nanoparticles”. In: *Proceedings of the National Academy of Sciences of the United States of America* 118.52 (2021).

[63] Esther J Lee, Caitlin M Guenther, and Junghae Suh. “Adeno-associated virus (AAV) vectors: Rational design strategies for capsid engineering”. In: *Current Opinion in Biomedical Engineering* 7 (2018), pp. 58–63.

[64] Hyunmin Lee, Zhaolei Zhang, and Henry M Krause. “Long Noncoding RNAs and Repetitive Elements: Junk or Intimate Evolutionary Partners?” In: *Trends in Genetics* 35.12 (2019), pp. 892–902. URL: https://doi.org/10.1016/j.tig.2019.09.006.

[65] Markus Sällman Almén et al. “Mapping the human membrane proteome: A majority of the human membrane proteins can be classified according to function and evolutionary origin”. In: *BMC Biology* 7 (2009), p. 50.

[66] Ian Dunham et al. “An integrated encyclopedia of DNA elements in the human genome”. In: *Nature* 489.7414 (2012), pp. 57–74.

[67] Ewan Birney et al. “Identification and analysis of functional elements in 1{\%} of the human genome by the ENCODE pilot project”. In: *Nature* 447.7146 (2007), pp. 799–816.

[68] Alexander F Palazzo and Eliza S Lee. “Non-coding RNA: What is functional and what is junk?” In: *Frontiers in Genetics* 5.JAN (2015).

[69] Carthew Richard and Sontheimer Erik. “Origins and Mechanisms of miRNAs and siRNAs”. In: *Cell* 136.4 (2009), pp. 642–655.

[70] Saba Valadkhan and Lalith S Gunawardane. “Role of small nuclear RNAs in eukaryotic gene expression”. In: *Essays in Biochemistry* 54.1 (2013), pp. 79–90.

[71] David Tollervey and Tamás Kiss. “Function and synthesis of small nucleolar RNAs”. In: *Current Opinion in Cell Biology* 9.3 (1997), pp. 337–342.

[72] Deniz M Ozata et al. “PIWI-interacting RNAs: small RNAs with big functions”. In: *Nature Reviews Genetics* 20.2 (2019), pp. 89–108. URL: http://dx.doi.org/10.1038/s41576-018-0073-3.

[73] Georges St.Laurent, Claes Wahlestedt, and Philipp Kapranov. "The Landscape of long noncoding RNA classification". In: *Trends in Genetics* 31.5 (2015), pp. 239–251. URL: http://dx.doi.org/10.1016/j.tig.2015.03.007.

[74] Raquel Fueyo, Julius Judd, and Cedric Feschotte. "Roles of transposable elements in the regulation of mammalian transcription". In: 23.July (2022), pp. 19–24.

[75] John Lonsdale et al. "The Genotype-Tissue Expression (GTEx) project". In: *Nature Genetics* 45.6 (2013), pp. 580–585.

[76] Federico Abascal et al. "Expanded encyclopaedias of DNA elements in the human and mouse genomes". In: *Nature* 583.7818 (2020), pp. 699–710.

[77] Robert C Jones et al. "The Tabula Sapiens: A multiple-organ, single-cell transcriptomic atlas of humans". In: *Science* 376.6594 (2022).

[78] Robert Y Yang et al. "A systematic survey of human tissue-specific gene expression and splicing reveals new opportunities for therapeutic target identification and evaluation Computational Sciences , Pfizer , Cambridge , MA 2 Neuroscience Research Unit , Pfizer , Cambridge , M". In: *bioRxiv* (2018).

[79] Thomas Derrien et al. "The GENCODE v7 catalog of human long noncoding RNAs : Analysis of their gene structure , evolution , and expression". In: (2012), pp. 1775–1789.

[80] Chunjie Jiang et al. "Identifying and functionally characterizing tissue-specific and ubiquitously expressed human lncRNAs". In: 7.6 (2016).

[81] Samir Ounzain et al. "Genome-wide profiling of the cardiac transcriptome after myocardial infarction identifies novel heart-specific long non-coding RNAs". In: *European Heart Journal* 36.6 (2015), pp. 353–368.

[82] Yueh Hua Tu, Hsueh Fen Juan, and Hsuan Cheng Huang. "Identification of cell states using super-enhancer RNA". In: *BMC Genomics* 22 (2021), pp. 1–8.

[83] Jiangping He et al. "Identifying transposable element expression dynamics and heterogeneity during development at the single-cell level with a processing pipeline scTE". In: *Nature Communications* 12.1 (2021), pp. 1–14. URL: http://dx.doi.org/10.1038/s41467-021-21808-x.

[84] Nicole Ludwig et al. "Distribution of miRNA expression across human tissues". In: *Nucleic Acids Research* 44.8 (2016), pp. 3865–3877.

[85] Yu Liang et al. "Characterization of microRNA expression profiles in normal human tissues". In: *BMC Genomics* 8 (2007), pp. 1–20.

