Harnessing Endogenous Human ADARs for RNA Repair

Dissertation

der Mathematisch-Naturwissenschaftlichen Fakultät

der Eberhard Karls Universität Tübingen

zur Erlangung des Grades eines

Doktors der Naturwissenschaften

(Dr. rer. nat.)

vorgelegt von

Philipp Reautschnig

aus Klagenfurt/Österreich

Tübingen

2019

Gedruckt mit Genehmigung der Mathematisch-Naturwissenschaftlichen Fakultät der Eberhard Karls Universität Tübingen.

Tag der mündlichen Qualifikation:	04.06.2019
Dekan:	Prof. Dr. Wolfgang Rosenstiel
1. Berichterstatter:	Prof. Dr. Thorsten Stafforst
2. Berichterstatter:	Prof. Dr. Ralf-Peter Jansen

Danksagung/Acknowledgments

Meinem Doktorvater Professor Thorsten Stafforst danke ich ganz herzlich für das hochinteressante und innovative Promotionsthema. Mit vielen konstruktiven Diskussionen und Ratschlägen stand er mir stets zur Seite und ermöglichte mir durch die Arbeit in seinem Labor, diese Dissertation zu verfassen.

Ebenso danke ich Professor Ralf Jansen für seine fachkundige und freundliche Unterstützung in allen Phasen meiner Promotion, auf welche ich jederzeit habe zählen können.

Ein besonderer Dank gebührt auch Jacqueline Wettengel, welche mich zu Beginn meiner Promotion bei vielen praktischen Fragen und dem Erlernen neuer Methoden unterstützte und mir im Laufe der Zeit zu einer guten Freundin wurde. Danke für die vielen aufrichtigen Gespräche! Sowohl die lockeren als auch die schwierigen, beide haben mein Leben bereichert!

Unseren großartigen Kooperationspartnern Dr. Sven Geisler, Professor Philipp Kahle und Professor Ulrike Naumann, die viele in dieser Arbeit beschriebene Projekte erst ermöglichten, danke ich für ihre unermüdliche Unterstützung in der Umsetzung unserer gemeinsamen Forschung.

Alfred Hanswillemenke, Gimi (Ngadhnjim Latifi), Anna Stroppel, Oliver Hess, Yvonne Füll, Karthika Devi, Paul Vogel und Tobias Merkle danke ich für erbauliche Diskussionen in unseren Labmeetings und für ihre konstruktiven Vorschläge. Meinen Studenten Madeleine Heep, Sophia Flad und Paul Bachmann danke ich für die erfolgreichen Arbeiten in unserem Labor, als auch für deren Beiträge zu meinem Projekt.

Roland Fritsch, Christoph Hoffmann und Jacqueline Wettengel danke ich ganz herzlich für ihre konstruktive Kritik an meiner Dissertation.

Meinen Eltern und Großeltern danke ich für ihren ständigen Beistand in allen Lebenslagen. Danke, dass ihr immer an mich glaubt und mich bedingungslos unterstützt!

Liebe Brigitte, danke für deine große Geduld und deine bedingungslose Unterstützung! Deine ermutigenden und liebevollen Worte halfen mir, selbst die härtesten Zeiten durchzustehen!

Table of Contents

ABBREVIATIONS

AA	amino acid
AAV	adeno associated virus
ACTB	β-actin
ADAR	adenosine deaminase acting on RNA
ADAR1 p110	ADAR1 110 kD isoform
ADAR1 p150	ADAR1 150 kD Isoform
ADAR-Flp cells	ADAR Flp-In T-REx cells
AdV / AdVs	adenovirus / adenoviruses
Ago2	argonaute 2 protein
APOBEC1	apolipoprotein B mRNA editing enzyme complex 1
ARNTL	aryl hydrocarbon receptor nuclear translocator-like protein 1
ASO	antisense oligonucleotide
A-to-I Editing	adenosine to inosine editing
BG	benzylguanine
BGH terminator	bovine growth hormone terminator
Cas	CRISPR associated protein
CCCP	carbonyl cyanide m-chlorophenyl hydrazone
CFTR	cystic fibrosis transmembrane conductance regulator
CLOCK	circadian locomotor output cycles kaput
Cluc	cypridina luciferase
CMV promoter	cytomegalovirus promoter
CPE	cytopathic effect
CRISPR	clustered regularly interspaced short palindromic repeats
C-to-U Editing	cytosine to uridine editing
DD	deaminase domain
DL	dual-luciferase
DSB	double strand break
dsRBD	double-stranded RNA binding domain
dsRBP	double-stranded RNA binding protein
dsRNA	double-stranded RNA
E1A	early region 1A protein
ECS	editing complementary sequence

ED50	effective dose 50
eGFP	enhanced green fluorescent protein
EIE	editing inducer element
FRT-site	flippase recognition target site
GAPDH	glyceraldehyde 3-phosphate dehydrogenase
GluA2 / GluR2	glutamate receptor ionotropic, AMPA 2 / glutamate Receptor 2
GPI	glucose-6-phosphate isomerase
GRIA2	glutamate ionotropic receptor AMPA type subunit 2
gRNA	guide RNA
GUSB	β-glucuronidase
hADAR	human ADAR
hAGT	human O6-alkylguanine-DNA alkyltransferase
HDR	homology directed repair
IFN-α	interferon alpha
IRES	internal ribosomal entry sites
IVT	in-vitro transcribed
KD	knock-down
KO	knock-out
KPNA1	karyopherin alpha 1
LOF mutation	loss-of-function mutation
MAVS	mitochondrial antiviral-signaling protein
MCP	MS2 coat protein
MDA5	melanoma differentiation antigen 5
miRNA	micro RNA
MOI	multiplicity of infection
MPP	matrix processing peptidase
NES	nuclear export sequence
NHEJ	non-homologous end joining
NLS	nuclear localization sequence
OD	optical density
ORF	open reading frame
PAM	protospacer adjacent motif
PARL-Protease	presenilin-associated rhomboid-like-protease
PFS	protospacer flanking site

PINK1	PTEN-induced putative kinase 1
PMA	phorbol 12-myristate 13-acetate
PPIB	peptidyl-prolyl cis-trans isomerase B
PRR	pattern recognition receptor
Q/R editing site	glutamine to arginine editing site
qPCR	quantitative polymerase chain Reaction
R/G editing site	arginine to glycine editing site
RC	recruitment cluster
RCF	recruitment cluster finder
RISC	RNA-induced silencing complex
RLU	relative light units
RNAi	RNA Interference
ROS	reactive oxygen species
SDRE	site-directed RNA editing
SINE	short interspersed nuclear element
siRNA	short interfering RNA
SNP	single nucleotide polymorphism
snRNA	small nuclear RNA
SV40 origin	simian virus 40 origin
TALEN	transcription-activator-like-effector nuclease
TRN1	transportin-1
tRNA	transfer RNA
US FDA	united states food and drug administration
UTR	untranslated region
VCP	valosin-containing protein
WWP2	WW domain containing E3 ubiquitin protein ligase 2
XPO5	exportin-5
ZFN	zinc-finger nuclease

LIST OF FIGURES

LIST OF TABLES

SUMMARY

Site-directed adenosine-to-inosine (A-to-I) RNA editing is a novel transcriptome engineering approach that allows to recode genetic information at RNA level by specific deamination of adenosines. Because the resulting inosine is read as guanosine by cellular machines and during translation, site-directed RNA editing (SDRE) does ultimately introduce A-to-G base substitutions into mRNAs. This opens up the opportunity to recode Start- and Stop-codons, splice signals, miRNA recognition sites, and 12 of the 20 canonical amino acids. Thus SDRE allows a wide range of interventions, including the manipulation of residues relevant for signaling or protein function, and the ability to correct G-to-A point mutations.

Our R/G guideRNA SDRE system allows to steer wild-type human ADAR (adenosine deaminase acting on RNA) enzymes towards selected target adenosines within user-defined mRNAs. The genetically encodable R/G gRNA design, which was developed during this thesis based on successive rounds of rational design, allows for the first time to recruit endogenous human ADAR enzymes. In contrast to all other SDRE approaches, the R/G gRNA system does therefore not require the overexpression of an editase. Transcriptome engineering can now be achieved by the simple application of a plasmid or adenovirus encoded short guideRNA. Up to 40% correction of a premature 5'-UAG stop codon into a 5'-UIG tryptophan codon could be achieved this way with a plasmid encoded R/G gRNA targeting a dual-luciferase reporter construct in HeLa cells.

During this thesis, several fundamental questions, which had to precede the further development of SDRE into a therapeutic application, could be solved. The R/G gRNA system became the first SDRE approach to prove editing of endogenous mRNAs. In addition, it was used to recode the recessive PINK1 W437X amber loss-of-function mutation in HeLa cells and could show that SDRE allows the functional rescue of PINK1-Parkin-mediated mitophagy, which is linked to the etiology of Parkinson's disease.

Our R/G gRNAs have high potential to become a next generation drug for the precise correction of disease-causing point mutations and the tuneable manipulation of protein function. In doing so, they would complement existing genome and transcriptome manipulation strategies by enabling interventions out of reach for currently available molecular tools.

ZUSAMMENFASSUNG

„Gerichtete A-nach-I RNA-Editierung“ ist ein neues Transkriptom-Manipulationsverfahren, das die Desaminierung ausgewählter Adenosine und damit deren Umwandlung zu Inosin ermöglicht. Da Inosin während der Transkription und Translation als Guanosin erkannt wird, können mit dieser Methode genetische Informationen auf der RNA-Ebene bearbeitet werden. Im Grunde fügt das Verfahren also gezielt A-nach-G Punktmutationen in mRNAs ein und erlaubt damit die Umcodierung von Start- und Stopp-Codons, Spleißstellen, miRNA-Bindestellen, sowie 12 der 20 kanonischen Aminosäuren. Hierdurch wird eine Vielzahl an Eingriffen möglich, wie die Korrektur von G-nach-A Punktmutationen, die Manipulation von Proteinfunktionen durch Änderung relevanter Aminosäuren, oder die Modulation zellulärer Signalkaskaden.

Das von uns entwickelte R/G guideRNA System dirigiert humanes Wildtyp-ADAR (adenosine deaminase acting on RNA) zu bestimmten Ziel-Adenosinen in ausgewählten mRNAs. Durch sukzessives rationales Design wurde im Zuge dieser Arbeit, ein genetisch codierbares R/G gRNA Design entwickelt, das es erstmals erlaubt endogenes humanes ADAR zu rekrutieren. Im Gegensatz zu allen anderen gegenwärtig verfügbaren gerichteten RNA-Editierungsverfahren, ist es daher nicht auf die Überexpression einer Editase angewiesen. Hierdurch kann nun genetische Information auf Ebene des Transkriptoms durch simples Verabreichen kurzer Plasmid- oder Adenovirus-codierter guideRNAs manipuliert werden. Bei der Reparatur eines vorzeitigen 5‘-UAG Stopp-Codons konnten mittels Plasmid-codierter gRNAs in HeLa Zellen bis zu 40% der Ziel-mRNAs korrigiert werden. Zusätzlich konnten in dieser Arbeit grundlegende Fragen beantwortet werden, die der Weiterentwicklung von gerichteter RNA-Editierung zu einer therapeutischen Anwendung vorangehen mussten. Es konnte erstmals gezeigt werden, dass durch das R/G gRNA-System auch die Editierung endogener mRNAs möglich ist. Des Weiteren wurde die rezessive PINK1-W437X-Amber Funktionsverlustmutation in HeLa Zellen korrigiert. Das Verfahren erlaubte es hierbei, die Funktion des durch die genannte Mutation gestörten PINK1-Parkin-vermittelten Mitophagie-Signalwegs, welcher mit der Ätiologie der Parkinsonkrankheit verknüpft ist, wieder herzustellen.

Unsere R/G gRNAs haben das Potential als völlig neue Generation von Medikamenten beispielsweise für die Behandlung von Punktmutationen oder therapeutische Regulierung von Proteinfunktionen eingesetzt zu werden. Hierbei würden bestehende Ansätze zur Genom- und Transkriptom-Manipulation durch bisher undurchführbare Eingriffe auf RNA-Ebene ergänzt.

LIST OF PUBLICATIONS AND PERSONAL CONTRIBUTIONS

Review articles (peer-reviewed)

PUBLICATION 1: Reautschnig P.*, Vogel P.*, Stafforst T. The notorious R.N.A. in the spotlight - drug or target for the treatment of disease. RNA Biol, 2017. 14(5): p. 651-668. * equal contribution

Personal contribution: I contributed to the conceptualization of the review article, prepared Table 1 and wrote the sections: “Therapeutic RNAi”, “Therapeutic mRNA” and “Oligonucleotides for vaccination and desensitization”.

Original research articles (peer-reviewed)

PUBLICATION 2: Wettengel J.*, Reautschnig P.*, Geisler S., Kahle P.J., Stafforst T. Harnessing human ADAR2 for RNA repair - Recoding a PINK1 mutation rescues mitophagy. Nucleic Acids Res, 2017. 45(5): p. 2797-2808. * equal contribution

Personal contribution: I contributed to the preparation and conceptualization of the study (generating the ADAR2 Flp-In T-Rex cell line; performing preliminary experiments to establish the PINK1-Parkin mitophagy assay). Additionally, I designed and conducted parts of the experimental work and data collection (datasets for the Figures 3, 4, S9, S15, S17, S19, S22 and S23 and cloning of the pEdit plasmid backbone, including the variants containing five copies of the R/G gRNAs). Other contributions include the preparation of supporting information, involvement in review and discussion of the manuscript, and the execution of follow-up experiments in response to reviewers’ comments on the manuscript.

PUBLICATION 3: Heep M., Mach P., Reautschnig P., Wettengel J., Stafforst T. Applying Human ADAR1p110 and ADAR1p150 for Site-Directed RNA Editing-G/C Substitution Stabilizes GuideRNAs against Editing. Genes (Basel), 2017. 8(1).

Personal contribution: I designed and cloned the novel gRNAs, planned and supervised the experiments performed by MH and gave scientific advice to PM for the generation of the ADAR1 Flp-In T-Rex cell lines.

PUBLICATION 4: Merkle T., Merz S., Reautschnig P., Blaha A., Li Q., Vogel P., Wettengel J., Li J.B., Stafforst T. Precise RNA editing by recruiting endogenous ADARs with antisense oligonucleotides. Nat Biotechnol, 2019. 37(2): p. 133-138.

Personal contribution: I designed and conducted cell culture experiments, contributed to writing the manuscript, and prepared parts of the supporting information. The sequences of the ASO v4, v9.4 and 9.5 are based on continuous rational design improvements of R/G gRNAs performed during my thesis. I created the luciferase reporter plasmid used for Figure S1, established the western-blot, as well as the siRNA-KO protocols, and generated data for the Figures 2, S1 and S4. Additionally, I obtained HeLa, SH-SY5Y, SK-N-BE, U87MG and A549 cells from cooperation partners and established suitable cell culture settings for them.

1. INTRODUCTION

1.1 RNA editing

The term RNA editing describes a process that can co- or post-transcriptionally change information encoded within RNA transcripts, compared to the original genomic template DNA. This allows the expression of functionally different proteins from a single gene, independent of alternative splicing, thus further increasing the transcriptome and proteome diversity [4]. RNA editing might therefore represent a force of adaptive evolution leading to more complex life [5].

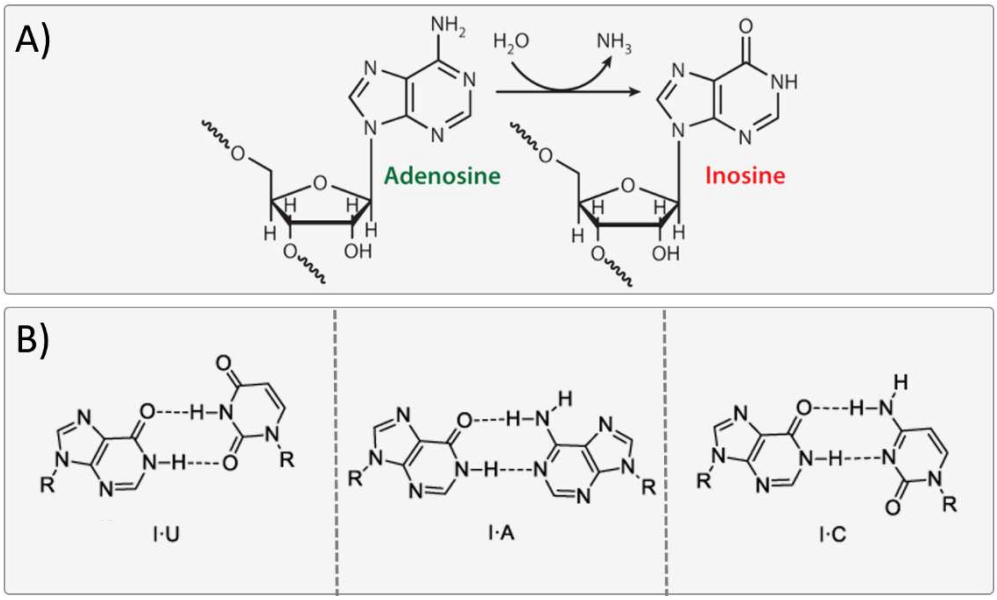


Figure 1-1: The inosine RNA-modification is introduced by deamination of adenosine and results in relaxed base pairing properties.

(A) A hydrolytic deamination reaction converts adenosine into inosine. (B) The relaxed base pairing properties of inosine allow it to form base pairs with uridine, adenosine and cytosine. The figure was assembled from images adopted from reference [6] and [7].

In general RNA editing can occur by insertion, deletion or by substitution of bases through RNA modification. Trypanosomes for example, recode their mitochondrial RNAs by insertion and deletion of uridine residues by TUTases [8]. However, the most common forms of RNA editing in mammals are adenosine-to-inosine (A-to-I) and cytidine-to-uridine (C-to-U) base substitutions. Both are introduced by specific deamination reactions, that change the base pairing properties of the modified bases and can be functionally distinguished. C-to-U RNA editing is performed by the multi-protein complex APOBEC-1 (apolipoprotein B mRNA editing enzyme complex 1), which is acting only on single-stranded RNA. The focus of this thesis lies on A-to-I RNA editing, which is performed by ADAR enzymes (adenosine deaminase acting on RNA). They are believed to form homodimers and require double-stranded RNA as substrate [9].

The hydrolytic deamination of adenine at its C6 position by ADAR enzymes results in hypoxanthine, the nucleobase of inosine (Figure 1-1A). Inosine forms stable Watson-Crick interactions with cytidine, but can also form weaker interactions with uridine and adenosine (Figure 1-1B) [21].

Until recently, it was commonly accepted that inosine-modifications within mRNAs are always read as guanosine by cellular machines and during translation [10]. It was believed that the relaxed base pairing with uridine and adenosine would be only relevant for the decoding of wobble base pairs at the first anticodon position of tRNAs, allowing a single tRNA to decode multiple codons [22, 23]. Although there is now *in-vitro* evidence that inosine within mRNAs could also be read as adenosine or uridine, and therefore could result in mostly rare codon-context dependent alternative decoding events, this needs to be verified *in-vivo* [10]. Until then, it can still be assumed that A-to-I editing of mRNA transcripts basically introduces A-to-G point mutations at the RNA level.

Consequently, naturally-occurring RNA editing is able to influence many important cellular functions. Co-transcriptional RNA editing at splice-sites, for example, can lead to the expression of alternative protein isoforms by the removal of natural, or creation of cryptic, splice-sites, causing intron-retention or exonization [11]. It can also modulate RNAi by A-to-I editing within precursor RNAs (e.g. pri-, pre-miRNAs, endo-siRNAs), thus affecting their maturation or specificity, as well as manipulate miRNA binding sites within mRNAs [12-15]. Some ADAR enzymes have important functions associated with innate immunity (section 1.1.1.1.4) and circadian regulation of gene expression (section 1.1.1.2.1) [16, 17]. Finally, and most importantly for this thesis, A-to-I editing events can recode transcripts. Well-characterized examples of naturally-occurring recoding with functional consequences mostly come from mRNAs encoding neuronal transporters and channel proteins in the CNS, like the human GRIA2-transcript, which encodes a subunit of the glutamate receptor 2 (GluR2). RNA editing within the GRIA2-transcript changes Glutamine (Q) to Arginine (R) at a specific site, called Q/R-Editing site, and drastically reduces the permeability of the resulting receptor channel for Ca^{2+} [18-20]. Further well-known examples for recoding events by A-to-I RNA editing include the changed velocity of a Na^{+}/K^{+} ATPase transporter [21], the modulation of G-protein/receptor interaction efficacies of the 5-HT(2C) serotonin receptor [22], and the inactivation of the human K^{+} channel K(V)1.1 [23].

Despite the obvious functional importance of recoding events, the majority of A-to-I RNA editing events is found in non-coding RNA, like Introns and 3’ UTRs [24, 25]. A-to-I conversion induced by natural RNA editing can range from barely detectable to nearly full conversion and can be different between cell types, tissues [26], and developmental stages [27]. The fact that abnormal A-to-I RNA editing and ADAR expression are associated with several diseases, as explained in section 1.1.1.4, does emphasise that RNA editing is a highly regulated process. A complex set of factors determines the selection of adenosines within dsRNA substrates, and affects the percentage of A-to-I conversion at

any given site, which will be discussed in more detail in section 1.1.2. In most cases, these factors can be tied to structural features within the functional domains of ADAR enzymes, which are summarized in the next section.

1.1.1 The human ADAR protein family

The human ADAR protein family consists of three proteins: ADAR1, ADAR2, and ADAR3, which contain similar functional domains. A deaminase domain at the C-terminus and two (ADAR2, ADAR3) or three (ADAR1) double-strand RNA binding domains (dsRBD) at the N-terminus [14]. While the binding to dsRNA-substrates is predominantly accomplished via dsRBDs, the catalytic activity of ADARs is executed by their deaminase domains.

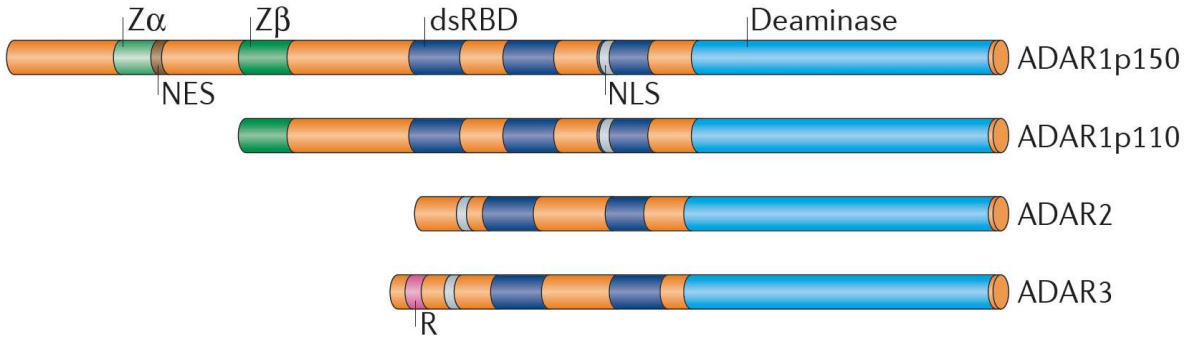


Figure 1-2: Domain structure of the human ADAR protein family.

The C-terminal deaminase domains are indicated in light blue. The double-stranded RNA binding domains (dsRBD) are indicated in dark blue. The Zα and Zβ parts of the Z-DNA/RNA binding domain are indicated in light and dark green. The nuclear localization sequences (NLS) are indicated in grey. The nuclear export sequences (NES) of ADAR1 p150 is indicated in brown. The arginine-(R)-rich domain of ADAR3 is indicated in red. This figure was adopted from reference [14].

1.1.1.1 ADAR1 exists in two distinct isoforms

ADAR1 has two reported isoforms, generated from the same locus: ADAR1 p110 and ADAR1 p150. While both are expressed from different promoters, their mRNAs differ only in the first exon and the utilized translation initiation codon, which is localized in exon 2 for the short ADAR1 p110 isoform, and in exon 1 for the long ADAR1 p150 isoform [28].

The short ADAR1 p110 isoform is expressed from a constitutively active promoter and consists of a deaminase domain, three dsRBDs, and the Zβ-Part of a Z-DNA/Z-RNA binding domain that is not functional when separated from its Zα-Part (Figure 1-2) [29]. ADAR1 p110 is expressed ubiquitously in most human tissues [30]. Under normal conditions, transportin-1 (TRN1) mediates its transport into the nucleus, where it predominantly accumulates in the nucleoli due to its dsRBDs [31, 32].

The long ADAR1 p150 isoform is expressed from an interferon(IFN)-inducible promoter and, compared to ADAR1 p110, includes an additional amino-terminal stretch with a nuclear export signal (NES), as

well as the missing Zα-Part of the Z-DNA/Z-RNA binding domain, rendering it functional (Figure 1-2) [14, 33, 34]. Exportin-5 (XPO5) mediates the transport of ADAR1 p150 into the cytosol.

Although both isoforms are known to shuttle between nucleus and cytosol, ADAR1 p110 is localized nearly exclusively in the nucleus, while ADAR1 p150 is predominantly localized in the cytosol [32]. Corresponding to their localization, both enzymes perform also different editing-dependent, as well as editing-independent, functions within the cell, which will be explained in the following sections.

1.1.1.1.1 ADAR1 p110 can suppress apoptosis under stress conditions

Stress factors, like UV radiation, or heat shocks, can stimulate the phosphorylation of ADAR1 p110 via the MAP kinase pathway (MKK6-p38-MSK MAP kinases). This promotes binding of exportin-5 (XPO5) and the subsequent transport of ADAR1 p110 into the cytosol (Figure 1-3).

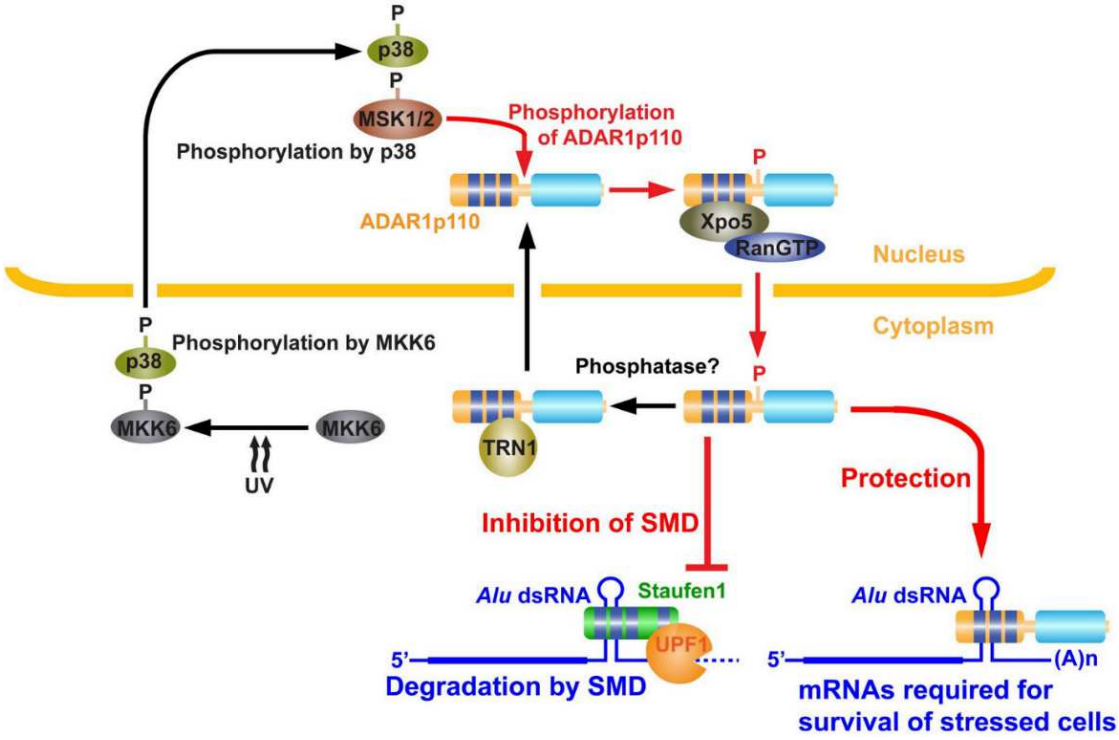


Figure 1-3: Stress response function of ADAR1p110 regulated by MKK6-p38-MSK1/2 MAP kinases.

Stress-response function of ADAR1p110 regulated by MKK6–p38–MSK1/2 MAP kinases. Stress, such as UV irradiation and heat shock, activates the MKK6–p38–MSK1/2 signaling cascade. Stress-induced phosphorylation of ADAR1p110 by MSK1 and MSK2 in the nucleus promotes ADAR1p110's binding to Xpo5 and export to the cytoplasm. ADAR1p110, perhaps dephosphorylated by a currently unknown phosphatase, binds TRN1 and is imported into the nucleus. After being translocated to the cytoplasm, ADAR1p110 inhibits binding of Staufen1 to Alu dsRNA present in the 3' UTRs of many antiapoptotic mRNAs and promotes survival of stressed cells. This figure and caption were adopted from reference [35].

Due to this localisation change, ADAR1 p110 can then compete with the cytosolic double-stranded RNA binding protein (dsRBP) Staufen-1 for its substrates. Staufen-1 binds dsRNA structures primarily composed of inverted Alu repeats, which are often present in the 3'UTRs of anti-apoptotic transcripts,

and facilitates their degradation in a process called Staufen-1-mediated decay. By preventing the decay of these nearly 500 different mRNAs, which are required for the survival of stressed cells, ADAR1 p110 can suppress apoptosis. This function of ADAR1 p110 is completely independent from its A-to-I editing activity [35, 36].

1.1.1.1.2 ADAR1 p110 regulates 3' UTR usage and alternative polyadenylation

Without stress factors, ADAR1 p110 stays within the nucleus, where it is involved in the regulation of alternative 3' UTR usage and alternative polyadenylation. Some of these events are editing-dependent, while others are not. It is assumed that ADAR1 p110 creates or destroys splice sites in the former and competes with relevant dsRBP in the latter cases. While the exact mechanism is not clear yet, the resulting change of miRNA- and regulatory protein binding sites represents a function of ADAR1 p110 in the regulation of the affected RNAs [37, 38].

1.1.1.1.3 The ADAR1 p150 Zαβ domain - An intensely discussed topic

The Z-DNA/Z-RNA binding domain (Zαβ domain) of ADAR1 p150 and its NES represent the most prominent differences compared to ADAR1 p110. While the purpose of the NES is self-explanatory, the function of the Zαβ domain is still widely discussed.

It was proposed that the Zαβ domain might facilitate dimerization of ADAR1 p150 and direct the enzyme to actively-transcribing genes to act upon the RNA prior to splicing [39], that it might allow preferential binding to dsRNA substrates containing Z-RNA forming motifs and thereby increases their specific editing [40], and that the Zαβ domain might play an important role in the recognition of foreign nucleic acids in the cytoplasm during host–pathogen interactions [41].

There is also evidence that the Zαβ domain promotes recruitment of ADAR1 p150 into stress granules. Under stress conditions, this results in the co-localisation with Tudor-SN, a nuclease involved in the decay of hyper-edited, inosine containing miRNA precursors [42-45]. The discovery that Tudor-SN promotes the activity of Endonuclease V, a ribonuclease that is specific for inosine-containing RNA, lead to the hypothesis that the Zαβ domain could be involved in the regulation of genes containing extensively-edited Alu dsRNAs [14, 46]. How this hypothesis would integrate with the involvement of ADAR1 p110 in Staufen1-mediated decay of Alu containing mRNAs under apoptotic stress is still open for debate. Important factors to better understand these interactions could be the specific kind of stressor, the expression level of ADAR1 p150 under non-pathogen stress, and the exact mechanism of the hypothetical Tudor-SN/Endonuclease V regulatory pathway.

While there is a lot of debate regarding the specific functions of the Zαβ domain, it was demonstrated that ADAR1 p150, as a whole, serves important functions in B cell development, intestinal homeostasis, and the innate immune response [47]. The latter is discussed next.

1.1.1.1.4 ADAR1 p150 enables self/non-self RNA sensing

To initiate innate immune responses only in the presence of pathogens, cells must have a way to discriminate cellular self-RNAs from extracellular non-self RNAs. The sensing of double-stranded, foreign nucleic acids (e.g. RNA viruses) by the cytosolic pattern recognition receptor MDA5 results in type-I interferon (IFN) production. IFN-I does in turn induce apoptosis of infected cells and causes an antiviral state in healthy cells by expression of hundreds of interferon-stimulated genes [48]. However, endogenous mRNAs containing dsRNA structures, like Alu repeats stay unnoticed by MDA5 [17, 47, 49]. This discrimination is achieved by A-to-I editing of endogenous dsRNAs through ADAR1 p150. A specific knock-out of the ADAR1 p150 isoform results in the accumulation of endogenous, immunostimulatory RNAs that are specifically detected by MDA5 and cause IFN-I production [47]. Consequently RNA editing by ADAR1 p150 enables cells to differentiate self-RNAs from non-self RNAs and to avoid autoinflammation [50, 51]. The fact that ADAR1 p150 prevents self-RNA sensing that would cause IFN production, but is expressed from an IFN inducible promoter, highlights that it is regulated in a negative feedback loop.

One of the mentioned interferon-stimulated genes is the protein kinase R (PKR). When activated by foreign dsRNAs, it induces translational shutdown and cell death. Like MDA5, the PKR can be activated by unedited endogenous dsRNAs as well. However, to prevent PKR activation by endogenous dsRNAs, the editing activity of both ADAR1 isoforms is required [52].

1.1.1.2 ADAR2

ADAR2 is most abundantly expressed in brain and heart tissues [53, 54]. It is an important mediator of recoding events that alter ion-channels in the CNS, like the Q/R as well as R/G editing of GRIA2 (manipulation of Ca^{2+} permeability and receptor desensitization) or the I/V editing of KCNA1 (K^{+} channel inactivation) [14].

ADAR2 contains an amino-terminal NLS that allows binding to importin-α and facilitates its transport into the nucleus, while its dsRBDs are necessary for its further dynamic enrichment at the nucleoli or at other sites of dsRNA accumulation within the nucleus [32, 55, 56]. After its import, ADAR2 is modified by the prolyl-isomerase PIN1 at proline 33 and phosphorylated by a still unknown kinase at threonine 32 to stabilize its nuclear localisation. In the cytosol, ADAR2 is rapidly degraded after being processed by the E3 ubiquitin ligase WWP2 [14, 57].

ADAR2 regulates its own expression in a negative feedback loop by auto-editing a proximal 3' splice-acceptor site within its own pre-mRNA, introducing a frameshift in the coding region that results in loss-of-function [11, 58].

1.1.1.2.1 ADAR2 is a key player in the circadian clockwork

ADAR2 is itself a circadian gene and as such linked to circadian rhythms, which are endogenously generated rhythms that synchronize major physiological processes with the environmental light-dark cycle of the solar day (e.g. the sleep/wake cycles or rhythms of hormonal secretion) [59].

It is regulated by the mammalian clock gene regulatory network (CLOCK–ARNTL) and in turn generates cyclic editing and mRNA rhythms for hundreds of downstream genes in the mouse liver. Thus, ADAR2 is important for various physiological aspects, like the speed of circadian oscillation rhythms or the light-induced phase-shift of the circadian clock, which makes A-to-I editing a key mechanism of post-transcriptional regulation in the circadian clockwork [16, 60, 61].

1.1.1.3 ADAR3 is assumed to negatively regulate A-to-I Editing by other ADARs

While ADAR1 and ADAR2 are catalytically active deaminases, ADAR3 is not only catalytically inactive, but also inhibits RNA editing performed by the other ADAR enzymes, indicating a regulatory function [62, 63]. The amino-terminal, arginine-rich single-stranded RNA binding domain (ssRBD) of ADAR3 functions as NLS [56] and enables ADAR3 to bind ssRNA, at least *in-vitro* [62]. A study suggested that ADAR3 has a significant impact on learning and memory in mice, presumably through regulation of genes involved in synaptic plasticity. The exact molecular mechanisms, however, remain unclear [64]. ADAR3 has been investigated far less than the catalytically active deaminases ADAR1 and ADAR2.

1.1.1.4 ADARs and their connection to disease

The participation of ADARs in numerous vital cellular processes, which are described above, highlights the importance of maintaining a homeostasis in cellular editing. Unphysiological changes in A-to-I RNA editing and/or ADAR expression levels are connected to disease.

The local abundance of ADAR2 in the brain supports its association with a broad spectrum of pathological CNS phenotypes. While ADAR2-KO leads to seizures and early death in mice [46, 47], aberrant editing of ADAR2 targets is connected to schizophrenia [48, 49], depression [50-53], autism [54], and amyotrophic lateral sclerosis [55-57]. Mice expressing deaminase deficient ADAR2 show adult onset obesity caused by hyperphagia, but not by metabolic derangement [58].

ADAR1 p150 deficiency results in accumulation of endogenous, immune-stimulatory RNAs that cause autoinflammation, while the absence of ADAR1 p110 might prevent the protection of anti-apoptotic transcripts from Staufen1-mediated decay (see section 1.1.1.1 for more details). The findings explain why a general ADAR1-KO, removing both isoforms, is embryonically lethal and accompanied by overproduction of type I IFNs and massive apoptosis in hematopoietic cells [65-67]. This fatality, which was also observed after a selective ADAR1-p150-KO in mice, can be neutralized by a simultaneous ADAR1-

p150-MAVS- or ADAR1-p150-MDA5-KO, which interrupts the MDA5 signaling pathway and prevents autoimmunity by accumulation of non-edited RNAs [47].

Although it is still unclear why edited RNAs are not detected by MDA5, it is indeed the RNA editing activity of ADAR1 p150 that is essential to prevent autoinflammation. Mice that expressed only deaminase deficient ADAR1^{E861A} died at embryonic day 13.5, like the mice with a full ADAR1-KO [68].

Accordingly, it is assumed that ADAR1 mutations are connected to autoimmune diseases. Mutations of ADAR1 are known to cause Aicardi-Goutières syndrome (AGS), an early onset childhood encephalitis with symptoms that mimic a congenital viral infection [69]. The disease leads in most cases to permanent neurological damage with severe intellectual and physical disability. It is accompanied by elevated concentrations of IFN-α in the cerebrospinal fluid (CSF)[70]. Another well-known example is dyschromatosis symmetrica hereditaria (DSH), an autosomal dominant pigmentary genodermatosis. It is characterized by a mixture of hyperpigmented and hypopigmented macules distributed on the dorsal aspects of the extremities [71] and appears to be a condition connected to the ADAR1 p150 isoform and IFN signaling [72]. Bilateral striatal necrosis (BSN), a dystonia caused by symmetrical degeneration of subareas of the striatum, is in some cases also caused by ADAR1 mutations and associated with a different IFN signature [73]. A further example is hereditary spastic paraplegia, a degenerative neurological disorder, which causes progressive spasticity and weakness in the legs [74]. A-to-I editing activity seems to be a very important factor in the mentioned diseases, since many patients share a common G1007R substitution within the ADAR1 deaminase domain that prevents editing completely [75, 76].

Finally, ADAR1 seems to also play a role in certain types of cancer. Aberrant editing and dysregulation of ADAR1 expression have been associated with basal cell carcinoma [77], breast cancer [78, 79], cervical cancer [80], colorectal cancer [81], esophageal squamous cell carcinoma [82], gastric cancer [83, 84], glioma [85], hepatocellular carcinoma [86], leukemia [87], lung cancer [88], medulloblastoma [89], melanoma [90], multiple myeloma [89], and prostate carcinoma [85].

1.1.2 The current mechanistic understanding of A-to-I RNA editing

The essential prerequisites for A-to-I RNA editing to happen are a dsRNA structure containing an adenosine as a substrate and the presence of an ADAR enzyme. The ADAR enzyme then binds to the dsRNA via its dsRBDs and positions its deaminase domain to remove the C6 amino group from the presented adenosine. Based on this oversimplified description, one would be tempted to assume the only relevant factors for RNA editing are the expression levels of enzyme and target RNA. Contrary to this, the expression levels play a minor role for the resulting percentage of A-to-I conversion (editing yield) at any given editing site. This is supported by the fact that the editing yield increases throughout neuronal

development in mice, while the ADAR expression levels stay the same [27]. Our mechanistic understanding of how specific adenosines are selected for RNA editing and which factors determine the rate of conversion into inosine is still very limited, but has significantly improved over the last decade. These findings include which editing modes exist, how and from which components editable dsRNA substrates are formed, how substrate specificity can be diverse between members of the ADAR protein family, how deaminase domains can target a specific adenosine, and which other factors enhance, compete with, or prevent A-to-I RNA editing.

1.1.2.1 The connection between ADARs, A-to-I editing modes and RNA substrate structure

Currently the literature distinguishes two modes of A-to-I RNA editing, which do not necessarily have to be mechanistically different. Site-selective RNA editing and non-selective RNA editing (formerly called hyper-editing) [91]. The term site-selective RNA editing describes an editing event that takes place at a specific adenosine of a particular transcript, while nearby adenosines stay unaffected. In most cases, these specifically-changed adenosines represent recoding events, which change the sense of the transcript and occur in dsRNAs with imperfect base pairing, bulges, and loops [91]. Among all RNA editing events that happen in the transcriptome, site-selective editing embodies the smallest fraction, but at the same time comprises some of the most important single editing events. A good example is the Q/R editing of GluR2, which recodes a C**A**G codon (Q) into a C**I**G codon (R) and is essential for survival in mammals [92]. Because ADARs can only edit dsRNAs, site-selective editing requires an editing complementary sequence (ECS), often localized within an adjacent intron [2]. The ECS allows the formation of a dsRNA structure as substrate for the deaminase domain and the dsRBDs of the different ADARs (see Figure 1-4). The highly efficient editing (up to 100% conversion) found at some mRNA (e.g. GluR2 Q/R site) or miRNA editing target sites, requires the additional presence of an adjacent long stem loop structure, referred to as editing inducer element (EIE), that facilitates the local enrichment of ADARs (see figure Figure 1-5) [3].

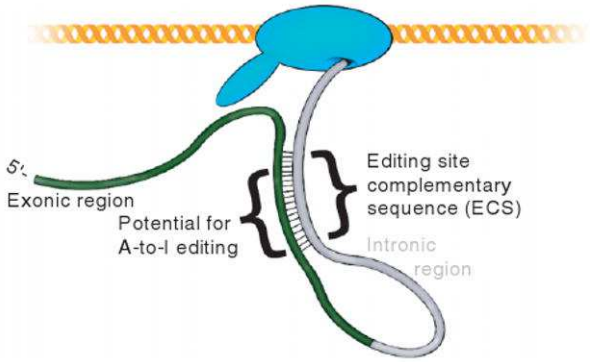


Figure 1-4: Co-transcriptional formation of a dsRNA substrate structure for A-to-I editing, between an exonic region, containing potential editing target sites, and an intronic ECS.