[86] Kimberly A Dittmar, Jeffrey M Goodenbour, and Tao Pan. "Tissue-specific differences in human transfer RNA expression". In: *PLoS Genetics* 2.12 (2006), pp. 2107–2115.

[87] Roberto Galizi et al. "Engineered RNA-Interacting CRISPR Guide RNAs for Genetic Sensing and Diagnostics". In: *CRISPR Journal* 3.5 (2020), pp. 398–408.

[88] Tanner Wiegand et al. "Structures and Strategies of Anti-CRISPR-Mediated Immune Suppression". In: *Annual Review of Microbiology* 74 (2020), pp. 21–37.

[89] Zhan Ming Ying et al. "Activatable CRISPR Transcriptional Circuits Generate Functional RNA for mRNA Sensing and Silencing". In: *Angewandte Chemie - International Edition* 59.42 (2020), pp. 18599–18604.

[90] Moe Hirosawa et al. “Cell-type-specific genome editing with a microRNA-responsive CRISPR-Cas9 switch”. In: *Nucleic Acids Research* 45.13 (2017).

[91] Jiao Lin et al. “Conditional guide RNA through two intermediate hairpins for programmable CRISPR/Cas9 function: building regulatory connections between endogenous RNA expressions”. In: *Nucleic Acids Research* 48.20 (2020), pp. 11773–11784.

[92] Jiao Lin et al. “Building Endogenous Gene Connections through RNA Self-Assembly Controlled CRISPR/Cas9 Function”. In: *Journal of the American Chemical Society* 143.47 (2021), pp. 19834–19843.

[93] Yang Liu et al. “Reprogrammed tracrRNAs enable repurposing RNAs as crRNAs and detecting RNAs”. In: *bioRxiv preprint* (2021), pp. 1–28.

[94] Ka-hei Siu and Wilfred Chen. “Riboregulated toehold-gated gRNA for programmable CRISPR – Cas9 function”. In: *Nature Chemical Biology* (2019). URL: http://dx.doi.org/10.1038/s41589-018-0186-1.

[95] Nicholas S McCarty et al. “Multiplexed CRISPR technologies for gene editing and transcriptional regulation”. In: *Nature Communications* 11.1 (2020), pp. 1–13. URL: http://dx.doi.org/10.1038/s41467-020-15053-x.

[96] Shiqi Xie et al. “Multiplexed Engineering and Analysis of Combinatorial Enhancer Activity in Single Cells”. In: *Molecular Cell* 66.2 (2017), pp. 285–299. URL: http://dx.doi.org/10.1016/j.molcel.2017.03.007.

[97] David J H F Knapp et al. “Decoupling tRNA promoter and processing activities enables specific Pol-II Cas9 guide RNA expression”. In: *Nature Communications* (2019). URL: http://dx.doi.org/10.1038/s41467-019-09148-3.

[98] Mikhail Hanewich-Hollatz et al. “Conditional Guide RNAs: Programmable Conditional Regulation of CRISPR/Cas Function in Bacterial and Mammalian Cells via Dynamic RNA Nanotechnology”. In: (2019).

[99] Victoria M Hunt and Wilfred Chen. “A microRNA-gated thgRNA platform for multiplexed activation of gene expression in mammalian cells”. In: *Chemical Communications* 58.42 (2022), pp. 6215–6218.

[100] Victoria M Hunt and Wilfred Chen. “Deciphering the Design Rules of Toehold-Gated sgRNA for Conditional Activation of Gene Expression and Protein Degradation in Mammalian Cells”. In: *ACS Synthetic Biology* 11.1 (2022), pp. 397–405.

[101] Noah Jakimo, Pranam Chatterjee, and Joseph M Jacobson. “SsRNA/DNA-Sensors via Embedded Strand-Displacement Programs in CRISPR/Cas9 Guides”. In: *bioRxiv preprint* (2018), pp. 1–12.

[102] Mike Jin et al. “Programmable CRISPR-Cas Repression, Activation, and Computation with Sequence-Independent Targets and Triggers”. In: *ACS Synthetic Biology* 8.7 (2019), pp. 1583–1589.

[103] Yue Li et al. “RNA Strand Displacement Responsive CRISPR/Cas9 System for mRNA Sensing”. In: *Analytical Chemistry* 91 (2019), pp. 3989–3996.

[104] Yanfang Fu et al. “Improving CRISPR-Cas nuclease specificity using truncated guide RNAs”. In: *Nature Biotechnology* 32.3 (2014), pp. 279–284.

[105] James E Dahlman et al. “Orthogonal gene knockout and activation with a catalytically active Cas9 nuclease”. In: *Nature Biotechnology* 33.11 (2015), pp. 1159–1161.

[106] Hanhui Ma et al. “CRI SPR-Cas9 nuclear dynamics and target recognition in living cells”. In: *Journal of Cell Biology* 214.5 (2016), pp. 529–537.

[107] Samuel D Perli, Cheryl H Cui, and Timothy K Lu. “Continuous genetic recording with self-targeting CRISPR-Cas in human cells”. In: *Science* 353.6304 (2016).

[108] Hanhui Ma et al. “Multiplexed labeling of genomic loci with dCas9 and engineered sgRNAs using CRISPRainbow”. In: *Nature Biotechnology* 34.5 (2016), pp. 528–530.

[109] Shipeng Shao et al. “Long-term dual-color tracking of genomic loci by modified sgRNAs of the CRISPR/Cas9 system”. In: *Nucleic Acids Research* 44.9 (2016).

[110] Jesse G Zalatan et al. “Engineering Complex Synthetic Transcriptional Programs with CRISPR RNA Scaffolds”. In: *Cell* 160.1-2 (2015), pp. 339–350. URL: http://dx.doi.org/10.1016/j.cell.2014.11.052.

[111] Qingkai Song et al. “Direct-seq: Programmed gRNA scaffold for streamlined scRNA-seq in CRISPR screen”. In: *Genome Biology* 21.1 (2020), pp. 1–15.

[112] Joseph M Replogle et al. “Combinatorial single-cell CRISPR screens by direct guide RNA capture and targeted sequencing”. In: *Nature Biotechnology* 38.8 (2020), pp. 954–961. URL: http://dx.doi.org/10.1038/s41587-020-0470-y.

[113] Fuguo Jiang and Jennifer A Doudna. “CRISPR – Cas9 Structures and Mechanisms”. In: *Annual Review of Biophysics* (2017), pp. 505–531.

[114] Martin Jinek et al. “Structures of Cas9 endonucleases reveal RNA-mediated conformational activation”. In: *Science* 343.6176 (2014).

[115] Hiroshi Nishimasu et al. “Crystal structure of Cas9 in complex with guide RNA and target DNA”. In: *Cell* 156.5 (2014), pp. 935–949. URL: http://dx.doi.org/10.1016/j.cell.2014.02.001.

[116] Chance M Nowak et al. “Guide RNA engineering for versatile Cas9 functionality”. In: *Nucleic Acids Research* 44.20 (2016), pp. 9555–9564.

[117] Hao Yin et al. “Partial DNA-guided Cas9 enables genome editing with reduced off-target activity”. In: *Nature Chemical Biology* 14.3 (2018), pp. 311–316.

[118] Hanhui Ma et al. “CRISPR-Cas9 nuclear dynamics and target recognition in living cells”. In: *Journal of Cell Biology* 214.5 (2016), pp. 529–537.

[119] Zachary J Kartje et al. “Chimeric Guides Probe and Enhance Cas9 Biochemical Activity”. In: *Biochemistry* 57.21 (2018), pp. 3027–3031.

[120] Zhiwei Hu et al. “Improving the precision of base editing by bubble hairpin single guide rna”. In: *mBio* 12.2 (2021), pp. 1–11.

[121] D Dewran Kocak et al. “Increasing the specificity of CRISPR systems with engineered RNA secondary structures”. In: *Nature Biotechnology* 37 (2019). URL: http://dx.doi.org/10.1038/s41587-019-0095-1.

[122] Liang-Zhong Yang et al. “Dynamic Imaging of RNA in Living Cells by CRISPR-Cas13 Systems.” In: *Molecular cell* (2019), pp. 1–17. URL: http://www.ncbi.nlm.nih.gov/pubmed/31757757.

[123] Baohui Chen et al. “Dynamic imaging of genomic loci in living human cells by an optimized CRISPR/Cas system”. In: *Cell* 155.7 (2013), pp. 1479–1491. URL: http://dx.doi.org/10.1016/j.cell.2013.12.001.

[124] Reza Kalhor and George M Church. “Rapidly evolving homing CRISPR barcodes”. In: *Nature Methods* 14.2 (2017).

[125] Thomas Gonatopoulos-Pournatzis et al. “Genetic interaction mapping and exon-resolution functional genomics with a hybrid Cas9-Cas12a platform”. In: *Nature Biotechnology* (2020), pp. 1–11. URL: http://www.nature.com/articles/s41587-020-0437-z.

[126] David M Shechner et al. “Multiplexable, locus-specific targeting of long RNAs with CRISPR-Display”. In: *Nature Methods* 12.7 (2015), pp. 664–670.

[127] Yaya Hao et al. “Programmable Live-Cell CRISPR Imaging with Toehold-Switch-Mediated Strand Displacement”. In: *Angewandte Chemie - International Edition* 59.46 (2020), pp. 20612–20618.

[128] Fan Hong and Petr Šulc. “Strand displacement: A fundamental mechanism in RNA biology?” In: *arXiv* (2019), pp. 1–18.

[129] Lukas Oesinghaus and Friedrich C Simmel. “Switching the activity of Cas12a using guide RNA strand displacement circuits”. In: *Nature Communications* (2019), pp. 1–11. URL: http://dx.doi.org/10.1038/s41467-019-09953-w.

[130] Victoria M. Hunt and Wilfred Chen. “A microRNA-gated thgRNA platform for multiplexed activation of gene expression in mammalian cells”. In: *Chemical Communications* 58.42 (2022), pp. 6215–6218.