This figure and caption were adopted from reference [2]

Excessive editing is called hyper-editing when it affects over 50% of all adenosines within >100 bp long, perfectly paired dsRNA structures [93]. This term was introduced after early *in-vitro* studies, however, and this kind of excessive editing has been rarely found *in-vivo*, since. This lead to the slightly more accurate term non-selective RNA editing for the description of multiple, locally adjacent editing events

in long, perfectly-paired dsRNAs [94]. But even seemingly non-selective, promiscuous editing events were shown to be site specific between different individuals [94]. Therefore, multiple locally adjacent editing events are not necessarily non-selective. There might be no mechanistic difference between the two mentioned editing modes, as non-selective editing could be just the presence of multiple site-specific editing events in direct proximity, but this remains to be verified.

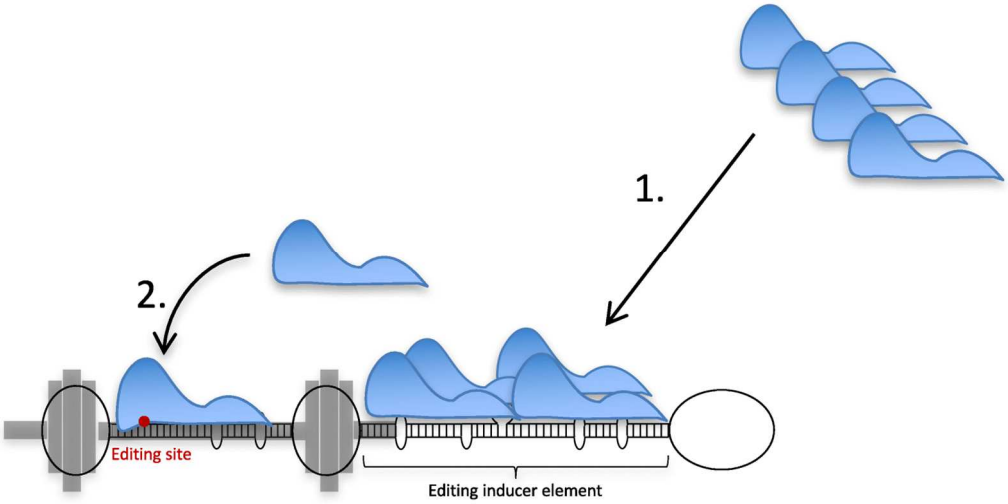


Figure 1-5: A model for efficient site selective A-to-I editing using an editing inducer element (EIE).

The process of efficient editing occurs as two consecutive events: 1) ADAR (in blue) recognizes a longer intronic stem by a non-specific interaction; 2) when the ADAR enzymes have been recruited, the catalytic domain of the protein interacts with a specific site, ideal for catalysis, situated in a shorter stem limited by a barbell-like structure (in grey). The site of selective editing is indicated in red. This figure and caption were adopted from reference [3].

However, editing events that are at present described as non-selective are predominantly found in non-coding RNA, especially enriched in Alu SINEs and other repetitive retrotransposon elements. Alu SINEs are ~300 bp long and can form the above indicated perfect dsRNA structures due to the presence of inverted repeats. While the number of edited individual adenosines within Alu repeats is high, the editing yield at each site, calculated over all Alu transcripts, is in most cases very low (below 1% conversion from A-to-I at each individual site) and is also decaying exponentially with the distance to the next reversely oriented neighbour Alu SINE [95].

Overall ADAR1 is the primary editor of repetitive sites like Alu SINEs and ADAR2 is the primary editor of nonrepetitive coding sites [96]. The function of ~80 mammalian genes might be regulated by side-selective (>20% editing yield) recoding events (a list of selected mRNA recoding events and their functional relevance can be found in review [14]). These, however, are a small minority relative to the over 100 million non-selective sites found in Alu repeats (~1% editing yield) [97].

This section clarifies that there is an association between site-selective editing, ADAR2, and imperfect dsRNA substrates on the one hand and non-selective editing, ADAR1, and perfect dsRNA substrates on the other.

1.1.2.2 ADAR domains and their individual contributions to selectivity and preference

The specificity of ADAR enzymes is defined by two factors: selectivity and preference [98]. In conjunction, they dictate which adenosines are selected for editing and to which extent these are converted to inosine.

1.1.2.2.1 Selectivity

The term selectivity describes the number of adenosines selected for editing within a given dsRNA and depends on substrate length, substrate structure (section 1.1.2.1), and interactions of the substrate with dsRBDs and deaminase specific 5' RNA binding loops of ADARs.

Double-stranded RNA binding domains of ADARs guide their catalytic deaminase domains towards dsRNA-substrates. They show a conserved αββββα-fold topology and bind to dsRNA sequences of 11-16 bp in length [99, 100], spanning from one minor groove to the next adjacent one in an A-form RNA helix. Two competing models try to explain how dsRBDs can affect adenosine selectivity, the RNA shape recognition model and the RNA sequence recognition model. The former proposes dsRNA recognition by shape [101, 102], and the latter extends the shape recognition by several sequence specific interactions [103, 104].

Because dsRBDs detect internal bulges and loops in dsRNA as sequence ends, the secondary structure of the dsRNA itself can regulate the binding modes available for dsRBDs, and does therefore affect selectivity [94, 105]. This mechanism might be fine-tuned by additional sequence recognition through readout of the minor grooves within A-form helix dsRNAs by dsRBDs [104].

While it was already discovered earlier that the exchange of the deaminase domains between ADAR1 and ADAR2 also switched the selectivity of the resulting chimeric protein [127], it was just recently revealed that this selectivity is mediated by deaminase specific 5' binding loops (aa457-479 in ADAR2, aa972-999 in ADAR1). These loops recognize the dsRNA structure 5' of the editing site through interactions with the phosphodiester backbone across the major groove [106]. It is assumed that the difference in selectivity between the two ADAR deaminase domains depends on a diverging bias for short (ADAR1) or long (ADAR2) stem structures 5' of the target adenosine. This bias is supposed to depend on the primary formation of contacts to both strands (ADAR2), or mostly the non-edited (ADAR1) strand, of the substrate [106]. Nevertheless, it is important to highlight that the structural details of the interaction between 5' binding loop and dsRNA substrate are only certain for ADAR2 [1] (see Figure 1-6).

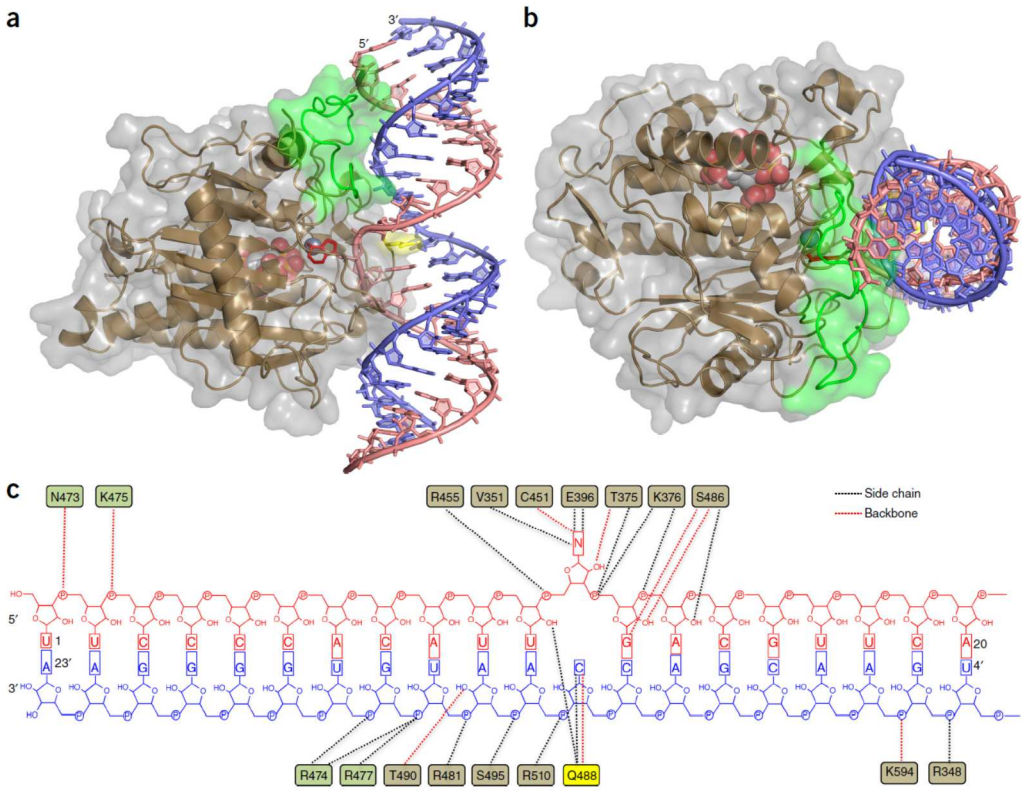


Figure 1-6: Structure of hADAR2d E488Q bound to the Bdf2-C RNA duplex at 2.75-Å resolution.

(a) View of structure perpendicular to the dsRNA helical axis. Colors correspond to those in Figure 1a and c. Red, flipped-out base N; gray space-filling sphere, zinc; yellow, Q488; green, previously disordered aa 454–477 loop; space filling, IHP. A transparent surface is shown for the hADAR2d protein. (b) View of structure along the dsRNA helical axis. (c) Summary of the contacts between hADAR2d E488Q and the Bdf2-C RNA duplex. Figure and caption were adopted from reference [1].

However, binding of an dsRNA by an ADAR enzyme does not necessarily result in RNA editing. Binding and editing are independent events, with independent functional relevance. This is demonstrated for instance by Staufen1-mediated decay, which is mediated by dsRNA binding (section 1.1.1.1.1) and by self-/non-self sensing, which is mediated by RNA editing (section 1.1.1.1.4) [94, 107, 108]. The factors that determine if editing occurs, and to which extent, are discussed next.

1.1.2.2.2 Preference

The term preference describes the extent of A-to-I conversion at a particular site [98]. Human ADAR1 and ADAR2 have a nearest neighbour nucleotide preference (also called triplet preference), ADAR1 even a 5’ second nearest neighbour preference [93, 106, 109]. The extent of editing for both enzymes is increased, if the 5’ neighbour next to the adenosine in question is U > A > C > G. In case of ADAR1 the second nearest 5’ neighbour is preferably a U [106]. Regarding their 3’ neighbour ADAR1 and ADAR2 have different preferences. ADAR1 prefers a 3’ G > C ~A > U, while ADAR2 prefers G > C > U ~A [110]. In addition, ADARs also have a counter base preference. Adenosines matched with a U or

mismatched with a C show high rates of A-to-I conversion, while A-A and A-G mismatches nearly abolish editing [111].

The ADAR2 crystal structure allows to explain the mechanistic background of the mentioned preferences. The deaminase domain uses a positively charged surface near the zinc-containing active site to contact both strands of the duplex on one side of the dsRNA covering a stretch of 20 bp (see Figure 1-6). To accomplish deamination, it needs to flip the target adenosine out of the duplex and into its catalytic centre. To stabilize the resulting complementary-strand orphan base, the residue 488, a part of the ADAR2 base-flipping loop, penetrates the helix from the minor-groove side, where it occupies the space vacated by the flipped-out base. There, it can then form hydrogen-bonds to the orphaned base and to the 2'-hydroxyl of the ribose immediately 5' to the editing site. This flip-out mechanism explains the counter base preference, because purines at the orphan base position would sterically clash with the penetrating 488 residue, while pyrimidines do not.

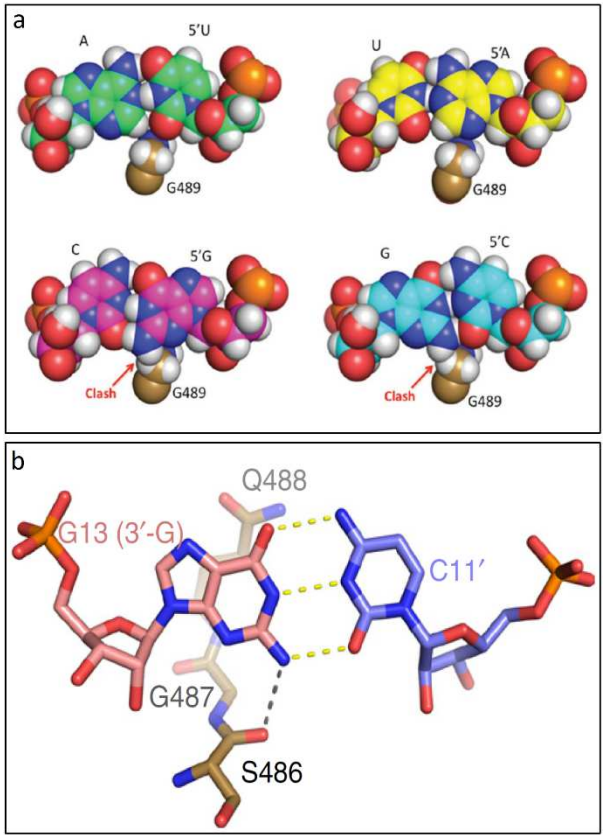


Figure 1-7: Structural reasons for nearest neighbour preferences.

(A) Guanosine 2-amino group clash with Gycine 489. (B) Guanosine 2-amino group hydrogen bound with Serine 486. This figure and caption were adopted from reference [1].

The 5' nearest neighbour preference for U (and A) can be explained in a similar manner, since the presence of a G or C at this position induces a destabilizing clash of an amino group in the minor groove with the protein backbone at Glycine 489 (see Figure 1-7A). However, this steric issue can be compensated by slight structural perturbations of the deaminase domain, resulting in moderately to dramatically lowered editing yields. Concerning the 3' nearest neighbour preference, the structure indicates quite the contrary. In this situation, the 2-amino group of the preferred 3' G functions as a hydrogen-bond donor for the backbone carbonyl oxygen at Serine 486 (see Figure 1-7B). Since G is the only common nucleobase that can do this within the RNA minor groove, the exchange of G for A, C or U reduces the rate of A-to-I editing and thereby explains the 3' nearest neighbour preference of ADAR2 [1].

1.1.2.3 The effect of other RNA-modifications on A-to-I RNA editing

A-to-I RNA editing takes place in an environment of non-canonical bases, surrounded by more than 160 other types of RNA modifications [112]. Despite the large number of RNA modifications that have been identified by now, N6-methyladenosine (m6A) is the only one known, so far, that does negatively correlate with inosine and therefore A-to-I RNA editing [113]. This effect is probably caused by a steric perturbation of the catalytic centre of ADAR enzymes by the additional methyl group. As m6A is only present in 0.2-0.6% of all adenosines, and, in most of these cases (~70%), within hard-to-edit GAC codons, the general impact of this negative correlation on natural A-to-I RNA editing is not very large, but could be relevant to fine-tune cellular processes [114-116]. Competition with known, or currently unknown, RNA-modifications, however, might explain rare and unexpected hard-to-edit sites in the context of site-directed RNA editing approaches, which will be discussed next.

1.2 Site-directed RNA editing

By redirecting adenosine deaminases to specific adenosines in selected mRNA targets, it is possible to harness the capabilities of natural RNA editing (Figure 1-8) and turn it into a powerful instrument for transcript manipulation, called “site-directed RNA editing” (SDRE).

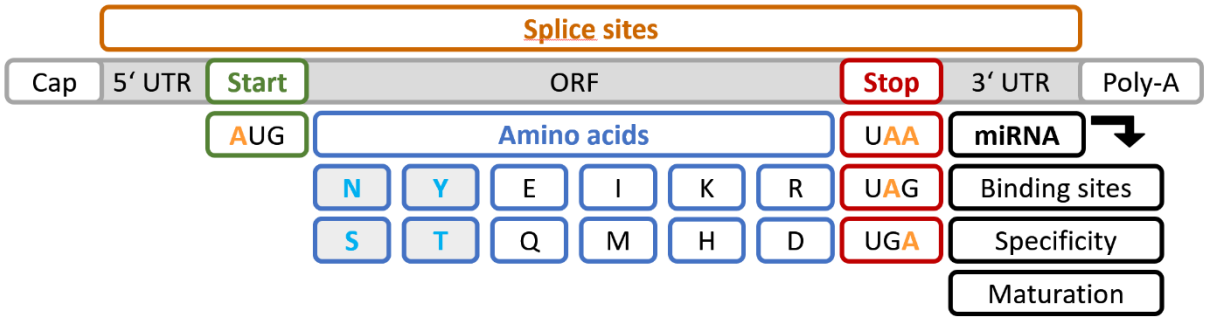


Figure 1-8: Conceivable direct and indirect manipulation of mRNAs by site-directed A-to-I RNA editing.

Splice sites, Start- and Stop-codons, 12 out of the 20 canonical amino acids and miRNA binding sites, mostly localized within the 3’ UTRs, could be altered by A-to-I RNA editing when directly targeting an mRNA. The RNA editing of Serine- (S), Threonine- (T), Tyrosine- (Y) and Asparagine (N)-residues, which are targets for phosphorylation or glycosylation, could manipulate signaling cascades, as well as structure and function of encoded proteins. It might also be possible to indirectly manipulate the regulation of a mRNA by RNA editing of corresponding pri-,pre- or even mature miRNAs.

As already mentioned earlier, RNA editing can affect important cellular processes, e.g. splicing and RNAi [14, 15]. But one of the most interesting capabilities of RNA editing, from a therapeutic point of view, is the ability to recode transcripts. The simple A-to-G base substitution can manipulate both start [117] and stop codons [118] as well as recode 12 of the 20 canonical amino acids [119, 120]. This allows to correct G-to-A transitions at the transcript level, which can cause splice-site, missense, and nonsense mutations. When including all types of point mutations, these make up the largest fraction (~63%) of all disease causing mutations in humans [121]. In addition, some of the 12 recodable canonical amino acids contain residues, which are targets for phosphorylation (serine, threonine, and tyrosine) or glycosylation (Asparagine). This allows the potential use of site-directed RNA editing to manipulate signaling cascades [120, 122-128], as well as influence protein structure and function [118, 129, 130].

The idea of site-directed RNA editing for the correction of mutated RNAs was first introduced by Tod Woolf and co-workers in 1995 [131]. They *in-vitro* hybridized a 52 nt long, single-stranded guideRNA, which was called “correcting RNA oligomer” at the time, to a target mRNA and injected the mixture into Xenopus eggs. A-to-I RNA editing within the created duplex, performed by endogenous Xenopus ADAR, resulted in the conversion of an amber stop codon (T**A**G) into a tryptophan codon (T**G**G), thereby restoring the full length reporter and its activity. However, issues like the relatively low editing yield and the strong off-target editing within the gRNA-mRNA duplex could not be solved back then.

Nevertheless, those pioneer experiments proved that RNA editing had the potential to be utilized as therapeutic approach for transcript repair in the future.

Nearly two decades after this proof-of-principle study, site-directed RNA editing was rediscovered and refined by the Stafforst group and others [132-136]. Several novel site-directed RNA editing systems utilizing engineered or wild-type deaminases, starting with the SNAP-ADAR system introduced in 2012 [133], were established since then and generated more and more interest in this field.

The goal of every SDRE approach is the efficient conversion of one certain target adenosine into inosine, while preventing the unspecific deamination of other non-target adenosines, called off-target editing. The general principle is that some form of antisense-oligonucleotide (ASO) guides a deaminase towards a specific adenosine within a selected target mRNA by Watson-Crick base pairing, thereby simultaneously forming the dsRNA substrate required by ADAR deaminase domains (DD). Due to their function, those ASOs are often called guideRNAs. Because ADAR DD prefer adenosine mismatched with cytosine for most efficient conversion (section 1.1.2.2.2), the target adenosine within the gRNA-mRNA duplex is normally arranged with the respective cytosine counter base [111]. Based on the same principle, adenosine-guanosine mismatches make the adenosine in question nearly editing-inert and can be utilized to prevent off-target editing [132]. Due to the triplet preference of wild-type ADAR DD already explained in section 1.1.2.2.2, the more esily edited 5'-UAG codons are often used as target when establishing SDRE approaches. Additionally, the codon scope of some SDRE systems was investigated partially (5'-UAN codons) or entirely (all editable codons) [120, 137, 138].

A SDRE system works efficiently when the A-to-I conversion of the target adenosine took place in most of the target transcripts. This can be quantified by comparison of adenosine-guanosine double peaks in Sanger sequencing reads of reverse-transcribed, edited mRNAs or by NGS. High specificity of a SDRE system is given when no off-target editing, or only few sites with low conversion yields, occur. The location of potential off-target editing is not limited to the duplex containing the target adenosine, but can also occur at other sites spread over the whole transcriptome. Which sites might be affected depends on the targeting strategy of the used SDRE approach. The duplex containing the target adenosine, as well as similar sequences that allow partial binding of the used gRNA, are always good candidates. Engineered deaminases, although their selectivity is dictated by the used targeting approach, can inherit the triplet preference of the original wild-type enzyme and may cause gRNA independent off-target editing at sites optimal for deaminase domain binding, if overexpressed excessively [120, 138].

1.2.1 Site-directed RNA editing systems that are using engineered deaminases

The natural dsRBD-dependent targeting mechanism of wild-type ADAR enzymes was initially believed to essentially require dsRNA-structures formed in cis, not trans [132], which were supposed to be difficult to create [138]. This, and the fact that engineering of deaminase domains allowed to extend their triplet preference, resulted in several SDRE approaches using fusion proteins consisting of mutated DD attached to alternative targeting systems, different from dsRBDs. These are explained in this section.

1.2.1.1 The SNAP-ADAR site-directed RNA editing system

The SNAP-ADAR site-directed RNA editing system was established in the Stafforst lab and first published in 2012 [133]. It is currently the best-characterized, site-directed RNA editing system [139]. The principle is based on covalently-bound RNA deaminase conjugates, which are guided towards a specific adenosine within a selected target mRNA by Watson-Crick base pairing. In this process, the RNA portion of the conjugate simultaneously guides it towards its target site and forms the necessary dsRNA structure as substrate for the deaminase (Figure 1-9). The RNA deaminase conjugate is assembled from the fusion protein SNAP-ADAR and a chemically modified guideRNA in a 1:1 stoichiometry.

The fusion protein part of the conjugate consists of a human wildtype or mutant ADAR deaminase domain and a SNAP-tag domain. The SNAP-tag is based on an engineered human O6-alkylguanine-DNA alkyltransferase (hAGT), which facilitates the irreversible transfer of the alkyl group from its substrate to one of its cysteine residues. The short 22 nt BG-gRNA contains 2'-O-Me and phosphorothioate backbone modifications similar to antagomirs [140], as well as a 5’ terminal O6-benzylguanine (BG), which can readily react as hAGT substrate [141] and facilitate the covalent bond of the guideRNA to the SNAP-ADAR fusion protein [142]. Due to the necessary BG-moiety, the SNAP-ADAR System cannot be genetically encoded. This represents a major disadvantage of the system, but comes with several advantages connected to the benefits of further chemical modifications, which can in turn not be realized in an encodable system. These include the potential to optimize BG-gRNA properties like potency, specificity, efficiency, stability, and most likely also immunogenicity [139].

By using photocaged BG-gRNAs the system did also allow the light driven assembly of the SNAP-ADAR conjugates. These could be used in cell culture to change the localisation of reporter proteins by inclusion of localisation signals [117] and even to facilitate SDRE in living organisms [143], proving the huge potential of the system to address biological questions in basic research.

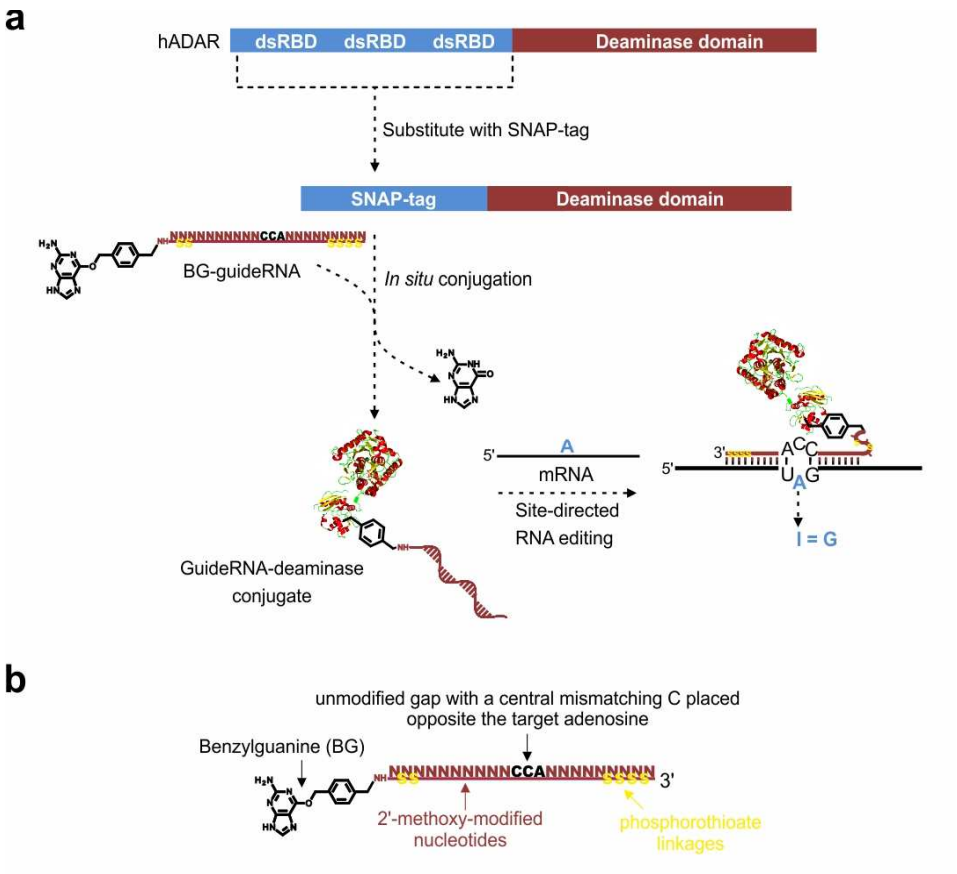


Figure 1-9: The SNAP-ADAR site-directed RNA editing system.

(A) General concept: The double-stranded, RNA-binding domains (dsRBDs) of hADAR have been substituted with the SNAP-tag. The latter is able to form a covalent bond to a guideRNA that is modified with benzylguanine (BG). When bound to the SNAP-ADAR, the guideRNA steers the attached SNAP-ADAR protein to the target RNA and forms the necessary secondary structure for A-to-I editing catalyzed by the deaminase domain. (B) A typical BG-guideRNA that targets a UAG site with a 5´-CCA anticodon. The guideRNA is 22-nt long and is densely chemically stabilized by 2'-methoxylation and terminal phosphorothioate linkages (commercially available). The first three 5'-terminal nucleotides do not base-pair with the target RNA, but serve as a linker. The sequence comprises an unmodified ribonucleotide gap (5´-CCA), which faces the target site and contains a central mismatching cytosine opposite the targeted adenosine for efficient deamination. A commercial C6-amino-linker is located at the 5'-end of the guideRNA to introduce the BG modification to the full length oligonucleotide. Modification of the guideRNA with OSu-activated BG can be performed in any reaction tube. Figure and caption were adopted from reference [120].

SNAP-ADAR variants consisting of different human ADAR1 and ADAR2 deaminase domains, including hyperactive EQ mutants, were characterized through targeting of endogenous mRNAs. The mutant DD allowed an extended triplet scope and achieved editing yields up to 90% in cell culture [120].

Due to the use of chemically modified BG-gRNAs, off-target editing within the gRNA-mRNA duplex was almost completely blocked, independent of the used DD. Global off-target editing was nearly absent when using wild-type DD and could be reduced for DD EQ mutants by using only shortly doxycycline-induced, SNAP-ADAR-expressing Flp-In T-REx cell lines. At the moment, the SNAP-ADAR SDRE system shows in every regard (gRNA-mRNA duplex, transcriptome wide) the lowest off-target editing of the available SDRE systems using engineered deaminases [139].

1.2.1.2 The λN-BoxB and the MCP-MS2 site-directed RNA editing systems

The λN-BoxB site-directed RNA editing system was established in the Rosenthal lab and first published in 2013 [132]. The approach uses a fusion protein consisting of the 22 amino acid long λN-peptide and a human ADAR2 deaminase domain (DD). The λN-peptide derives from the λ-phage N-protein, were it normally facilitates the nanomolar-affinity binding to boxB RNA hairpin motifs, thereby causing the anti-termination of λ-phage mRNA transcription [144]. The guideRNAs containing one or more 17 nt long boxB motifs and a ~50 nt antisense part were used to steer the fusion protein to its target mRNA and to create the dsRNA substrate necessary for the deaminase domain activity (Figure 1-10).

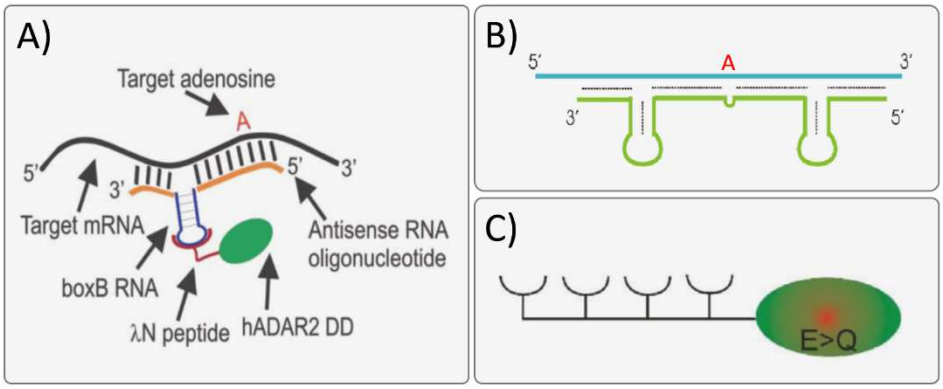


Figure 1-10: The λN-BoxB site-directed RNA editing system.

(A) Schematic showing the basic component of the λN-BoxB site-directed RNA editing system. (B) Improved 2xboxB gRNA. (C) Improved 4xλN-DD fusion protein with EQ mutation. This figure was assembled from images adopted from reference [145] and [138].

The system allowed the correction of the CFTR W496X amber mutation both *in-vitro* and in *Xenopus* oocytes after, their injection with guideRNA and a mRNA encoding the λN-DD fusion protein. In addition, the eGFP W58X amber mutation could be corrected up to 20% after transfection of a plasmid encoded version of the system in HEK-293T cells [132]. By using a fusion protein containing four λN-peptides and an ADAR2 E488Q deaminase domain, together with a gRNA containing two boxB motifs (~85 nt in total length), editing yields up to ~70% could be achieved for triplets preferred by the ADAR2 DD, but resulted also in strong off-target editing [138].

Even though off-target editing could be largely abolished by reduction of the transfected gRNA amount, this diminished the editing yields to 25% [138, 139, 144]. A localisation change of the fusion proteins into the nucleus by incorporation of an NLS lowered off-target editing by 50% in general, while keeping the on-target editing unchanged. Anyway, even the NLS-4xλN-DD, which resulted in the lowest off-target editing, caused 8800 events of >10% conversion, as detected by NGS [145]. The system can perform very efficient on-target editing, but its specificity remains a major issue.

Although the λN-BoxB system was so far mainly characterized by targeting overexpressed reporter constructs, experiments published by the Mandel lab in 2017 achieved editing of an endogenous Mecp2 mRNA in cultured mouse neurons. In their study, SDRE restored 72% of the mutant mRNAs, which cause Rett syndrome, resulting in functional protein repair [139, 146].

A SDRE system similar to λN-BoxB was established in the Tsukahara lab and first published in 2017. It used MS2 coat proteins (MCP) instead of the λN-peptide, and MS2-loops instead of boxB motifs, but achieved only ~5% A-to-I conversion when editing the same eGFP W58X amber reporter [134].

1.2.1.3 The CRISPR-Cas13 site-directed RNA editing system

The CRISPR-Cas13 site-directed RNA editing system was established in the Zhang lab and first published in 2017 [137]. It is again similar to the λN-BoxB and the MS2 site-directed RNA editing approaches, but, in contrast to those, uses a Cas13-DD fusion protein (Figure 1-11). The hyperactive deaminase domain mutants ADAR1 E1008Q and ADAR2 E488Q were used for the fusion protein again, as they result in higher editing yields. The used Cas13 variant PspCas13b is lacking nuclease activity due to directed mutations of conserved residues in HEPN domains.

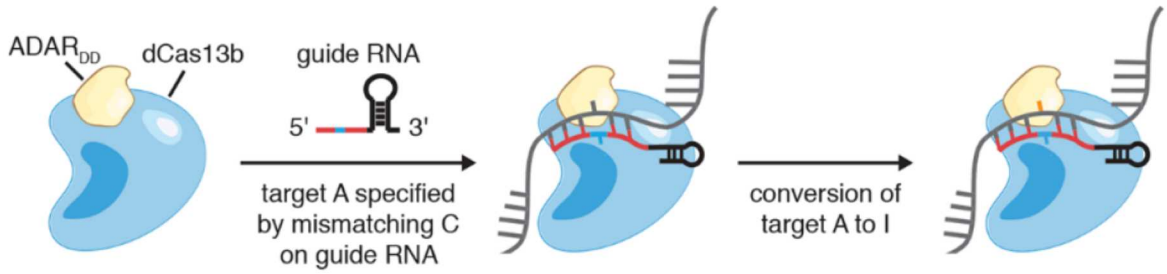


Figure 1-11: The CRISPR-Cas13 site-directed RNA editing system

Catalytically dead Cas13b (dCas13b) is fused to the deaminase domain of human ADAR (ADAR$_{DD}$), which naturally deaminates adenosines to inosines in dsRNA. The gRNA specifies the target site by hybridizing to the bases surrounding the target adenosine, creating a dsRNA structure for editing, and recruiting the dCas13b-ADAR$_{DD}$ fusion protein. A mismatched cytidine in the gRNA opposite the target adenosine enhances the editing reaction, promoting target adenosine deamination to inosine, a base that functionally mimics guanosine in many cellular reactions. Figure and caption were adopted from reference [136].

It has no sequence constraints by protospacer adjacent motif (PAM) or protospacer flanking site (PFS) like other CRISPR-Cas9 or -Cas13 systems. In regard to targeting limitations, this brings the system on-pair with the other discussed SDRE approaches that utilize deaminase domain EQ mutants within their fusion proteins. The used guideRNA is expressed from a U6 promoter and consists of a 35 nt 3'-terminal hairpin, called direct repeat (DR) sequence, which recruits the Cas13 and a 50 nt antisense part that itself targets the chosen mRNA.

The system was tested by targeting 200-bp long excerpts of 34 disease relevant target genes, each cloned into a reporter construct, and achieved editing yields distributed evenly between <5% and 35%

in this setup. At a *Cluc*-reporter target, a maximum of ~90% editing could be achieved. Also 5'-UAG sites of endogenous KRAS and PPIB mRNAs were edited with efficiencies from 15% to 45%.

In contrast to the λN-BoxB system, off-target edits were in case of the CRISPR-Cas13 system mainly caused by overexpression of the fusion protein, not the gRNA. Thus focusing on the fusion protein, a screen of several residues involved in the DD-RNA contact surface formation resulted in the ADAR2 DD E488Q/T375G double-mutant. The T375 residue normally contacts the ribose-phosphate backbone of the target adenosine (Figure 1-6C). This new mutant allowed to lower the number of off-target edits from more than 18.000 to only 20 events. This was accompanied in many cases by a reduction of on target editing, e.g. at the *Cluc*-reporter target from ~90% to ~45%, but also at other targets (Figure 6F and Figure S19 in [136]).

1.2.2 The R/G gRNA site-directed RNA editing system

The R/G gRNA site-directed RNA editing system was established in the Stafforst lab and first published in 2016 [118]. It was the first SDRE system able to harness overexpressed human wild-type ADAR2 and represents the foundation for this thesis.

The system was based on a naturally-occurring and well characterized ADAR2 dsRNA substrate, the R/G-site of the human GRIA2-transcript, which is encoding a subunit of the glutamate receptor 2 (GluR2). In the natural GRIA2-transcript the R/G dsRNA motif that is formed in cis with an intronic ECS, facilitates the binding of ADAR2 via its dsRBDs. This results in an amino acid change from Arginine (R) to Glycine (G) though an **A**GA to **G**GA recoding event at a specific exonic site. Accordingly, the process was named R/G-editing. The natural R/G editing of the GluR2 ion channel results in faster recovery rates from desensitization [19]. The R/G motif was selected as starting point for the development of the R/G gRNA SDRE system due to the high conversion rate of up to 90% reported at the natural R/G-site [92], its very small size (only 45 nt), the distinct secondary structure with a terminal pentaloop believed to be relevant for ADAR2 recognition [147], and its already characterized and well-defined interaction with ADAR2 dsRBDs [103]. The Figure 1-12 shows how the trans-acting R/G gRNA was engineered based on the natural cis-acting motif.

A R/G gRNA binds its target mRNA by Watson-Crick base pairing via its single-stranded antisense part. The antisense-part-mRNA duplex represents the substrate for the ADAR2 deaminase domain, while the R/G motif represents the dsRNA structure necessary for the interaction with the ADAR2 dsRBDs.

How the system was first optimized *in-vitro,* and then successfully transferred to cell culture and early *in-vivo* experiments, using the annelid *Platynereis dumerilii*, is summarized in Dr. Wettengel's dissertation [148].

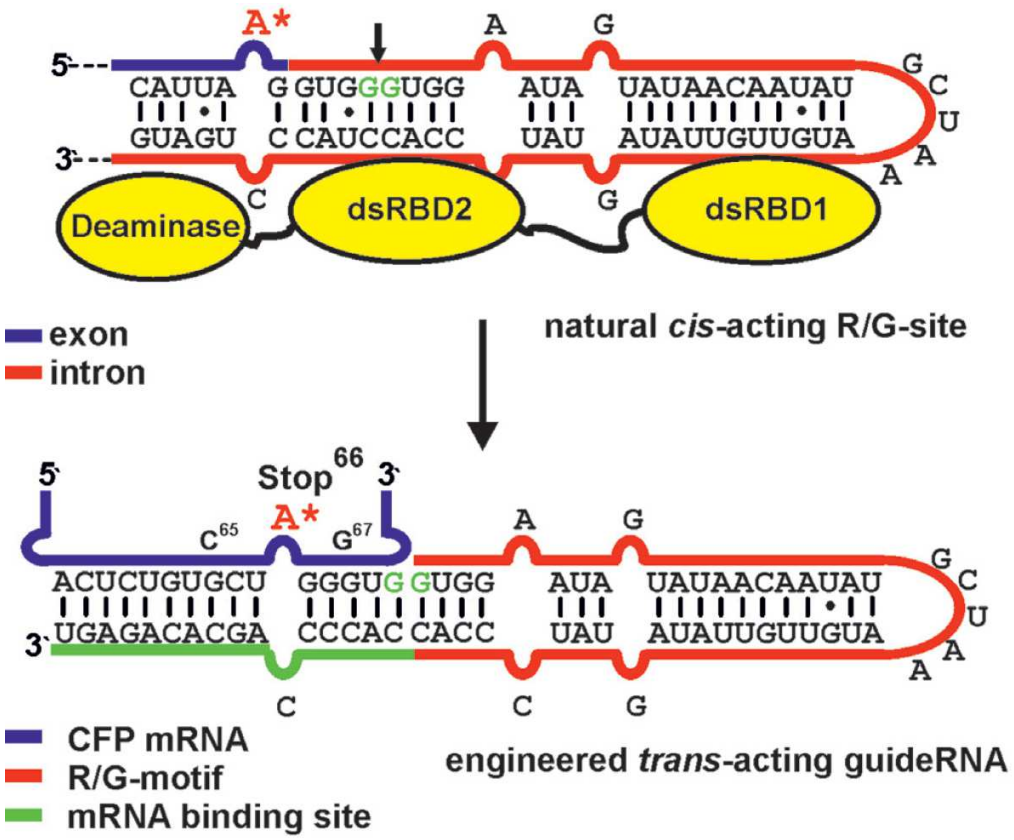


Figure 1-12: Engineering of the trans-acting R/G gRNA from the cis-acting R/G-site within the GRIA2-transcript for the recruitment of wild-type human ADAR2 for site-directed RNA editing.

The top image shows the natural R/G-site bound by the two dsRBD and the deaminase domain of ADAR2. The blue and red lines indicate intronic and exonic parts of the GRIA2-transcript. The arrow indicates the site between two guanosines, indicated in green, were the R/G motif was cut *in-silico*. The bottom image shows how the resulting hairpin (Red) was then fused to a 16 nt long single-stranded antisense part (Green), reverse complementary to a selected target mRNA (indicated in blue, and in this example being a mutant W66X amber CFP mRNA), and containing a cytosine counterbase at the intended editing site. The editing site of the trans-acting gRNA within the antisense-part-mRNA duplex was initially positioned 6 nt away from the just mentioned *in-silico* cutting site, to resemble the distance of the natural R/G motif in cis. The resulting R/G gRNA allows to steer wild-type human ADAR2 to the CFP reporter mRNA to repair the stop codon at position 66 (U**A**G) to tryptophan. The figure was adopted from reference [118].

When this project began, the state of the art allowed successful recruitment of human wild-type ADAR2, overexpressed in HEK-293T cells, to specifically edit G-to-A mutations like the eGFP W58X amber nonsense mutation. The correction of disease-promoting point-mutations, like the p53 mutations R175H and R282Q, had also been shown at RNA level. The cell culture transfection settings for HEK-293T had already been optimized and allowed high-yield editing at the eGFP W58X amber target site with up to 50% correction. Off-target editing concerning the target mRNA had been avoided by reduction of the used ADAR2 plasmid amount and by restricting the length of the R/G gRNA antisense part to a length of 16 nt. Dr. Wettengel had also shown that simultaneous editing at two different target sites in the same cell culture experiment was possible and did not negatively affect the editing yield [148].

Work performed during this thesis that contributed to the initially mentioned 2016 publication, and to the further development of the R/G gRNA SDRE system thereafter, are discussed in section 3.1. A less extensive study with similar claims regarding the recruitment of overexpressed wild-type human ADAR2 was published in 2017 by the Nakagawa lab [135]. A very recent study utilizing AAV8-encoded guide RNAs based on our early R/G designs achieved ~2% *in-vivo* correction of a dystrophin mutant in mdx-mice and ~20% *in-vivo* correction of an OTC deficiency in spf^{ash} mice by using overexpressed hyperactive ADAR2 E488Q. However, ADAR2 wt overexpression achieved only <1% and ~3% correction [149].

1.2.3 Pros and contras resulting from functional components of the described SDRE systems in the context of their intended therapeutic application

In principle, all above SDRE systems can be dissected into functional components, granting the higher-level system unique properties but also limitations, which might become relevant for their potential utilization as future therapeutic applications. These functional components are the deaminase domain, which facilities the deamination reaction, the targeting domain, which facilitates the interaction between the guideRNA and the deaminase domain, and the guideRNA itself, which steers the whole SDRE complex towards its target mRNA by forming Watson-Crick interactions.