[131] Jinjin Shi et al. “MicroRNA-responsive release of Cas9/sgRNA from DNA nanoflower for cytosolic protein delivery and enhanced genome editing”. In: *Biomaterials* 256 (2020), p. 120221. URL: https://doi.org/10.1016/j.biomaterials.2020.120221.

[132] Lisa M Hochrein, Heyun Li, and Niles A Pierce. “High-Performance Allosteric Conditional Guide RNAs for Mammalian Cell-Selective Regulation of CRISPR/Cas”. In: *ACS Synthetic Biology* (2021).

[133] Chunlei Jiao et al. “Noncanonical crRNAs derived from host transcripts enable multiplexable RNA detection by Cas9”. In: *Science* 372.6545 (2021), pp. 941–948.

[134] Lisa M Hochrein et al. “Conditional dicer substrate formation via shape and sequence transduction with small conditional RNAs”. In: *Journal of the American Chemical Society* 135.46 (2013), pp. 17322–17330.

[135] Vanessa Chong-Morrison et al. “Re-purposing Ac/Ds transgenic system for CRISPR/dCas9 modulation of enhancers and non-coding RNAs in zebrafish”. In: *bioRxiv* (2018), pp. 1–33.

[136] Le A Trinh et al. “Biotagging of Specific Cell Populations in Zebrafish Reveals Gene Regulatory Logic Encoded in the Nuclear Transcriptome”. In: *Cell Reports* 19.2 (2017), pp. 425–440.

[137] Chang Liu et al. “Delivery strategies of the CRISPR-Cas9 gene-editing system for therapeutic applications”. In: *Journal of Controlled Release* 266.June (2017), pp. 17–26. URL: https://doi.org/10.1016/j.jconrel.2017.09.012.

[138] Lior Nissim et al. "Multiplexed and Programmable Regulation of Gene Networks with an Integrated RNA and CRISPR/Cas Toolkit in Human Cells". In: *Molecular Cell* 54.4 (2014), pp. 698–710. URL: http://dx.doi.org/10.1016/j.molcel.2014.04.022.

[139] Liang Qu et al. "Programmable RNA editing by recruiting endogenous ADAR using engineered RNAs". In: *Nature Biotechnology* 37.9 (2019), pp. 1059–1069. URL: http://dx.doi.org/10.1038/s41587-019-0178-z.

[140] Sumitaka Hasegawa, Gayatri Gowrishankar, and Jianghong Rao. "Detection of mRNA in mammalian cells with a split ribozyme reporter". In: *ChemBioChem* 7.6 (2006), pp. 925–928.

[141] Silvana Konermann et al. "Genome-scale transcriptional activation by an engineered CRISPR-Cas9 complex". In: *Nature* 517.7536 (2015), pp. 583–588. URL: http://dx.doi.org/10.1038/nature14136.

[142] Marlene Oeffinger and Daniel Zenklusen. "To the pore and through the pore: A story of mRNA export kinetics". In: *Biochimica et Biophysica Acta - Gene Regulatory Mechanisms* 1819.6 (2012), pp. 494–506.

[143] Abdul-rahman Allouche. "Software News and Updates Gabedit — A Graphical User Interface for Computational Chemistry Softwares". In: *Journal of computational chemistry* 32 (2012), pp. 174–182.

[144] Ysobel R Baker et al. "An LNA-amide modification that enhances the cell uptake and activity of phosphorothioate exon-skipping oligonucleotides". In: *Nature Communications* 2022 (), pp. 1–11.

[145] Gunter Meister and Thomas Tuschl. "Mechanisms of gene silencing by double-stranded RNA". In: *Nature* 431.7006 (2004), pp. 343–349.

[146] Sheng K Pong and Monika Gullerova. "Noncanonical functions of microRNA pathway enzymes – Drosha, DGCR8, Dicer and Ago proteins". In: *FEBS Letters* 592.17 (2018), pp. 2973–2986.

[147] Saurabh Singh, Ajit S Narang, and Ram I Mahato. "Subcellular Fate and Off-Target Effects of siRNA , shRNA , and miRNA". In: (2011), pp. 2996–3015.

[148] Witold Filipowicz, Suvendra N Bhattacharyya, and Nahum Sonenberg. "Mechanisms of post-transcriptional regulation by microRNAs: Are the answers in sight?" In: *Nature Reviews Genetics* 9.2 (2008), pp. 102–114.

[149] Kaspar Burger et al. "Nuclear phosphorylated Dicer processes doublestranded RNA in response to DNA damage". In: *Journal of Cell Biology* 216.8 (2017), pp. 2373–2389.

[150] Seung Min Kim et al. "Cancer-derived exosomes as a delivery platform of CRISPR/Cas9 confer cancer cell tropism-dependent targeting". In: *Journal of Controlled Release* 266.April (2017), pp. 8–16. URL: https://doi.org/10.1016/j.jconrel.2017.09.013.