GuideRNAs can be applied in a chemically-modified or a genetically-encoded form. Both harbour their own pros and contras. Genetically encodable gRNAs can be expressed from polymerase III promoters, like the U6-promoter, and can be encoded on plasmids, mini-circles, or viruses. Through the use of, e.g., AAVs it might therefore be possible to facilitate their long term expression within transduced cells and tissues [150]. Chemically-modified gRNAs on the other hand are only transiently available, but come with a broad spectrum of potential advantages, which depend on the actually used modification(s) and can include, as mentioned earlier, improved potency, specificity, efficiency, stability, and potentially also lowered immunogenicity [139]. While many systems (λN-BoxB, MCP-MS2, CRISPR-Cas13 and R/G gRNA) are in theory able to utilize chemically-modified as well as encoded gRNAs, the use of chemically-modified gRNAs would deprive them of one of their biggest advantages, the ability to be completely encodable, and therefore not transient in nature. Some other techniques are limited to modified gRNAs due to specific prerequisites of the used targeting domain, or the used deaminase domain. In case of the SNAP-ADAR system this would be the BG-moiety, which is essential for the interaction with the SNAP targeting domain [142]. Examples for deaminase domain prerequisites would be the ADAR2 E488F or E488Y deaminase domain mutants, which can only edit efficiently if the gRNA contains an abasic site at the counterbase position, which is called the “bump-hole approach” [151].

The targeting domain as link between the deaminase domain and the gRNA differs between all mentioned SDRE systems and has therefore been eponymous for most of the systems. The only exception is the R/G gRNA approach, which is named after its gRNA. However, the basic principle is the same for all SDRE systems. The important factor to consider in this context is the origin of the proteins/peptides utilized as targeting domain. In case of the CRISPR-Cas13, the λN-BoxB, and the MCP-MS2 SDRE systems, these proteins/peptides are foreign to human cells. The CRISPR-Cas13 originates from bacteria (most commonly *Staphylococcus aureus/pyogenes*) [152], while the λN-peptide, as well as the MS2 coat protein (MCP), originate from bacteriophages [132, 153], which creates the problem of potential immunoreactions against the targeting domains of these engineered editases [152]. The R/G gRNA and the SNAP-ADAR systems, however, were developed based on proteins and RNA-motifs of human origin. As already mentioned, the used SNAP-tag is based on the human O6-alkylguanine-DNA alkyltransferase (hAGT). In case of the R/G gRNA system the targeting domain of the protein was not exchanged in the first place. The natural dsRBDs of human ADAR2 were used to bind to the R/G gRNA. This makes the SNAP-ADAR and the R/G gRNA systems a safer choice in regard to unintended immunereactions.

The last essential functional component to address is the deaminase domain. Systems using engineered editases allow for rational design or directed evolution of the used deaminase domains to improve specific desired features, or to diminish undesired ones. Popular examples are the EQ mutants of human ADAR1 or ADAR2 deaminase domains, which lower their specificity, but simultaneously improve the editing yield drastically [98]. Even the codon scope, which is defined by the nearest neighbour preferences of the used deaminase (section 1.1.2.2.2), can be extended to some degree by engineering of the DD [120, 136]. After the crystal structure of the ADAR2 deaminase domain, bound to a dsRNA substrate, was published in 2016, it can be expected that several new mutants will emerge that will allow further improvement to SDRE systems using engineered deaminases or mutant ADARs. A very good recent example for this is the earlier mentioned “bump-hole approach”. This new approach uses ADAR2 E488F or E488Y deaminase domain mutants, which are unable to edit natural dsRNA substrates. Due to sterical hindrance of the natural interaction between the counterbase and the amino acid 488 by the introduced bulky aromatic residues at this position, these mutants are only able to efficiently edit dsRNA duplexes containing an abasic site instead of a counterbase. Thus the use of chemically-modified guideRNAs containing a defined abasic site should facilitate more specific editing of target mRNAs and reduce unspecific off-target events often caused by overexpression of the engineered deaminases [151]. As these deaminase mutants can be used independently from the utilized targeting strategy, all SDRE systems can profit from this novel idea, as long as they switch to overexpressed mutant deaminases and chemically-modified gRNAs, thereby introducing already mentioned

limitations. The only system that can benefit from this without any repercussions is obviously the SNAP-ADAR system, as it is already restricted to chemically-modified gRNAs containing a BG-moiety.

Currently all SDRE strategies have in common that they require the overexpression of an editase (an engineered one or wild-type ADAR2). In general, the overexpression of proteins is undesirable as it can be toxic to cells and lead to adverse effects. Examples are the exhaustion of cellular resources (protein burden), protein aggregation, unintended modulation of biological pathways, and the disruption of natural regulatory processes [154, 155]. If, and to which extent, artificial proteins could cause such problems, needs to be evaluated for each one individually, as the biological role of a protein cannot be used to predict how it might harm the cell [154]. In the particular context of SDRE, overexpressed editases cause off-target editing, in addition to the already mentioned problems [139]. Also the overexpression of wild-type ADARs, like ADAR2, which have a broad spectrum of natural functions (section 1.1.1), might have strong negative effects on the editing homeostasis of human cells. But while all the SDRE systems that are using artificial enzymes are fundamentally restricted to the overexpression of their individual editase-construct, the R/G gRNA system has the prospect to potentially use the available endogenous wild-type ADARs in the future.

Finally, it needs to be considered, how the components of the SDRE systems can be administered into living organisms. Because the therapeutic delivery of both DNA and RNA into living organisms are extensive, and rapidly improving, fields of research by themselves, the following is not intended to comprise all available options within these fields, but to give a brief impression what delivery options could potentially be used for the SDRE systems. The λN-BoxB, the MCP-MS2, the CRISPR-Cas13 and the R/G gRNA systems are fully genetically encodable, and could be administered, e.g., in the form of AAVs, which are the current state of the art in clinical gene-therapy approaches [156]. This gives them a broad spectrum of tissues as potential targets, just limited by the tropism of available AAV serotypes, including CNS, heart, liver, lung, eye, and skeletal muscle [157]. While the SNAP-ADAR editase can also be encoded in an AAV, the corresponding BG-gRNA and every other type of chemically-modified gRNA has to be administered separately. A problem of the whole ASO field is the rapid excretion of naked small RNAs through renal clearance [158]. However, an increasing spectrum of novel nanoparticles [159], cell penetrating peptides (CPP) [160], or targeting-ligands like N-Acetylgalactosamine (GalNAc) [161] are becoming available and can facilitate the efficient uptake of chemically-modified RNAs, although still mostly limited to the liver, when applied intravenous or subcutaneous. Direct application into the cerebrospinal fluid (intrathecal injection), the eye (intravitreal injection or eye drops), or intratracheal nebulisation into the lung are already used in clinical studies, too, and allow chemically-modified RNAs to reach other organs than the liver, as well [162, 163].

While systems using chemically-modified gRNAs might be limited to become treatments due to their transient nature, this limitation could also be a benefit in regard to ethical issues and safety, as all changes performed by these systems would be entirely reversible. The fully encodable systems on the other hand harbour the potential to cure diseases permanently.

To summarize, the functional components of a SDRE system and how they can be administered are important factors to consider in the context of their intended therapeutic application.

1.2.4 Site-directed RNA editing in the context of genome engineering, RNAi, and mRNA transcript therapies

Beside the mentioned SDRE approaches, there are several other techniques in active development that are focused on the manipulation of the transcriptome or the genome for therapeutic purposes. Due to their very specific abilities and limitations, it is important to position SDRE in their context, to make clear how it complements and extends the current reach of these other strategies.

1.2.4.1 Genome engineering

Genome engineering methods have been in development for several decades now. Starting with meganucleases, then zinc-finger nucleases (ZFN), transcription-activator-like-effector nucleases (TALEN), and now CRISPR-associated protein-9 nucleases (CRISPR-Cas9), the targeting strategy has improved to become more modular, as well as faster and easier to engineer. Today the time-consuming assembly of zinc-fingers or TALE-modules is not necessary anymore. The Cas9 nuclease can be steered towards a DNA-target site by simple Watson-Crick base pairing via a short guide RNA, similar to how SDRE approaches are targeting mRNA. However, the final goal is still the induction of double strand breaks (DSB) at specific genomic sites to either knock-out a gene by causing indels (Heterogenic mixtures of insertions or deletions) via the error prone non-homologous end joining (NHEJ) repair pathway, or to correct a gene via homology directed repair (HDR) using a simultaneously applied DNA-repair template [164].

While CRISPR-Cas9 theoretically allows the correction of any type of mutation, it comes also with limitations. One is the protospacer adjacent motif (PAM), which derives from the origin of the CRISPR-Cas9 system as adaptive immune system of bacteria and archaea. The PAM is a short 2-5 bp long recognition motif that helps bacteria to distinguish their own DNA from the DNA of invading bacteriophages and must be present within a target DNA for Cas9 to induce a DSB [165]. While tremendous efforts are underway to engineer CRISPR-Cas systems with broader PAM compatibility, it is still not feasible to target any possible site in the human genome [166]. SDRE systems on the other hand are limited to the manipulation of transcripts and can currently only perform A-to-G transitions within their range of

triplet preferences. While there is great interest to extend this by C-to-U transitions through site-directed recruitment of deaminases from the APOBEC-family, or even non-enzymatic, site-directed C-to-U RNA editing, we are not there yet [167]. In the context of potential editing sites, the CRISPR-Cas9 system and the current SDRE approaches have obviously different limitations, which allows them to complement each other in this regard.

Concerning safety and efficiency, however, the SDRE approaches show some clear advantages. It is important to highlight that the following arguments are focused on the *in-vivo* application of CRISPR-Cas9, not on *ex-vivo* treatments, which allow more control over the genetic constitution of the re-administered cells. While NHEJ can knock-out target genes, the much less frequent HDR and the simultaneous presence of a DNA repair template are necessary to actually correct a mutation. This is a problem, as NHEJ is a much faster and more efficient process than HDR [168]. It becomes even more challenging if post-mitotic cells, like neurons, are targeted, as HDR happens nearly exclusively in late S and G2 stages of the cell cycle [169]. Due to the necessity of HDR, the simple presence of all CRISPR-Cas9 components within a cell is not sufficient to repair a mutant gene and could as well just cause additional damage at the target site through NHEJ. Therefore, most current *in-vivo* applications of CRISPR-Cas9 in basic research are focused on the disruption of their target genes, resulting in knock-out, but not repair [170, 171]. However, if HDR does indeed occur, it can restore 100% of the wild-type gene product within the cell in question. SDRE approaches on the other hand result directly in correction at RNA level in every cell positive for all SDRE system components. While SDRE approaches might never reach 100% correction, a smaller fraction of functional gene-product could be restored in a much larger fraction of cells this way and might result in a larger patient benefit, at least in the context of loss-of-function mutations. The option to control the time span as well as the extent of the manipulation could also enable interesting new therapeutic options as some interventions on, e.g., kinases, apoptosis-, transcription-, or translation-factors might be cytotoxic if enforced permanently [172]. This opens up the possibilities to modulate signaling cascades, including local inflammation [136]. In addition, there is no reason why SDRE should not work efficiently in post-mitotic cells.

Also concerning safety, the CRISPR-Cas9 genome editing approach has still some problems to solve. As already mentioned in the previous section, the Cas-proteins originated from bacteria. This brings along the problem of immunoreactions against an overexpressed foreign protein [173]. While this is also true for most of the mentioned SDRE approaches, at least the SNAP-ADAR and the R/G gRNA approach do not contain foreign proteins or peptides. Off-target editing of genome engineering approaches is also widely discussed as a safety hazard in the literature, especially in the context of *in-vivo* patient applications. The off-target induction of DSB can cause large scale genomic alterations and oncogenic mutations with potentially pathological effects on the treated patients [174, 175]. SDRE approaches are

promising in this regard, as they do not require any interaction with the genome. While off-target editing is also an issue that SDRE approaches have to deal with, they could never result in the complete knock-out of a gene and, depending on the used SDRE system, are transient and reversible.

A novel type of CRISPR-based genome editing using engineered DNA base editors is currently evaluated. These systems use CRISPR guided deaminases based on TadA- or APOBEC-enzymes to deaminate DNA-bases, allowing A-to-I and C-to-U DNA editing. While these systems do not require DSBs, the most efficient base-correction is performed when nicks are introduced near the editing site, resulting again in increased indel frequency. At the same time the CRISPR base editors are still limited by their PAM, induce immunogenicity, and result in all or nothing changes at the DNA level. This is important as one big issue of these systems is currently their lack of precision. For example, it is not possible to edit a specific cytosine at the target site while leaving surrounding cytosines unaffected [176]. The previously described SDRE systems can achieve this kind of precision to some degree through the use of chemically-modified gRNAs [120], the introduction of defined mismatches within the gRNA-mRNA duplex, or by utilization of the specific counterbase preferences of the used deaminase [118, 138].

1.2.4.2 Therapeutic RNA interference

Sequence-specific gene silencing through targeted mRNA cleavage or translational repression is called RNA interference (RNAi). The elegance to use RNAi for the realization of a therapeutic goal is based on the fact that an endogenous pathway is utilized, requiring only the application of an exogenous small RNA to work.

The endogenous RNAi pathway regulates whole gene expression networks. It does so by using accurately processed short RNAs to guide the RNAi machinery, in form of the RNA-induced silencing complex (RISC), towards target mRNAs. If the short RNA matches its target perfectly, the target mRNA is cleaved, and if not, its translation is repressed. Perfect matching short RNAs targeting only one mRNA are called endogenous short interfering RNAs (endo-siRNA), while short RNAs that contain mismatches, but affect several mRNA targets are called microRNAs (miRNA) [177]. The application of artificial miRNAs, the supplementation with miRNA mimics, or the specific modulation of miRNA expression (miRNA knock-out or knock-down) could be used for therapeutic purposes in the future, but are currently limited by our understanding of miRNA regulatory networks [178-181]. The application of exogenous short interfering RNAs (siRNA) is currently further developed and will soon enter the clinic as a next generation treatment. The first siRNA-based drug, Patisiran, was approved by the US FDA in August 2018 [182]. Further background knowledge about clinical studies in the RNAi field can be found within Pub.1, in the chapter “Therapeutic RNAi”. In general, siRNAs are 21-23 bp in length, often with two nucleotide overhangs at the 3’ ends, although variations in length and overhangs are tolerated [183]. They can be administered as chemically-modified dsRNAs, or in the form of genetically encoded

precursors termed short hairpin RNAs (shRNA). The chemically-modified dsRNAs are already synthesised in the correct length to be directly loaded into the RNA-induced silencing complex (RISC). The encoded precursors have to be processed in advance. First by the nuclear microprocessor complex, and then, after transport into the cytosol via Exportin-5, by the endoribonuclease dicer into mature siRNAs [184]. After loading into RISC, one strand of the siRNA, termed the passenger strand, is cleaved by the Argonaut-2 (Ago2) endoribonuclease, which is a major component of the RISC, and then ejected from the complex. The remaining guide strand steers RISC towards the target mRNA, and thereby assists its cleavage by Ago2 [185]. The knock-down of a target gene *in-vivo* caused by chemically-modified siRNAs can last from several days up to weeks, depending on the targeted tissue [186]. The selection of the guide strand is not fully understood, but depends on several factors, including the composition of the RISC and the thermodynamic properties, 5' nucleotide identity, and structure of the siRNA duplex [187]. Through specific design of the guide strand against SNPs, or disease-causing nucleotide variations, different alleles of the same gene can even be targeted independently [188].

In conclusion, miRNA and siRNA therapeutics could be used to knock-down specific genes or to regulate whole networks of them, which makes RNAi an extremely powerful tool for the manipulation of gene expression. Especially in the context of dysregulated genes in cancer [189], repeat expansion disorders [190], or regarding gain-of-function mutations [191], RNAi has a lot of therapeutic potential. In contrast to all current SDRE and genome engineering systems it does only require a short RNAs, but no additional overexpressed protein or DNA-repair template, to work. This makes it easier to administer, as only the uptake of one agent is necessary to achieve the effect. Although the utilization of the endogenous pathway resulted in toxicity, accelerated tumorigenesis, organ failure, and even death in early animal experiments, it was discovered that this was caused by the saturation of the RNAi machinery with exogenous siRNAs. After it had been revealed that the rate-limiting factors were Exportin-5 and the Argonaut proteins, solutions for adequate dosing were found [192]. However, while RNAi approaches can modify expression levels in a revolutionary way, they are unable to alter mRNA sequences. Similarly, SDRE approaches cannot efficiently knock-down genes, but allow transcriptome engineering in an unprecedented way. Both methods manipulate the transcriptome in completely different ways, thus they complement each other and extend the bandwidth of realizable interventions by available molecular tools.

1.2.4.3 mRNA transcript therapy

In-vitro transcribed (IVT) mRNAs can be directly applied as a drug to achieve transient therapeutic effects. They can encode artificial proteins, like engineered Cas9 nucleases for genome editing, but also wild-type proteins for replacement or supplementation strategies. By encoding specific antigens or antibodies, even immunisation strategies against infectious diseases or cancer might be feasible and

corresponding clinical studies are already underway. Massive initial problems of mRNA transcript therapy approaches, including high immunogenicity and low translational efficiency, could be solved through the development of novel purification methods, synthetic cap analogs, the use of chemically-modified nucleotides, and *in-silico* codon optimization [172]. While mRNA transcript therapies hold great promise for every area of application that benefits from a transient burst-like protein expression, like genome engineering or immunisation strategies, the problem of replacement and supplementation strategies lies in the short lifespan of mRNAs [172]. In addition, proteins replaced by IVT-mRNA application lack one important feature of their endogenous counterparts, which is the accurate and tissue-specific regulation of their gene expression level. While regulatory elements, like miRNA binding sites, can be implemented into UTRs of IVT-mRNAs to control expression in a cell-specific manner, this does not automatically adjust their expression towards natural endogenous levels [162]. A good example for the relevance of expression levels is the Mecp2 protein. An underexpression leads to Rett syndrome and an overexpression to autism spectrum disorders [132, 193].

SDRE approaches aim to manipulate endogenous mRNAs and are therefore sustaining their natural regulation. Depending on the used SDRE method they have also the potential to achieve long term correction. This makes SDRE approaches favourable options over IVT-mRNA based protein replacement strategies, at least within the range of editable triplets and mutations. However, complete loss-of-transcript cannot be addressed by SDRE. The SDRE approaches do therefore complement the existing IVT-mRNA strategies in some specific areas, but not in others.

1.3 The goal of this project

Within the spectrum of existing SDRE approaches (section 1.2.1), our R/G gRNA site-directed RNA editing system excels in one important premise. It uses human, wild-type ADAR2 instead of an engineered editase. The necessity to overexpress ADAR2 represented one of the most problematic limitations of the R/G gRNA system at the time. As explained, earlier ADAR2 has important functions in the circadian gene regulation of several hundred target genes (section 1.1.1.2.1). Therefore, the overexpression of ADAR2 could cause imbalances in gene regulation and editing homeostasis with unforeseeable consequences. More general, adverse effects, like protein burden (section 1.2.3), also have to be considered. By avoiding overexpression, the R/G gRNA system could solve one of its main shortcomings and simultaneously obtain a significant advantage over the other SDRE systems, which are fundamentally restricted to the overexpression of their artificial editases. Additionally, the number of components that must be present within the same cell to achieve editing would drop to one, allowing transcriptome engineering through the simple application of a small RNA, similar to RNAi.

To enable editing with endogenous ADAR enzymes, rational improvements to the R/G gRNA architecture became one major challenge that had to be solved throughout this thesis. To accomplish this, a fast and accurate SDRE reporter system had to be established to rapidly screen novel R/G gRNA designs for superior editing yields in cell culture. Advanced R/G gRNAs had to be tested in cell lines from different tissues, to assess if SDRE with endogenous ADARs would, in principle, be possible in all of them. A panel of ADAR-expressing cell lines from tissues that might represent interesting targets for a therapeutic approach in the future were selected for this. As many of these cell lines are hard to transfect, a transition from plasmid to viral delivery of the encoded R/G gRNAs was necessary as well.

Although the potential application of SDRE as novel therapeutic strategy had already been discussed by Tod Woolf in his original 1995 article, several important questions in this regard remained afterwards unanswered for more than two decades [131]. When this study began, the essential question, if endogenous mRNAs could be targeted by SDRE, was not solved, as only overexpressed plasmid-encoded mRNA targets had been tested at the time. There was also no proof that recoding events caused by SDRE in human cells could actually result in functional repair of signaling cascades connected to disease. The answers to both questions represent a basic prerequisite for the therapeutic utilization of SDRE in the future, and therefore became major goals of this thesis. To tackle the first question, human cells had to be treated with R/G gRNAs targeting a panel of endogenous mRNAs. Then, the A-to-I conversion at the target site had to be quantified to investigate if editing could be achieved. The second question required the development of a reporter assay that allowed the quantification of a pathological cellular phenotype as readout. The signaling cascade behind this phenotype had to be well-characterized to clearly pinpoint any phenotype corrections to our SDRE approach. Finally, the targeted G-to-A point mutation had to be connected to disease.

By addressing these fundamental questions, while simultaneously enabling the recruitment of endogenous ADAR enzymes by the encodable R/G gRNA SDRE system, this thesis intents to pave the way for the development of site-directed RNA editing as a next generation therapeutic strategy.

2. MATERIAL AND METHODS

2.1 Material

2.1.1 Chemicals & solutions

Table 2-1: Used chemicals & solutions

Chemicals & Solutions	Supplier
1kb Plus DNA Ladder (0.1-10.0 kb)	NEB
Agarose HR Plus	Carl Roth
Agarose NEEO	Carl Roth
Ampicillin	Life Technologies
APS (Ammonium peroxidisulfate)	Carl Roth
Avidin from eggwhite	Sigma-Aldrich
Blasticidin S	Life Technologies
Blotting-Grade Blocker	BioRad
Bradford reagent	BioRad
Brillant Blue G 250	Carl Roth
Bromophenol Blue sodium salt	Carl Roth
Caesium chloride	Carl Roth
Calcium chloride dihydrate	Carl Roth
CCCP (Carbonyl cyanide m-chlorophenylhydrazone)	Sigma-Aldrich
Clarity™ Western ECL Substrate	BioRad
DMEM (High Glucose, L-Glutamine, Phenol Red)	Life Technologies
DMSO (Dimethylsulfoxid)	Carl Roth
dNTP solutation mix 40x	NEB
DTT (1,4-Dithiothreit)	Carl Roth
EDTA (Ethylenediaminetetraacetic acid)	Carl Roth
Ethanol	Sigma-Aldrich
FBS (Fetal Bovine Serum)	Life Technologies
Fibronectin human plasma	Sigma-Aldrich
FuGene6	Promega
Glycerol 86%	Carl Roth
Glycine	Carl Roth
HiPerFect	Quiagen

Chemicals & Solutions	Supplier
Hygromycin B	Life Technologies
Interferon-Alpha-A, Human Recombinant	Sigma-Aldrich
Kanamycin	Life Technologies
LB-Agar (Lucia/Miller)	Carl Roth
LB-Medium (Lucia/Miller)	Carl Roth
Lipofectamine 2000	Life Technologies
Lipofectamine 3000	Life Technologies
Loading Dye (6x)	NEB
Methanol	Sigma-Aldrich
Opti-MEM®	Life Technologies
Penicillin-Streptomycin 100x	Life Technologies
PFA (Paraformaldehyde)	Sigma-Aldrich
Phusion HF Buffer 5x	NEB
Poly-D-lysine hydrobromide (150,000-300,000 kD)	Sigma-Aldrich
Potassium acetate	Carl Roth
Potassium chloride	Carl Roth
rNTPs mix (25 mM each)	NEB
Rotiphorese SDS-PAGE 10x	Carl Roth
Rotiphorese® Gel 30 (37,5:1)	Carl Roth
Roti-Safe	Carl Roth
SDS (Sodiumdodecylsulfat)	Carl Roth
Sodium acetate	Carl Roth
Sodiumdihydrogenphosphate	Carl Roth
Spectinomycin	Life Technologies
TEMED (Tetramethylethylenediamine)	Sigma-Aldrich
ThermoPol Reaction Buffer 10x	NEB
TRIS PUFFERAN®	Carl Roth
Triton X-100	Carl Roth
Trypan Blue 5x	Biochrom
Trypsin/EDTA	Biochrom
Urea	Carl Roth
Water (Nuclease free)	VWR
Zeocin	Life Technologies

2.1.2 Media & buffers

Table 2-2: Used media and their composition

Media	Components	Composition
Culture medium [HEK293T, HeLa, HeLa PINK-1 KO, A549, SK-N-BE, Huh7, SY5Y, U87MG, HepG2, MEF, Hepa1-6]	DMEM FBS	1x 10 %
Culture medium [THP1]	RPMI-1640 FBS	1x 10 %
Induction medium [Flp-In T-REx 293]	DMEM FBS Doxycycline	1x 10% 10 ng/ml
LB-Agar	LB-Agar	40 g/L
LB-Media	LB-Media	25 g/L
Parental cell medium [Flp-In T-REx 293]	DMEM FBS Zeocin Blasticidin S	1x 10% 100 µg/ml 15 µg/ml
Selection medium [Flp-In T-REx 293]	DMEM FBS Hygromycin B Blasticidin S	1x 10 % 100 µg/ml 15 µg/ml
SOC medium	Tryptone Yeast-extract NaCl KCl $MgCl_2$ $MgSO_4$ Glucose	2 % 0.5 % 10 mM 2.5 mM 10 mM 10 mM 20 mM

Table 2-3: Used buffer and their composition

Buffer	Components	Composition
Dialysis-buffer	TRIS-base Glycerol	10 mM 0.5 % pH 8.0 (Adjusted with HCL)
Gibson-mix	5x ISO buffer T5 Exonuclease (10 U/µl)	320 µl 0.64 µl

Buffer	Components	Composition
	Phusion DNA-Polymerase (2 U/µl) Taq DNA-Ligase (40 U/µl)	20 µl 160 µl Ad 1.2 mL
ISO-buffer 5x	PEG-8000 Tris-HCl pH 7.5 MgCl2 DTT (1,4-Dithiothreitol) dNTPs NAD	25% 500 mM 50 mM 50 mM 1 mM each 5 mM
Laemmli-buffer 6x	SDS (Sodiumdodecylsulfat) Bromophenol blue Glycerol 86% Tris solution (0.5M, pH6.8) Millipore water DTT (1,4-Dithiothreit)	1.2 g 6 mg 4.7 ml 1.2 ml 2.1 ml 0.93gr DTT
PBS 10x	NaCl KCl Disodium hydrogen phosphate (Na2HPO4) Potassium dihydrogen phosphate (KH2PO4)	8 g 0.2 g 1.42 g 0.27 g Ad to 1 L
TAE 1x	TRIS-base Acetic acid EDTA, pH 8.3	4.84g/L 1.14 mL/L 1mM
TE-buffer 1x	EDTA TRIS-base	1 mM 10 mM pH 8.0 (Adjusted with HCL)
Transfer-buffer 10x	TRIS-base Glycine	30,28 g 144,08 g Ad to 1 L
Urea lysis-buffer 1x	Urea Sodiumdihydrogenphosphate (NaH2PO4) TRIS-base	8 M 100 mM 10 mM pH 8,0 (Adjusted with HCL)

2.1.3 Commercial kit systems

Table 2-4: Used commercial kits

Commercial Kits	Supplier
Adeno-X™ Rapid Titer Kit	Clonetech
Dual-Glo® Luciferase Assay System	Promega
Dual-Luciferase® Reporter Assay System	Promega
Fast SYBR™ Green Master Mix	Thermo Fisher Scientific
High Pure miRNA Isolation Kit	Sigma-Aldrich
High-Capacity cDNA Reverse Transcription Kit	Thermo Fisher Scientific
NucleoBond Xtra Mini- and Midiprep Kit	Macherey-Nagel
NucleoSpin Gel and PCR Clean-Up Kit	Macherey-Nagel
NucleoSpin Plasmid Miniprep Kit	Macherey-Nagel
PARIS Kit	Thermo Fisher Scientific
Quiagen RNeasy MinElute Cleanup Kit	Quiagen
Turbo DNA-free™ Kit	Thermo Fisher Scientific

2.1.4 Enzymes

Table 2-5: Used enzymes

Enzyme	Supplier
AgeI	NEB
Antarctic Phosphatase	NEB
ApaI	NEB
AvrII	NEB
BamHI	NEB
BbsI	NEB
BstXI	NEB
DNase-I (RNase free)	NEB
EagI	NEB
EcoRI	NEB
EcoRV	NEB
HincII	NEB
HindIII	NEB
HpaI	NEB
KpnI	NEB

Enzyme	Supplier
MluI	NEB
M-MuLV Reverse Transcriptase	NEB
Murine RNase Inhibitor	NEB
NheI	NEB
NotI	NEB
PacI	NEB
Phusion® High-Fidelity DNA-Polymerase	NEB
PmeI	NEB
SacI	NEB
SpeI	NEB
T4 DNA-Ligase	NEB
T5 Exonuclease	NEB
Taq DNA-Ligase	NEB
Taq DNA-Polymerase	NEB
Turbo DNase (RNase free)	Thermo Fisher Scientific
XbaI	NEB
XhoI	NEB
XmnI	NEB

2.1.5 Antibodies

Table 2-6: Used Antibodies

Antibody	Supplier
Goat α-Mouse HRP	Jackson Immuno Research Lab (115-035-003)
Mouse Anti-ADAR1 15.8.6	Santa Cruz Biotech. (sc-73408)
Mouse Anti-ADAR2 1.3.1	Santa Cruz Biotech. (sc-73409)
Mouse Anti-Adenovirus Type 5 E1A M58 (RUO)	BD Pharmingen (554155)
Mouse Anti-β-Actin AC15	Sigma Aldrich (A5441)

2.1.6 Oligonucleotides

Oligo #	Oligonucleotide name	Oligonucleotide Sequence [5'->3']
1336	SV40 reverse	TAAGATACATTGATGAGTTTGGACAAACCAC
144	BGH backward	CTAGAAGGCACAGTCGAGGC

Oligo #	Oligonucleotide name	Oligonucleotide Sequence [5'->3']
2157	qPCR_snRNA-U6_fw	GCTTCGGCAGCACATATACTAAAAT
2158	qPCR_snRNA-U6_bw	CGCTTCACGAATTTGCGTGTCAT
2252	BetaActin_qPCR-Adapt-erD1+guide+RG_bw	GAAATCGTCGCTGACTATCGTATGACTTGGTT-GCGTGTGGGATAC
2253	BetaActin_qPCR-AdapterD2+RG-V1_bw	CAGCTGTACCGTTGAATCGAGTGGAA-TAGTATAACAATATGCTAA
2329	FLuc_qPCR-Adapt-erD1+guide+RG-V20_bw	GACTATCGGGACGGCTGGCTGCACGTGGTCGAG
2330	qPCR-AdapterD2+RG-V20_bw	CAGCTGTACCGTTGAATCGAGTGGTCGAGAAGAG-GAGAACAATAT
2374	Malat1_qPCR_fw	AGGCGTTGTGCGTAGAGGA
2375	Malat1_qPCR_bw	GGATTTTTACCAACCACTCGC
238	psilencer_1_fw	CTCGCGCGTTTCGGTGATGACGGTG
2408	HPRT1_V2_qPCR_fw	TGGCGTCGTGATTAGTGATG
2409	HPRT1_V2_qPCR_bw	ACCCTTTCCAAATCCTCAGC
46	Luci_seq_mid_fw	CTATTCTTGCGCAGCTTGCAAGAC
Not as-signed	GAPDH_fw	CAACAGCCTCAAGATCATCAG
Not as-signed	GAPDH_bw	CCTTCCACGATACCAAAGTTG

2.1.7 qPCR primer sets

Table 2-7: Used qPCR primer sets

Primer Set	Primer
GAPDH (cytoplasmic reference gene)	GAPDH_fw + GAPDH_bw
HPRT1 (cytoplasmic reference gene)	2408 + 2409
Malat1 (nuclear reference gene)	2374 + 2375
R/G gRNA design #2 β-Actin 3' UTR TAG #1	2252 + 2253
R/G gRNA design #3 Firefly Luciferase W417X Amber	2329 + 2330
U6 snRNA (nuclear reference gene)	2157 + 2158

2.1.8 Adenoviruses

Insert	Purification	IFU/µl
Dual-Luciferase Firefly W417X Amber R/G gRNA design #3	Crude	$7.42*10^6$
Dual-Luciferase wt/amb Reporter	Crude	$3.85*10^7$
Dual-Luciferase wt/wt Reporter	Crude	$1,70*10^5$
β-Actin 3' UTR TAG#1 R/G gRNA design #2	CsCl-Gradient/Dialysis	$2.88*10^7$

2.1.9 Plasmids

pTS#	Insert	Backbone	Resistance
pTS247	Hind-III cutting-side, Spacer, Bbs-I cutting side	pSilencer	Amp
pTS296	ADAR2 with His-tag, Gibson-Linker-Element with EcoRV-side	pEdit1.2	Amp
pTS363	ADAR2 with His-tag, 5x RG-V1 gRNA PINK-1 W437X amber CCAp8 with BoxB	pEdit1.2	Amp
pTS554	Dual-Luciferase (Renilla wt + 2A peptide + Firefly wt)	pShuttle-CMV	Kan
pTS555	Dual-Luciferase (Renilla wt + 2A peptide + Firefly W417X Amber)	pShuttle-CMV	Kan

2.1.10 Strains

Table 2-8: Used bacterial and cell culture strains

Strain	Supplier
A549	University Hospital Tübingen (Kormann Lab)
BJ5183 electroporation-competent cells	Agilent Technologies
***E. coli* Stbl3**	Thermo Fisher Scientific
***E. coli* XL1 blue**	ETH Zurich
hADAR1p110 Flp-In T-REx 293	Generated in our lab
hADAR1p150 Flp-In T-REx 293	Generated in our lab
hADAR2 Flp-In T-REx 293	Generated in our lab
HEK293T	DSMZ
HeLa	Hertie-Institute (Kahle Lab)
HeLa PINK-1 KO	Generated in collaboration with the Kahle Lab
Hepa1-6	Nordheim Lab
HepG2	DSMZ
Huh7	CLS GmbH
MEF	Nordheim Lab
Parental Flp-In T-REx 293	AG Jung
SK-N-BE	Hertie-Institute (Kahle Lab)
SY5Y	Hertie-Institute (Kahle Lab)
THP1	Hertie-Institute (Kahle Lab)
U87MG	Hertie-Institute (Kahle Lab)

2.1.11 Laboratory equipment

Table 2-9: Used laboratory equipment

Equipment	Supplier
7500 Fast Real-Time PCR System	Applied Biosystems
Agarose gel electrophoresis chamber horizontal	BioRad
ARKTIK thermal cycler	Thermo Fisher Scientific
AxioCam HRc microscope camera	Zeiss
Axio-Imager equipped with ApoTome	Zeiss
Axiovert 200M microscope	Zeiss
Heraeus™ Megafuge™ 8R	Thermo Fisher Scientific
MIKRO 120	Hettich Zentrifugen
MIKRO 220 R	Hettich Zentrifugen
Multipart pipette set	Gilson
Neubauer counting chamber	Carl Roth
POWER PAC 300	BioRad
POWER PAC basic	BioRad
Spark 10M Plate reader	Tecan
ThermoMixer C	Eppendorf
Tube revolver / rotator	Thermo Fisher Scientific
Vortexer	NeoLab

2.1.12 Consumables

Table 2-10: Used consumables

Equipment	Supplier
10K Slied-A-Lyzer Dialyis Cassette G2 (87730)	Thermo Fisher Scientific
LumiNunc™ F96-MicroWell™ plate, with lid, sterile (732-2696)	VWR

2.2 Methods

Methods not mentioned in this section, but referred to in the results and discussion chapter, are described in the method parts of the corresponding, accepted publications, which are listed in section 6, or were performed as described in [148].

2.2.1 Design of *in-silico* optimized R/G gRNAs using the recruitment cluster finder python tool

Several R/G gRNAs used in this thesis contain functional elements called recruitment clusters, which were *in-silico* optimized using our custom-made recruitment cluster finder (RCF) python tool. Background knowledge regarding recruitment clusters (RCs) and their functional relevance is described in section 3.12.6. The RCF tool was programmed in python 3.7.2 and its core functions are listed in the appendix. To compute new RCs, the necessary inputs were researched online and entered into the RCF. This includes the search of target gDNAs in online databases like uniprot [194] or the ensembl genome browser [195]. The Figure 2-1 explains in a conceptual way how the tool selects recruitment clusters. Figure 2-2 shows the input panel of the recruitment cluster finder and explains how it was used. The folding of the R/G gRNAs was performed using ViennaRNA [196].

After the algorithm had calculated the new R/G gRNAs, they were exported into a text-file. The R/G gRNAs that resulted in the weakest secondary structure formation within the antisense part were double-checked for their structure using the Nupack web application [197]. The most promising ones were then blasted using the “blastn” NCBI web application against the human genomic and transcript database optimized for “somewhat similar sequences” [198]. R/G gRNAs containing RCs with off-target editing potential were excluded. Optimal R/G gRNAs were then back engineered into two DNA oligonucleotides with the necessary overhangs for the cloning procedure shown in Figure 2-3C. This was followed by the cloning procedure explained in section 2.2.2.

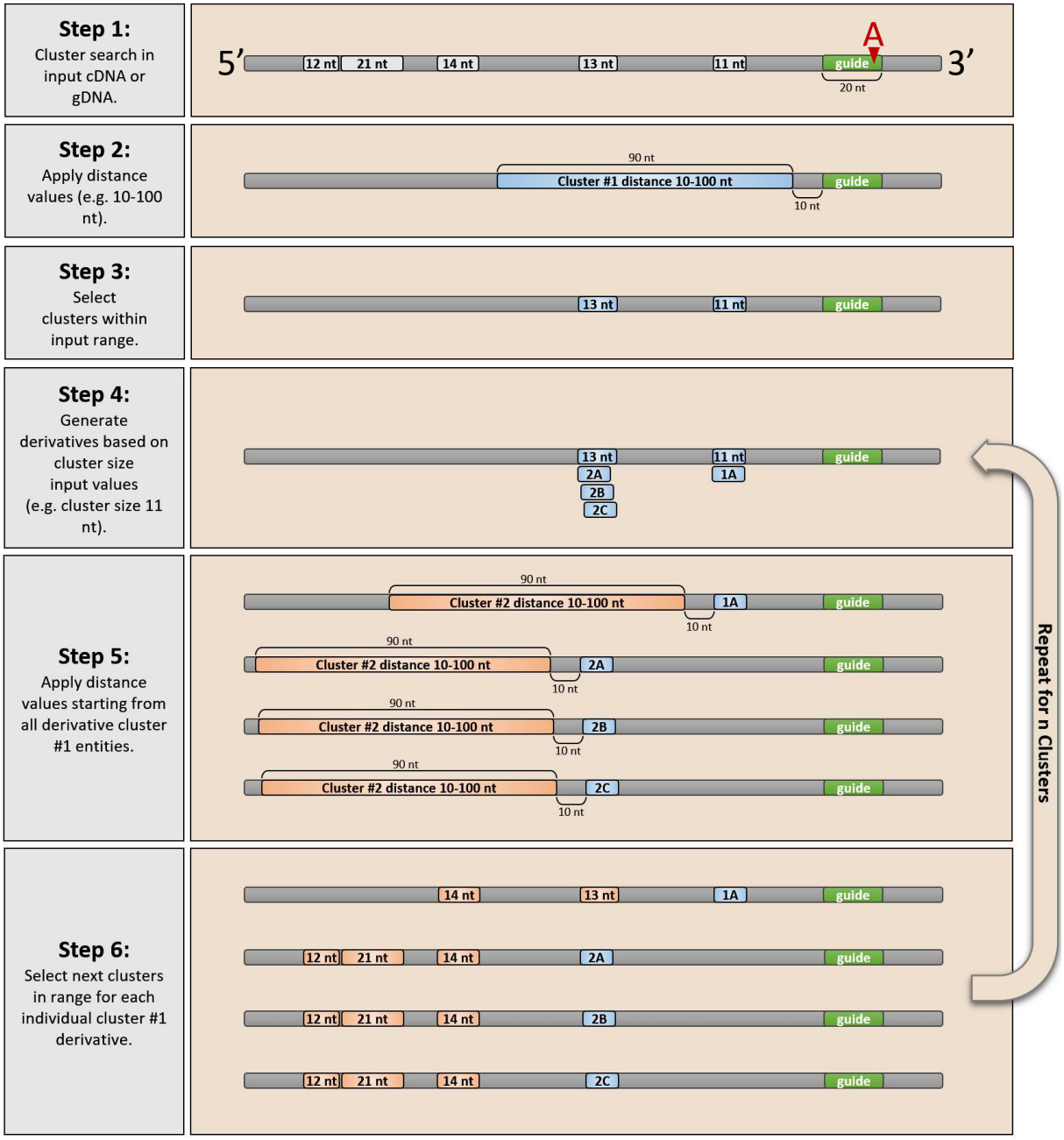


Figure 2-1: Conceptual description explaining how the recruitment cluster finder (RCF) selects RCs.

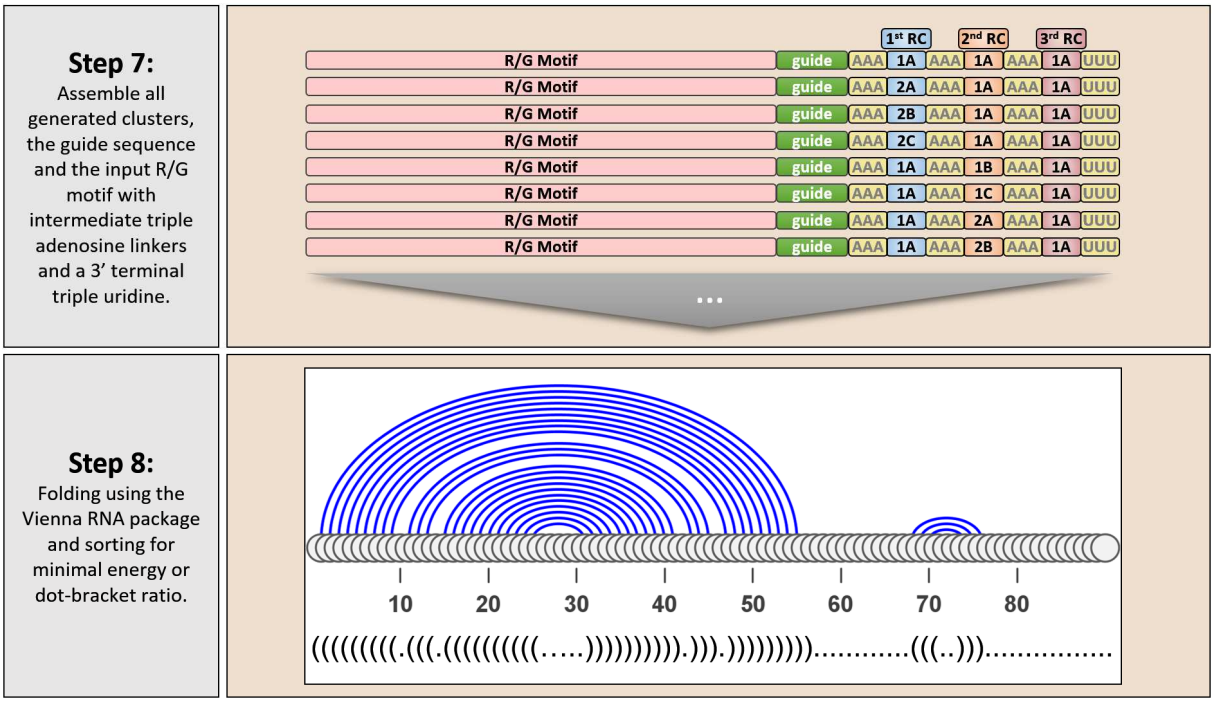


Figure 2-1 continued: Conceptual description explaining how the recruitment cluster finder (RCF) selects RCs.