[151] Ross C Wilson and Jennifer A Doudna. "Molecular Mechanisms of RNA Interference". In: *Annual Review of Biophysics* 42.1 (2013), pp. 217–239.

[152] Akihiro Urasaki, Kazuhide Asakawa, and Koichi Kawakami. "Efficient transposition of the Tol2 transposable element from a single-copy donor in zebrafish". In: *Proceedings of the National Academy of Sciences of the United States of America* 105.50 (2008), pp. 19827–19832.

[153] Donglei Zhang et al. "The piRNA targeting rules and the resistance to piRNA silencing in endogenous genes". In: *Science* 359.6375 (2018), pp. 587–592.

[154] Colin D. Malone and Gregory J. Hannon. "Small RNAs as Guardians of the Genome". In: *Cell* 136.4 (2009), pp. 656–668. URL: http://dx.doi.org/10.1016/j.cell.2009.01.045.

[155] Reinhard Koster and Wiebke A. Sassen. "A molecular toolbox for genetic manipulation of zebrafish". In: *Advances in Genomics and Genetics* (2015), p. 151.

[156] Hanhui Ma et al. "CRISPR-Cas9 nuclear dynamics and target recognition in living cells". In: *Journal of Cell Biology* 214.5 (2016), pp. 529–537.

[157] Baekgyu Kim, Kyowon Jeong, and V. Narry Kim. "Genome-wide Mapping of DROSHA Cleavage Sites on Primary MicroRNAs and Noncanonical Substrates". In: *Molecular Cell* 66.2 (2017), pp. 258–269. URL: http://dx.doi.org/10.1016/j.molcel.2017.03.013.

[158] Daniel Rosenblum et al. "CRISPR-Cas9 genome editing using targeted lipid nanoparticles for cancer therapy". In: *Science Advances* 6.47 (2020).

[159] G J Hannon. "RNA interference". In: *Nature* 418.6894 (2002), pp. 244–251.

[160] Anastasia Khvorova and Jonathan K Watts. "The chemical evolution of oligonucleotide therapies of clinical utility". In: *Nature Biotechnology* 35.3 (2017), pp. 238–248.

[161] Peter H Hagedorn et al. "Locked nucleic acid: modality, diversity, and drug discovery". In: *Drug Discovery Today* 23.1 (2018), pp. 101–114. URL: http://dx.doi.org/10.1016/j.drudis.2017.09.018.

[162] Amritraj Patra et al. "2-Fluoro RNA Shows Increased Watson-Crick H-Bonding Strength and Stacking relative to RNA: Evidence from NMR and Thermodynamic Data". In: *Angew Chem Int Ed Engl.* (2012).

[163] Michael Petersen et al. "The conformations of locked nucleic acids (LNA)". In: *Journal of Molecular Recognition* 13.1 (2000), pp. 44–53.

[164] Qiubing Chen, Ying Zhang, and Hao Yin. "Recent advances in chemical modifications of guide RNA, mRNA and donor template for CRISPR-mediated genome editing". In: *Advanced Drug Delivery Reviews* 168 (2021), pp. 246–258. URL: https://doi.org/10.1016/j.addr.2020.10.014.

[165] Cheng Qiong Mao et al. "Synthetic lethal therapy for KRAS mutant non-small-cell lung carcinoma with nanoparticle-mediated CDK4 siRNA delivery". In: *Molecular Therapy* 22.5 (2014), pp. 964–973.

[166] Christopher R Cromwell et al. "Incorporation of bridged nucleic acids into CRISPR RNAs improves Cas9 endonuclease speci fi city". In: *Nature Communications* 2018 (), pp. 1–11. URL: http://dx.doi.org/10.1038/s41467-018-03927-0.

[167] Daniel O'Reilly et al. "Extensive CRISPR RNA modification reveals chemical compatibility and structure-activity relationships for Cas9 biochemical activity". In: *Nucleic Acids Research* 47.2 (2019), pp. 546–558.

[168] Irina E Catrina, Salvatore A E Marras, and Diana P Bratu. "Tiny molecular beacons: LNA/2'-O-methyl RNA chimeric probes for imaging dynamic mRNA processes in living cells". In: *ACS Chemical Biology* 7.9 (2012), pp. 1586–1595.

[169] Andrew Tsourkas, Mark A Behlke, and Gang Bao. "Hybridization of 2-O-methyl and 2-deoxy molecular beacons to RNA and DNA targets". In: *Nucleic Acids Research* 30.23 (2002), pp. 5168–5174.

[170] Karen Martinez et al. "Locked nucleic acid based beacons for surface interaction studies and biosensor development". In: *Anal Chem* (2009).

[171] Chelsea Hull, Corinne Szewcyk, and Pamela M St. John. "Effects of locked nucleic acid substitutions on the stability of oligonucleotide hairpins". In: *Nucleosides, Nucleotides and Nucleic Acids* 31.1 (2012), pp. 28–41.

[172] Won Jong Rhee et al. "Target accessibility and signal specificity in live-cell detection of BMP-4 mRNA using molecular beacons". In: *Nucleic Acids Research* 36.5 (2008), pp. 1–10.