This figure does not describe the exact combinatorial implementation, which is used by the algorithm to process input data, but explains the idea behind the tool in a conceptual way for better understanding. **(Step 1)** The gDNA/cDNA corresponding to the target mRNA is screened for clusters that only contain G, C, T and GA. Each cluster is defined by base distance indices relative to the beginning of the cDNA/gRNA input sequences. In the current version (1.0.1), the screening is performed 5' of the guide sequence. **(Step 2)** Starting from the guide sequence, clusters in 5' range are detected. **(Step 3)** These clusters are selected and analyzed for their size. **(Step 4)** The cluster size input is used to generate cluster derivatives. If an uninterrupted cluster is, e.g., 13 nt long, then three derivative clusters can be created if the input cluster size was set to 11 nt. **(Step 5)** Starting from the first set of derivative clusters, the next clusters in range are detected. **(Step 6)** These clusters are selected and analyzed for their size. Step 4-6 are repeated until n clusters are selected. **(Step 7)** The resulting list of clusters, which are matching all input variables are recombined and assembled with the input R/G motif, triple adenosine linkers and the three terminal uridines, which result from the U6 termination sequence. **(Step 8)** The Vienna RNA package is used to fold all R/G gRNAs within the list and to generate dot-bracket representations of these folds. The RCF allows to sort the structures by their free energy or by their dot-bracket ratio (ratio of the dot-bracket notation). Structures with good dot-bracket ratios (minimal base pairing within the antisense part of the R/G gRNA), are further sorted for the shortest number of brackets in a row within the antisense part. The shortest ones, which represent the weakest secondary structures, get the highest ranking. The selected clusters of resulting R/G gRNAs with good secondary structures can then be manually blasted to exclude recruitment clusters which might cause off-target effects within the transcriptome.

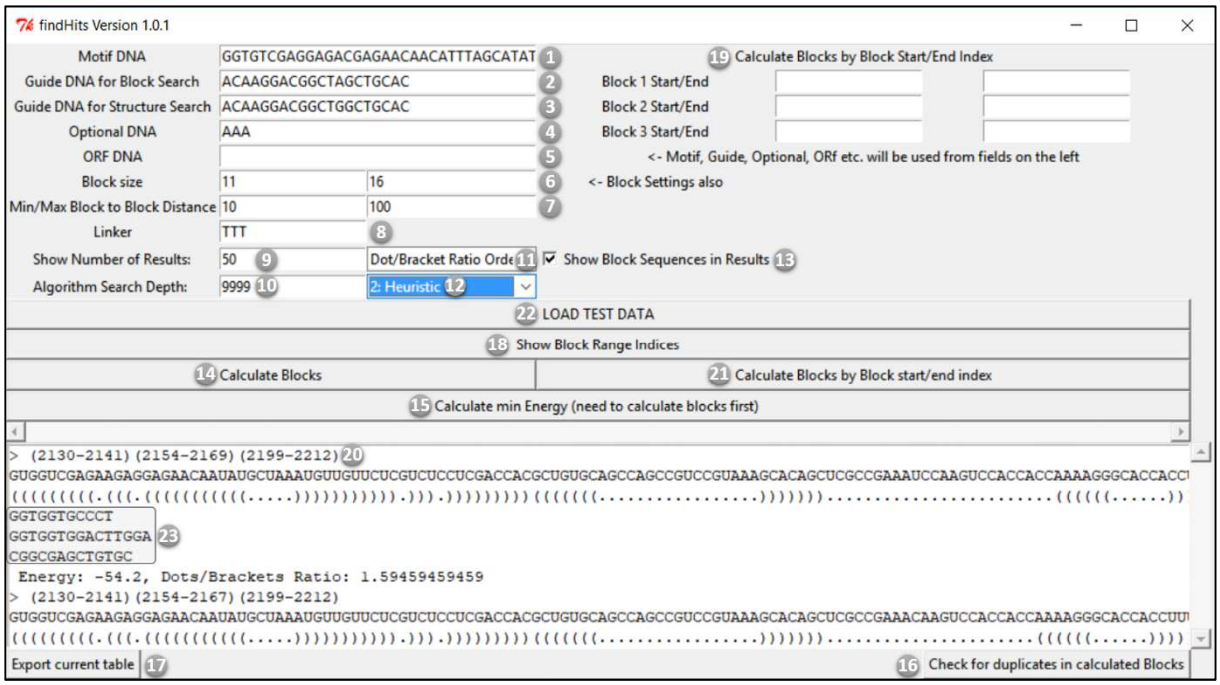


Figure 2-2: Input panel of the recruitment cluster finder python tool (Version 1.0.1).

(1) Input field for the R/G gRNA motif sequence. **(2)** Input field for the guide sequences without the counterbase mismatch, allowing the tool to find the starting point for the cluster search. **(3)** Input field for the guide sequences with the counterbase mismatch used by the tool to assemble the final R/G gRNA. **(4)** Input for 3' terminal sequences, e.g. a triple A, resulting in a triple U for the structure calculation, and resembling the residues of the U6 termination sequence. **(5)** Input field for the target cDNA or gDNA. **(6)** Input field for recruitment cluster size range, e.g. 11-16 nt. **(7)** Input field for minimal and maximal cluster-to-cluster distance, e.g. 10-100 bp. **(8)** Input field for linker sequence, e.g. TTT, which becomes AAA in the final R/G gRNA. **(9)** Input field for the number of resulting R/G gRNAs that the program should print (Length of the results list). **(10)** Algorithm search depth. The 9999 input results in the analysis of each possible cluster combination. Numerically lower inputs limit the search to the n clusters closest to the guide sequence. **(11)** Selection field for the used R/G gRNA sorting algorithm ("Sorting by numerical order", "dot/bracket ratio order", or "minimal energy order"). **(12)** Type of algorithm search pattern ("Exact" searches the whole sequences, "Heuristic" uses grid sampling and skips adjacent/similar cluster sequences). **(13)** Field that can be ticked to print each resulting cluster separately as DNA sequence as shown in (23). **(14)** Button click starts the calculation of the clusters and filters them for combinations that apply to the other input variables, like cluster size in (6), or cluster-to-cluster distance in (5). **(15)** After clusters calculation, this button starts the recombination of the clusters followed by assembly and folding of the new R/G gRNAs. **(16)** After the list of new R/G gRNAs was printed, it can be screened for duplicates. This function might be used in subsequent versions of the tool to remove potential duplicates. **(17)** This button allows the export of the sequence list as text file. **(18)** This button prints the indices of all possible clusters and allows to handpick clusters for manual input into the fields under (19). **(19)** In these fields, handpicked clusters can be manually entered. **(20)** The output field lists all calculated R/G gRNAs, up to the limit chosen in (9), and is always showing the cluster indices, followed by sequence, dot/bracket representation, and also the separate clusters as DNA sequence, if (13) was ticked. **(21)** Allows to calculate R/G gRNAs based on input in (19). **(22)** This button loads a correct set of input data and allows the user to experiment with input variables. **(23)** Output field that shows the clusters as DNA sequences, to allow a fast and simple blast of each cluster.

2.2.2 Cloning of R/G gRNA expression vectors

The R/G gRNA expression vectors, based on the "pSilencer 2.1 U6 hygro" backbone, were initially cloned using primer extension PCR (Figure 2-3A and B). Therefore, two or more overlapping forward primers were used in successive PCRs with the same reverse primer until the new insert was extended to its full size, containing terminal restriction sites. The insert, as well as the corresponding pSilencer expression vector, were then cut with the restriction enzymes Hind-III and Sap-I (Figure 2-3A), or Xba-I and Sap-I (Figure 2-3B), according to the manufacturers protocol. The ligation was performed using the T4 DNA ligase, again after manufacturer protocol. This was followed by heat shock transformation into XL-1 blue *e. coli*, colony PCR, and sequencing of positive clones.

To circumvent the need of multiple extension PCRs, a new procedure was established. A vector (pTS247) was created that contained the recognition sequence for the golden-gate cloning enzyme Bbs-I arranged in a way that the cutting site was positioned at the intersection between the R/G motif and the antisense part of the gRNA (Figure 2-3C). It allowed the generation of a 5' GTGG-overhang at the beginning of the R/G motif. After the pSilencer backbone was cut with the restriction enzymes Hind-III and Bbs-I, this system allowed the use of two hybridized, phosphorylated oligonucleotides as inserts, which were already synthesized with the necessary overhangs. The lyophilized oligonucleotides were re-suspended in Millipore water and to a concentration of 1 µg/µl. Then, 1.3 µl of each oligo were added to 197.4 µl Millipore water. 39 µl of this mixture were then combined with 5 µl T4-PNK buffer, 5 µl ATP (10 mM), and 1 µl T4 polynucleotide kinase. The phosphorylation (30 min at 37°C) and hybridization (heating to 95°C and stepwise cooldown by 1°C per minute for 90 minutes) of the oligonucleotides was performed in a thermocycler. Based on the used amounts of oligonucleotides, the insert was expected to have a concentration of 10 ng/µl and was directly used for the ligation, without further clean-up steps. Due to the high efficiency of the new procedure (>90% positive colonies), the colony PCR could be skipped. Otherwise, the remaining cloning procedure was performed as explained before. By using the Bbs-I and the BamHI restriction sites, the R/G motif could also be specifically exchanged, as the BamHI restriction site on the pSilencer expression vector was located directly between the U6 promoter and the beginning of the R/G motif. By using the Hind-III and BamHI restriction sites and inserts based on long oligonucleotides up to 180 nt, it was also possible to simultaneously clone novel antisense parts and R/G motifs (Figure 2-3C).

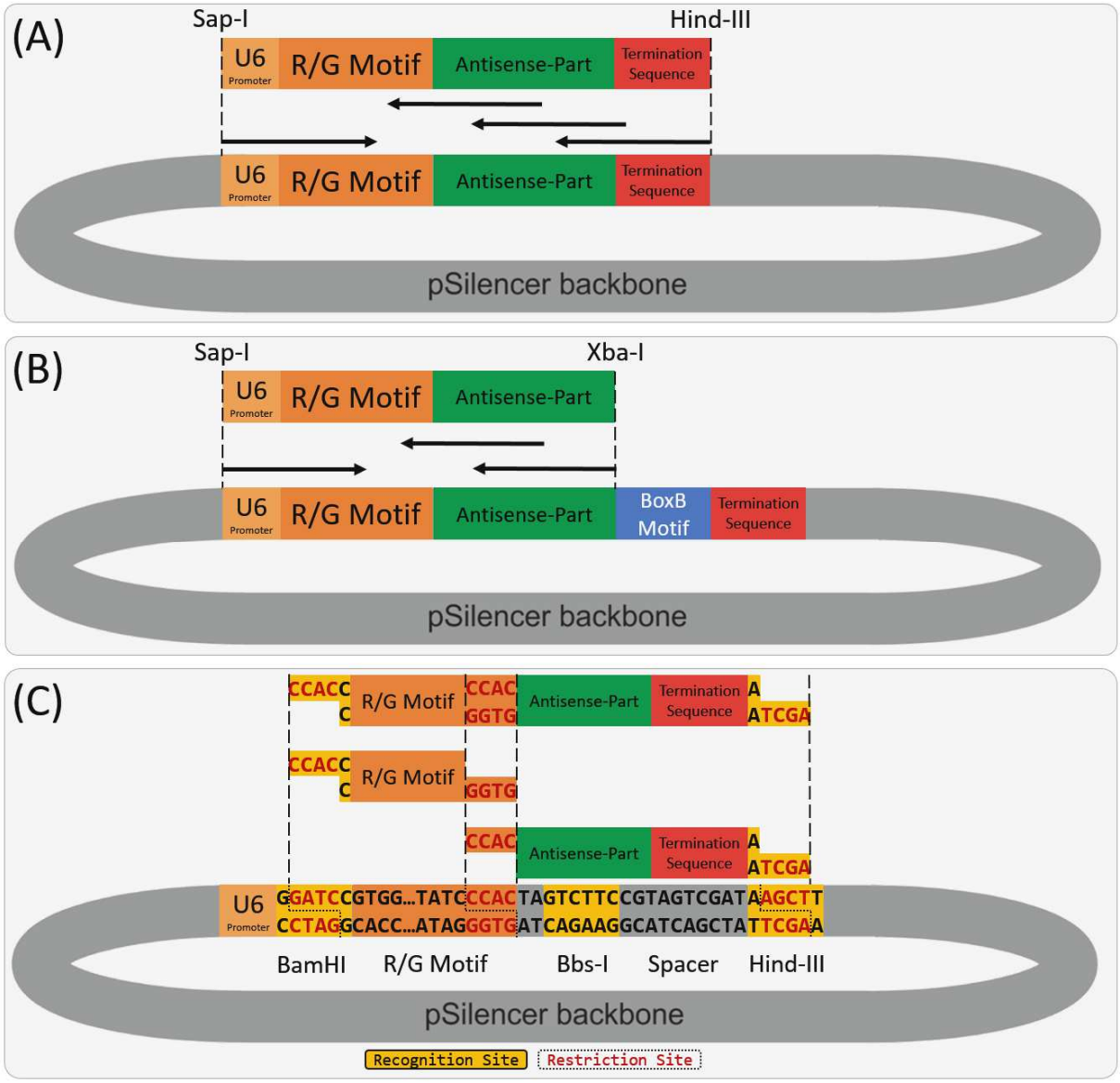


Figure 2-3: Cloning procedures for R/G gRNA expression vectors.

(A) Primer extension PCR cloning of R/G gRNA antisense parts containing no BoxB motif. The arrows indicate forward and reverse primers. The terminal 3’ and 5’ primers contain the restriction sites for Hind-III and Sap-I, allowing the cloning of the insert into the backbone used as PCR-template. (B) Primer extension PCR cloning of R/G gRNA antisense parts containing the BoxB motif. The arrows indicate forward and reverse primers. The terminal 3’ and 5’ primers contain the restriction sites for Xba-I and Sap-I, allowing the cloning of the insert into the backbone used as PCR-template. Due to the Xba-I restriction site near the BoxB motif, the U6 termination sequence is already part of the backbone, reducing the necessary amount of extension PCRs. (C) Modular oligonucleotide hybrid-insert cloning of R/G gRNA antisense parts using a modified pSilencer backbone. Two phosphorylated and hybridized oligonucleotides can be used as insert, using the restriction site Hind-III and the newly introduced golden gate restriction site Bbs-I, which creates a 5’-GTGG overhang at the intersection between the antisense part and the R/G motif. By using Bbs-I and BamHI, the R/G motif can be exchanged. By using Hind-III and BamHI, the antisense part and the R/G motif can be exchanged simultaneously.

2.2.3 Quantification and localization of R/G gRNAs by qPCR

To evaluate the relative quantity and subcellular localisation (nucleus, cytoplasm) of R/G gRNAs, a RT-qPCR workflow based on the ΔΔC(t)-method was established. After the cell culture experiment was performed and the cells had been harvested by trypsinization, the RNA was isolated using either the High Pure miRNA Isolation Kit from Roche, or the PARIS Kit from Invitrogen, both after manufacturer protocol. For the High Pure miRNA Isolation Kit the binding enhancer was used and the elution was performed with a total of 50 µl elution buffer per sample. For the PARIS Kit 300 µl fractionation buffer, disruption buffer, and 2x Lysis/binding solution were used. Centrifugation of fractionated samples was performed for 5 min at 4°C and 500g. The optional washing step of the nuclear fraction was performed with an additional 300 µl fractionation buffer. The filter cartridges were centrifuged at 15.000g. Elution was performed two times, first with 40 µl, followed by 10 µl of elution-buffer heated to 95°C. Beginning with the RNA isolation, all steps were performed on ice, except for the 2xLysis/binding solution treatment in the PARIS Kit protocol. RNase-free filter-tips were used for all pipetting steps. The High Pure miRNA Isolation Kit was used for quantification experiments including only plasmid and adenovirus (AdV) encoded R/G gRNAs (e.g. Figure 3-16A, B and C). The PARIS Kit was used for quantification and localisation experiments which also included *in-vitro* transcribed (IVT) R/G gRNAs (e.g. Figure 3-16D and Figure 3-21). The TURBO DNA-free kit from Invitrogen was used to remove traces of R/G gRNA expression plasmid from the isolated total RNAs. The Turbo-DNase treatment was performed as indicated in the section “rigorous DNase treatment” of the manufacturer protocol. Two successive 30 min incubations at 37°C, using 1 µl of Turbo-DNase each, were performed in a total volume of 50 µl.

The high-capacity cDNA reverse transcription kit from Applied-Biosystems was used to reverse-transcribe the total RNA after manufacturer protocol (protocol without RNase inhibitor). 500 ng of DNase treated total RNA were used per sample. If samples had a low concentration, multiple reverse transcriptions were performed until 500 ng RNA had been used in total. Such samples were then pooled in the PCR clean-up using the MN NucleoSpin Gel and PCR clean-up kit.

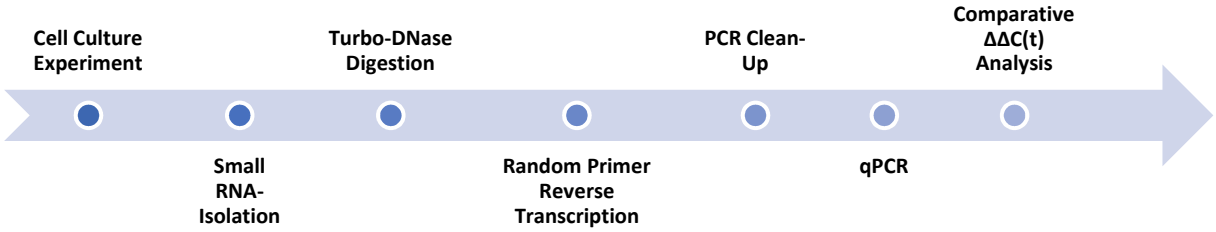


Figure 2-4: qPCR workflow to quantify R/G gRNA expression after cell culture experiments.

The cDNA of each sample was eluted in 18 µl TE-buffer and the concentration measured in triplicate using a Nanodrop 1000 spectrophotometer. The cDNA was then diluted to 10 ng/µl using TE-buffer. In the qPCR, 2 µl of this dilution (20 ng) and 18 µl of mastermix containing the appropriate primer pair (see Table 2-7) were used per well of the 96-well qPCR-plate.

The mastermix consisted of 10 µl Sybr-Green-Mix, 7.2 µl Nuclease-free water, as well as 0.4 µl of each primer, and was scaled up with the performed number of technical replicates. The cDNA samples were usually measured with two to three technical replicates, while TE-buffer negative controls were measured only in duplicate. The pipetting was performed at a pre-PCR workplace with dimmed light, due the light sensitivity of the Sybr-Green-Mix. After 18 µl of mastermix and 2 µl of sample per well had been distributed, the plate was sealed with a clear adhesive film and spun down in a plate centrifuge. From the time the adhesive film was attached, until the plate was inserted into the Applied-Biosystems 7500 qPCR machine for the run, it was covered with aluminium foil for light protection. The run method is shown in Table 2-11.

Table 2-11: RT-qPCR run method

Stage	Temperature [°C]	Duration [s]	Cycle #
Initial denaturation	95	20	1
Denaturation	95	3	40
Annealing/Elongation	60	30	
Melt curve	95	15	1
	60	60	
	95	15	
	60	15	

After the run, the melt curves were analysed to ensure specific amplicons, and a baseline correction was applied for each dataset before the C(t) values were acquired. The 7500 data analysis software was used for that purpose. The start cycle number was changed to the cycle number where the first background fluorescence occurred. The end cycle number was changed to the cycle number were the first specific amplification took place. After this, the ΔΔC(t)-calculations were performed with the following equations and based on the references [199-201].

Calculation of ΔΔC(t) for the target gene between two different samples:

$$\Delta C(t)_{\text{treatment A (e.g. plasmid)}} = \text{Treatment A mean } C(t)_{\text{target gene (e.g. R/G gRNA)}} - \text{Treatment A mean } C(t)_{\text{reference gene (e.g. U6 snRNA)}}$$

$$\Delta C(t)_{\text{treatment B (e.g. AdV)}} = \text{Treatment B mean } C(t)_{\text{target gene (e.g. R/G gRNA)}} - \text{Treatment B mean } C(t)_{\text{reference gene (e.g. U6 snRNA)}}$$

$$\Delta\Delta C(t) = \Delta C(t)_{\text{treatment A (e.g. plasmid)}} - \Delta C(t)_{\text{treatment B (e.g. AdV)}}$$

Calculation of ΔΔC(t) between nucleus and cytoplasm for one reference gene:

$\Delta C(t)_{nuclear}$ = Nuclear fraction mean $C(t)_{reference\ gene}$

$\Delta C(t)_{cytoplasmic}$ = Cytoplasmic fraction mean $C(t)_{reference\ gene}$

$\Delta\Delta C(t)$ = $\Delta C(t)_{nuclear} - \Delta C(t)_{cytoplasmic}$

Calculation of ΔΔC(t) between nucleus and cytoplasm for one target gene:

$\Delta C(t)_{nuclear}$ = Nuclear fraction mean $C(t)_{target\ gene\ (e.g.\ R/G\ gRNA)}$
– Nuclear fraction geometric mean $C(t)_{reference\ genes\ (e.g.\ HPRT1,\ GAPDH,\ U6\ snRNA,\ Malat1)}$

$\Delta C(t)_{cytoplasmic}$ = Cytoplasmic fraction mean $C(t)_{target\ gene\ (e.g.\ R/G\ gRNA)}$
– Cytoplasmic fraction geometric mean $C(t)_{reference\ genes\ (e.g.\ HPRT1,\ GAPDH,\ U6\ snRNA,\ Malat1)}$

$\Delta\Delta C(t)$ = $\Delta C(t)_{nuclear} - \Delta C(t)_{cytoplasmic}$

Calculation of fold change based on ΔΔC(t):

Fold change = $2^{-\Delta\Delta C(t)}$
Maximum fold change = $2^{-\Delta\Delta C(t) - SD\ of\ target\ gene\ C(t)}$
Minimum fold change = $2^{-\Delta\Delta C(t) + SD\ of\ target\ gene\ C(t)}$

Calculation of the nuclear percentage for one target gene:

$$\%\ Nuclear\ localization = \frac{nuclear\ fold\ change}{(nuclear\ fold\ change\ + cytoplasmic\ fold\ change)} * 100$$

For each new primer pair, the amplification efficiency and the melt curve were analysed using a cDNA dilution series, to ensure high quality datasets (see Figure 3-15). To normalize the R/G gRNA data in quantification experiments using plasmid and AdV encoded R/G gRNAs, the pol III gene product U6 snRNA was used as reference gene accounting for the fact that R/G gRNAs are pol III gene products. For the quantification and localization experiments also using IVT R/G gRNAs, the U6 snRNA and Malat1 reference genes were used as nuclear reference genes and the GAPDH and HPRT1 reference genes were used as cytoplasmic ones.

2.2.4 Recombinant adenovirus production

This procedure for the production of recombinant serotype 5 adenoviruses (AdVs) is based on the original protocol from our cooperation partner Prof. Ulrike Naumann, who kindly supported Dr. Wettengel and myself in our adenovirus (AdV) projects. We used the first generation pAdEasy-1 system from Agilent technologies [202, 203]. The production process is shown in Figure 2-6. Our GOIs were cloned into pShuttle, pShuttle-CMV, pAd-Track, or pAd-Track-CMV backbones (Table 2-12).

Table 2-12: Shuttle vector features.

This table was adopted from the "AdEasy Adenosirus Vector System" manual, revision D1, from Agilent technologies. © Agilent Technologies, Inc. Reproduced with Permission, Courtesy of Agilent Technologies, Inc.

Vector	Cloning capacity	Promoter	Poly A	MCS restriction sites	Description
pShuttle	7.5 kb	—	—	*Kpn* I, *Not* I, *Xho* I, *Xba* I, *EcoR* V, *Hind* III, *Sal* I, *Bgl* II	Ligate an entire expression cassette into MCS
pShuttle-CMV	6.6 kb	CMV	+	*Kpn* I *Sal* I, *Not* I, *Xho* I, *Hind* III, *EcoR* V	Ligate gene of interest into MCS between the CMV promoter and poly A
pShuttle-IRES-hrGFP-1 or -2	5.2 kb	CMV	+	*Bgl* II*, *Not* I*, *Sca* I*, *Nhe* I* *Spe* I*, *EcoR* V, *Pvu* I, *Sal* I, Srf I, *Xho* I	Ligate gene of interest into MCS, in-frame with the FLAG or HA tag. Dicistronic transcript encoding hrGFP allows monitoring of the expression of the gene of interest by GFP fluorescence

Then, they were linearized with Pme-I and purified by ethanol precipitation to remove salts, which might interfere with the electroporation. A mixture of 500 ng Pme-I linearized shuttle vector and 100 ng of pAdEasy-1 in a total volume of 6 µl were then electroporated together with 20 µl of BJ5183 *E. coli* bacteria suspension using a BioRad Genpulser at 1.6 kV, 200 Ohm, and 25 µF. It was important to constantly work on ice during this process. There was no additional mixing of bacteria and plasmid solution beside the transfer of the 6 µl plasmid solution into the bacteria suspension, as competent bacteria are extremely sensitive. After the settings were applied to the BioRad Genpulser, the total 26 µl were transferred from the Eppendorf tube into the precooled electroporation cuvette and moved to the bottom of it by a single firm shake. It was important for high transformation efficiency to keep the time between the transfer into the cuvette and the electroporation pulse as short as possible.

Immediately after the pulse the cuvette was cautiously rinsed with 500 µl of SOC medium and the cells were transferred to a fresh Eppendorf tube. This was followed by 45 min of incubation at 37°C and the subsequent production of different dilutions, again in SOC medium (undiluted, 1:2, 1:10, 1:100). 100 µl dilution per plate were distributed on kanamycin LB-agar plates followed by over-night incubation at 37°C.

On the next day, the smallest colonies were picked and used to inoculate 6 ml overnight cultures. Small colonies are more likely to contain positive clones, as the over 36 kb large constructs resulting from the recombination with the pAdEasy-1 AdV genome need more time in replication, compared to the small shuttle vectors containing only the GOI and a kanamycin resistance. We also picked a single large colony as negative control.

After another 24h, the plasmids were isolated from the overnight cultures using the MN NucleoBond® mini gravity-flow columns, because spin columns could damage the large plasmids by generating shear forces. This was followed by a Pac-I control digestion for a duration of 1h, using 20.2 µl isolated plasmid, 2.3 µl CutSmart buffer, and 0.5 µl Pac-I restriction enzyme. The digested plasmids were analyzed using a 0.5% agarose gel, which ran for 40 min at 100V. Restriction of positive, recombinant pAd-Easy constructs with Pac-I resulted in a large ~30 kb fragment and a smaller fragment, either 3 kb or 4.5 kb in size. If the recombination of shuttle vector and pAdEasy-1 took place between the left homology arms, the fragment was 3 kb in size. If the recombination took place at the Ori, then the fragment was 4.5 kb in size. The reason was that the distance between the left homology arm and the Ori was different between the pAd-Easy and the shuttle vectors (see Figure 2-6). Negative clones showed either only the pAd-Easy, or only the shuttle vector bands (see Figure 2-5).

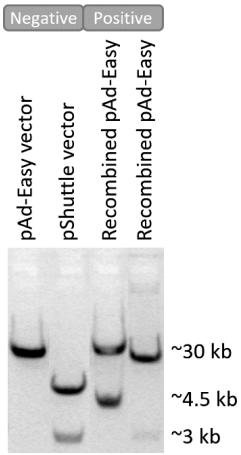


Figure 2-5: Examples for negative controls and positive clones after the recombination of the pAd-Easy1 adenovirus genome and a shuttle vector in BJ5183 *E. coli*.

The indicated plasmids were digested using Pac-I and then resolved in a 0.5% agarose-gel. Shown are empty pAd-Easy and pShuttle backbones as negative controls and two positively recombined clones.

The plasmid corresponding to a positive clone was then heat-shock transformed into chemically competent XL-1 blue *E. coli*. Therefore, 100 µl of TE-buffer and 10 µl of the undigested plasmid were added to 100 µl of competent XL-1 blue *E. coli* and incubated on ice for 30 minutes. This was followed by a heat shock at 42°C for 1 minute and another 5 minutes of incubation at room-temperature. Then 800 µl LB-medium were added, followed by another incubation, this time for 1h at 37°C. The XL-1 blue *E. coli* were then pelleted by centrifugation at 5000 rpm for 2 minutes.

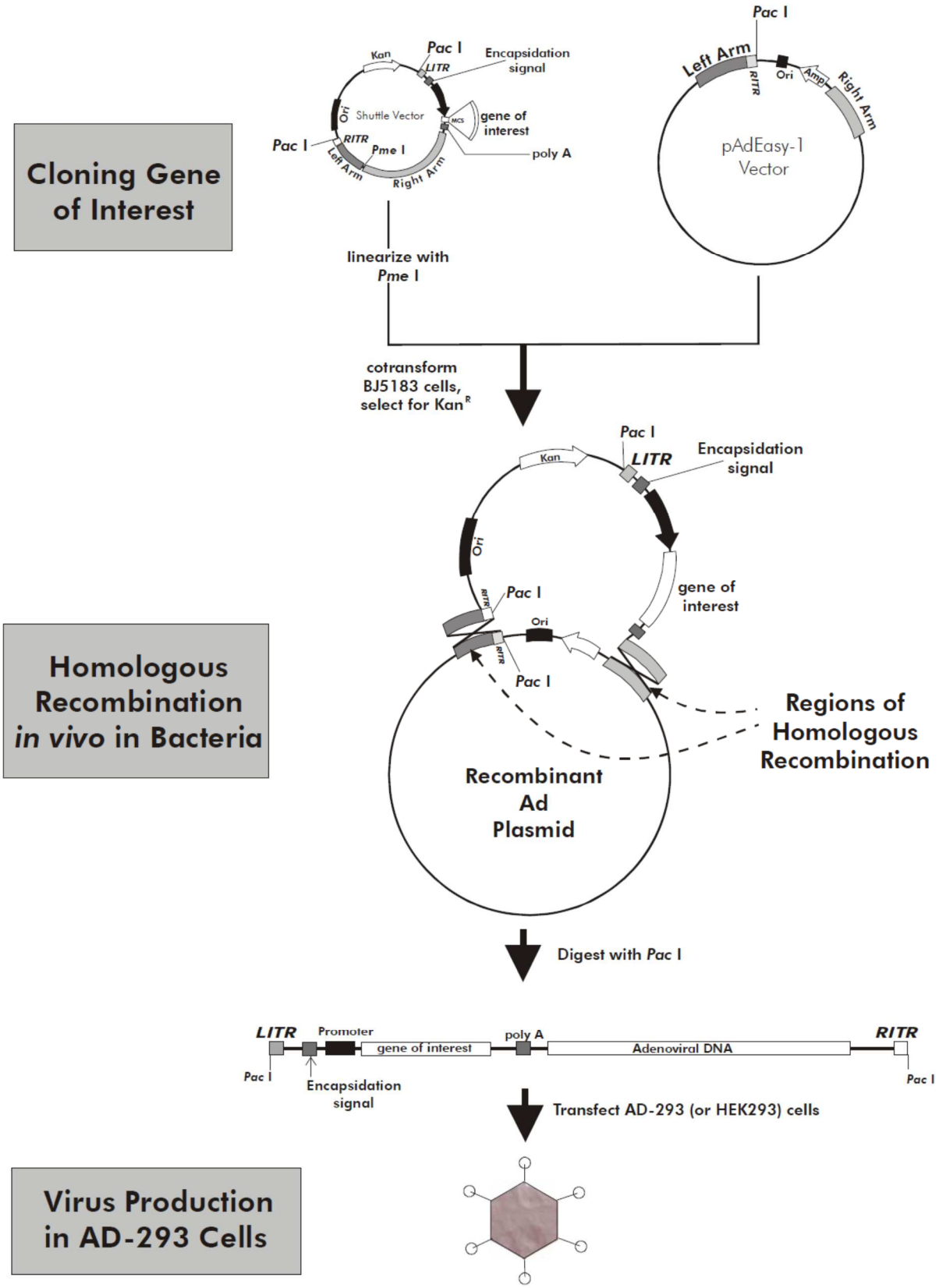


Figure 2-6: Production of recombinant adenovirus.

A Pme-I linearized shuttle vector, containing the GOI, is recombined with the pAd-Easy1 vector in BJ5183 *E.coli*. The resulting recombinant Ad plasmid is digested with Pac-I and then transfected into Ad293 producer cells. This figure was adopted from the "AdEasy Adenovirus Vector System" manual, revision D1, from Agilent technologies. © Agilent Technologies, Inc. Reproduced with Permission, Courtesy of Agilent Technologies, Inc.

After decantation of the supernatant, the pellet was re-suspended in the remaining supernatant and distributed over a kanamycin LB-agar plate, followed by incubation over-night at 37°C.

The next day, a colony was picked from the plate to inoculate a 100 ml overnight culture. After 24h, the recombinant plasmids were isolated using a MN NucleoBond® midi gravity-flow column. Then 30-100 µg of recombinant plasmid were digested using Pac-I (see Figure 2-6). The digestion product was again purified by ethanol precipitation. After Ad293 producer cells had been seeded in a 15 cm cell culture dish, they were transfected with 10 µg of the Pac-I digested recombinant plasmid at a confluency of 40-80%, using a plasmid (µg) to Lipofectamine 2000 (µl) ratio of 1:3. From this point on, all work steps, which could potentially expose the experimenter to the AdV, were performed in a S2(German biohazard level 2)-certified safety cabinet. The cells were cultivated 7-10 days after transfection. During this time, medium was only changed when necessary to prevent acidic conditions. If possible, medium was not removed, but only added, to prevent dilution of the newly produced AdV within the medium. The cells were harvested when widespread hotspots of cytopathic effect (CPE) could be detected within the plate. The assessment of the optimal time point for the harvest, only based on the microscopically visual adenoviral cell lysis, was the most critical decision for high titers in the production process. As no high-tech methods for evaluation of the cell layer were used, it was highly subjective.

At the chosen time point, the cells were detached using a cell scraper and then, still within the medium, transferred into 50 ml falcons using a 25 ml pipettor. The cells were pelleted at 500g for 5 minutes at room temperature. After the medium was removed and 1 ml of PBS containing 10% glycerol was added to the cell pellet, it was frozen in a -80°C freezer three times (freeze-thaw cycles) to open up the cells and release the AdV.

After another centrifugation at 500g for 5 minutes at room temperature, the AdV containing PBS/Glycerol mixture was transferred to Eppendorf tubes in 100 µl aliquots and then frozen and stored at -80°C.

To produce AdV quantities sufficient for multiple experiments, it was necessary to use the frozen aliquots for re-infection of 15 to 30 plates with a 15 cm diameter. To estimate the volume needed per 15 cm plate, a 24-well plate was seeded with $2x10^5$ Ad293 cells per well and infected 24h after seeding with a dilution series of the crude AdV. The required amount per 15 cm plate was then calculated from the dilution used in wells showing optimal CPE after 48h. This was followed by re-infection and another round of AdV production, this time infecting Ad293 cells at >90% confluency. If the amount of crude AdV was not sufficient to infect at least 15 plates in the second round, an additional third round was performed.

After enough crude AdV had been produced, e.g. a 3 ml concentrate collected from 30 plates of 15 cm diameter, the absence of E1A expression, and thereby replication deficiency of the AdV in E1A-negative cells, was verified using western blot. Therefore, protein-extracts of AdV-infected HeLa cells were compared to protein-extracts of untreated HeLa cells as negative control and protein-extracts of Ad293 cells as positive control using the "mouse anti-adenovirus type 5 E1A clone M58" antibody (see Figure 2-7). The titer of the crude AdV was then measured using the Adeno-X™ Rapid Titer Kit from CloneTech, performed after manufacturer protocol. After replication deficiency was verified and the infectious units per µl (IFU/µl) were determined, the crude AdV was used for cell culture experiments. For higher transduction efficiencies in cell culture or to reach the purity standards for animal experiments, the AdV could also be further purified before E1A and titer determination were performed. This was done by CsCl density gradient ultra-centrifugation to remove cell debris, followed by dialysis to remove the CsCl salts introduced by the first step.

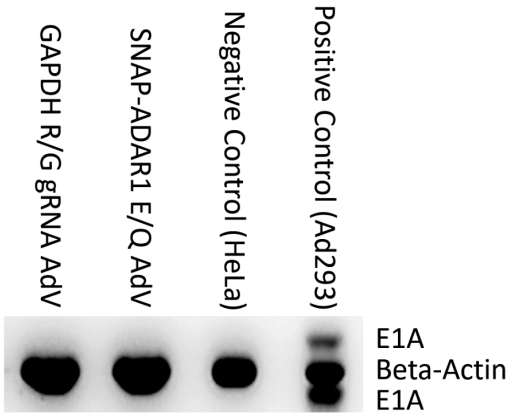


Figure 2-7: Exemplary E1A western blot result.

Protein-extracts from untreated HeLa cells (negative control), untreated Ad293 cell (positive control), as well as HeLa cells treated with two different AdVs were analyzed for E1A expression. The western blot was performed as described in Pub. 4, Material and Methods, Western Blot, except for the skipped Bradford assay. A 1:1000 dilution of the E1A antibody was used. Two E1A bands are located directly above and below the β-actin loading control. The E1A protein species range between 30 and 46 kD, while the mass of β-actin is 42 kD, both according to the technical datasheets of the used antibodies. The GAPDH 3' UTR TAG#1 R/G gRNA of design #3 and the SNAP-ADAR1-E/Q-encoding AdVs were E1A negative and could therefore be used for further experiments.

The gradient for the CsCl density-gradient ultra-centrifugation was layered in a sealable ultra-centrifuge tube, starting with 3 ml of 1.41 g/ml CsCl (27.42g CsCl in 50ml PBS) at the bottom. To prevent mixing the layers, each subsequent layer was added slowly (~10 µl per second) and directly to the surface of the previous one. The next layer consisted of 4.5 ml 1.27 g/ml CsCl (18.47g CsCl in 50ml PBS). After another regular centrifugation of the crude AdV at 500 g for 10 minutes at room temperature, 3 ml of the supernatant were added on top of the 1.27 g/ml CsCl layer. The ultra-centrifuge tube was then filled up to 3 mm below the top of the tube with crude oil. Thereafter, it was balanced on an

analytical balance with other tubes, using crude oil again. After vacuum was applied, the ultra-centrifugation ran for 2h at 32.000 rpm, with an acceleration and deceleration of 1. The tubes were then removed very carefully, as not to disturb the gradient, and fixed in an already prepared ultra-centrifuge tube stand. To improve the visibility of the narrow, white AdV layer, a black sheet of plastic was fixed in another stand behind the ultra-centrifuge tube. To extract the virus, the ultra-centrifuge tube was penetrated below the boundary between the 1.41 g/ml and the 1.27 g/ml CsCl layer, with the needle directed towards the white AdV layer. This was done without disturbing the gradient. To proactively contain a potential spill of the highly infectious solution during penetration of the ultra-centrifuge tube, the area below the stand was covered with paper tissues.

It was important to never remove the needle or syringe during this process. In case the needle became clogged up, another one could be used to penetrate the ultra-centrifuge tube from another angle. Only when the AdV layer was completely extracted and the remaining content of the ultra-centrifuge tube had thereafter been poured into the liquid waste under the S2 bench, the syringe containing the AdV could be removed.

Another ultra-centrifugation step followed immediately, in which a 10-12 ml 1.34 g/ml CsCl layer (22.71g CsCl in 50ml PBS) was overlaid with the just extracted AdV solution, and again balanced using crude oil. The ultra-centrifuge was started as before, but ran overnight at 32.000 rpm this time.

On the next day, after 18-24h, the run was stopped and the AdV layer extracted as previously explained. The extract was inserted into a dialysis cassette using the injection site #1. As each injection site can only be used once, it was important to mark which one had been used in this step. While injecting the virus into the dialysis cassette, it was also important to remove some air from the cassette from time to time, using the same syringe, to prevent a rupture of the membrane due to increasing pressure. After the whole virus solution was transferred, as much air as possible was removed from the cassette to increase the contact surface between the membrane and the AdV solution. The cassette was then transferred into a glass beaker filled with 1L of dialysis buffer and a magnetic stir bar. The dialysis was started and performed in the fridge at 4°C. At 1h, 2h, 3h, 5h, and 7h after the start, the whole dialysis buffer was exchanged. After the 7h step, the dialysis ran overnight without further buffer exchanges. After the dialysis was performed, the purified AdV was extracted from the injection site #2 of the dialysis cassette. Roughly estimated 50-100 µl aliquots of the virus were immediately transferred to already prepared 0.5 ml Eppendorf tubes and frozen at -80°C. Two extra tubes, containing only 10 µl, were prepared for the upcoming E1A western blot and the titer determination. The aliquots were never thawed more than two times to ensure consistent AdV quality between experiments.

2.2.5 Dual-Luciferase reporter assay

The principle behind the assay is explained in Figure 2-8. The substrate solutions from the Promega "Dual-Glo® Luciferase Assay System" and the Promega "dual-luciferase® Reporter Assay System" kits could both be used for the dual-luciferase reporter assay. The firefly luciferase signals of both kits were similar, while the renilla luciferase signals were different for both kits. This did not affect the final results after normalization. Only if very weak signals were expected it was beneficial to use the "dual-luciferase® reporter assay system" kit, as it produced stronger renilla luciferase signals.

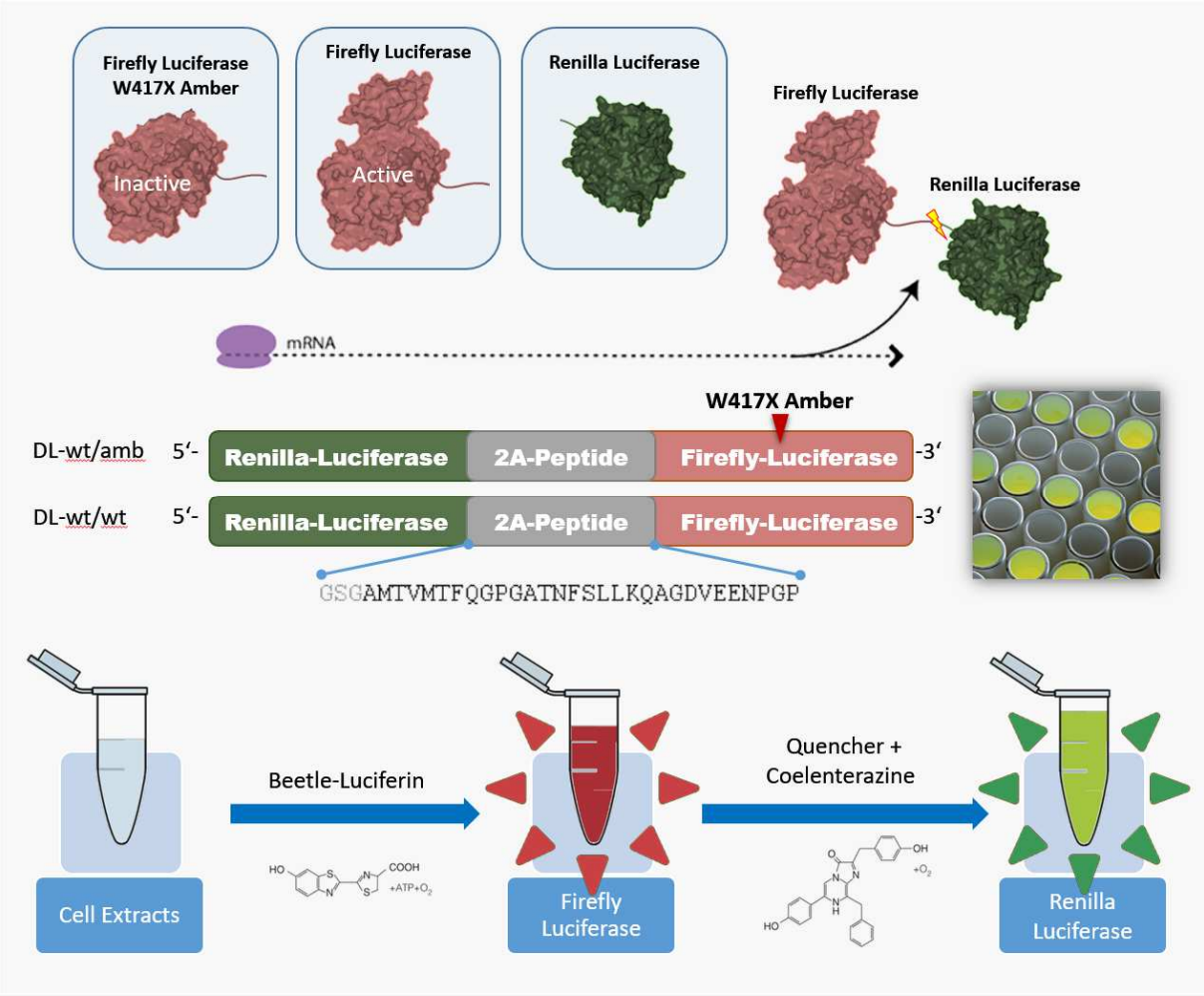


Figure 2-8: The dual-luciferase (DL), site-directed RNA editing reporter system.

The 2A peptide allows the equimolar expression of both luciferases (firefly and renilla) from one bicistronic construct. It causes a ribosomal "skip" in translation, resulting in the separation of both enzymes between the last glycine and proline of the 2A peptide. This allows in turn to normalize the RLU of the firefly by the RLU of the renilla luciferase. The assay is performed in a plate-reader by automated successive addition of the substrates, followed by RLU measurement. First the firefly substrate is added, followed by a firefly quencher and the renilla substrate. By defining the normalized DL wt/amb results as 0% and the DL wt/wt results as 100% the percentage of restored firefly luciferase activity can be calculated.

After the cell culture part of the experiments was performed, the medium was removed and the cells were washed with PBS. After that, the cells were lysed with an appropriate amount of 1xPLB (Passive

lysis buffer), which originated from a 5x concentrate diluted with Millipore water. 30 µl per well were used for 96-well plates, and 100 µl for 24-well plates. After the 1xPLB was added, the lysis was performed at room temperature and 700 rpm for 15 min in a plate shaker. After this, the plate could stay at room temperature for 4h, or could be put in the fridge, or on ice, for later use on the same day. It could also be frozen at -20°C for short-term storage under one month. Long-term storage of lysates was performed at -70°C.

The Spark 10M plate reader was tempered to 22°C, if room temperature differed significantly from the standard values. The substrates were solved in their respective buffer according to the manufacturers protocol. Frozen substrate aliquots were allowed to reach room temperature before the experiment. The injectors were washed for 10x piston strokes with Millipore water. After the washing step, another 5x strokes were performed dry to remove the Millipore water from the system. Then, the substrate-priming was performed, always using injector #1 (left side) for the firefly substrate and injector #2 (right side) for the renilla substrate, as the renilla substrate contained a highly efficient quencher for the firefly luciferase.

The injectors were primed until no air bubbles were visible in the syringes of both injectors, and the program was adjusted to the wells used in the current experiment. For the dual-luciferase experiments, LumiNunc 96-well plates were used. 7 µl of lysate per well were used when HEK-293T or Flp-In T-REx cells had been transfected with plasmid, or when HeLa, A549, SK-N-BE, Huh7, or THP1 had been transduced with AdV. If HeLa cells had been transfected with plasmid, then 30 µl lysate per well were used, due to weak signal intensities. Before the actual experiment, a single read was always performed using a small amount of the positive control lysate to adjust the used OD filter for the experiment at hand. At signals higher than 6 million photons per second, the machine could not detect individual photons anymore, resulting in an “over”-error. In this case OD None was changed to OD1 (90% signal reduction), OD2 (99% signal reduction), or OD3 (99.9% signal reduction). The measurement was always performed 5 seconds after injection for a period of 10 seconds (see Figure 3-9).

If possible, the corner wells of the LumiNunc 96-well plates were not used, due to diminished signal intensity. The blank samples (lysate from untreated cells) were always used in the top row, the negative controls (dual-luciferase wt/amb) in the row below, all treated samples (dual-luciferase wt/amb and a R/G gRNA) in the rows after that and the positive control (dual-luciferase wt/wt) in the bottom row. As the machine started to inject and read at the top row, this way, there was no chance that scattered light from a sample influenced the blank measurements or negative controls, or that a signal from the positive control might influence the signals from the samples.

Due to the strong variance between luciferase assays, at least three, in most cases five, biological replicates were performed simultaneously in the same experiment, and not, as common, in successive experiments. Each biological replicate was then measured in one to three technical replicates.

Finally, the resulting data were processed. The measured blank values (background) were subtracted from the samples and the controls. Then, all firefly values were divided by the corresponding renilla values. The resulting normalized activity of all samples was then either scaled towards the positive control, to get the restored normalized firefly activity in percent, or towards a specific sample, to get the firefly activity as fold change relative to this sample.

3. RESULTS AND DISCUSSION

This chapter describes the progress of the R/G gRNA approach achieved during this thesis. Personally generated data from the listed publications, as well as supporting, unpublished results, are discussed in this section. Work performed by others is indicated in the continuing text.

This thesis is based on the preceding work by Dr. Jacqueline Wettengel, outlined in section 1.2.2, and was intended to answer open questions that are crucial to the further development of the R/G gRNA SDRE system into a therapeutic strategy. When this project started, there was no proof if complex protein functions, e.g. in essential signaling cascades, could be restored by SDRE. The important question, if endogenous mRNAs could be affected, too, had not been solved either. The answers to these questions would determine, if a therapeutic use of SDRE would be feasible (section 3.1 and 3.2).

The liberation of our R/G gRNA system from its dependence on overexpressed ADAR2 was similarly important. To accomplish this, it was first necessary to identify a panel of human cell lines positive for endogenous ADAR expression (section 3.4). Because many human cell lines are hard to transfect with plasmids, viral vectors for efficient gene transfer of encoded R/G gRNAs into these cell lines had to be established and characterized (section 3.7). The question had to be solved, if other ADAR enzymes, besides the previously used ADAR2, principally could be recruited as well. Endogenous ADAR2 is prominently expressed in the CNS. The ability to recruit the constitutively and ubiquitously expressed ADAR1 p110, or the IFN-α inducible ADAR1 p150, promised to extend the spectrum of accessible tissues (section 3.5). Consequently, it was also necessary to determine, which endogenous ADAR enzyme would be primarily recruited by our encodable R/G gRNAs (section 3.10). As in the early phase of this project only very low editing yields could be expected, a reliable detection method for minimal amounts of SDRE had to be established (section 3.6), which could then be used as baseline for incremental optimizations to the R/G gRNA architecture (section 3.12).

3.1 Restoration of the PINK1-Parkin signaling cascade by site-directed RNA editing

At the beginning of this thesis, the R/G gRNA system allowed the conversion of premature amber stop-codons (T**A**G) into tryptophan codons (T**G**G), and could thus restore the fluorescence of an eGFP reporter containing a nonsense mutation (Pub. 2, Figure 2B). It was important to demonstrate that it could not only restore the protein function of artificial fluorescence reporters, but also repair key components in cellular signaling pathways connected to human disease and especially that the correction would restore not only the target protein, but the whole signaling cascade in question.

Together with Dr. Sven Geisler (Kahle Group), our collaboration partner from the Hertie-Institute for clinical brain research, we modified an assay, which allowed to model the pathological consequences

of PINK1 loss-of-function mutations in HeLa cells for the use in conjunction with our R/G gRNA approach. The detailed assay procedure is described in the supplement materials of publication 2.

The kinase PINK1 is a key component of the PINK1-Parkin signaling pathway, which functions like a mitochondrial quality control system. If mitochondria are damaged and lose their proton gradient, this could lead to the release of reactive oxygen species (ROS) and harm the entire cell. Therefore, mitochondria with failing proton gradient are removed by mitophagy, a mitochondria-specific type of autophagy. The selection for mitophagy is based on the accumulation of PINK1, followed by Parkin, on the surface of these damaged mitochondria and results in their translocation towards the nucleus (perinuclear clustering), followed by mitophagy. The mechanism is described in more detail in Figure 3-1.

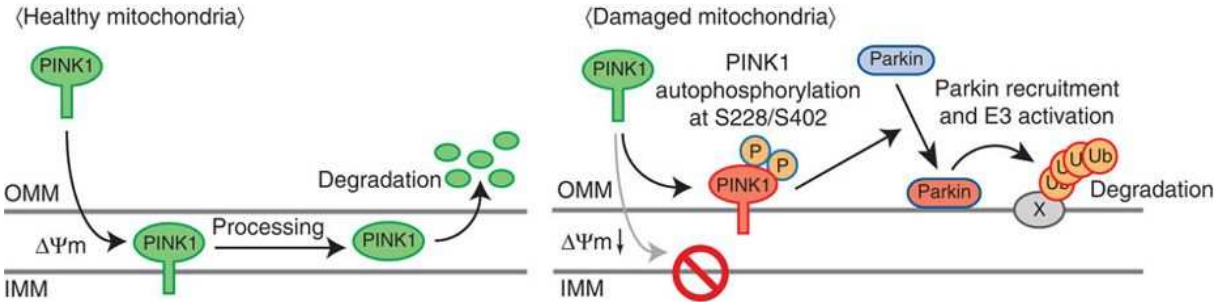


Figure 3-1: A model for Parkin recruitment to damaged mitochondria.

The kinase PINK1 (PTEN-induced kinase 1) is an essential component of a signaling pathway regulating mitochondrial quality control. Healthy mitochondria import PINK1 to their inner mitochondrial membrane (IMM) in a proton-gradient-dependent manner. The release of PINK1 into the cytosol is facilitated by two subsequent IMM-protease cleavages (PARL-Protease, MPP), and results in degradation of PINK1 through the N-end rule pathway [132]. Damaged mitochondria, without a proton gradient, can recruit PINK1 only to their outer mitochondrial membrane (OMM). There, it is stabilized and becomes fully active after homodimerization and autophosphorylation [133]. Active PINK1 phosphorylates the E3 ubiquitin ligase Parkin, leading to translocation of Parkin to the OMM and subsequent ubiquitination of mitochondrial surface proteins. This causes microtubule dependent transport of damaged mitochondria to the endoplasmic reticulum, which is located around the nucleus (perinuclear clustering), followed by autophagosome formation and mitophagy [204, 205]. This figure was adopted from reference [206].

This behaviour was utilized to establish an assay which allowed to verify the functional rescue of mutant PINK1 by site-directed RNA Editing using our R/G gRNA system. To achieve this, it was first necessary to remove wild-type PINK1 from the used HeLa cells.

Preliminary experiments with the goal to knock-down PINK1 in wild-type HeLa cells, using specific siRNAs, showed that even small amounts of functional PINK1 protein could rescue mitophagy (data not shown). As a collaborative effort, a PINK1 knock-out HeLa cell line was generated, using CRISPR-Cas9 (Pub. 2, Figure S20). It was tested for its reaction to mitochondrial depolarisation using the proton gradient uncoupling agent CCCP. In healthy cells the treatment with CCCP induces perinuclear clustering of all mitochondria followed by their mitophagy ~24h later (Figure 3-2). This was not true for the PINK1 KO HeLa cells.

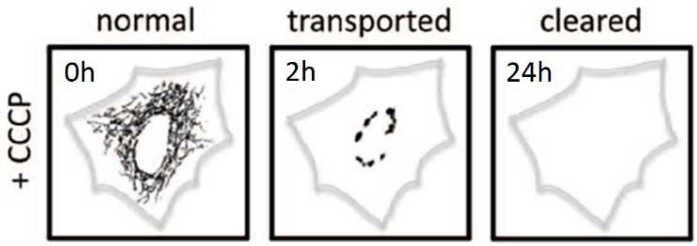


Figure 3-2: Effect of CCCP treatment on mitochondria.

Without treatment mitochondria are distributed all over the cell. Two hours after CCCP treatment, mitochondria are perinuclear clustered. Within 24h, all mitochondria are cleared by mitophagy. This figure was adopted from reference [205].

To quantify perinuclear clustering, it was necessary to track mitochondria, but only in cells which were positive for all components of the PINK1-Parkin signaling pathway, as well as all components of the R/G gRNA editing system. To achieve the former, the cells were transfected with Parkin-GFP, immunostained for PINK1, and their mitochondria visualized by MitoTracker Red CMXROS. The experiments showed that the PINK1 KO HeLa cells were not able to react to mitochondrial depolarisation (Pub. 2, Figure 4A panel A and Figure S23B), while HeLa wt cells could (Pub. 2, Figure 4A panel B and Figure S23A). This allowed the screening of several different PINK1 mutations, transfected as plasmids, for their activity, or the lack thereof, in the induction of perinuclear clustering and mitophagy. The PINK1 W437Stop mutation was finally selected. It results in a non-functional PINK1 protein with truncated kinase domain (Figure 3-3) and is linked to an hereditary, monogenetic form of Parkinson's disease [207]. It was assumed that all three stop codons at position 437 were disease-relevant, and therefore the easiest to edit of the three, the amber codon, was chosen.

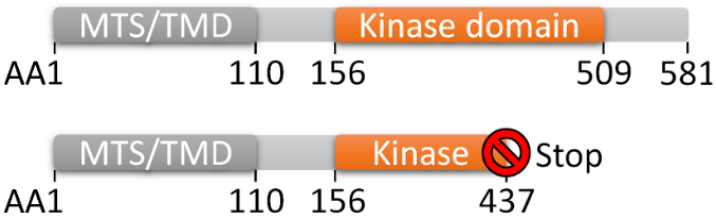


Figure 3-3: The PINK1 W437Stop mutation truncates the kinase domain, which renders it inactive.

The mitochondrial targeting sequence (MTS), as well as the transmembrane domain (TMD) remain intact. The presence of the truncated PINK1 after overexpression could be verified by Sven Geisler using western blot (Pub. 2, Figure S21).

In the next step, it was necessary to detect the presence of our ADAR2 enzyme and the R/G gRNA intended to correct the PINK1 W437Stop mutation. Overexpressed ADAR2 could be detected by simple immunostaining (Pub. 2, Materials and Methods, Mitophagy assay). The R/G gRNAs should be detected indirectly by utilizing a plasmid containing ADAR2 and five copies of the R/G gRNA construct (Pub. 2,

Figure 3C). Multiple R/G gRNA copies were assumed to be necessary due to the fact that a ~1:5 stoichiometry of ADAR2 to R/G gRNA had produced the most efficient editing in previous cell culture experiments (Pub. 2, Figure S12).

Similar constructs, containing ADAR2 and four copies of the R/G gRNA, had been evaluated before, but their comparibly large size (9061 bp) led to low transfection efficiency and stability. Thus, it was not possible to use them for reliable and consistent editing experiments (section 4.2.2.6.1 in [148]). To reduce the total plasmid size, a much smaller novel backbone was engineered. It contained all necessary components of a high-copy plasmid and was named pEdit. It was assembled in a stepwise process out of PCR fragments acquired from several commonly known expression plasmids (pcDNA3.1, pUC57, pSilencer2.1-U6 hygro). The large spacer DNA sequences commonly present within the stock backbones were omitted in this process. The basic pEdit contained human ADAR2, extended with a His-Tag and *in-silico* optimized Gibson linker elements with a central EcoRV cutting site, to allow the simple cloning of pEdit vectors with different R/G gRNAs. As a result, the final pEdit1.2 backbone was only 5043 bp in size, already containing ADAR2. The additional introduction of five copies of the PINK1 W437Stop R/G gRNA via Gibson cloning extended its size to 7463 bp. This was even smaller than our regular pcDNA3.1 ADAR2 expression vector, which was 7478 bp in size. Due to the logarithmic correlation between plasmid size and transfection efficiency, the new backbone could compensate the issues of previously assessed constructs [208].

The R/G gRNA used within the pEdit construct was similar to the later discussed design #2 (Figure 3-23, Figure 3-24), but contained an additional BoxB motif at its 3' terminus (Pub. 2, Figure 2C). While the BoxB motif had been initially intended to improve stability, it had been shown by Dr. Wettengel that it had no effect on the editing yield (section 4.2.2.6 and Figure 4-76 in [148]). Because the simplified cloning procedure with oligonucleotide hybrids as inserts (Figure 2-3C) was not available at this time, and the effect of the BoxB motif was neutral, the cloning procedure shown in Figure 2-3B was used for convenience. The procedure resulted in the presence of the 3' terminal BoxB hairpin. All other R/G gRNA designs discussed in this thesis were generated using the new cloning procedure (Figure 2-3C) and did not contain a 3' terminal BoxB motif.

Preliminary experiments in HEK-293T cells showed that the strong overexpression of ADAR2 allowed unspecific PINK1 W437X amber repair in the absence of R/G gRNAs. This could be circumvented by reduction of the ADAR2 plasmid amount (and the corresponding pEdit amount in the editing experiment) from 300 ng to 100 ng per well (see Figure 3-4). All subsequently discussed results therefore stem from experiments performed with 100 ng ADAR2 or pEdit plasmid.

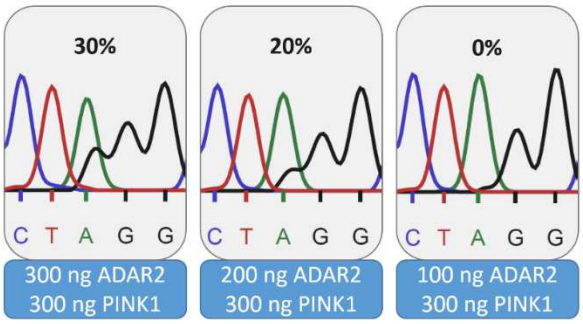


Figure 3-4: Three different PINK1 W437X amber, ADAR2 co-transfection settings lacking the R/G gRNA compared in HEK-293T cells.

The unspecific editing happened probably due to an internal secondary structure within the PINK1 mRNA, as this kind of independent editing was unprecedented in numerous, previous editing experiments. Fortunately, a very strong overexpression of ADAR2 was necessary for this to happen, and the reduction of the ADAR2 plasmid amount prevented the problem. $2x10^5$ HEK-293T cells were seeded in 24-well scale. After 24h they were transfected with the indicated amounts of ADAR2 and PINK1 W437X amber plasmid, at 70-90% confluence, with a 1:3 ratio of Lipofectamine 2000. The cells were harvested 48h post transfection. This was followed by RNA isolation, DNase-I digestion, RT-PCR, and Sanger sequencing.

After the pEdit1.2 backbone was completed and its amount adjusted, it allowed the indirect detection of the mentioned R/G gRNA through fluorescent microscopy of immunostained ADAR2. This allowed in turn to calculate the exact percentage of HeLa PINK1 KO cells positive for perinuclear clustering after treatment with the complete R/G gRNA system. The transfection bias could be removed from the data by excluding all cells negative for one of the necessary components.

The final datasets revealed visually (Pub. 2, Figure 4A panel C and Figure S23C) and statistically (Pub. 2, Figure 4B) that the presence of all R/G system components allowed the rescue of the perinuclear clustering phenotype nearly to wild-type PINK1 levels: 84.7%±7.9% restored by the R/G editing system in HeLa PINK1 KO cells vs. 89.2%±3.2% in HeLa wt cells. The restored perinuclear clustering also resulted in mitophagy 24h after CCCP treatment, as described in the literature for healthy cells [205] (Pub. 2, Figure 4C and Figure S23C). If one component (ADAR2, R/G gRNA, PINK W437Stop substrate) was missing, perinuclear clustering was not efficiently restored (Pub. 2, Figure 4B, Figure 4A, panel D, E, F, and Figure S23D, E, and F). While still extremely low, the percentages of perinuclear clustering detected in the controls that excluded ADAR2 (8.2%±5.9%), or the R/G gRNA (10.8%±2.7%), were minimally higher than in those excluding the PINK W437Stop substrate (2.5%±1.5%). The minor increase for the control missing the R/G gRNA could be explained by ADAR2 still being overexpressed, even though the setting was limited to 100 ng of plasmid. Due to the high sensitivity of the assay, it was probably detecting residual gRNA-independent editing caused by ADAR2 overexpression (see Figure 3-4). The small percentage of perinuclear clustering found in the control missing ADAR2, but containing the R/G gRNA, was much more interesting. It might have been induced by the very weak recruitment of an endogenous ADAR. ADAR1 p110 was later found to be expressed in HeLa cells (see Figure 3-6)

and could be recruited to a very limited degree by early R/G gRNA designs very similar to the ones used in these experiments (See R/G gRNA design #4 in Figure 3-23 and Figure 3-24).

While over 30% editing of the PINK1 W437X amber mRNA could be achieved in HEK-293T cells under co-overexpression with ADAR2 (Pub. 2, Figure 3D), only ~12% could be measured in HeLa PINK1 KO cells when using the pEdit plasmid (Pub. 2, Figure 4B, sequencing read C). This low editing yield resulted probably from measuring a mixture of cells, including only partially-transfected ones. It is likely that cells, showing perinuclear clustering would also exhibit higher levels of correction, if measured independently, e.g. after FACS.

In conclusion, under the indicated settings, the R/G system was able to repair mutant PINK1, a key component in the PINK1-Parkin mitophagy signaling pathway connected to Parkinson's disease. It was the first proof of principle for functional protein rescue in a cellular signaling cascade by site-directed RNA editing using human ADAR2.

3.2 Successful site-directed RNA editing of endogenous mRNAs

Because so far only overexpressed mRNAs had been targeted by SDRE, it was important to assess if the R/G gRNA system could also be used to edit endogenous mRNAs. The housekeeping genes ACTB, GAPDH, GPI, GUSB, RAB7A, and VCP were selected as targets for a cooperative set of editing experiments with Dr. Wettengel. The targets were selected from a list of housekeeping genes to assure that all of them would actually be expressed in the used HEK-293T cells. After RT-PCR settings had been established for each of them, 13 different 5'-UAG-3' triplets within the 3' UTRs of these 6 transcripts were edited, using transiently overexpressed ADAR2. The target sites for this proof of principle study were selected with three premises in mind:

1. The ADAR2 deaminase domain preference makes 5'-UAG-3' triplets favourable editing targets (section 1.1.2.2).
2. Base substitution in an UTR cannot cause recoding events and is therefore unlikely to be functionally relevant or toxic for the cells.
3. The selection of the 3' UTR prevents competition of R/G gRNA binding with ribosomal activity.

Without further optimization of the experimental settings, it was possible to edit 12 out of 13 selected target sites in the first attempt, with editing yields ranging between 9% and 32% (Pub. 2, Figure 3B and Figure S19A). Due to the unspecific effects caused by the excessive ADAR2 overexpression in the preliminary PINK1 editing experiments (300 ng vs. 100 ng ADAR2 plasmid), it was important to know how drastically reduced ADAR2 expression levels would affect editing. Therefore, a Flp-In T-REx cell line was engineered that contained only a single genomic copy of ADAR2, stably integrated at a genomic FRT-site and expressed from a doxycycline inducible CMV promoter. The ADAR2 expression level of the

new cell line was analysed by my co-worker Dr. Vogel, using qPCR, and was found to be 20-fold lower compared to transient overexpression from a plasmid (Figure S16E).

Then, the editing experiments targeting endogenous mRNAs were repeated under ectopic expression of the guideRNA alone, using the ADAR2-Flp cells. Again, 12 out of 13 selected target sites could be edited. Compared to the earlier ADAR2 overexpression experiment, the editing yield of ACTB and GAPDH decreased, while RAB7A increased. The editing yields of GPI, GUSB and VCP did not change. Overall, the editing ranged between 9% and 38% (Pub. 2, Figure 3B and Figure S19B). The used R/G gRNA was of design #2 (see Figure 3-24).

With these results, a major question of the SDRE field was solved. The R/G system was able to site-specifically edit the 3' UTRs of endogenous mRNAs in HEK-293T cells and ADAR2-Flp cells, even under 20-fold lower ADAR2 expression in the latter.

3.3 The reduced ADAR2 overexpression in stable cell lines decreases off-target editing and increases on-target editing within the eGFP W58X amber reporter mRNA

A lower, but more homogeneous, expression of ADAR2 in Flp-In T-REx cells increased the editing yield, compared to transient overexpression, for the editing of the eGFP W58X amber reporter. Probably because fewer plasmids had to be co-transfected. Editing yields of 67% were reached using 1300 ng of co-transfected R/G gRNA, compared to only 50% when performing the same experiment under transient ADAR2 overexpression (Pub. 2, Figure 2E, Figure 3A panel A, and Figure S17). When reducing the amount of R/G gRNA plasmid, the editing yield was still better (800 ng) or at least equal (400 ng) compared to transient overexpression (Pub. 2, Figure 3A panel C, D and Figure S17). Additionally, no off-target editing could be observed under these conditions in a search through the entire ORF of the eGFP reporter gene, including the known off-target site adenosine 381 (Pub. 2, Figure S18). This result is consistent with the earlier observation that a reduction of the ADAR2 plasmid amount lowers off-target editing when using transient overexpression settings. The beneficial effect of reduced ADAR2 overexpression on off-target editing was a good sign for the intended use of endogenous ADARs.

3.4 Detecting the endogenous expression of ADAR family proteins

To liberate the R/G gRNA system from its dependence on overexpressed ADAR2, it was necessary to screen for cell lines positive for endogenous ADAR expression. Cell lines deriving from tissues which might represent interesting targets for a future therapeutic approach were selected.

These included the neuroblastoma cell lines SK-N-BE and SH-SY5Y, the glioblastoma cell line U-87 MG, the liver cell lines Huh7 and Hep G2, the PBMC line THP-1, and the lung alveolar cell line A549. Due to their common use in basic research, the human cervical cell line HeLa was also analysed. In addition,

the murine liver cell line Hepa 1-6 and the murine fibroblast cell line MEF were screened for murine ADAR2 as a step towards potential future animal experiments.

To evaluate the ADAR2 expression on protein level, it was necessary to extend the spectrum of methods available in our lab with western blotting. The established protocol can be found in the material and methods section of publication 4. Unfortunately, ADAR2 expression could not be detected in any of the mentioned cell lines even when excessive amounts of total protein, up to 10 µg per lane, were loaded onto the polyacrylamide gels (see Figure 3-5). In 2017, at the 1st symposium on nucleic acid modifications in Mainz, it was stated in a scientific discussion that ADAR2 expression is silenced shortly after extraction of cells from their host organism. Although there are some rare reports of endogenous ADAR2 detection in HeLa [209] and PMA-differentiated THP-1 cells [210], this statement explained why it was not possible to detect ADAR2 in any of the tested cell lines in our lab.

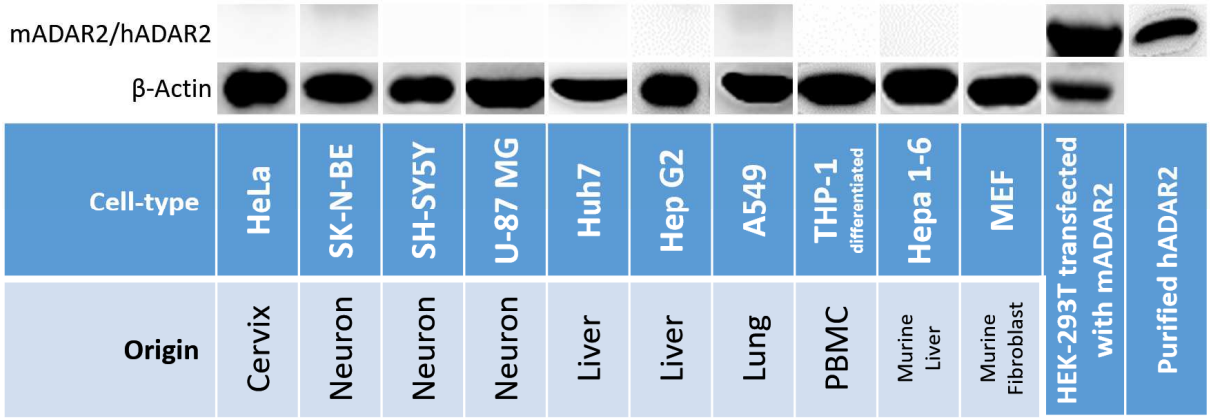


Figure 3-5: Western Blot screening of several cell lines for human or murine ADAR2 expression.

The mADAR2 transfected HEK-293T positive control and the purified hADAR2 control showed that the antibody is able to detect both murine and human ADAR2. However, ADAR2 expression could not be detected in any of the assessed cell lines. 10 µg of protein-extract were loaded per lane for HeLa, SK-N-BE, SH-SY5Y, U-87 MG, HepG2, Huh7, and A549 cells. 6 µg for Hepa 1-6, MEF, and mADAR2-transfected HEK-293T. 4 µg for THP-1 cells. 1 ng purified hADAR2 was loaded as positive control. The figure was assembled from five individual western blots. The western-blots were performed as described in Pub. 4. ADAR2 has a mass of 90 kD according to the technical datasheet of the used antibody.

Another round of screening, this time for ADAR1 p110 and ADAR1 p150 (section 1.1.1.1), was successful. The constitutively expressed p110 isoform of ADAR1 could be readily detected in all human cell lines, which had been tested for ADAR2 before. The Interferon-inducible p150 isoform, on the other hand, could only be positively detected in HeLa, SK-N-BE, Huh7, A549, and PMA-differentiated THP-1 cells (see Figure 3-6).

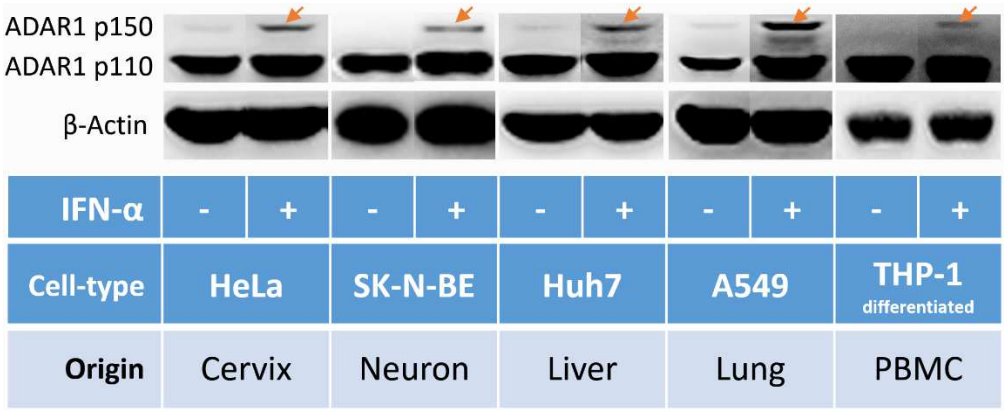


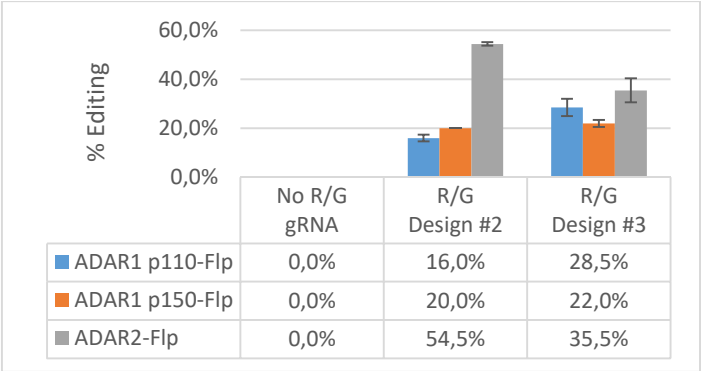
Figure 3-6: Western Blot screening of several cell lines for human ADAR1 expression.

All analyzed celltypes express ADAR1 p110. ADAR1 p150 expression can also be induced by IFNα treatment, indicated by the orange arrow. The IFNα-treated cells were induced for 24h (HeLa, A549, THP-1, SK-N-BE) or 48h (Huh7), using 1000 units (A549, THP-1, SK-N-BE) or 5000 units (HeLa, Huh7) IFN-α in 1 ml of medium. All cells were seeded to be 70-90% confluent at the time of transfection/transduction, using 12-well-scale cell culture plates. 20 µg (HeLa), 15 µg (A549), 14 µg (Huh7), 10 µg (SK-N-BE) or 4 µg (THP-1) of protein-extract were loaded per lane. The figure was assembled from five individual western blots. The western-blots were performed as described in Pub. 4. Both ADAR1 isoforms have a mass corresponding to their name (110 kD and 150 kD), according to the technical datasheets of the used antibodies.

3.5 Demonstrating the principle recruitability of both human ADAR1 isoforms by the R/G gRNA system

After the screening experiments for endogenous ADARs had been completed, ADAR1 became the prime target for the planned endogenous recruitment experiments. The R/G gRNA system, however, had been established based on a classical ADAR2 substrate [103]. It was therefore mandatory to assess, if our R/G gRNA architecture would be able to harness ADAR1 as well.

By restoring a luciferase reporter with W417X amber nonsense mutation in both ADAR1 p110- and p150-Flp cells, the principle recruitability of ADAR1 isoforms could be demonstrated. However, the yield was lower than in comparable ADAR2-Flp cell experiments (see Figure 3-7 R/G design #2). At the same time, a bachelor student whom I supervised, Madeleine Heep, showed that this was also true for the editing of our traditional eGFP W58X amber target (Pub. 2, Figure 1C). Incremental design optimizations, which will be discussed in more detail below (see Figure 3-23 and Figure 3-24), allowed to nearly double the editing yield when harnessing human ADAR1 p110 in Flp-In T-REx cells, but at the same time significantly reduced editing by ADAR2 (see Figure 3-7 R/G design #3).



	No R/G gRNA	R/G Design #2	R/G Design #3
ADAR1 p110-Flp	0,0%	16,0%	28,5%
ADAR1 p150-Flp	0,0%	20,0%	22,0%
ADAR2-Flp	0,0%	54,5%	35,5%

Figure 3-7: Site-directed RNA editing using R/G gRNAs and human ADAR1.

R/G gRNAs are able to harness human ADAR1, too, although this results in a reduced editing yield, compared to ADAR2. Incremental design optimizations, discussed below (Figure 3-23, Figure 3-24), allowed to nearly double the editing yield achieved with ADAR1 p110. The editing yield was quantified using Sanger sequencing reads from n=2 experiments targeting firefly luciferase W417X amber. $2.5x10^5$ ADAR1 p110 or p150 Flp-In T-REx and $3x10^5$ ADAR2 cells per well were seeded in 24-well scale on PDL coated plates. 24h post seeding, at 70-90% confluence, the cells were transfected with 300 ng dual-luciferase reporter plasmid and 1300 ng R/G gRNA plasmid using Lipofectamine 2000 in a 1:3 ratio. The cells were harvested 72h post transfection. This was followed by RNA isolation, using the Quiagen RNeasy Mini Kit, DNase-I digestion, RT-PCR, and Sanger sequencing.

The observation that ADAR1 isoforms can actually be recruited, in conjunction with the solid endogenous expression of ADAR1 p110 in HeLa cells, supported the earlier hypothesis that the unanticipated low frequency of perinuclear clustering witnessed in the ADAR2-lacking negative control of the PINK-Parkin mitophagy assay, could have indeed been caused be the recruitment of endogenous ADAR1. The corresponding sequencing reads acquired in those early experiments, however, did not show any editing peaks (Pub. 2, Figure 4B, sequencing read D). Therefore, the assumed editing must have been below the detection limit of Sanger sequencing (<5-10% depending on read quality), while the PINK1-Parkin mitophagy assay was sensitive enough to detect it in some rare occasions (Pub. 2, Figure 4B, bar diagram D with 2h CCCP). This hypothesis had to be assessed with a simpler and more reliable SDRE detection method.

3.6 The dual-luciferase reporter system - A highly sensitive method for the detection of RNA repair by site-directed RNA editing

Based on previous experiments, it was expected that the furthest developed R/G gRNA architecture at that time (design #3) could only achieve SDRE yields below the Sanger sequencing detection limit, when restricted to endogenous ADAR1. To confirm the existence of these anticipated editing events, it was necessary to establish a new reporter system with the ability to detect minimal amounts of site-directed RNA editing with high certainty. This system should allow to assess novel R/G gRNA designs at physiological ADAR1 expression levels to ultimately enable efficient SDRE, independent from over-expressed editases.

A mutant W417X amber firefly luciferase, which had already been established as editing target in our lab, was utilized to create a novel reporter system, custom-tailored for the low-level detection and quantification of site-directed RNA editing. The new system should also bypass the biggest issue of the original firefly reporter, its high signal variability. To achieve this, the signal of the restored firefly luciferase should be normalized through the signal of a reference luciferase. In general, this is performed by co-transfection of both reporters, which makes the readout vulnerable to pipetting errors. To achieve higher accuracy, two bicistronic constructs were cloned, each combining a renilla luciferase, a 2A-peptide, and either a mutant W417X amber or a wild-type firefly luciferase (see Figure 2-8). The 2A peptide facilitated the efficient separation of the bicistronic construct by performing a ribosomal "skip" in translation (for details regarding 2A-peptides see [211-214]). The 19 aa long 2A-peptide from the porcine teschovirus-1, with an additional N-terminal GSG-motif (Glycine-Serine-Glycine), had a high cleavage efficiency (>90%) in human cell lines [214] and was thus used for the dual-luciferase (DL) reporter system (Figure 3-8).

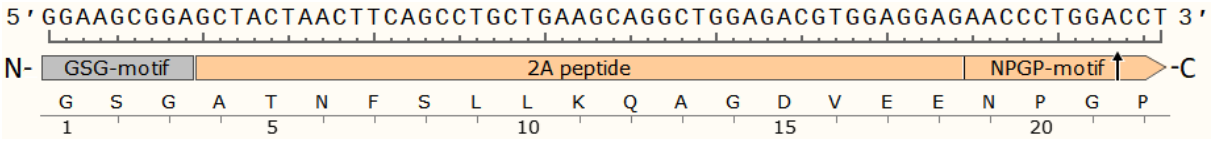


Figure 3-8: 2A peptide with GSG-motif, based on the porcine teschovirus-1.

The 2A motif is in total 22 amino acids long. It consists of the 3 aa long N-terminal GSG-motif and the 19 aa long original 2A peptide, that derives originally from the porcine teschovirus-1. The C-terminal NPGP-motif, containing the translational “skip”-site (indicated by an arrow), is a conserved feature among variants of 2A peptides [215].

The resulting constructs allowed to express both luciferases in equimolar amounts more accurately than co-transfection or internal ribosomal entry sites (IRES) would allow [216]. The percentage of restored firefly luciferase activity could be calculated by defining the normalized results of the dual-luciferase wild-type/amber construct (DL wt/amb) as 0% and the results of the dual-luciferase wild-type/wild-type construct (DL wt/wt) as 100%. Before the new reporter system could be used, it was necessary to determine optimal settings for signal acquisition and to characterize its performance in cell culture experiments. Our Tecan Spark 10M plate reader is equipped with two attenuation filters, which can be set independently, thus reducing the signal either not at all (OD-None), by 90% (OD-1), 99% (OD-2), or both together by 99.9% (OD-3). These filters, as well as different luciferase lysate dilutions, were analysed in kinetic measurements to generate first settings for the use of this new reporter system. The results showed that dilutions up to 1:100 could still be measured with high accuracy, indicated by the fact that RLU (relative light units) values showed a linear dynamic range. Undiluted cell lysate resulted in ~4 million RLU, while 1:5 dilution resulted in ~800.000 RLU, 1:10 in ~400.000 RLU and 1:100 in ~40.000 RLU (Figure 3-9A, Firefly, OD-None curve values 5 seconds after injection, indicated by red line).