[173] Brenda F Baker et al. "2'-O-(2-methoxy)ethyl-modified anti-intercellular adhesion molecule 1 (ICAM-1) oligonucleotides selectively increase the ICAM-1 mRNA level and inhibit formation of the ICAM-1 translation initiation complex in human umbilical vein endothelial cells". In: *Journal of Biological Chemistry* 272.18 (1997), pp. 11994–12000.

[174] Annemieke Aartsma-Rus et al. "Guidelines for antisense oligonucleotide design and insight into splice-modulating mechanisms". In: *Molecular Therapy* 17.3 (2009), pp. 548–553.

[175] John Goodchild, Eddie Carroll, and Jay R Greenberg. "Inhibition of rabbit {β}-Globin synthesis by complementary oligonucleotides: Identification of mRNA sites sensitive to inhibition". In: *Archives of Biochemistry and Biophysics* 263.2 (1988), pp. 401–409.

[176] Lisa M Hochrein et al. "Signal Transduction in Human Cell Lysate via Dynamic RNA Nanotechnology". In: *ACS Synthetic Biology* 7.12 (2018), pp. 2796–2802.

[177] Michael Ma et al. "Intracellular mRNA cleavage induced through activation of RNase P by nuclease-resistant external guide sequences". In: *Nature Biotechnology* 18.1 (2000), pp. 58–61.

[178] Y Yuan, E S Hwang, and S Altman. "Targeted cleavage of mRNA by human RNase P". In: *Proceedings of the National Academy of Sciences of the United States of America* 89.17 (1992), pp. 8006–8010.

[179] Omar O Abudayyeh et al. "RNA targeting with CRISPR-Cas13". In: *Nature* 550.7675 (2017), pp. 280–284. URL: http://dx.doi.org/10.1038/nature24049.

[180] Abdul-rahman Allouche. "Software News and Updates Gabedit — A Graphical User Interface for Computational Chemistry Softwares". In: *Journal of computational chemistry* 32 (2012), pp. 174–182.

[181] Andreas R Gruber et al. "The Vienna RNA websuite." In: *Nucleic acids research* 36 (2008), pp. 70–74.

[182] Michael Kertesz et al. “Genome-wide measurement of RNA secondary structure in yeast”. In: *Nature* 467.7311 (2010), pp. 103–107.

[183] Julius B Lucks et al. “Multiplexed RNA structure characterization with selective 2-hydroxyl acylation analyzed by primer extension sequencing (SHAPE-Seq)”. In: *Proceedings of the National Academy of Sciences of the United States of America* 108.27 (2011), pp. 11063–11068.

[184] Byron Lee et al. “Comparison of SHAPE reagents for mapping RNA structures inside living cells”. In: *Rna* 23.2 (2017), pp. 169–174.

[185] Tobias Merkle et al. “Precise RNA editing by recruiting endogenous ADARs with antisense oligonucleotides”. In: *Nature Biotechnology* 37.2 (2019), pp. 133–138. URL: http://dx.doi.org/10.1038/s41587-019-0013-6.

[186] Dhruva Katrekar et al. “Robust RNA editing via recruitment of endogenous ADARs using circular guide RNAs”. In: *bioRxiv preprint* (2021).

[187] Kaiyi Jiang et al. “Programmable eukaryotic protein expression with RNA sensors”. In: *bioRxiv* (2022), p. 2022.01.26.477951. URL: https://www.biorxiv.org/content/10.1101/2022.01.26.477951v1%0Ahttps://www.biorxiv.org/content/10.1101/2022.01.26.477951v1.abstract.

[188] Simone Sciabola et al. “PFRED: A computational platform for siRNA and antisense oligonucleotides design”. In: *PLoS ONE* 16.1 January (2021), pp. 1–11. URL: http://dx.doi.org/10.1371/journal.pone.0238753.

[189] Sook Hee Ku et al. “Chemical and structural modifications of RNAi therapeutics”. In: *Advanced Drug Delivery Reviews* 104 (2016), pp. 16–28.

[190] Kwan T Chow, Michael Gale, and Yueh Ming Loo. “RIG-I and Other RNA Sensors in Antiviral Immunity”. In: *Annual Review of Immunology* 36 (2018), pp. 667–694.

[191] Tomoko Takahashi and Kumiko Ui-Tei. “Mutual regulation of RNA silencing and the IFN response as an antiviral defense system in mammalian cells”. In: *International Journal of Molecular Sciences* 21.4 (2020).

[192] Beeke Wienert et al. “In vitro – transcribed guide RNAs trigger an innate immune response via the RIG-I pathway”. In: *PLoS biology* (2018), pp. 1–18.

[193] Raymond V Fucini et al. “Adenosine modification may be preferred for reducing siRNA immune stimulation”. In: *Nucleic Acid Therapeutics* 22.3 (2012), pp. 205–210.