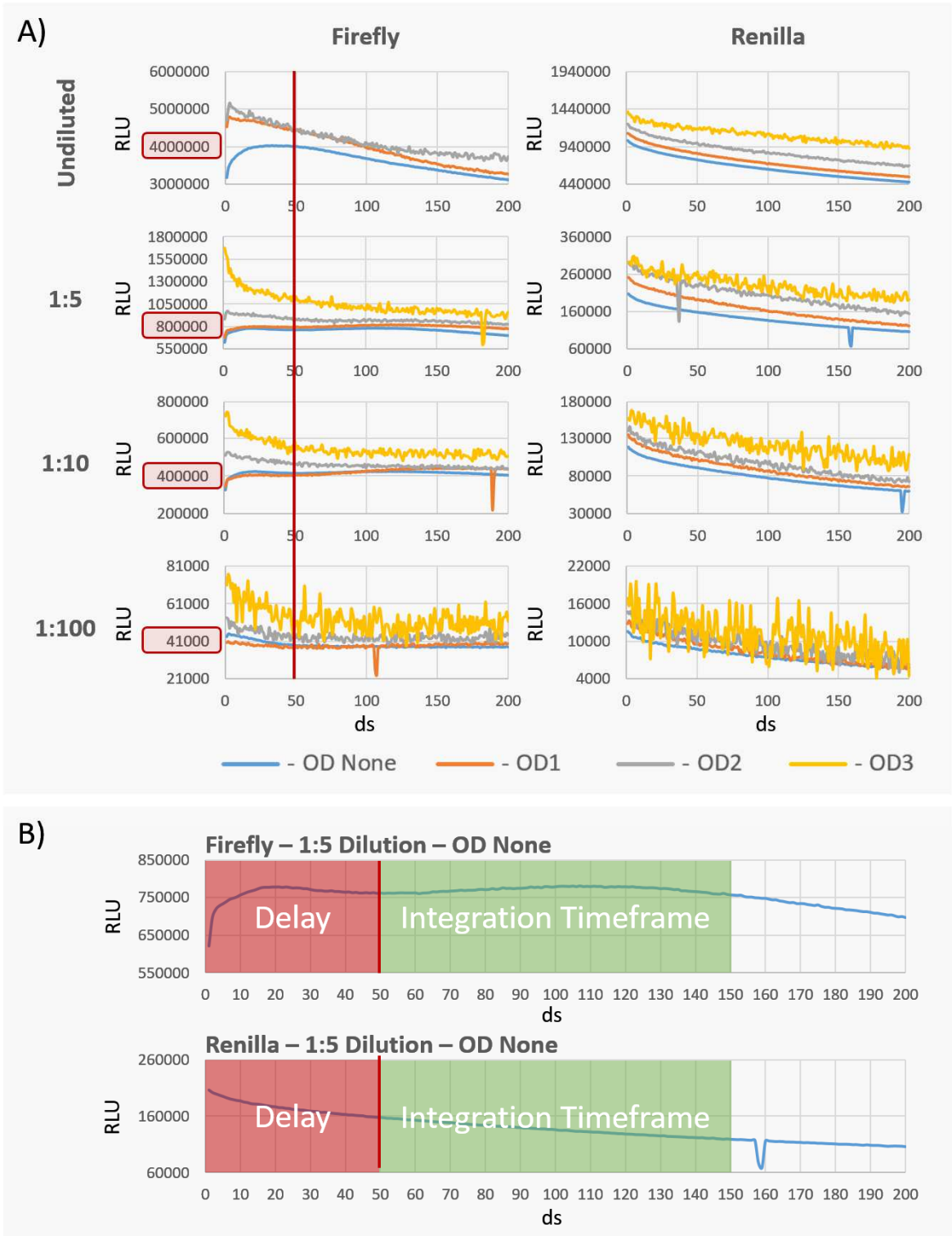


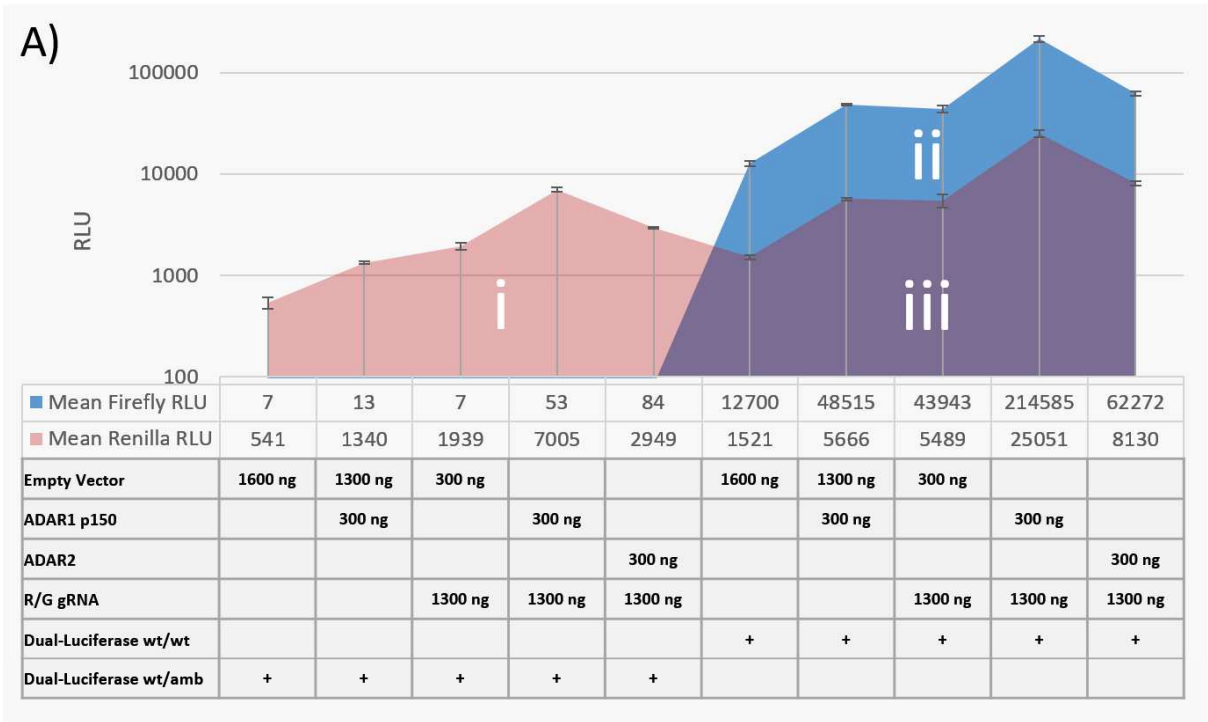
Figure 3-9: The impact of cell lysate dilutions and OD attenuation filter settings on the dual-luciferase assay.

(A) Kinetic RLU measurement over 20 seconds, collecting one data point every deci-second (ds), with different attenuation settings (OD-None, -1, -2, -3), as well as different cell lysate dilutions (undiluted, 1:5, 1:10, 1:100). The red line at 50 ds (5 seconds) and the corresponding, indicated RLU values show that all tested dilutions can be performed with linear dynamic range. (B) Determination of the optimal delay and integration timeframe after substrate injection. For panel A and B $2x10^5$ HEK-293T cells were transfected with 300 ng of the DL wt/wt construct in 24-well scale. 72h after transfection, dual-luciferase assay data were collected, using kinetic measurements. Seven µl of cell lysate (100 µl 1xPLB per well had been used to harvest each well of the 24-well plate) with 35 µl of each substrate reagent were used per well in a 96-well LumiNunc plate.

As the plate reader automatically extrapolated the RLU-values based on the applied OD-filter(s), the noise in the resulting curves increased with attenuation of the signal. This was indicated by increased fluctuation of the measured curves, when comparing OD-None with OD-1, -2, and -3. Therefore, the selection of adequate OD-settings was very important to generate robust data, and needed to be adjusted relative to the signal intensities of individual experiments.

The kinetic RLU curves of the undiluted firefly measurements, as well as all measurements of the different renilla dilutions, seemed to decline over time. The firefly RLU curve could be stabilized by using a 1:5 dilution and OD-None (Figure 3-9B, Firefly), resulting in a plateau after 5 seconds delay post substrate injection. The renilla RLU on the other hand kept declining, independent of the applied settings (Figure 3-9B, Renilla). While the firefly RLU decline in the undiluted cell lysate might have been caused by substrate depletion, the constant decline of the renilla RLU was probably caused by accumulation of inhibitory side products of the reaction. Five seconds delay after substrate injection, followed by 10 seconds of signal integration time, were deduced as optimal time settings from the results in Figure 3-9B. Due to the variability of the luciferase signal intensity between experiments, it seemed best to perform a single kinetic measurement of the positive control previous to the actual dual-luciferase assay. The result could then be used to determine the individual optimal settings for signal acquisition, regarding cell lysate dilution and attenuation.

The next step was to characterize the system regarding its capability to normalize the resulting data. Our lab had already experienced that the RLUs of luciferase experiments are drastically, and differently, influenced by the type of co-transfected plasmid. This enhancement effect, which is a current object of research, resulted in more than 15-fold differences between RLUs, merely by changing a co-transfected plasmid to another one (see Figure 3-10A). Therefore, it was of particular importance to prove that the new reporter system would be able to compensate for this. Figure 3-10B shows that the RLUs of renilla wild-type and firefly wild-type enzymes, expressed from the same bicistronic DL wt/wt plasmid, correlate perfectly into a regression curve (R^2-value of 0.99). While the enhancement effect obviously still applied to the resulting RLUs, it did so consistently for both enzymes. Therefore, the normalization of the firefly RLUs, with the corresponding renilla RLUs, allowed to eliminate this bias from all calculations performed thereafter. Although the values were correlating, it was obvious that the firefly RLUs were ~8.6-fold higher than the renilla RLUs (see regression curve function in Figure 3-10B). This was probably caused by differential levels of renilla and firefly enzyme activity. In addition, the photon production could be affected by the used substrate solutions (commercial kits) and cell types.



Mean Firefly RLU	7	13	7	53	84	12700	48515	43943	214585	62272
Mean Renilla RLU	541	1340	1939	7005	2949	1521	5666	5489	25051	8130
Empty Vector	**1600 ng**	**1300 ng**	**300 ng**			**1600 ng**	**1300 ng**	**300 ng**		
ADAR1 p150		**300 ng**		**300 ng**			**300 ng**		**300 ng**	
ADAR2					**300 ng**					**300 ng**
R/G gRNA			**1300 ng**	**1300 ng**	**1300 ng**			**1300 ng**	**1300 ng**	**1300 ng**
Dual-Luciferase wt/wt						+	+	+	+	+
Dual-Luciferase wt/amb	+	+	+	+	+					

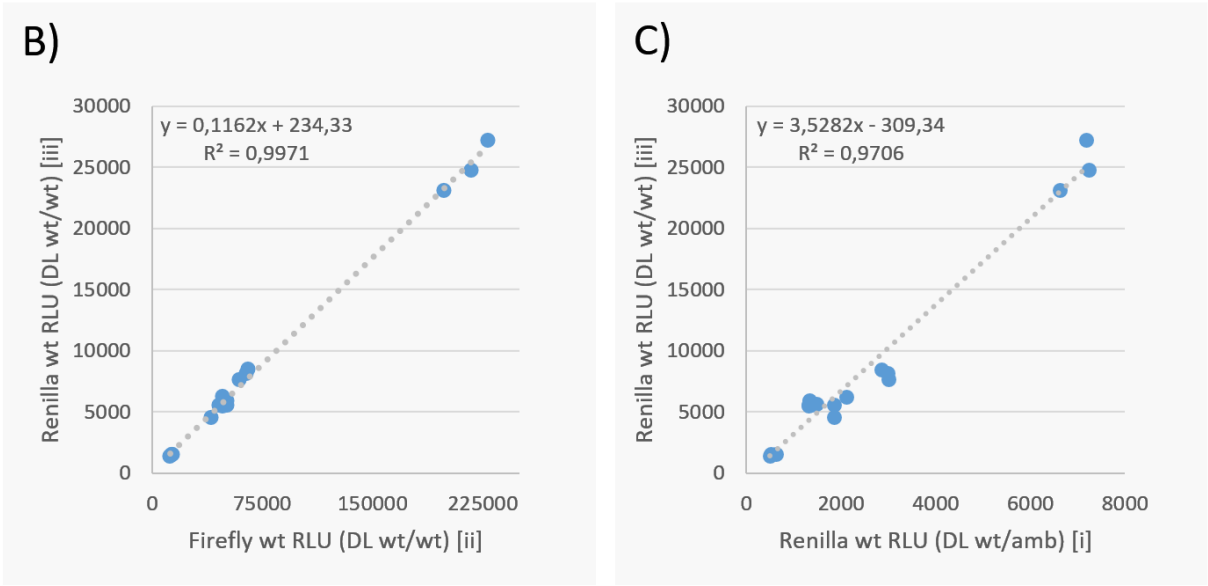


Figure 3-10: Characterization of the dual-luciferase reporter system.

(A) Area diagram showing the renilla and firefly luciferase RLU raw data resulting from several different dual-luciferase wt/wt or wt/amb co-transfection settings. (B) Linear correlation between the renilla wild-type and the firefly wild-type activity (RLU), both produced from the same bicistronic DL wt/wt plasmid. (C) Linear correlation of activity's (RLU) between the renilla wild-type produced from the DL wt/wt plasmid and the renilla wild-type produced from the DL wt/amb plasmid. $2x10^5$ HEK-293T cells were seeded in 24-well scale for the results in panel A, B and C. 24h post seeding, the cells were transfected with the indicated amounts of plasmid, using Lipofectamine 2000 in a 1:3 ratio including either 300 ng dual-luciferase wt/wt or wt/amb reporter plasmid. The luciferase assay was performed 72h post transfection. They were measured at OD-None using 7 µl lysate per well of the LumiNunc 96-well plate (100 µl 1xPLB per well had been used to harvest each well of the 24-well plate).

The Figure 3-10C shows that the RLU of the renilla wild-type enzyme was ~3.5-fold lower when expressed from the DL wt/amb construct, compared to the DL wt/wt construct (see regression curve function in Figure 3-10C). This was probably caused by some form of NMD, induced by the nonsense mutation present within the DL wt/amb construct. While this difference came unexpected, it did not represent an issue, as the renilla wild-type RLU values of both constructs still correlated very well (R^2-value of 0.97).

Because of this correlation, the restored firefly luciferase activity found in R/G-gRNA-treated DL wt/amb samples can be expressed, after normalization, as percentage relative to the normalized firefly values of untreated DL wt/amb (negative control) and DL wt/wt (positive control). Alternatively, the difference between samples can be expressed as relative fold change.

$$\%\ \boldsymbol{Restored\ Luciferase\ Activity\ (Normalized)} =$$

$$\frac{(\text{Mean } Firefly\ RLU\ Sample - \text{Mean } Firefly\ RLU\ Untreated\ cells)}{(\text{Mean } Renilla\ RLU\ Sample - \text{Mean } Renilla\ RLU\ Untreated\ cells)} \Big/ \frac{(\text{Mean } Firefly\ RLU\ Positive\ Control - \text{Mean } Firefly\ RLU\ Untreated\ cells)}{(\text{Mean } Renilla\ RLU\ Positive\ Control - \text{Mean } Renilla\ RLU\ Untreated\ cells)} * 100$$

$$\boldsymbol{Fold\ Change} =$$

$$\frac{(\text{Mean } Firefly\ RLU\ Sample\ A - \text{Mean } Firefly\ RLU\ Untreated\ cells)}{(\text{Mean } Renilla\ RLU\ Sample\ A - \text{Mean } Renilla\ RLU\ Untreated\ cells)} \Big/ \frac{(\text{Mean } Firefly\ RLU\ Sample\ B - \text{Mean } Firefly\ RLU\ Untreated\ cells)}{(\text{Mean } Renilla\ RLU\ Sample\ B - \text{Mean } Renilla\ RLU\ Untreated\ cells)}$$

The high numerical differences between the firefly RLUs of the negative control (DL wt/amb construct) and the positive control (DL wt/wt construct) indicate that this system is extremely sensitive and allows to detect the correction of firefly proteins far below 1% of their wild-type activity (Figure 3-10A).

Although the system was obviously capable of accurate quantification at the protein level, it was necessary to show that the measured percentages did also correlate to the editing at RNA level. This was mandatory to exclude that unspecific, not RNA-editing-related effects would lead to restored firefly activity. Therefore, diverse R/G gRNA designs were individually co-expressed with the dual-luciferase reporter system in ADAR1p110-, ADARp150, and ADAR2-Flp cells, resulting in a spectrum of different editing yields. These were analyzed simultaneously at the RNA level, using RT-PCR followed by Sanger-sequencing, and at the protein level, using the dual-luciferase assay.

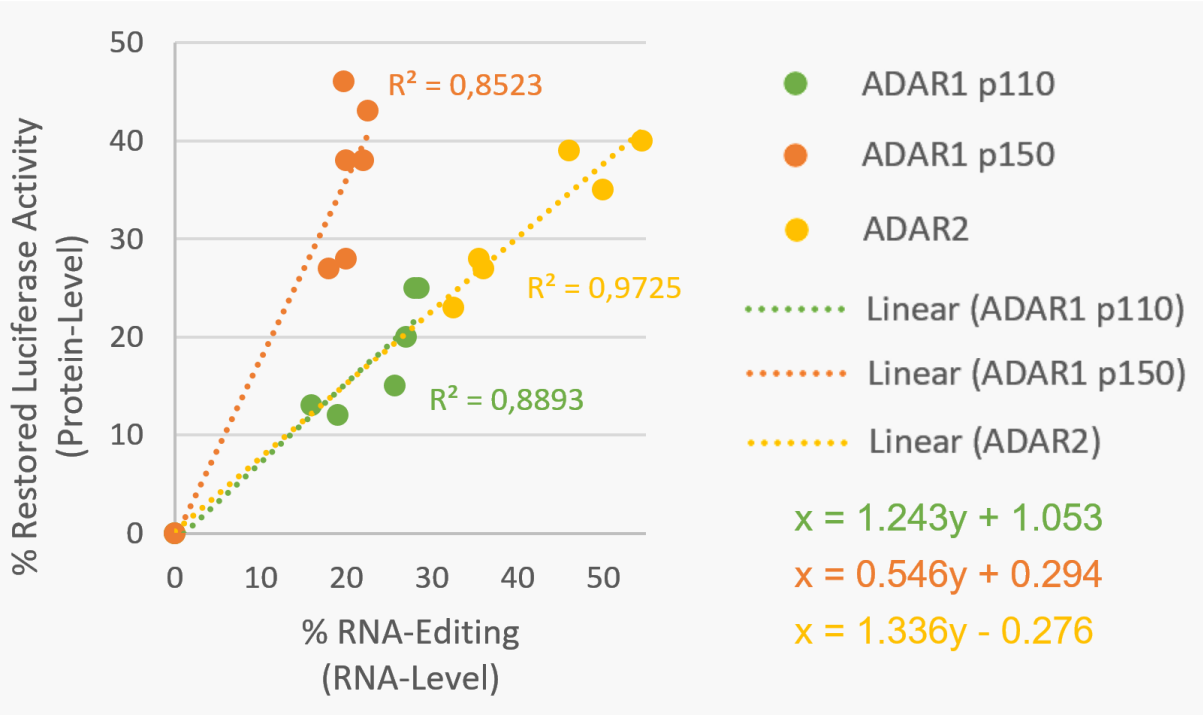


Figure 3-11: Correlation of dual-luciferase data at protein level, with editing data at RNA level.

The formulas on the lower right can be used to extrapolate the editing percentage at the RNA level from the data generated at the protein level. The formulas might only be valid within the displayed linear range of 0% to ~50% editing. $2x10^5$ ADAR1 p110, p150 Flp-In T-REx, or $3x10^5$ ADAR2 Flp-In T-REx cells per well were seeded in 24-well scale. The cells were induced with 10 ng/ml doxycycline at seeding. 24h post seeding, at 70-90% confluence, the cells were transfected with 300 ng dual-luciferase reporter plasmid and 1300 ng R/G gRNA plasmid using Lipofectamine 2000 in a 1:3 ratio. This was performed in duplicate. All cells were harvested 72h post transfection. For one of the duplicates, this was followed by the dual-luciferase reporter assay at OD-None using 7 µl lysate per well of the LumiNunc 96-well plate (100 µl 1xPLB per well had been used to harvest each well of the 24-well plate). For the other duplicate, it was followed by RNA isolation using the Quiagen RNeasy Mini Kit, DNase-I digestion, RT-PCR, and Sanger sequencing.

Both datasets (protein- & mRNA level) correlated nicely, even though there were differences between the ADAR enzymes (see Figure 3-11). The ADAR1 p110 and ADAR2 correlations were nearly identical, while the ADAR1p150 correlation diverged from the other two. However, by using the formulas displayed in Figure 3-11, it was thereafter possible to estimate the percentage of site-directed RNA editing based on results measured at protein level. Consequently, the system could now also be used as fast screening tool for novel gRNA designs avoiding the slow and cumbersome RT-PCR procedure that was used before.

After the assay was established and characterized, it could finally be used for its intended purpose: The detection of site-directed RNA editing using endogenous ADARs. Therefore the R/G gRNA design #3 was utilized, which had already been shown to be much more effective in the recruitment of ADAR1 (see Figure 3-7). As expected the achieved editing was extremely low, ~0.15% correction at the protein level in HeLa cells. Nevertheless, even these low values could be unequivocally and reliably detected, as indicated by the significant ($p = 0.0022$, unpaired, two-tailed, Mann-Whitney U test), six-fold difference between samples treated with R/G gRNA design #3 and the negative controls (see Figure 3-12). This was the starting point for further optimization, because the new readout allowed to generate meaningful SDRE data at physiological ADAR1 expression levels.

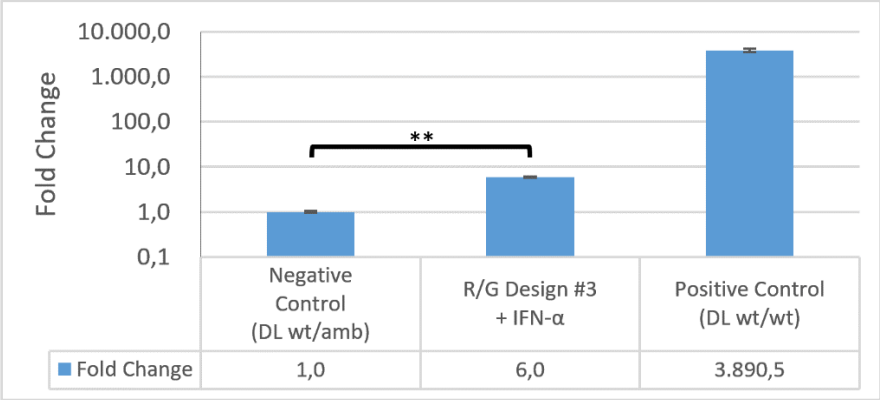


Figure 3-12: Dual-luciferase assay performed in IFN-α induced HeLa cells using plasmid-encoded dual-luciferase and R/G gRNA design #3 gRNAs.

Endogenous ADAR1 can be recruited in HeLa cells. Relative to the positive control, the 6-fold increase represents 0.15% correction at protein level. 1.2x104 HeLa cells were seeded in 96-well scale. 24h after seeding the cells were transfected with 100 ng per well of either empty vector, or design #3 R/G gRNA targeting firefly W417X amber, using Lipofectamine 3000 in a 1:1.5 ratio. Simultaneously, they were infected with 50 MOI dual-luciferase reporter adenovirus per well. The cells were treated with 200U IFN-α per well, beginning on the day of transfection/infection, and continuing for 48h, until the dual-luciferase assay was performed. They were measured at OD-None using 7 µl lysate per well of the LumiNunc 96-well plate (100 µl 1xPLB per well had been used to harvest each well of the 24-well plate). Medium was changed every 24h. The assay was performed as biological duplicate, each one measured as technical triplicate. The data are displayed as logarithmic fold change relative to the negative control. The results of the negative control and the R/G gRNA design #3 sample are significantly different with $p = 0.0022$ in a unpaired, two-tailed, Mann-Whitney U test.

3.7 Establishing the adenoviral transduction and qPCR quantification of R/G gRNAs

The low transfection efficiency in HeLa cells compared to HEK-293T cells (see Figure 3-13), which had been witnessed in many earlier experiments (e.g. using eGFP plasmids), lead to the conclusion that the low editing yield in HeLa cells might be caused by non-uniform and low R/G gRNA expression, besides other potential reasons. Consequently, it was attempted to simultaneously increase the R/G gRNA expression and find a way to quantify it. In cooperation with Prof. Naumann, an expert for recombinant adenoviruses (AdVs) from the Hertie-Institute Tübingen, we established the transport of our R/G gRNAs into hard-to-transfect cell types.

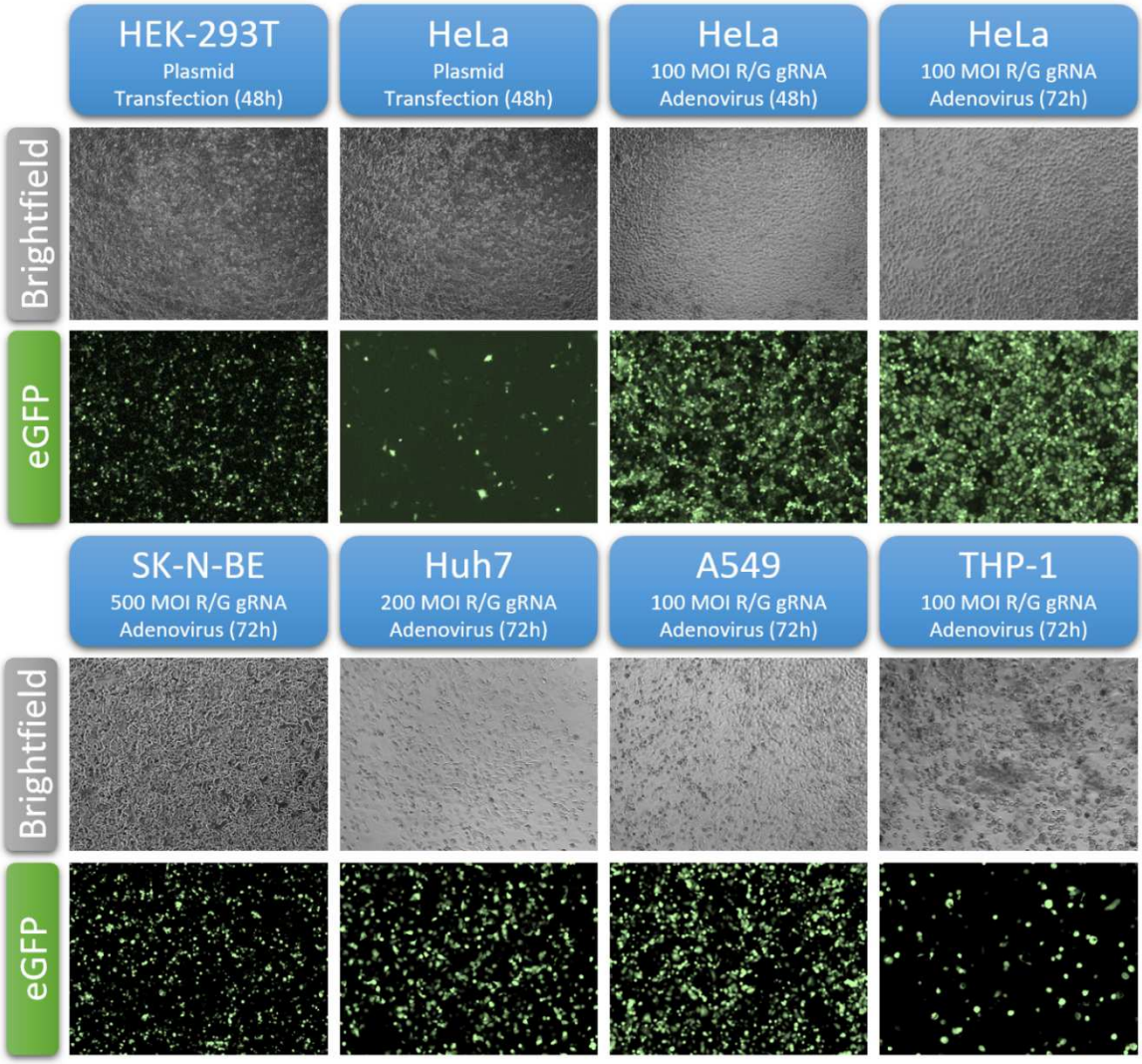


Figure 3-13: eGFP expression levels after plasmid transfection or adenovirus transduction.

The images were captured using the optimal settings and the optimal time points for each vector and cell type, as established in previous experiments and known at the time when the experiment was performed. It was aimed at the highest possible level of transgene expression, which was achived 48h post transfection for the plasmids, and 72h post transduction for the AdVs. The plasmid-transfected HEK-293T cells were treated with 1300 ng R/G gRNA and 300 ng eGFP using Lipofectamine 2000 at a 1:3 ratio (3 µl Lipofectamine 2000 per 1000 ng of plasmid). The plasmid-transfected HeLa cells were treated with 1300 ng R/G gRNA and 300 ng eGFP using Lipofectamine 3000 at a 1:1.5 ratio (1.5 µl Lipofectamine 3000 per 1000 ng of plasmid). The cells were infected with the indicated MOI of adenovirus, that encoded the β-actin 3' UTR targeting R/G gRNA of design #2 and an eGFP-tag. All cells were seeded to be 70-90% confluent at the time of transfection/transduction, using 24-well scale cell culture plates. The THP-1 cells had been PMA-differentiated previous to the experiment, using a 200 nM PMA treatment for 3 days, followed by 5 days of rest in medium without PMA.

Due to the fact that earlier experiments had required large amounts of plasmid and harsh transfection settings to achieve high editing yields, we decided to use AdVs instead of adeno-associated viruses (AAV) at this stage of development. In contrast to AAVs, AdV are known for high levels of transgene expression as well as high production titers.

AdVs are non-enveloped dsDNA viruses with an icosahedral capsid structure. The used AdV of serotype 5 is mechanistically the best-characterized and the most widely used virus over all clinical studies [217].

We used a first generation recombinant AdV system containing a E1A gene deletion, which prevents unintended viral replication. Ad293 cells were used for the AdV production, as they supply E1A *in trans* [218]. For details regarding the AdV production, see section 2.2.4.

Dr. Wettengel kindly provided an AdV for preliminary transduction optimization experiments. It encoded a eGFP reporter and a R/G gRNA of design #2, targeting the first 5'-UAG-3' triplet in the 3' UTR of endogenous β-actin mRNAs (hence referred to as β-actin R/G gRNA AdV). Initially, these optimizations were performed for several cell types by using the eGFP fluorescence of the β-actin R/G gRNA AdV as approximation for the R/G gRNA expression level. However, the eGFP expression from the CMV promoter did in fact not represent a valid approximation for the R/G gRNA expression from the U6 promoter, and lead to the false conclusion that the R/G gRNA expression might be already near its peak after ~48h (see Figure 3-13, 100 MOI AdV at 48h and 72h in HeLa cell).

As a means to directly asses the R/G gRNA quantity, a RT-qPCR workflow using the comparative ΔΔC(t) method was established (see Figure 2-4 and section 2.2.3) [200]. The qPCR primer pairs (see Table 2-7) were optimized regarding their melt temperatures, secondary structures, as well as off-target binding using primer-blast [219]. They required 5' strand-extensions to allow successful amplification of the only 61-71 nt long R/G gRNAs (length depending on the used R/G gRNA design). The resulting amplicons were 101-111 nt in size. In addition, the qPCR primer pairs were characterized regarding their efficiency (standard curve and amplification plot of dilution series), and specificity (melt curve and agarose gel). The results of this characterization for all primer pairs used in this thesis are summarized in Figure 3-14 and Figure 3-15.

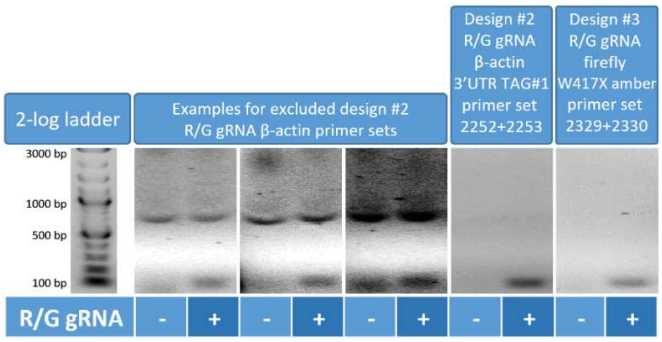


Figure 3-14: Evaluation of R/G gRNA qPCR primer pairs for unspecific by-products via agarose gel.

RNA was isolated from HEK-293T cells transfected with 1300 ng of the appropriate R/G gRNA, or from untransfected HEK-293T cells, using the high pure miRNA isolation kit. The RNA samples were reverse transcribed with the high-capacity cDNA reverse transcription kit, followed by phusion PCR, using several different primer sets. The products were resolved on a 1.4% agarose gel. The β-actin 3' UTR TAG#1 R/G gRNA primer set 2252+2253 and the firefly W417X amber primer set 2329+2330 showed no by-products and were used for R/G gRNA quantification and localization experiments.

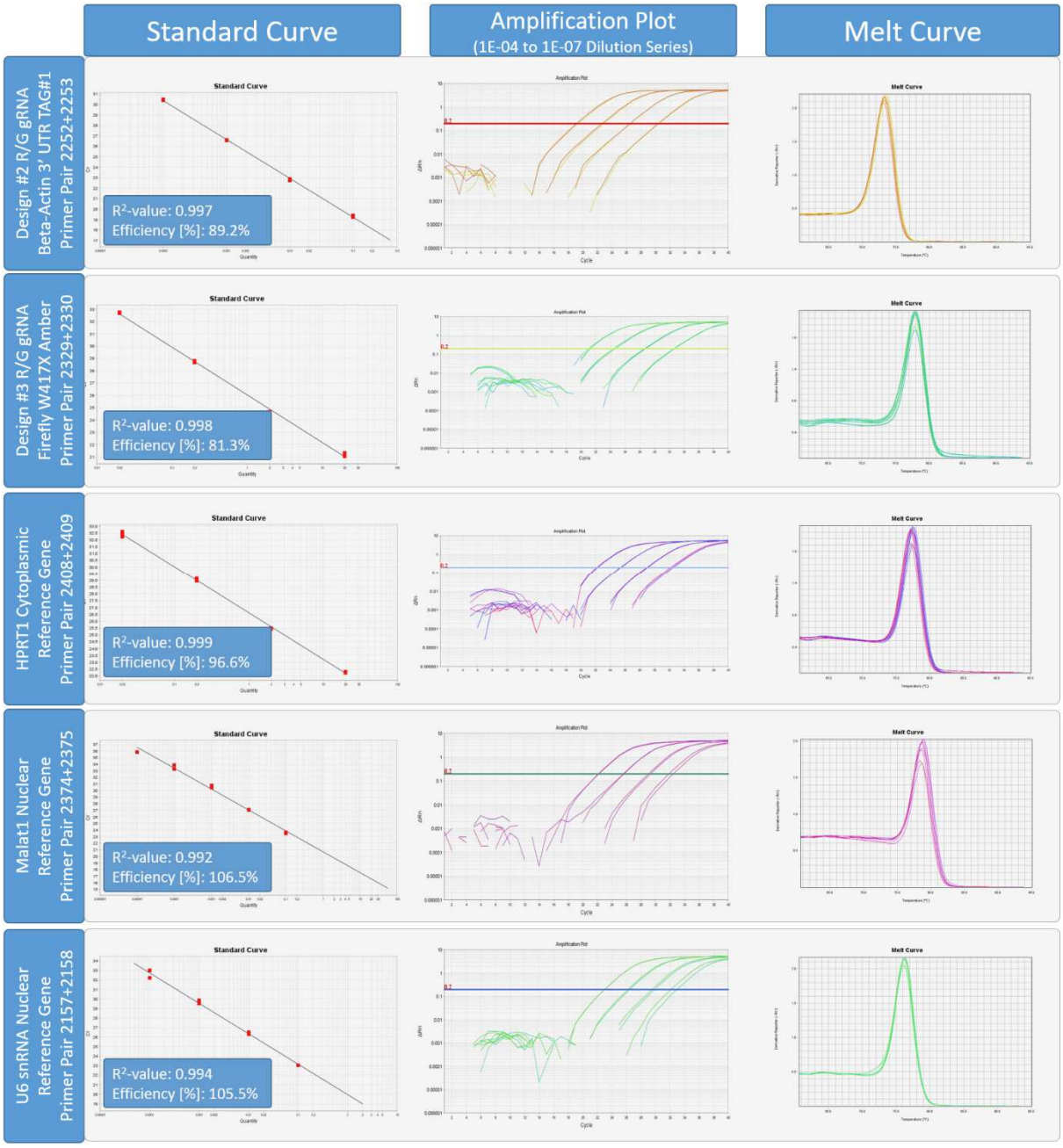


Figure 3-15: Characterization of qPCR primer pairs.

The efficiency of all measured primer pairs was in the acceptable range between 80% to 110%, with good standard curve fits represented by R2-values >0.99. The dilution series ($1x10^{-4}$, $1x10^{-5}$, $1x10^{-6}$, $1x10^{-7}$) resulted in a linear dynamic range up to dilutions of $1x10^{-7}$ in the amplification plots, and the melt curves showed no by-products. The R/G gRNA primer pairs were established in our lab. The primer pair for HPRT1 came from [220]. The primer pair for Malat1 came from [221]. The primer pair for U6 snRNA came from [222]. The GAPDH primer pair had already been characterized in our lab by Dr. Vogel and was not analyzed again. 2x105 HEK-293T cells were transfected with 1300 ng plasmid, encoding either a design #2 R/G gRNA targeting β-actin 3' UTR TAG#1, or a design #3 R/G gRNA targeting firefly-luciferase W417X amber. Lipofectamine 2000 in a 1:3 ratio was used. cDNA based on total RNA isolated from the former was used for the β-actin R/G gRNA, HPRT1, Malat1 and U6 snRNA dilution series, while cDNA based on total RNA isolated from the latter was used for the firefly-luciferase R/G gRNA dilution series. The cDNAs were diluted in nuclease-free water. The qPCRs were performed as described in section 2.2.3 and Figure 2-4.

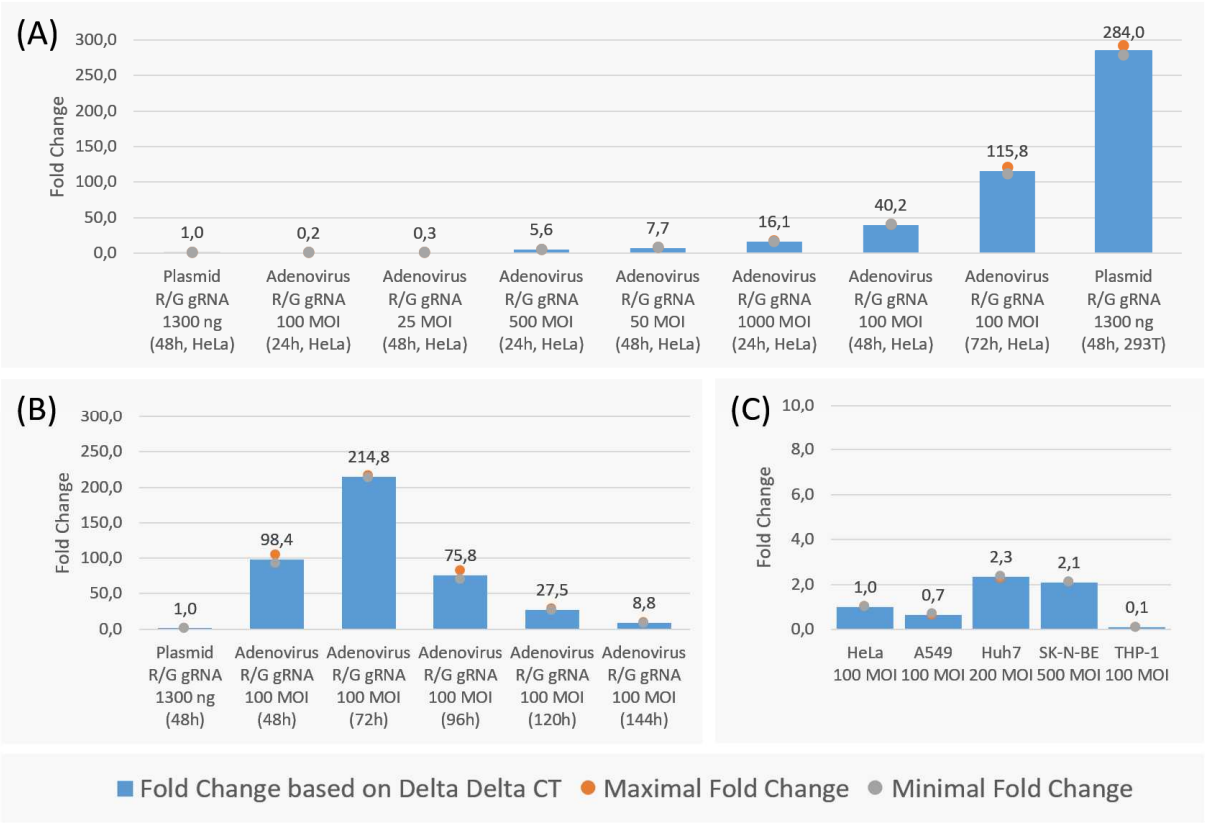


Figure 3-16: Relative quantification of plasmid- and AdV-encoded R/G gRNA expression levels using RT-qPCR.

(A) Comparison of several MOIs and time points to achieve high level R/G gRNA expression in HeLa cells. The transfections of HEK-293T and HeLa cells with identical plasmid encoded R/G gRNAs were used as reference values. (B) Verifying the optimal time-point, using 100 MOI of AdV and an extended time course in HeLa cells. The maximum yields differ from other experiments due to a less successful plasmid transfection. However, this does not influence the optimal time-point. (C) R/G gRNA expression of several different cell lines transduced with individual MOI's relative to HeLa cells transduced with 100 MOI. The U6 snRNA was used as reference gene for normalization. The used R/G gRNA was identical sequence-wise for each used vector (R/G design #2 gRNA targeting the first 5'-UAG-3' triplet in the 3' UTR of human β-actin). The following settings were used for the panels in this figure: $6x10^4$ HeLa cells, $2x10^5$ HEK-293T, $1x10^5$ A549, $5x10^4$ Huh7, $2x10^5$ SK-N-BE, or $1.5x10^5$ THP-1 (differentiated) cells were seeded in 24-well scale, and were then either transfected 24h post seeding with the indicated amount of plasmid, or transduced with an AdV using the indicated MOI. Previously established, optimal transfection/transduction settings were used for each vector. Lipofectamine 2000 was used in a ratio of 1:3 for the transfection of HEK-293T cells with plasmids. For the transfection of all other celltypes in this figure with plasmids, Lipofectamine 3000 was used in a ratio of 1:0.8. All cells were harvested by trypsination after the indicated time. RNA isolation, Turbo-DNase digestion, reverse transcription, and qPCRs were performed as described in section 2.2.3 and Figure 2-4.

The R/G gRNAs used for the quantification experiments in Figure 3-16 were sequence-wise completely identical, allowing the direct comparison of different gene transfer vectors. The used R/G gRNA was always of design #2, targeting the first 5'-UAG-3' triplet in the 3' UTR of the human β-actin mRNA.

The qPCR-workflow was then used to assess the β-actin R/G gRNA expression with several different combinations of MOI and time point. Using 500 MOI and 1000 MOI of the β-actin R/G gRNA AdV, both amounts already being moderately toxic for HeLa cells, brought only 5.1- and 16.1-fold increases, respectively, when quantified 24h post transduction and compared to plasmid-transfected HeLa cells

using the optimal transfection settings. When the quantification was performed after 72h, however, just 100 MOI could lead to a >100-fold increase. The 72h time point could be confirmed as optimal by an extended time course (Figure 3-16B). Although the R/G gRNA expression level of plasmid-transfected HEK293T cells could not be reached entirely via transduction in HeLa cells, the yields improved considerably, compared to the earlier transfection (Figure 3-16A).

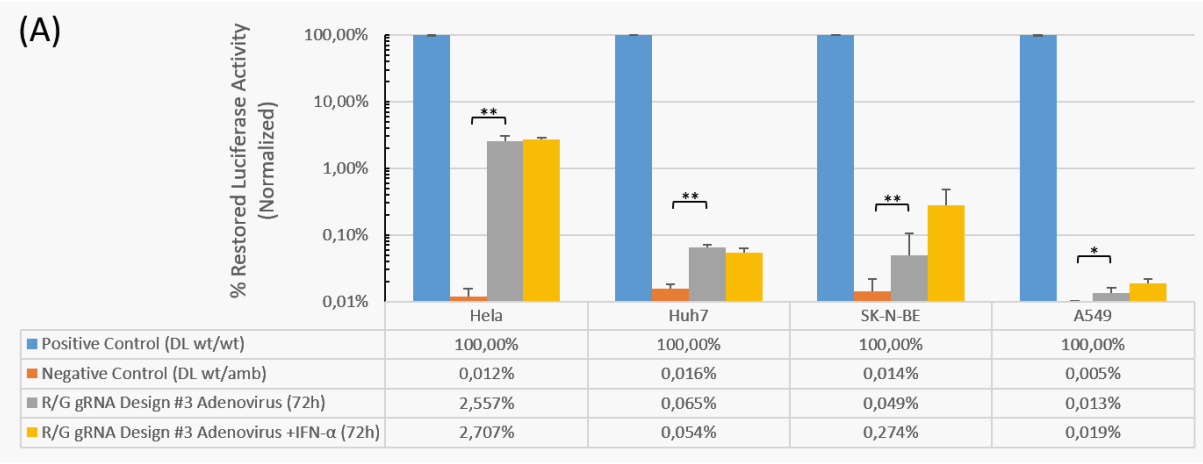
A relative comparison between the panel of ADAR1-positive and hard-to-transfect cell lines showed that in most of them similar R/G gRNA expression levels could be achieved. Only the THP-1 cells showed an expression level that was one order of magnitude lower and were thus not pursued any further (Figure 3-16C).

However, early experiments performed in HeLa and A549 cells using the β-actin R/G gRNA AdV failed to achieve SDRE detectable by Sanger sequencing (data not shown). Nevertheless, it was important to evaluate if SDRE in hard-to-transfect cell lines at endogenous ADAR1 expression levels would principally be feasible. In Addition, the impact of switching from plasmid to AdV-encoded R/G gRNAs on the editing yield had to be assessed.

3.8 Demonstration of site-directed RNA editing in hard-to-transfect cell lines using only endogenous ADAR1 and adenovirus-encoded R/G gRNAs

While Sanger sequencing had failed in this regard, the dual-luciferase reporter was sensitive enough to quantify the extent to which the adenoviral gene transfer had improved the editing yield, compared to the earlier used plasmids. Adenoviral vectors encoding the dual-luciferase reporters and a corresponding firefly W417X amber R/G gRNA of design #3 were utilized in this endeavor.

To achieve detectable RLU signals, reverse transduction of the cells with increased MOIs was required. For the reverse transduction, premixed AdV dilutions were distributed in 96-well plates, followed by the addition of cell suspensions with defined amounts of counted cells. Using the newly established time point of 72h and the harsher transduction settings, it was possible to detect significant levels (p-values < 0.03) of site-directed RNA editing, based purely on endogenous ADAR1, in HeLa, Huh7, SK-N-BE and A549 cells (Figure 3-17A). The measured correction yields were still very low, in case of A549 cells only slightly above the detection limit of the dual-luciferase assay, ranging from 0.02% to 2.6%. However, the achieved editing percentage using AdV-encoded R/G gRNAs in HeLa cells was significantly increased (2.6%), compared to the previous plasmid experiment (0.15%). By extending the experiment to an endpoint of 96h post transduction, the editing yield in HeLa cells could be increased to 5.3% (Figure 3-17B). While the R/G gRNA expression level was indeed highest after 72h, the editing reaction itself required also time, wherefore the delayed response came not completely unexpected.



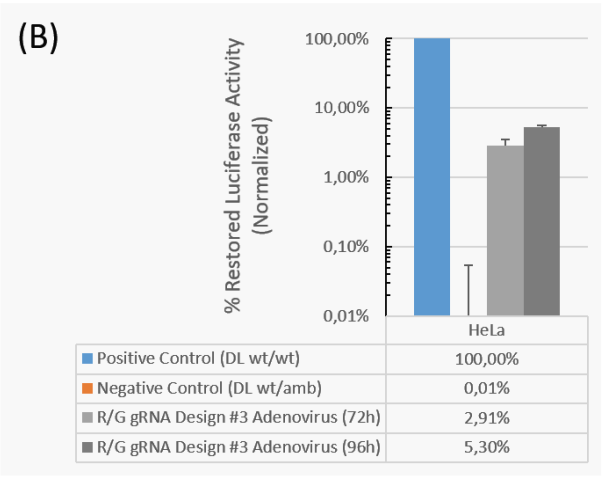


Figure 3-17: Recruitment of endogenous ADAR1 using AdV-encoded R/G gRNAs of design #3.

(A) Correction of the dual-luciferase reporter in several cell lines, using only endogenous ADAR1. $2x10^5$ HeLa cells were reverse-infected with 300 MOI R/G gRNA AdV and 50 MOI dual-luciferase wt/amb AdV. $1x10^5$ Huh7 cells were reverse-infected with 800 MOI R/G gRNA AdV and 100 MOI dual-luciferase wt/amb AdV. $5x10^4$ SK-N-BE cells were reverse-infected with 2000 MOI R/G gRNA AdV and 500 MOI dual-luciferase wt/amb AdV. And $2x10^5$A549 cells were reverse-infected with 400 MOI R/G gRNA AdV and 200 MOI dual-luciferase wt/amb AdV. This was performed in 96-well scale for all cell types. The positive control wells for all cell types were treated with 50 MOI dual-luciferase wt/wt AdV. Due to the normalization, the wt/wt AdV amount is insignificant for the analysis, while the wt/amb AdV amount determines the percentage of positive cells, and therefore the probability that R/G gRNA and wt/amb AdVs infect the same cell. The dual-luciferase AdV was applied 48h before the endpoint, while the R/G gRNA AdV was applied 72h before the endpoint. The statistical significances were calculated using an unpaired, two-tailed Mann-Whitney U test. The p-values are 0.0043 (HeLa), 0.0043 (Huh7), 0.0043 (SK-N-BE), and 0.0303 (A549). The identical p-values are a coincidence. The experiment was performed as biological triplicate, each one measured as technical duplicate. 7 µl of lysate were used per well of the LumiNunc 96-well plate and measured at OD-1. 30 µl 1xPLB per well had been used to harvest each well of the 96-well cell culture plate. (B) Delaying the endpoint from 72h to 96h nearly doubled the correction of the dual-luciferase reporter in HeLa cells, using only endogenous ADAR1. $2x10^5$ HeLa cells in 96-well scale were reverse-infected with 300 MOI R/G gRNA AdV and 50 MOI dual-luciferase wt/amb AdV. The dual-luciferase AdV was applied 48h before the endpoint, while the R/G gRNA AdV was applied 72h or 96h before the endpoint. The experiment was performed as biological triplicate, each one measured as technical duplicate. 7 µl of lysate were used per well of the LumiNunc 96-well plate and measured at OD-1. 30 µl 1xPLB per well had been used to harvest each well of the 96-well cell culture plate.

Overall, the switch from plasmid- to AdV-encoded R/G gRNAs had increased SDRE ~35-fold in HeLa cells. In addition, the low correction of the dual-luciferase reporter in all analyzed cell types explained why editing of endogenous β-actin mRNAs with AdV-encoded R/G gRNAs could not be detected by Sanger sequencing.

Finally, it could be deduced from these experiments that mainly endogenous ADAR1 p110 had been recruited, because the IFN-α treatment had no pronounced beneficial effect on the editing yield. The only exception was the SK-N-BE sample, which showed a high standard deviation, rendering this specific difference questionable.

3.9 The R/G gRNA copy-number is not the determining factor for superior editing yields achieved with chemically modified antisense oligonucleotides

While the adenoviral R/G gRNA vectors were characterized, our co-worker Tobias Merkle explored an alternative editing approach that combined *in-vitro*-transcribed (IVT) design #3 R/G motifs with chemically modified antisense parts (Pub. 4, Figure 1B). The resulting ASOs were directly transfected into cells with Lipofectamine 2000. Using only the endogenous ADAR1 in HeLa cells, even without IFN-α induction, these modified ASOs worked very well (~35% editing yield) for the editing of the same β-actin target, which AdV-encoded R/G gRNAs with identical sequence had failed to edit before (Pub. 4, Figure 2A).

The most obvious difference between the two systems are the chemical modifications. And indeed, Tobias Merkle could show that IVT R/G gRNAs without modifications achieved lower editing yields than sequence-identical but phosphorothioate and 2´-O-methyl-modified ASO R/G gRNAs in ADAR-Flp-In T-REx cells (Pub. 4, Figure S2).

Such chemical modifications, however, cannot be reproduced in a genetically encodable system. Therefore, other factors that might contribute to the superior editing yields of ASO R/G gRNAs, compared to AdV R/G gRNAs, were evaluated. As a first step, it was important to understand how the encodable and the ASO R/G gRNAs would compare regarding their relative transcript copy-number.

If the AdV-encoded R/G gRNA copy-numbers would, despite their improved expression levels compared to plasmids, be orders of magnitude below those of ASO R/G gRNA, it would probably be impossible to achieve editing, even when an alternative solution for the beneficial effect of chemical modifications would be found. A ED_{50} determination of ASO induced editing yields, performed by Tobias Merkle, was used as reference. It indicated that AdV-encoded R/G gRNAs would need to be expressed in a copy-number range similar to ~25 pmol ASO per well (24-well scale), to render efficient editing possible (see ED_{50} diagram in Pub. 4, Figure 2E). The previously used β-actin R/G gRNA sequence was

utilized again to compare the R/G gRNA copy-numbers of the different vectors in HeLa cells (Figure 3-18A).

In-vitro transcribed (IVT) R/G gRNAs without chemical modifications were used as substitute for ASOs in these experiments, to make sure that all compared samples would exhibit similar qPCR amplification efficiencies. For this reason, the improved nuclease resistance of 2'-O-Me and phosphorothioate backbone-modified ASOs, compared to completely *in-vitro* transcribed R/G gRNAs, could not be taken into account. However, it was assumed that an amplification bias would be more sever in a RT-qPCR experiment.

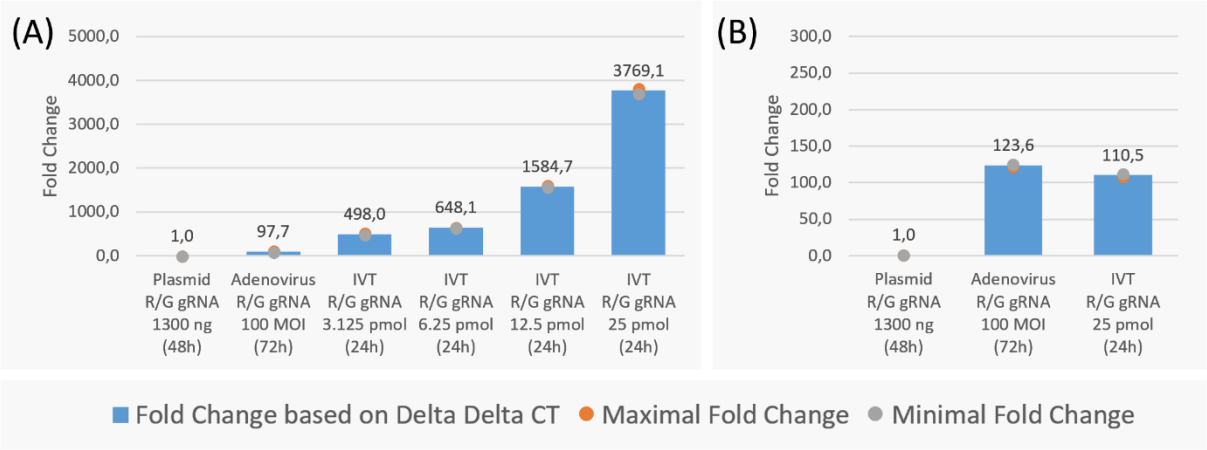


Figure 3-18: Relative quantification of R/G gRNA copy-numbers by RT-qPCR after treatment of HeLa cells with plasmid-encoded, AdV-encoded or *in-vitro* transcribed R/G gRNAs.

(A) Comparison of R/G gRNA copy-numbers between HeLa cells treated with plasmid-encoded, AdV-encoded or *in-vitro* transcribed R/G gRNAs using RT-qPCR. The total RNA was isolated using the High Pure miRNA Isolation Kit. (B) Comparison of R/G gRNA copy-numbers between HeLa cells treated with plasmid-encoded, AdV-encoded or *in-vitro* transcribed R/G gRNAs, using RT-qPCR. The total RNA was isolated using the PARIS Kit. The experiments were performed with three (A), or two (B), technical replicates. In panel A, the U6 snRNA was used as reference gene for normalization. In panel B, the geometric mean of C(t) values acquired from the reference genes HPRT1, Malat1, GAPDH, and U6 snRNA was used for normalization. The used R/G gRNA was identical sequence-wise for each used vector (R/G design #2 gRNA targeting the first 5'-UAG-3' triplet in the 3' UTR of human β-actin). The following settings were used for the panels in this figure: $6x10^4$ HeLa cells, seeded in 24-well scale, were transfected 24h post seeding with the indicated amount of plasmid or were transduced with an AdV using the indicated MOI. Alternatively they were reverse-transfected at seeding, using *in-vitro* transcribed R/G gRNAs. The previously established optimal transfection/transduction settings were used for each vector. Lipofectamine 3000 was used in a ratio of 1:0.8 for the plasmid transfection. For the transfection with IVT R/G gRNAs, 2.5 µl Lipofectamine 2000 were used per well. All cells were harvested by trypsination after the indicated time. RNA isolation, Turbo-DNase digestion, reverse transcription, and qPCRs were performed as described in section 2.2.3 and Figure 2-4.

The initial result obtained with the High Pure miRNA Isolation Kit showed again a ~100-fold increase of the AdV-encoded R/G gRNAs, compared to plasmid. Interestingly, it also showed a ~3800-fold increase in samples treated with our usual 24-well scale dose of 25 pmol IVT R/G gRNA (Figure 3-18A). Other studies in the siRNA field, however, already demonstrated that the largest fraction of transfected small RNAs are normally trapped within endosomes and are not readily available in cytosol or

nucleus, were editing should take place [223]. Data generated using the PARIS RNA Isolation Kit with milder buffer conditions, in conjunction with the earlier produced data, allowed a more accurate comparison of IVT- and AdV-encoded R/G gRNA levels.

While the High Pure miRNA Kit uses a harsh initial treatment to lyse the cells and all their compartments, the PARIS Kit uses very mild buffer conditions to selectively lyse the plasma membrane and the nucleus, allowing fractionation. The PARIS Kit cell fractionation buffer, which was used to lyse the plasma membrane, does not lyse endosomes, according to the Invitrogen customer service. The endosomes are trapped on top of the silica membrane, which is used to bind and wash the isolated RNAs. The measured relative expression of IVT R/G gRNAs using the PARIS Kit corresponded to 2.93% of the value obtained from the High Pure miRNA Kit (110.5-fold from Figure 3-18B divided by 3769.1-fold from Figure 3-18A). This comes very close to the average percentage (3.5%) of non-endosomally-trapped siRNAs [223] and supports the conclusion that most IVT R/G gRNAs are trapped in endosomes after their cellular uptake. The actual relative IVT R/G gRNA expression could now be quantified as ~100-fold higher than plasmids, exactly in the same range as the AdV-encoded R/G gRNAs (Figure 3-18B).

The switch from plasmids to AdVs did therefore close the gap between encodable and IVT R/G gRNAs in regard to their transcript copy-numbers. The different editing yields achieved by both types of R/G gRNA transfer when recruiting endogenous ADAR1, however, could not be explained by their quantities.

3.10 AdV-encoded and chemically modified ASO R/G gRNAs selectively recruit different ADAR1 isoforms

After the R/G gRNA copy-number had been ruled out as potential cause for the different editing yields of AdV-encoded and ASO R/G gRNAs, it was evaluated next, if selective recruitment of ADAR1 isoforms could be an alternative explanation.

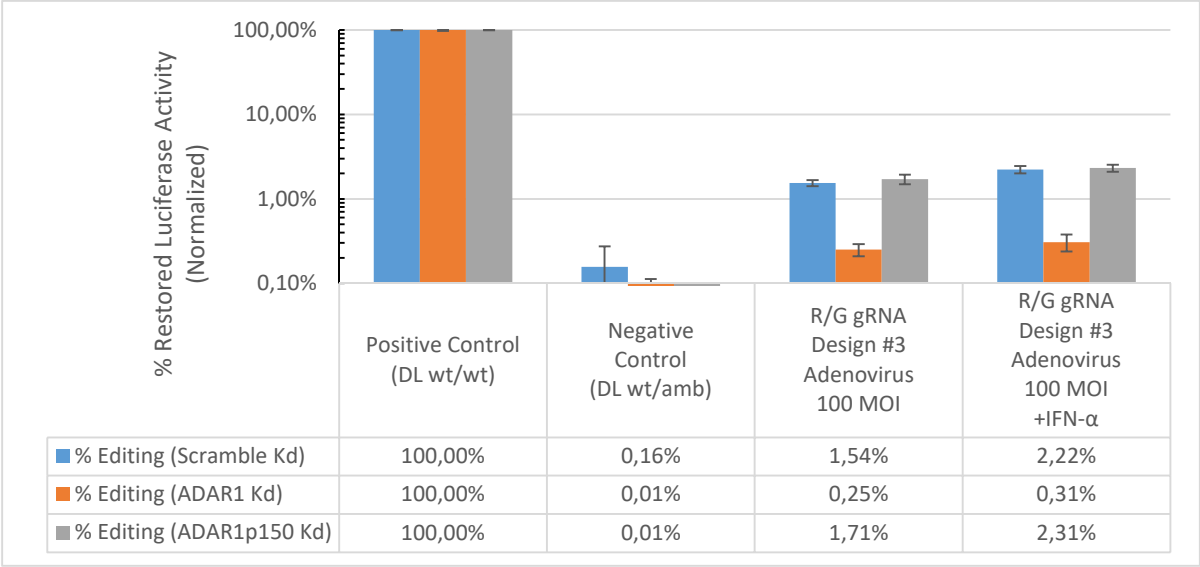
The enzymes responsible for the observed editing yields were determined by selective siRNA knockdown (KD) of endogenous ADAR1 isoforms in HeLa cells. The efficient simultaneous KD of both ADAR1 isoforms, or the selective KD of the p150 isoform alone, were confirmed by western blot (Pub.4, Figure 2D).

The considerably beneficial effect of IFN-α induction, in many cases nearly doubling the editing yield, which had been observed by Tobias Merkle when using ASO R/G gRNAs, already indicated the recruitment of ADAR1 p150 (Pub. 4, Figure 2A). In a cooperative experiment, we could show that completely

modified ASO R/G gRNAs (Pub. 4, Figure 1B, ASO v9.5) indeed recruited mainly ADAR1 p150 (Pub. 4, Figure 2C and D), when targeting the 3' UTR of endogenous GAPDH mRNAs in HeLa cells.

A similar experiment performed with AdV-encoded R/G gRNAs, targeting the dual-luciferase reporter, showed that these recruited mainly ADAR1 p110. Both this experiment and previous ones using AdV-encoded R/G gRNAs, consistently found no pronounced beneficial effect of IFN-α induction on the editing yield (Figure 3-17 and Figure 3-19).

Thus, the different ways of R/G gRNA application somehow affected which ADAR1 isoform was selectively recruited for SDRE. ASOs recruited mainly ADAR1 p150, while AdV-encoded R/G RNAs recruited mainly ADAR1 p110.



	Positive Control (DL wt/wt)	Negative Control (DL wt/amb)	R/G gRNA Design #3 Adenovirus 100 MOI	R/G gRNA Design #3 Adenovirus 100 MOI +IFN-α
% Editing (Scramble Kd)	100,00%	0,16%	1,54%	2,22%
% Editing (ADAR1 Kd)	100,00%	0,01%	0,25%	0,31%
% Editing (ADAR1p150 Kd)	100,00%	0,01%	1,71%	2,31%

Figure 3-19: Effect of general (p110 & p150) and specific (p150) ADAR1 knock-down on the editing yield achieved by AdV-encoded design #3 R/G gRNAs targeting the dual-luciferase reporter in HeLa cells.

Seeding of $1.2x10^5$ HeLa cells in 12-well scale and reverse transfection with scramble siRNA, ADAR1 siRNA, or ADAR1p150 siRNA using 2.5 pmol siRNA per well (3 µl HighPerfect + 2.5µl 10 nM siRNA ad 200 µl with OptiMem). 6 wells per siRNA were seeded. Each used well of the 12-well plate contained the indicated 200 µl transfection mix for the corresponding siRNA. 24h post siRNA transfection, similarly treated wells were harvested and the cells pooled. Thereafter $5x10^4$ cells per well in 96-well scale were reverse-infected with 150 MOI R/G gRNA AdV targeting firefly W417X amber. 100 µl of HeLa cell suspension were added to 50 µl AdV dilution in PBS, already distributed in the 96-well plate. 24h post R/G gRNA infection, the cells were infected with 50 MOI dual-luciferase AdV per well of the 96-well plate. Either wt/wt for the positive control, or wt/amb for the negative control and the editings were used. 48h post dual-luciferase infection, the luciferase assay was performed using OD-1 and 7 µl of the lysate per well of the LumiNunc 96-well plate. 30 µl 1xPLB per well had been used to harvest each well of the 96-well cell culture plate. Three individually executed experiments gave similar results. Each experiment was performed as biological triplicate, with each biological sample measured as technical duplicate.

3.11 The regulation of selective ADAR1 isoform recruitment by differently applied R/G gRNAs goes beyond simple co-localization in nucleus or cytoplasm

As mentioned earlier, ADAR1 p110 is known to be localized in the nucleus, and ADAR1 p150 in the cytosol (section 1.1.1.1). Consequently, the subcellular distribution of differently applied R/G gRNAs over these compartments had to be assessed as potential reason for their selective ADAR1 isoform recruitment. Thus, the localizations of plasmid-encoded, AdV-encoded, and IVT R/G gRNAs were determined by RT-qPCR (section 2.2.3 and Figure 2-4) after subcellular fractionation of nucleus and cytoplasm using the PARIS RNA isolation kit (Figure 3-21). Again, sequence-identical β-actin R/G gRNAs of design #2 were used for all treatments (plasmid, AdV, IVT). To calculate the relative percentage of R/G gRNAs between nucleus and cytoplasm, the geometric mean of the cytoplasmic reference genes HPRT1 and GAPDH, as well as the nuclear reference genes Malat1 and U6 snRNA, was used to normalize the target gene. The calculations, which are based on the ΔΔC(t) method, are explained in section 2.2.3. The qPCR primer pairs of the mentioned reference genes were characterized as explained before (see Figure 3-15). The marked difference in fold change between the cytoplasmic and the nuclear fraction, especially of Malat1, confirmed that the fractionation had worked efficiently (Figure 3-21A).

Small RNAs expressed from U6 promoters are in general assumed to be mainly localized in the nucleus [224-228], while the literature regarding ASO localization differs widely, depending on various factors, like its size and structure [229], the type of incorporated chemical modifications [230], and RNA-protein interactions [231, 232].

Because it had already been shown that R/G gRNA ASOs recruit the cytosolic ADAR1 p150 isoform (Pub. 4, Figure 2C and D), it was expected that they would be present in the cytosol. In addition, the high editing yields achieved by ADAR1 p150 made its selective recruitment desirable (see Pub. 4, Figure 2C and D). Accordingly, it was relevant to find out if the inability of the encodable R/G gRNAs to recruit the seemingly more active ADAR1 p150 was caused by their expected nuclear retention. If true, their editing yield should be improved by causing a localization change into the cytosol, e.g. by attaching a tRNA to the R/G gRNA. The hypothesis was therefore, that ASOs would be present in the cytosol and thus would be able to recruit cytosolic ADAR1 p150, while the encodable R/G gRNAs, expressed from a U6 promoter, would be localized exclusively in the nucleus, limited to the nuclear ADAR1 p110.

The analyzed IVT R/G gRNAs were again unmodified to allow efficient qPCR amplification. Therefore, the IVT R/G gRNA localization would not necessarily resemble the localization of chemically modified ASO R/G gRNAs. It was nevertheless interesting to see how unmodified IVT R/G gRNAs would be localized, as they recruited the cytosolic ADAR1 p150 better than the nuclear ADAR1 p110, or ADAR2, in

ADAR Flp-In T-REx cells (Pub. 4, Figure S2). Consequently, they were also expected to be at least partially localized in the cytoplasm.

As indicator for the subcellular distribution of ASO R/G gRNAs, my co-worker Tobias Merkle kindly provided a set of fluorescence images that show the localization of an Atto594-coupled, modified ASO of design #3 in HeLa cells. They demonstrate that ASOs are localized mainly in endosomes, but otherwise all over the cell, including the cytosol (Figure 3-20).

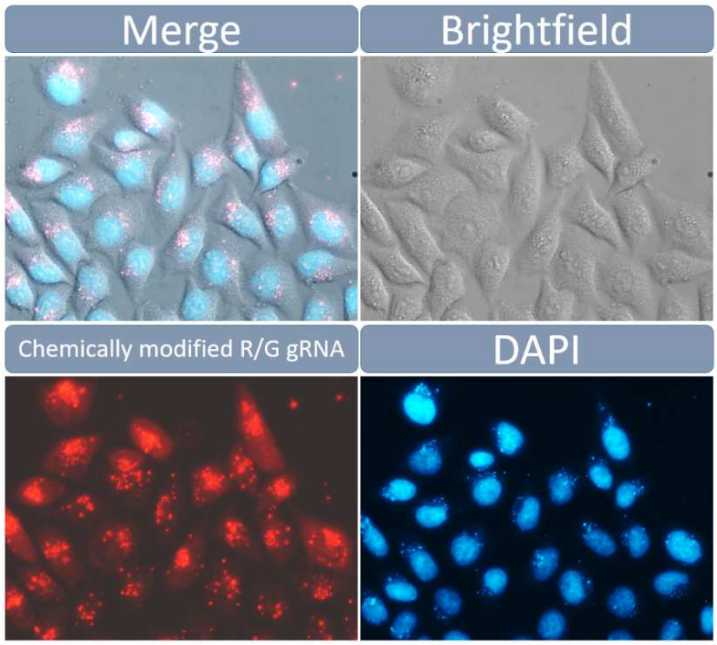


Figure 3-20: Localization of chemically modified ASO R/G gRNAs in HeLa cells 24h post transfection.

The R/G gRNAs are mainly localized within endosomes, visible as small red dots all over the cell. Some of these vesicles might also be extracellular lipoplexes. The presence of released R/G gRNAs within the cells is visible as faint red signal filling the contour of the cells, as demonstrated by comparing the R/G gRNA channel to the bright-field channel. Comparing the R/G gRNA channel to the DAPI channel shows that the R/G gRNAs are present in cytosol and nucleus. The R/G gRNAs were chemically modified as explained in Pub. 4, Figure 1A, Version 9.4. In addition, they had an amino-linker that was used to tag them with an Atto594-N-Hydroxysuccinimide ester. For the experiment $1x10^4$ HeLa cells were reverse-transfected in 96-well scale with RNAi Max (5 pmol R/G gRNA, 0.3 l RNAi Max). 24h post transfection, the fluorescent imaging was performed. These images were kindly provided by my co-worker Tobias Merkle.

The previously stated hypothesis proved to be only partially accurate (Figure 3-21B). While the plasmid-encoded R/G gRNAs were indeed nearly exclusively localized within the nucleus in all tested cell types (HeLa, A549, HEK-293T), the AdV-encoded R/G gRNAs, which were also expressed from an U6 promoter, were distributed evenly over both compartments in all tested cell types.

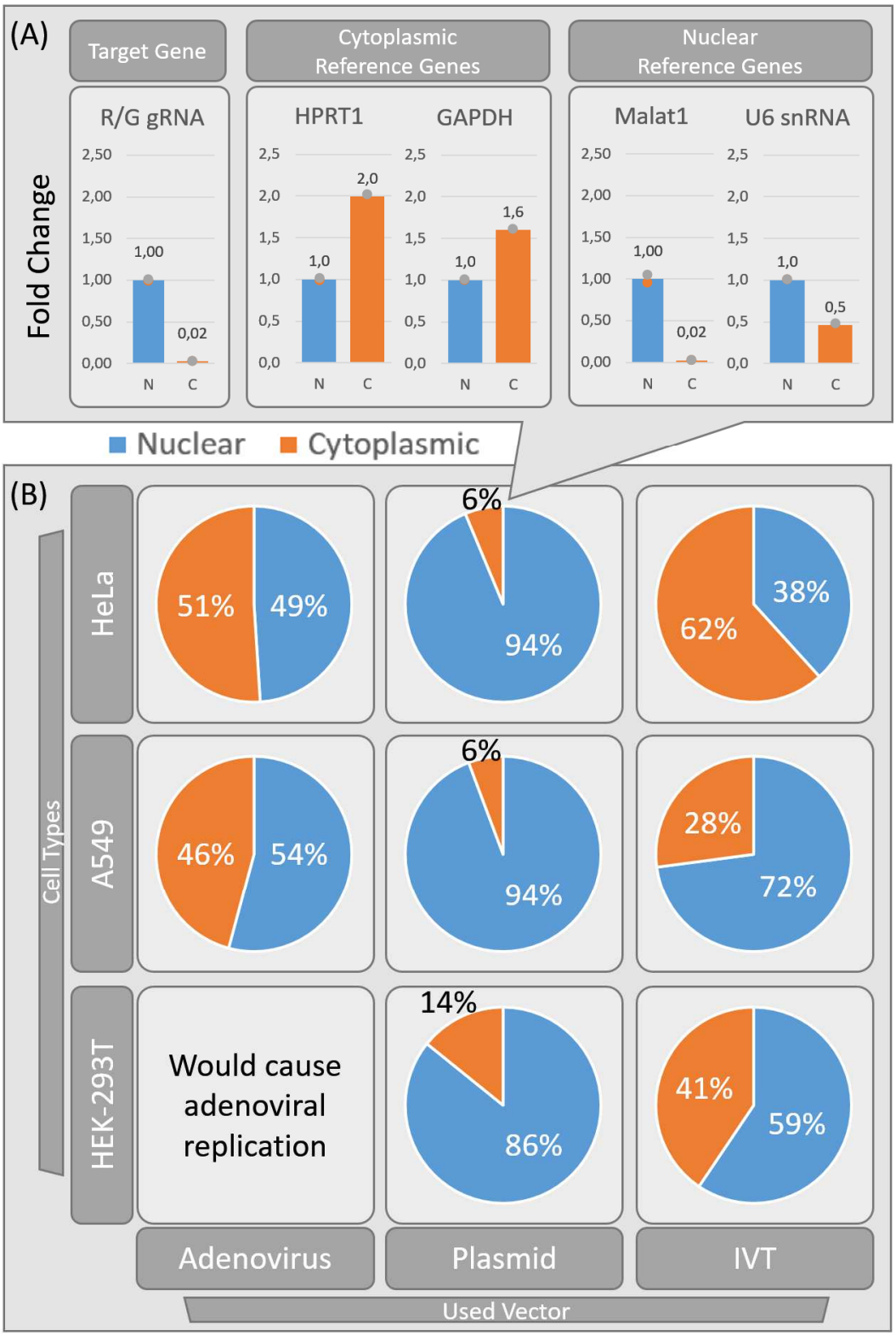


Figure 3-21: Sub-cellular localization of R/G gRNAs applied to different cell lines using adenovirus, plasmid, or IVT R/G gRNAs.

(A) Exemplary fold change between the nuclear and cytoplasmic fractions of the target gene and the used reference genes. The geometric mean of all four reference genes was used to normalize the target gene results. The example data derived from HeLa cells transfected with a plasmid encoding the β-actin 3' UTR gRNA design #2. (B) R/G gRNA distribution between nucleus and cytoplasm depending on the used vector and cell type.

Figure 3-21 continued: Sub-cellular localization of R/G gRNAs applied to different cell lines using adenovirus, plasmid, or IVT R/G gRNAs.

HEK-293T cell could not be transduced using an AdV, because these cells are E1A positive and would facilitate AdV replication, causing the characteristic cytopathic effect (cell death), which can be observed in Ad293 cells during AdV production. Each result was normalized using the geometric mean of all four reference genes and then converted into a relative percentage value. The used R/G gRNA was identical sequence-wise for each used vector (R/G design #2 gRNA targeting the first 5'-UAG-3' triplet in the 3' UTR of human β-actin). $2x10^5$ HEK-293T cells per well were seeded in 24-well scale. 24h post seeding, the cells were transfected with 1300 ng R/G gRNA plasmid using a Lipofectamine 2000 ratio of 1:3 or were reverse-transfected at seeding with 5 pmol IVT R/G gRNA using 2.5 µl Lipofectamine 2000. $1x10^5$ A549 cells, or $6x10^4$ HeLa cells, per well were seeded in 24-well scale. 24h post seeding, the cells were transfected with 1300 ng R/G gRNA plasmid using a Lipofectamine 3000 ratio of 1:0.8, or were reverse-transfected at seeding with 5 pmol IVT R/G gRNA using 2.5 µl Lipofectamine 2000. Transduction of A549 and HeLa cells was performed 24h post seeding using 100 MOI AdV. IVT-R/G-gRNA treated cells were harvested 24h post transfection. The plasmid-transfected cells were harvested 48h post transfection. The AdV-transduced cells were harvested 72h post infection. RNA isolation, Turbo-DNase digestion, reverse transcription, and qPCR were performed as explained in section 2.2.3 and Figure 2-4. The experiments were performed with two technical replicates.

The unsuspected even distribution, might have been caused by structural damage of the nuclear pore complex by the adenoviral entry mechanism [233]. The unmodified IVT R/G gRNAs were present in both compartments, as expected, but showed different tendencies for nucleus or cytosol, depending on the analyzed cell type. As explained earlier, the PARIS Kit, which was used for RNA isolation in these experiments, excludes the large endosomally-trapped portion and quantifies only IVT R/G gRNAs that were already released from the endosomes. The IVT R/G gRNAs detected in the cytosol were therefore no artefact of the utilized RNA-isolation method, but actually functional, cytosolic R/G gRNAs. In conclusion, this means that AdV-encoded, ASO, and IVT R/G gRNAs exhibit a similar localization profile, being distributed over nucleus and cytosol, but do somehow individually recruit ADAR1 isoforms, which are present in only one of those compartments (Figure 3-19, Figure 3-21B and Pub. 4, Figure 2C and D). The exact factors that determine the selective recruitment of ADAR1 isoforms by differently applied R/G gRNAs are a topic of ongoing research.

3.11.1 Successful localization-change of an encodable R/G gRNA into the cytoplasm

The possibility of changing the localization of plasmid-encoded R/G gRNAs was investigated simultaneously with the previous experiments. Although the R/G gRNA localization was eventually not the only determining factor for selective ADAR1 isoform recruitment, interesting insights for future projects have been learned in the endeavor. To enforce the localization-change of a design #4 R/G gRNA, a valine tRNA was fused to it. The approach was based on a publication which used the same valine tRNA to export ribozymes [234]. As expected, the R/G gRNA-$tRNA^{Val}$ hybrid was exported into the cytosol with more than 95% efficiency (Figure 3-22A). In comparison, the regular plasmid-encoded design #4 R/G gRNA (Figure 3-22B) was mainly localized within the nucleus, as seen in previous experiments. Due to the activity of the tRNA as internal promotor, the R/G gRNA $tRNA^{Val}$ hybrid was expressed nearly 300 times higher than the regular design, which was using a U6 promoter (Figure 3-22D). Despite this,

SDRE caused by hybrid R/G gRNAs could only be detected using the highly sensitive dual-luciferase reporter system in ADAR1 Flp-In T-REx cells. And although SDRE could be detected, it was more than one order of magnitude lower than with a similar R/G gRNA without the tRNA (Figure 3-22E). This was assumed to be caused by steric problems with the tRNA. By incorporation of an intermediate linker between the R/G gRNA and the tRNA, it was possible to restore some of the editing activity (Figure 3-22C), but not to the levels which could be reached with a regular design #4 R/G gRNA (Figure 3-22B). Although the tRNA hybrid approach was ultimately not successful in increasing the editing yield, the localization of the R/G gRNA could be changed and its expression level increased drastically. This knowledge might be useful for follow-up projects.

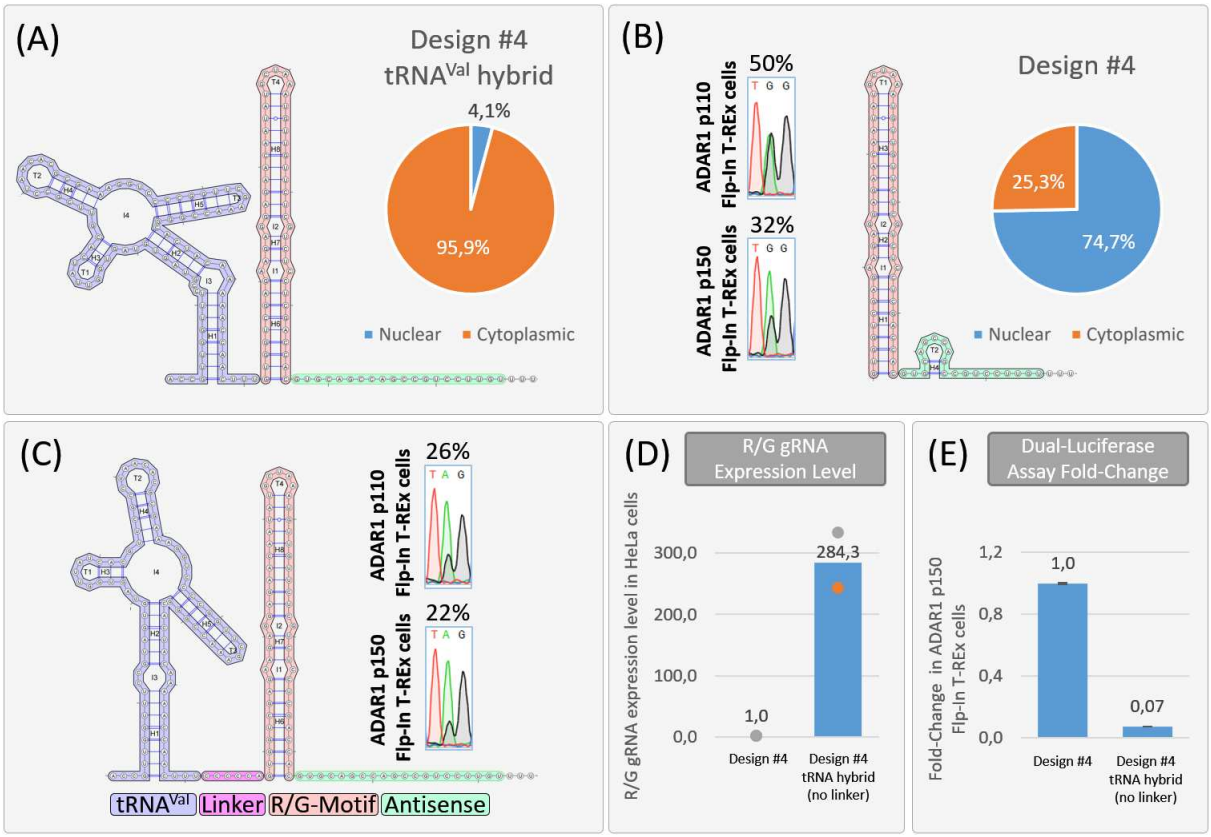


Figure 3-22: Generation of a R/G gRNA tRNA^{Val} hybrid and its effect on localization, R/G gRNA expression level and editing yield measured at RNA level, using Sanger sequencing, and at protein level, using the dual-luciferase reporter system.

(A) Structure of a tRNA^{VAl} fused 5' to a design #4 R/G gRNA and its subcellular localization. (B) The editing yield on RNA level in ADAR1p110 and ADAR1 p150 Flp-In T-REx cells and subcellular localization of a regular design #4 R/G gRNA. (C) Structure of a tRNA^{VAl} fused 5' to a design #4 R/G gRNA with an intermediate linker region and its editing yield on RNA level in ADAR1p110 and ADAR1 p150 Flp-In T-REx cells. (D) Comparison of the R/G gRNA expression level between a U6-promoter-expressed design #4 R/G gRNA and a R/G gRNA tRNA^{Val} hybrid (no linker) expressed from the internal tRNA^{Val} promoter. Quantification was performed as explained in Figure 3-16. (E) Fold change comparison of a regular design #4 R/G gRNA and the R/G gRNA tRNA^{Val} hybrid (no linker) regarding their editing yield on protein level, using the dual-luciferase reporter system. For the qPCRs resulting in the localization data for panel A and B, and the quantification data for panel D, 6x10^4 HeLa cells were seeded in 24-well scale. 24h post seeding, they were transfected with 1300 ng R/G gRNA plasmid, using Lipofectamine 3000 at a 1:0.8 ratio. Quantification and localization were performed as explained in section 2.2.3. The geometric mean of the reference genes HPRT1, GAPDH, Malat1, and U6 snRNA was used for normalization.

Figure 3-22 continued: Generation of a R/G gRNA tRNA^{Val} hybrid and its effect on localization, R/G gRNA expression level and editing yield measured at RNA level, using Sanger sequencing, and at protein level, using the dual-luciferase reporter system.

For the Sanger sequencing in panel B and C, 2x10^5 ADAR1 p110 or p150 Flp-In T-REx cells per well were seeded in PDL coated 24-well plates. 24h post seeding, at 70-90% confluence, the cells were transfected with 300 ng dual-luciferase reporter plasmid and 1300 ng R/G gRNA plasmid, using Lipofectamine 2000 in a 1:3 ratio. 72h post transfection, the cells were harvested. This was followed by RNA isolation with the Quiagen RNeasy Mini Kit, DNase-I digestion, RT-PCR, and Sanger sequencing. The experiment was performed as biological triplicate. For the dual-luciferase assay results in panel E, 4x10^4 p150 Flp-In T-REx cells per well were seeded in PDL coated 96-well plates. 24h post seeding, at 70-90% confluence, the cells were transfected with 60 ng dual-luciferase reporter plasmid and 260 ng R/G gRNA plasmid using Lipofectamine 2000 in a 1:0.8 ratio. 72h post transfection, the dual-luciferase assay was performed using OD-None and 7 µl of the lysate per well of the LumiNunc 96-well plate. 30 µl 1xPLB per well had been used to harvest each well of the 96-well cell culture plate. The dual-luciferase assay was performed as biological triplicate, each one measured again as technical triplicate.