[194] Yale S Michaels et al. “Precise tuning of gene expression levels in mammalian cells”. In: *Nature Communications* 2019 (), pp. 1–12. URL: http://dx.doi.org/10.1038/s41467-019-08777-y.

[195] Catherine L Jopling. “Liver-specific microRNA-122: Biogenesis and function”. In: *RNA Biology* 9.2 (2012), pp. 137–142.

[196] Alessia Baccarini et al. “Kinetic analysis reveals the fate of a MicroRNA following target regulation in mammalian cells”. In: *Current Biology* 21.5 (2011), pp. 369–376. URL: http://dx.doi.org/10.1016/j.cub.2011.01.067.

[197] Saba Parvez et al. “MIC-Drop: A platform for large-scale in vivo CRISPR screens”. In: *Science* 373.6559 (2021), pp. 1146–1151.

[198] Karim Shalaby, Mustapha Aouida, and Omar El-Agnaf. "Tissue-specific delivery of crispr therapeutics: Strategies and mechanisms of non-viral vectors". In: *International Journal of Molecular Sciences* 21.19 (2020), pp. 1–27.

[199] Intellia Therapeutics. "Intellia Therapeutics Receives Grant to Develop Curative CRISPR / Cas9 In Vivo Sickle Cell Disease Treatments External Affairs {\&} Communications Investors :" in: *Keystone Symposia* 2020 (2020).

[200] Dylan Kotliar et al. "Identifying gene expression programs of cell-type identity and cellular activity with single-cell RNA-Seq". In: *eLife* 8 (2019), pp. 1–26.

[201] Farzad Abdolhosseini et al. "Cell Identity Codes: Understanding Cell Identity from Gene Expression Profiles using Deep Neural Networks". In: *Scientific Reports* 9.1 (2019), pp. 1–14. URL: http://dx.doi.org/10.1038/s41598-019-38798-y.

[202] Lilith V J C Mannack, Sebastian Eising, and Andrea Rentmeister. "Current techniques for visualizing RNA in cells". In: *F1000Research* 5.0 (2016), p. 775. URL: http://f1000research.com/articles/5-775/v1.

[203] Alexander A Green et al. "Complex cellular logic computation using ribocomputing devices". In: *Nature* 548.7665 (2017), pp. 117–121. URL: http://dx.doi.org/10.1038/nature23271.

[204] Alexander A Green et al. "Toehold switches: De-novo-designed regulators of gene expression". In: *Cell* 159.4 (2014), pp. 925–939. URL: http://dx.doi.org/10.1016/j.cell.2014.10.002.

[205] Philip J Santangelo et al. "Single molecule-sensitive probes for imaging RNA in live cells". In: *Nature Methods* 6.5 (2009), pp. 347–349.

[206] Sheel Shah et al. "Dynamics and Spatial Genomics of the Nascent Transcriptome by Intron seqFISH". In: *Cell* 174.2 (2018), pp. 363–376. URL: https://doi.org/10.1016/j.cell.2018.05.035.

[207] Xavier Pichon et al. "A Growing Toolbox to Image Gene Expression in Single Cells: Sensitive Approaches for Demanding Challenges". In: *Molecular Cell* 71.3 (2018), pp. 468–480. URL: https://doi.org/10.1016/j.molcel.2018.07.022.

[208] Feng Chen et al. "Click-encoded rolling FISH for visualizing single-cell RNA polyadenylation and structures". In: *Nucleic Acids Research* 47.22 (2019).

[209] Chatarina Larsson et al. "In situ detection and genotyping of individual mRNA molecules". In: *Nature Methods* 7.5 (2010), pp. 395–397.

[210] Harry M.T. Choi et al. "Mapping a multiplexed zoo of mRNA expression". In: *Development (Cambridge)* 143.19 (2016), pp. 3632–3637.

[211] Han Wu et al. "RNA imaging in living mice enabled by an: In vivo hybridization chain reaction circuit with a tripartite DNA probe". In: *Chemical Science* 11.1 (2020), pp. 62–69.

[212] Richard Weiszmann, Ann S Hammonds, and Susan E Celniker. "Determination of gene expression patterns using high-throughput RNA in situ hybridization to whole-mount drosophila embryos". In: *Nature Protocols* 4.5 (2009), pp. 605–618.

[213] Qilong Wu et al. "Sticky-flares for in situ monitoring of human telomerase RNA in living cells". In: *Nanoscale* 10.19 (2018), pp. 9386–9392.

[214] Kun Huang et al. "FASTmiR: an RNA-based sensor for in vitro quantification and live-cell localization of small RNAs". In: *Nucleic acids research* 45.14 (2017), e130.

[215] Hao Huang et al. "A G-quadruplex-containing RNA activates fluorescence in a GFP-like fluorophore". In: *Nature Chemical Biology* 10.8 (2014), pp. 686–691.

[216] Rita L Strack, Matthew D Disney, and Samie R Jaffrey. "A superfolding Spinach2 reveals the dynamic nature of trinucleotide repeat RNA". In: *Nature Methods* 10.12 (2013), pp. 1219–1224.

[217] Grigory S Filonov et al. "Broccoli: Rapid selection of an RNA mimic of green fluorescent protein by fluorescence-based selection and directed evolution". In: *Journal of the American Chemical Society* 136.46 (2014), pp. 16299–16308.

[218] Elena V Dolgosheina et al. "RNA Mango aptamer-fluorophore: A bright, high-affinity complex for RNA labeling and tracking". In: *ACS Chemical Biology* 9.10 (2014), pp. 2412–2420.

[219] Alexis Autour et al. "Fluorogenic RNA Mango aptamers for imaging small non-coding RNAs in mammalian cells". In: *Nature Communications* 9.1 (2018). URL: http://dx.doi.org/10.1038/s41467-018-02993-8.

[220] Wenjiao Song et al. "Imaging RNA polymerase III transcription using a photostable RNA-fluorophore complex". In: *Nature Chemical Biology* 13.11 (2017), pp. 1187–1194.

[221] Ranjan Batra et al. "Elimination of Toxic Microsatellite Repeat Expansion RNA by RNA-Targeting Cas9". In: *Cell* 170.5 (2017), pp. 899–912.

[222] Yuchen Liu et al. "Targeting cellular mRNAs translation by CRISPR-Cas9". In: *Scientific Reports* 6 (2016), pp. 2–10.

[223] D B T Cox et al. "RNA editing with CRISPR-Cas13". In: *Science* 0180 (2017), pp. 1–15.

[224] Silvana Konermann et al. "Transcriptome Engineering with RNA-Targeting Type VI-D CRISPR Effectors". In: *Cell* 173.3 (2018), pp. 665–676. URL: https://doi.org/10.1016/j.cell.2018.02.033.

[225] Steven C Strutt et al. "RNA-dependent RNA targeting by CRISPR-Cas9". In: *eLife* 7 (2018), pp. 1–17.

[226] Melissa Lopez-jones, Xiuhua Meng, and Robert H Singer. "Visualization of Dynamics of Single Endogenous mRNA Labeled in Live Mouse". In: *Science* 343.6169 (2014), pp. 422–424.

[227] Bin Wu, Jiahao Chen, and Robert H Singer. "Background free imaging of single mRNAs in live cells using split fluorescent proteins". In: *Scientific Reports* 4 (2014), pp. 11–13.

[228] C.-G. Cheong and T M T Hall. "Engineering RNA sequence specificity of Pumilio repeats". In: *Proceedings of the National Academy of Sciences* 103.37 (2006), pp. 13635–13639. URL: http://www.pnas.org/cgi/doi/10.1073/pnas.0606294103.

[229] David A Nelles et al. "Programmable RNA Tracking in Live Cells with CRISPR/Cas9". In: *Cell* 165.2 (2016), pp. 488–496. URL: http://dx.doi.org/10.1016/j.cell.2016.02.054.

[230] Luke A Gilbert et al. "CRISPR-Mediated Modular RNA-Guided Regulation of Transcription in Eukaryotes". In: *Cell* 154.2 (2012), pp. 442–451. URL: http://dx.doi.org/10.1016/j.cell.2013.06.044.

[231] Bastiaan Spanjaard et al. "Simultaneous lineage tracing and cell-type identification using CrIsPr-Cas9-induced genetic scars". In: *Nature Biotechnology* 36.5 (2018), pp. 469–473.

[232] Chiao Lin Chen et al. "Proteomic mapping in live Drosophila tissues using an engineered ascorbate peroxidase". In: *Proceedings of the National Academy of Sciences of the United States of America* 112.39 (2015), pp. 12093–12098.

[233] Jonathan S Gootenber and Omar O Abudayyeh. "Nucleic acid detection with CRISPR-Cas13a/C2c2". In: *Science* 356.6336 (2017), pp. 241–250. URL: http://science.sciencemag.org/content/sci/356/6336/438.full.pdf.

[234] Maturada Patchsung et al. "Clinical validation of a Cas13-based assay for the detection of SARS-CoV-2 RNA". In: *Nature Biomedical Engineering* (2020). URL: http://dx.doi.org/10.1038/s41551-020-00603-x.

[235] Cameron Myhrvold et al. "Field-deployable viral diagnostics using CRISPR-Cas13". In: *Science* 360.6387 (2018), pp. 444–448.

[236] Jonathan S Gootenberg et al. "Multiplexed and portable nucleic acid detection platform with Cas13, Cas12a and Csm6". In: *Science* 360.6387 (2018), pp. 439–444.

[237] James P Broughton et al. "CRISPR – Cas12-based detection of SARS-CoV-2". In: *Nature Biotechnology* (2019).

[238] Jonathan S Gootenberg. "CRISPR diagnostics". In: 372.6545 (2021), pp. 914–916.

[239] Sebastian M Castillo-Hair et al. "FlowCal: A User-Friendly, Open Source Software Tool for Automatically Converting Flow Cytometry Data from Arbitrary to Calibrated Units". In: *ACS Synthetic Biology* 5.7 (2016), pp. 774–780.