3.12 Enabling the efficient recruitment of human ADAR1 p110 by subsequent rational design of the R/G gRNA architecture

Even small amounts of ADAR1 p150, hardly detectable by western blot without IFN-α induction, caused strikingly high editing yields of ~40% correction, when using chemically modified R/G gRNAs (see Pub. 4, Figure 2C and D). This initially lead to the previously described efforts to shift the selectivity of encoded R/G gRNAs from ADAR1 p110 towards ADAR1 p150. ADAR1 p110, however, was expressed in all analyzed cell types to a much higher degree (Figure 3-6). In addition, it is constitutively expressed in most tissues and does not require IFN-α induction in the first place. This would make it the perfect ADAR for site-directed RNA editing, if it could be recruited more efficiently. Furthermore, nuclear-localized editases are supposed to cause less off-target editing, compared to cytosolic ones [145]. While this was only shown for artificial deaminases of the λN-boxB system, so far, it would be an additional advantage, if also true for the nuclear ADAR1 p110. Consequently, the next goal was to find a R/G gRNA architecture that could recruit ADAR1 p110 more efficiently and would thereby increase the editing yield. To achieve this, a multitude of novel R/G gRNA designs was conceived and then screened in ADAR1-Flp cells.

The new cloning procedure, explained in section 2.2.1 and Figure 2-3C, and the dual-luciferase reporter system allowed to rapidly test new designs by circumventing the cumbersome primer extension PCR cloning procedure, as well as the time consuming RT-PCR and Sanger sequencing protocol.

This section summarizes all improvements to the R/G gRNA architecture that were achieved during this thesis. It starts with design #1, originally created by Dr. Wettengel, and explains step-by-step the rational design decisions that enabled the encodable R/G gRNA system to efficiently recruit endogenous ADAR1 p110 for site-directed RNA editing.

At the most basic level, a R/G gRNA consists of two parts. The double-stranded R/G motif and the single-stranded antisense part (Figure 3-23, A).

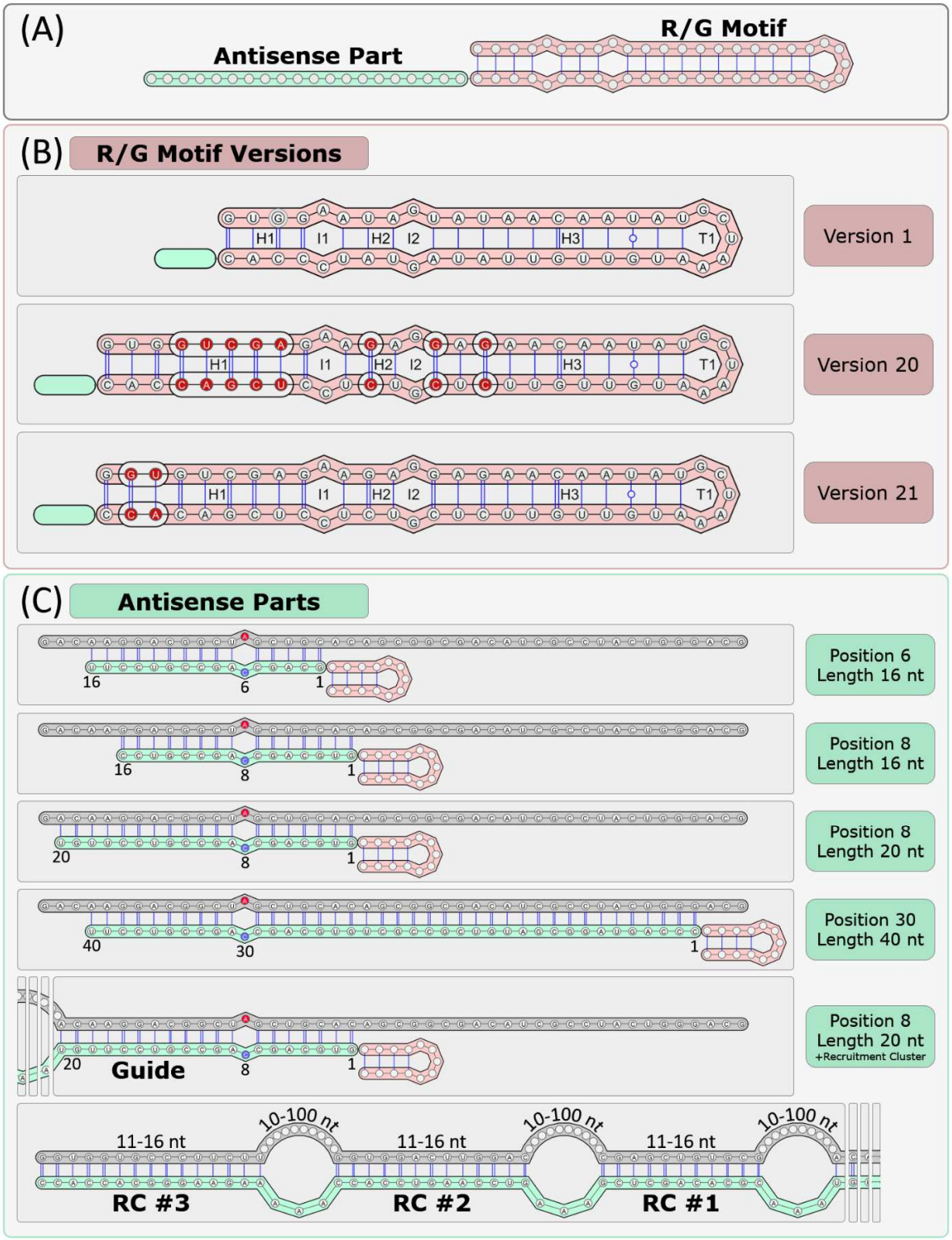
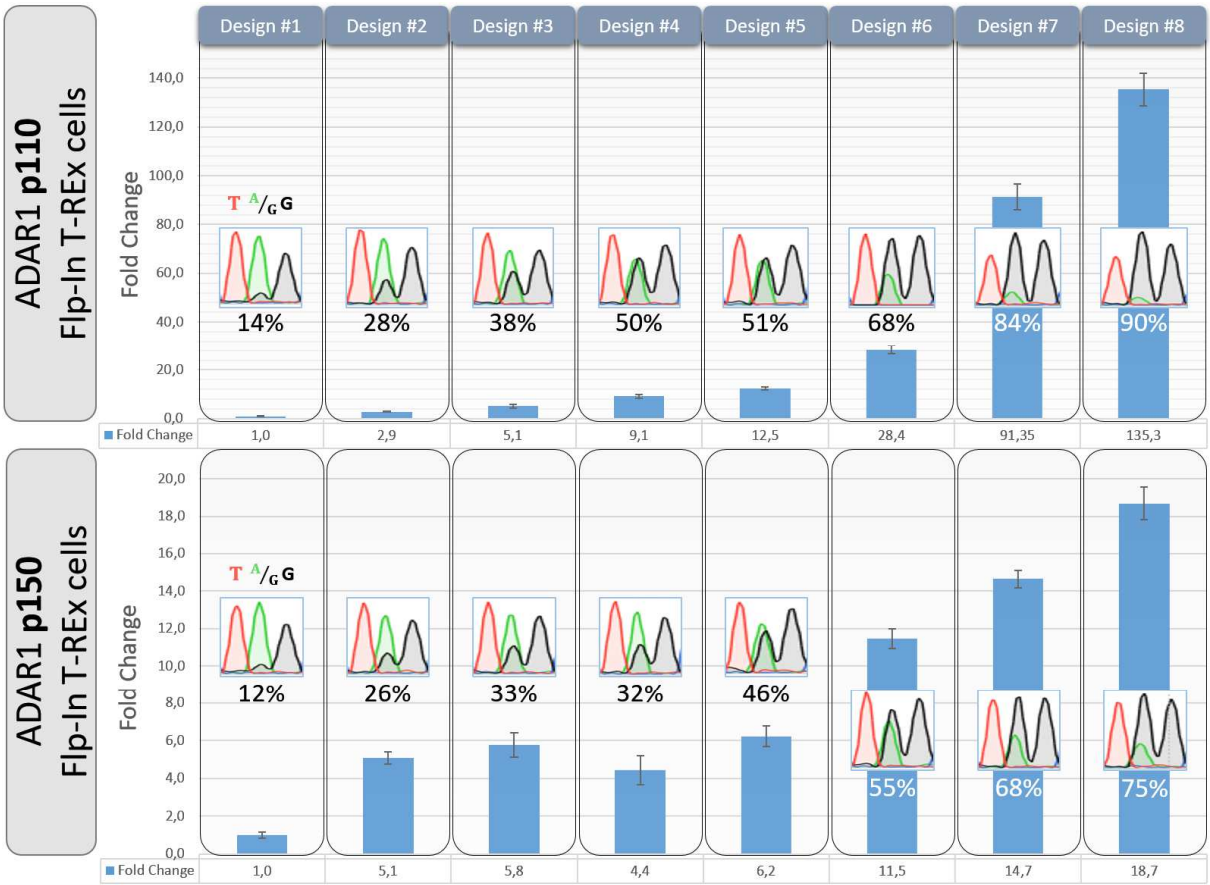


Figure 3-23: Description of R/G gRNA design improvements performed throughout this thesis.

(A) Simplified model of an R/G gRNA. It contains a dsRNA R/G motif at its 5' end to recruit ADAR enzyme via their dsRBD and a single-stranded antisense part at its 3' end, which allows target recognition by simple Watson-Crick base pairing. (B) Different R/G gRNA motifs generated by rational design. Red-circled bases indicate differences to the previous version. (C) Different types of antisense parts also generated by rational design. The grey RNA sequence represents the mRNA target. The green RNA represents the antisense parts of the R/G gRNAs. The red-circled adenosine is the editing target site. The opposite blue circled nucleotide is the counterbase (often a cytosine). The number below the target adenosine and counter base mismatch indicates the editing position.

Figure 3-23 continued: Description of R/G gRNA design improvements performed throughout this thesis.

The numbers left and right of the green antisense part indicate its length in nucleotides counted starting at the R/G motif. In the most advanced design, the antisense part was further differentiated into the "guide sequence", describing the gRNA-mRNA duplex containing the target site, and into "recruitment clusters" (RCs), which do not contain editable adenosines. The range displayed above each recruitment cluster indicates an exemplary length range, which is defined in the cluster selection algorithm (e.g. 11-16 nt) and can vary depending on the input variables. The range above each loop represents the distance between the cluster target sites on the target mRNA and can also vary depending on the input variables. The adverse sequence within R/G gRNA contains a 3-nt-long adenosine linker between the recruitment clusters. All exemplary sequences displayed are targeting the dual-luciferase reporter system.



Editing Position	Pos 6	Pos 8	Pos 8	Pos 8	Pos 8	Pos 30	Pos 8	Pos 8
R/G-Version	#1	#1	#20	#20	#21	#21	#21	#21
Antisense Length	16 nt	16 nt	16 nt	20 nt	20 nt	40 nt	20 nt + RC	20 nt + RC
Recruitment Cluster	None	None	None	None	None	None	3x (11-16 nt)	8x (11-16 nt)

Figure 3-24: Comparison of the most successful incremental R/G gRNA improvements performed throughout this thesis.

An explanation of the various modified design parameters (editing position, R/G version, antisense length, and recruitment clusters) is shown in Figure 3-23. The used readouts were the editing at RNA level (double peak at the target 5'-UAG-3' triplet in the shown Sanger sequencing reads) and the relative fold change at protein level, normalized based on the original R/G gRNA design #1 (dual-luciferase assay). The only difference between design #7 and #8 is an increase in the used number of recruitment clusters. For the Sanger sequencing, $2.5x10^5$ ADAR1 p110 or p150 Flp-In T-REx cells per well were seeded in 24-well scale on PDL coated plates. The cells were induced with 10 ng/ml doxycycline at seeding. 24h post seeding, at 70-90% confluence, the cells were transfected with 300 ng dual-luciferase reporter plasmid and 1300 ng R/G gRNA plasmid, using Lipofectamine 2000 in a 1:3 ratio. The cells were harvested 72h post transfection. This was followed by RNA isolation with the Quiagen RNeasy Mini Kit, DNase-I digestion, RT-PCR, and Sanger sequencing.

Figure 3-24 continued: Comparison of the most successful incremental R/G gRNA improvements performed throughout this thesis.

For the dual-luciferase assay, $4x10^4$ ADAR1 p110 or p150 Flp-In T-REx cells per well were seeded in 96-well scale on PDL coated plates. The cells were induced with 10 ng/ml doxycycline at seeding. 24h post seeding, at 70-90% confluence, the cells were transfected with 60 ng dual-luciferase reporter plasmid and 260 ng R/G gRNA plasmid, using Lipofectamine 2000 in a 1:0.8 ratio. 72h post transfection, the dual-luciferase assay was performed using OD-None and 7 µl of the lysate per well of the LumiNunc 96-well plate. 30 µl 1xPLB per well had been used to harvest each well of the 96-well cell culture plate. The dual-luciferase assay was performed with triplicates for each sample, each one measured with two technical replicates.

The concept is that the antisense part of the R/G gRNA binds to the target mRNA by simple Watson-Crick base pairing. This allows the ADAR enzymes to be recruited to the target mRNA by binding to the double-stranded R/G motif with their dsRNA-binding domains. This process positions the deaminase domain of the respective ADAR in a way that allows the deamination of a mRNA target adenosine within the antisense-part-mRNA-duplex.

As explained earlier, both the 5' and 3' nearest-neighbor preference, as well as the counterbase preference, of the ADAR enzymes play an important role regarding the achievable editing yield at the target site (section 1.1.2). Therefore, the optimal 5'-UAG-3' triplet and a cytosine as counterbase were used for most of the optimization experiments.

The initial R/G motif, version 1 (Figure 3-23, B), was based on the naturally occurring R/G motif found in the GluR2 mRNA (Figure 1-12). The first R/G gRNA design consisted of the R/G motif version 1 and a 16 nucleotides long antisense part with the editing site positioned 6 nucleotides away from the R/G motif (Figure 3-23, B, C). While the used editing position of this first R/G gRNA design was based on the distance of the naturally occurring R/G motif to its target adenosine in cis (Figure 1-12), the antisense part length was optimized by Dr. Wettengel using ADAR2 plasmid overexpression settings in HEK-293T cells (Pub. 2, Figure S10).

It allowed a correction of the W417X amber non-sense mutation in the dual-luciferase reporter mRNA between 12% (ADAR1 p110 Flp-In T-REx cells) and 14% (ADAR1 p150 Flp-In T-REx cells) at RNA level. Design #1 was the starting point for all optimizations and thus provided the reference value, which was used to normalize and compare all other designs, when performing a dual-luciferase reporter assay, that measures editing success at the protein level (Figure 3-24). It is important to highlight that the transfection settings used for the two readouts displayed in Figure 3-24 were different.

To measure editing success at the RNA level via Sanger sequencing, the ADAR1 Flp-In T-REx cells were always treated with much harsher conditions (µg plasmid to µl Lipofectamine 2000 ratio of 1:3), to find the highest achievable editing yield. When transfecting ADAR1 Flp-In T-REx cells for the dual-luciferase assay, the cells were treated with milder conditions (µg plasmid to µl Lipofectamine 2000 ratio of

1:0.8), to reduce the amount of plasmid uptake. This should render the amount of R/G gRNA more scarce and allow therefore to screen for R/G gRNAs with high potency.

3.12.1 Design #2 – Shifting the editing position within the gRNA-mRNA duplex

As a first optimization the position of the editing site within the antisense-part-mRNA duplex was varied (Figure 3-23, C), leading to R/G gRNA design #2 (Figure 3-24). It was assumed that the deaminase domains of ADAR enzymes might have some kind of optimal arrangement when sitting on top of the duplex of antisense part and mRNA. Thus, several positions, ranging from position three to nine, were tested comparatively. These experiments were performed in cooperation with Dr. Wettengel. While she tested different positions using the eGFP W58X amber construct as target (Pub. 2, Figure S11), I used the PINK1 R407Q construct (Pub. 2, Figure S15). Together, we concluded that position 8, corresponding to 7 intervening nucleotides between the R/G motif and the editing site, resulted in the highest editing yield. When targeting the dual-luciferase reporter mRNA, this optimization resulted in editing yields between 26% (ADAR1 p110 Flp-In T-REx cells) and 28% (ADAR1 p150 Flp-In T-REx cells) at RNA level, doubling the results obtained with R/G gRNA design #1. At the protein level, this optimization resulted in a three- to five-fold increase (ADAR1 p110 and ADAR1 p150 Flp-In T-REx cells, respectively) of restored firefly luciferase protein, compared to R/G gRNA design #1 (Figure 3-24, design #2). A crystal structure of the ADAR2 deaminase domain (see Figure 1-6) published after these experiments had been performed, suggested that a target adenosine at position 9 was sterically preferable. This was very near to the experimentally determined optimal editing position 8 and explained why much higher or lower editing positions worked less efficiently.

3.12.2 Design #3 – Rational design changes to the R/G Motif

Previous *in-vitro* experiments performed by Dr. Wettengel (Pub. 3, Figure 2) had shown that the R/G motif itself can be the target for deamination by ADARs. This resulted in the idea that auto-editing of the AU-rich R/G motif of version 1 might be destabilizing its secondary structure in cell culture, too, and that this might result in lowered editing yields. Thus, novel R/G motifs with increased G/C content were designed (Pub. 3, Figure 3A). The R/G motif version 2 (Pub. 3, Figure 2) only contained base exchanges that prevented editing at sites previously found *in-vitro* to show the strongest auto-editing.

Auto-editing in version 2 was attempted to be prevented either by base substitution of potential off-target adenosines with other nucleotides, or by generation of hard-to-edit 5’-GAN-3’ triplets through exchange of selected nucleotides with guanosines, as well as a combination of both. In case of version 3 and 4, an increasingly excessive base substitution for G/C base pairs was performed (Pub. 3, Figure 3A). Building on this initial experiments, I supervised Madeleine Heep, who worked on the topic of novel R/G motifs to obtain her Bachelor’s degree [235].

Ms. Heep could show that auto-editing was also happening *in-vivo*. R/G motifs embedded into the 3' UTR of eGFP showed 16% editing at adenosine 8 when performed in ADAR2 Flp-In T-REx cells (Yellow circle), and 8% editing at adenosine 13 when performed in ADAR1 p110 Flp-In T-REx cells (Blue circle) (Pub. 3, Figure 3A, R/G version 1). Similar experiments with the R/G versions 2-4, also in *cis*, showed that even version 2 could already completely prevent auto-editing within the R/G motif (Pub. 3, Figure 3A, version 2-4). The *trans* R/G gRNA with motif version 2 lead to a slight increase of on-target editing, in ADAR1 p110 Flp-In T-REx cells, and to no change in ADAR2 Flp-In T-REx cells. Version 3 and 4 caused no further improvements in ADAR1 p110 Flp-In T-REx cells, but nearly halved editing in ADAR2 Flp-In T-REx cells (Pub. 3, Figure 3B).

Several other versions were tested by Ms. Heep during her Bachelor's degree project [235]. An increase in editing yield similar to the one achieved by increasing the G/C-content, was accomplished in ADAR1 Flp-In T-REx cells by extension of the first helix (H1) of the R/G motif (Figure 3-23B, Version 1, H1). The initial idea had been to generate more contact surface for the second dsRBD of ADAR2 to efficiently bind to the R/G motif. As mentioned earlier dsRBDs need between 11 and 16 base pairs contact surface to efficiently bind to an dsRNA (section 1.1.2.2.1). When the original R/G motif version 1 was designed based on the natural substrate of the GluR2 mRNA, it was cut between the two guanosines 5 and 6 nt downstream of the editing site (Pub. 2, Figure 1A). Thereby, the assumed contact surface with the dsRBD2 of ADAR2 was cut in half [103]. Several extensions of different length were tested by Ms. Heep. Extensions of 2-6 bp did not improve the editing yield in ADAR2 Flp-In T-REx cells, but did surprisingly improve the editing yield in ADAR1 Flp-In T-REx cells (see Figure 3.27 in [235]).

By combining the stabilizing base substitutions from version 2 (Pub. 3, Figure 3B, Version 2) and the five base pair extension of the first helix, the R/G motif version 20 was created. The R/G motif version 20 (Figure 3-23, B), combined with a 16 nt antisense part, and with the target adenosine at position 8 (Figure 3-23, C), resulted in the design #3. It considerably improved the editing yield at RNA and protein level. When targeting the dual-luciferase reporter mRNA the R/G gRNA design #3 resulted in editing yields between 33% (ADAR1 p150 Flp-In T-REx cells) and 38% (ADAR1 p110 Flp-In T-REx cells) at RNA-level, better than both design #1 and design #2. At protein level, this optimization resulted in a five- to six-fold increase (ADAR1 p110 and ADAR1 p150 Flp-In T-REx cells, respectively) of restored functional firefly luciferase protein, compared to R/G gRNA design #1 (Figure 3-24, design #3). Due to my contributions to the development of the R/G motif version 20, used in RG gRNA design #3, I became co-inventor of patent application PCT/EP2018/067718.

3.12.3 Design #4 – Deductions from the ADAR2 deaminase domain crystal structure

A nuclease footprinting study showed that the ADAR2 deaminase domain covers 23 bp of dsRNA in an asymmetric fashion around the editing site. It was assumed to cover 18 nt on the 5' side and ~5 nt on

the 3' side [236]. The crystal structure of the ADAR2 deaminase domain, published in 2016, showed that it was in contact with the dsRNA over a distance of 20 nt [1]. The residues N473 and K475 facilitate the contact to the 5’ end of the strand that contains the edited adenosine, while the residues R348 and K594 contact the 5’ end of the unedited strand (Figure 1-6C). In the same publication, it was shown that changes of the amino acids R348 or K594 to alanine were resulting in a decrease of the editing yield by more than one order of magnitude (Figure 6C in [1]), revealing that these interactions were extremely important for efficient editing. Referring to our editing system, the strand that is highlighted in red in Figure 1-6C would represent the target mRNA, while the blue strand would represent the antisense part of the R/G gRNA. Consequently, these structural features of the ADAR2 deaminase domain imply that the target-adenosine-containing duplex, consisting of the target mRNA and the antisense part of the R/G gRNA, should be at least 20 bp long to allow efficient binding, utilizing all possible contact points between both interaction partners.

The transition of the project towards ADAR1 posed the question, whether the structural ADAR2 observations could also be applied to ADAR1, as ADAR1 crystal structures were not available. When aligning the protein sequences of the deaminase domains of human ADAR1 (UniProtKB - P55265) and ADAR2 (UniProtKB - P78563), using “NCBI Align Sequences Protein BLAST”, the sequence identity was 37%, close to the ~40% guidance value defined by the enzyme commission as likely to result in functional similarity [237]. Thus, ADAR1 might function in a similar way and prefer a duplex that was at least 20 bp long. The requirement to assess an extended antisense part was further supported by emerging hypotheses about the preference of ADAR1 for longer duplexes 5’ of the target adenosine, caused by its specific 5’ binding loop (section 1.1.2.2.1).

Previous experiments by Dr. Wettengel showed that antisense parts longer than 16 nt did not further benefit the editing yield, and they could even become disadvantageous above 20 nt (Pub. 2, Figure S10). While this was correct for the settings used at the time, HEK-293T cells with ADAR2 overexpressed from a plasmid, the correlation may change under different conditions.

To find out, whether an antisense part extension might have a positive effect on the editing yield, a R/G gRNA of design #3 with a 16 nt long antisense part was compared to the new R/G gRNA design #4 with a 20 nt long antisense part, using ADAR1 Flp-In T-REx cells that express ADARs to a much lower degree. Regarding their R/G motif version and the editing position, both designs were identical. The result of this slight change was a strong positive effect on the editing yield, when using ADAR1 p110 Flp-In T-REx cells, but interestingly a minimally negative effect when using ADAR1 p150 Flp-In T-REx cells. As ADAR1 p110 and p150 are identical, regarding their deaminase domains, but differ in their N-termini (Figure 1-2), this difference might be connected to the N-terminal Z-DNA/RNA binding domain of ADAR1 p150. The editing yield at RNA level in ADAR1 p150 Flp-In T-REx cells was 38%, using design

#3, relative to 50%, when using design #4. Compared to the original design #1, the new design #4 resulted in a 9-fold improvement of restored firefly luciferase activity on protein level. In ADAR1 p150 Flp-In T-REx cells, the editing at RNA level was similar between design #3 and #4, with 33% and 32%, respectively. At protein level, when comparing design #3 and #4, the new design #4 resulted in a slight reduction of restored firefly luciferase activity from 5.8- to 4.4-fold, compared to design #1 (Figure 3-24).

3.12.4 Design #5 – Minor R/G motif base-exchanges improve ADAR1 p150 recruitment

My co-worker, Tobias Merkle, who had used a variant of the R/G motif version 20 for experiments with IVT and chemically modified ASO R/G gRNAs, had introduced a small change at the start of the motif, to improve its *in-vitro* transcription by the used T7 promoter. In his nomenclature, this variant was named ASO v9.4 (see Pub. 4, Figure 1B). He swapped the second and the third base pairs in helix 1 (H1) of the R/G motif duplex from 5'-GUG to 5'-GGU, and 3'-CAC to 3'-CCA, respectively (see Figure 3-23). Two guanines at the beginning of a T7 transcript resembled the optimal T7 sequence more closely and increased the produced R/G gRNA yield [238]. Out of curiosity for the effects of such slight changes, the R/G motif version 21 was created, mimicking this swap in the encodable R/G gRNA system. If the nucleotide change would generally influence the expression level of the R/G gRNA from the U6 promoter *in-vivo*, similar to the T7 promoter *in-vitro*, then a general increase in editing could be expected. This was not the case, as ADAR1 p110 and ADAR1 p150 reacted differently to the change. The editing yield at RNA level in ADAR1 p110 Flp-In T-REx cells was 51% when using design #5 with the R/G motif version 21, and thus nearly identical to the 50% achieved with design #4. The increase was also slim at the protein level, changing from 9-fold (design #4) to 12-fold (design #5), compared to design #1. When using ADAR1 p150 Flp-In T-REx cells, the editing at RNA level went up from 32% to 46%. At protein level, it changed from 4.4-fold to 6.2-fold, relative to design #1. Because the swap resulted in an overall improvement, the R/G motif version 21 was used for all further designs.

3.12.5 Exemplified limitations of unmodified antisense parts

Despite all the mentioned design improvements, which were analyzed in ADAR1 Flp-In T-REx cells, it had been impossible to achieve editing above the detection limit of Sanger sequencing when recruiting endogenous ADARs in other cell lines, e.g. in HeLa cells. At least, when the encodable R/G gRNA-system was used, and not the chemically modified one.

Adenoviral transduction, optimized in regards of MOI and time point, as well as transcript copy-number comparison and localization-change of R/G gRNAs did not solve this issue (see Figure 3-16, Figure 3-17, Figure 3-21, Figure 3-22). This narrowed the potential causes for the superior editing of chemically modified R/G gRNAs down to their modifications, something that could not directly be reproduced in an encodable R/G gRNA system.

The 2'-O-Me modifications were known to improve potency, nuclease resistance and overall pharmacokinetic and pharmacodynamic properties. Importantly, they conferred RNAs into their most energy-favorable conformation, the C3'-endo (N-type) sugar pucker, increasing their Watson-Crick binding affinity [239].

In contrast to most of the properties mentioned above, the Watson-Crick binding affinity could also be increased in an encodable R/G gRNA system, through further extension of the antisense part. But while a 20 nt long antisense part represented the optimal length for full interaction between the ADAR2 (and maybe also the ADAR1) deaminase domain and the gRNA-mRNA duplex, a substantial further extension can introduce several problems.

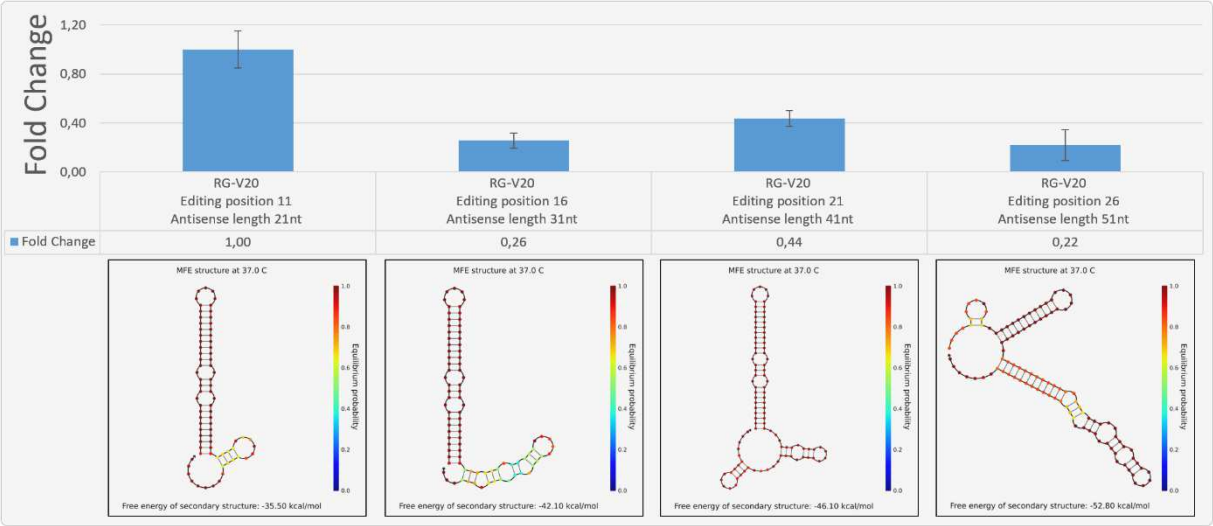


Figure 3-25: Comparison of R/G gRNAs with increasing antisense part length, regarding their secondary structure and the archived relative editing yield on protein level, using the dual-luciferase reporter system.

The decrease in relative editing in this example was partially caused by the suboptimal positioning of the target adenosine, which was always placed in the center of the antisense part. Anyway, the relative editing seems to be impaired when the antisense part length is increased. The secondary structure of the antisense part also increased in strength (base pairs with high equilibrium probability are indicated by their color) and/or length, thereby masking the hybridization surface for interaction with the target mRNA. The 51 bp long antisense part even misfolded the entire R/G motif with high probability. In this experiment, $4x10^4$ ADAR1 p150 Flp-In T-REx cells per well were seeded in 96-well scale. The cells were induced with 10 ng/ml doxycycline at seeding. 24h post seeding, at 70-90% confluence, the cells were transfected with 60 ng dual-luciferase reporter plasmid and 260 ng R/G gRNA plasmid using Lipofectamine 2000 in a 1:0.8 ratio. 72h post transfection, the dual-luciferase assay was performed using OD-None and 7 µl of the lysate per well of the LumiNunc 96-well plate. 30 µl 1xPLB per well had been used to harvest each well of the 96-well cell culture plate. The experiment was performed as biological triplicate with one technical measurement per sample. The R/G gRNA secondary structures were generated using Nupack [197]. Background information on the used R/G motif version 20 and a general explanation of antisense parts can be found in Figure 3-23.

With increasing length, the single-stranded antisense part can become more likely to form intramolecular secondary structures by folding back to itself, becoming partially double-stranded and thereby masking the contact surface to the target mRNA. Above a certain length threshold, there is even a

chance that the whole R/G motif might become misfolded by forming an unwanted secondary structure with the extended antisense part. This was exemplified in Figure 3-25. The impact of this effect depends strongly on the individual R/G gRNA sequence.

Additionally, a conventionally extended antisense part does not allow rational design, as it is limited to the nucleotides that come next in line on the corresponding target mRNA. If further adenosines in the target mRNA would be covered by this elongated antisense part, it could lead to unwanted off-target editing within the resulting duplex (see Figure 3-28). The extended antisense part might also bind to other targets in the transcriptome and facilitate off-target editing in completely different mRNAs, which could not be avoided without the possibility of rational design.

3.12.6 Recruitment clusters – Novel *in-silico*-optimized antisense parts

To be able to increase the R/G gRNA affinity, while trying to circumvent the issues mentioned above, I devised a novel type of antisense part, which could be optimized *in-silico*. After the desired functionalities had been outlined, Mr. Nicolai Wahn, a bioinformatics expert, contributed decisively to the project by programming the tool that enabled these *in-silico* optimizations. On top of the basic implementation, he improved the program with a heuristic search function.

The new architecture consisted of a 20 nt antisense part, hence named the “guide”-sequence, which was connected to the R/G motif at its 5’ end and to several smaller *in-silico*-optimized antisense parts at its 3’ end (Figure 3-23C). These *in-silico*-optimized parts are hence called “recruitment clusters” (RCs). RCs bind to the target mRNA, like the guide sequence, but were designed to target only sequences that do not contain editable adenosines. The RCs are connected to each other and to the guide sequence by 3 nt long adenosine linkers, and were selected *in-silico* using the custom built “recruitment cluster finder” (RCF) tool.

Preliminary experiments with hand-selected RCs, showed no or only a slight increase in editing (Figure 3-26). This might have been caused by strong secondary structures as well as contact surface masking, which could not be prevented efficiently, when handpicking the RCs. Therefore, the *in-silico* approach was introduced, which then lead to significant improvements (Figure 3-24, design #7 and #8).

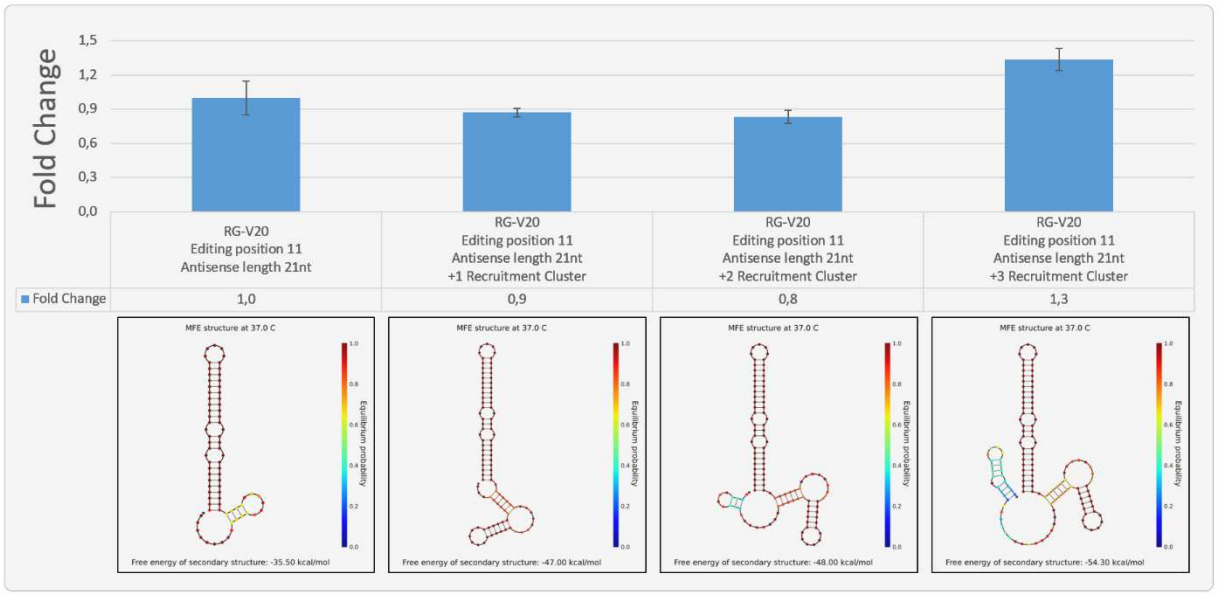


Figure 3-26: Comparison of a regular R/G gRNA with variants containing one to three hand-selected recruitment clusters.

The designs with +1 or +2 RCs have a secondary structure with higher probability (orange to red) in their antisense part, compared to the design without RCs (yellow). Both designs show slightly lower editing. In comparison, the secondary structure of the design with +3 RCs has a longer single-stranded section within its antisense part, as well as a lower probability for strong secondary structures at the 3' end (blue). It shows slightly increased editing. In contrast to later designs, the early R/G gRNAs used in this experiment contained the R/G motif version 20, had the editing site located at position 11, and contained a 21 nt long antisense part. In this experiment, $4x10^4$ ADAR1 p150 Flp-In T-REx cells per well were seeded in 96-well scale. The cells were induced with 10 ng/ml doxycycline at seeding. 24h post seeding, at 70-90% confluence, the cells were transfected with 60 ng dual-luciferase reporter plasmid and 260 ng R/G gRNA plasmid using Lipofectamine 2000 in a 1:0.8 ratio. 72h post transfection, the dual-luciferase assay was performed using OD-None and 7 µl of the lysate per well of the LumiNunc 96-well plate. 30 µl 1xPLB per well had been used to harvest each well of the 96-well cell culture plate. The experiment was performed as biological triplicate with one technical measurement per sample. The R/G gRNA secondary structures were generated using Nupack [197]. Background information on the used R/G motif version 20 and a general explanation of antisense parts can be found in Figure 3-23.

The function of the "recruitment cluster finder" (RCF) tool is to scan an input DNA, corresponding to the target pre-mRNA/mRNA, for clusters of uninterrupted sequences that contain only guanosine, thymidine, cytosine, but not adenosine, except if the adenosine sat 3' of a directly adjacent guanosine.

Allowing such an adenosine was possible because the 2-amino group of the guanosine sterically interferes with the activity of the ADAR deaminase domain, lowering and in most cases preventing editing of the adenosine in question [1]. The specific arrangement of adenosines in codon contexts inefficient for RNA editing had already worked before, e.g. in the prevention of R/G gRNA autoediting of R/G motif version 2 (Pub. 3, Figure 3A).

Including these specific adenosines in the search function increased the possible number of detected RC, as not every RC sequence containing an adenosine had to be filtered out.

A conceptual description of how the recruitment cluster finder (RCF) selects RCs is displayed in Figure 2-1. The theoretical benefits of these *in-silico*-optimized RCs are the following:

1. The selection of sequences as RCs, which are free of editable adenosines (only G, T, C and 5'-GA), allowed a controlled increase of the "gRNA-to-mRNA" affinity by adding the RCs as 3' extensions to an R/G gRNA with a constant, 20 nt long guide sequence. The sequence composition of the RCs rendered off-target editing within the target mRNA nearly impossible. An increase in antisense length should consequently also increase specificity for the target mRNA.
2. The availability of several attachment points (RCs) to spatially remote locations on the target mRNA, compared to only one attachment point in the former design, could increase the likelihood of spatial gRNA/mRNA convergence.
3. By using a brute force combinatorial approach, the RCs could be *in-silico*-selected for minimal secondary structure within the total single-stranded part of the R/G gRNA (RCs and guide-sequence). RCs which might form potential secondary structures with the R/G motif could also be excluded. This could circumvent the masking and misfolding problems mentioned earlier.
4. The clusters could be defined in their length. This allowed the guide sequence to be 20 nt long, which was assumed optimal for deaminase domain binding, while the RCs could be as short as 11 nt, which was the minimal length of a dsRNA to be still recognized by ADAR dsRBDs. Therefore, off-target editing within the RCs was not only minimized by the sequence composition (G, T, C, 5'-GA), but also sterically by creating an inferior substrate for deaminase domain binding, while editing within the guide sequence was not hampered. Consequently, this should allow editing-inducer-element-like (section 1.1.2.1) local accumulation of ADARs by creating a substrate for dsRBD binding (RCs), while preventing editing within these RCs by sequence composition and length.
5. The *in-silico*-selected RCs could additionally be blasted to predict if they might potentially bind to other mRNAs in the transcriptome and also if the resulting duplex would contain editable adenosines. RCs found to allow off-target editing in non-target mRNAs, could thus already be excluded in the R/G gRNA design process.
6. The distance of the RCs to each other on the mRNA could be defined by simple inputs into the "recruitment cluster finder". This could be of potential interest for editing with ADAR1p150 in the cytosol, where gRNAs have to compete with ribosomes, which are assumed to constantly detach gRNAs from their target mRNAs when targeting the ORF. Longer distances between RCs could potentially allow the gRNA to stay on target in such situations, due to several spatially separated attachment points.

Although many of these theoretical ideas sound intriguing, some of them will remain speculative until the extent of transcriptome-wide off-target editing in this system has been evaluated using Next-generation-sequencing (NGS). Unfortunately, this has not yet been done at the time of submitting this thesis. However, significant improvements in editing yield and a decrease in off-target editing within the gRNA-mRNA duplex, compared to extended antisense parts not optimized *in-silico* as described above, could already be shown and will be discussed shortly.

To characterize the new RC system, several types of RCs were calculated using the RCF. The tool allows a very granular control over all components of the calculated R/G gRNA, like the used R/G motif, the guide sequence, the linker length and sequence, as well as potential 3'-terminal nucleotides (like the three uridines, which are a normal residue of the U6 termination signal).

All of these can be customized and are then included into the calculation. To test the optimal RC size and distance between the RCs on the target mRNA, several RC types were compared, targeting the dual-luciferase reporter system in ADAR1 Flp-In T-REx cells. Sanger sequencing and the dual-luciferase assay were used as readout again.

The cluster size range input for the RC type A, B, E and F were 11-16 nt, resembling the size optimal for dsRBD. The cluster size range input for the RC type C and D were 18-22 nt, resembling the size of miRNAs. The RC set A-E contained three RCs, one for each dsRBD of ADAR1. The set F contained eight RCs thereby providing multiple dsRBD attachment points similar to an EIE. The distance range inputs (distance of the clusters to each other and to the guide sequence on the reporter plasmid or gene in bp) were also selected individually. The following input ranges were used: Type E 10-100 bp, Type C 10-250 bp, Type F 10-500 bp, Type B 100-250 bp, Type D 100-500 bp, Type A 250-500 bp. For details on cluster size, cluster-to-cluster distance, achieved editing at RNA- and protein-level, as well as secondary structures, see Figure 3-27.

The results in Figure 3-27A and B show that short clusters in close vicinity (Type E) are superior to short clusters with larger cluster-to-cluster distance (Type A and B). This can be compensated to some degree by an increase in cluster size (Type C and D). At least in ADAR1 Flp-In T-REx cells, the addition of more than three clusters (Type F) does improve the editing yield even further, especially in situations with decreased R/G gRNA quantities, like the dual-luciferase assay with mild transfection conditions (Figure 3-27A). Figure 3-27C shows how efficiently the RCF selects RCs with extremely weak and short secondary structures within the antisense part, compared to the hand-selected RCs displayed in Figure 3-26.

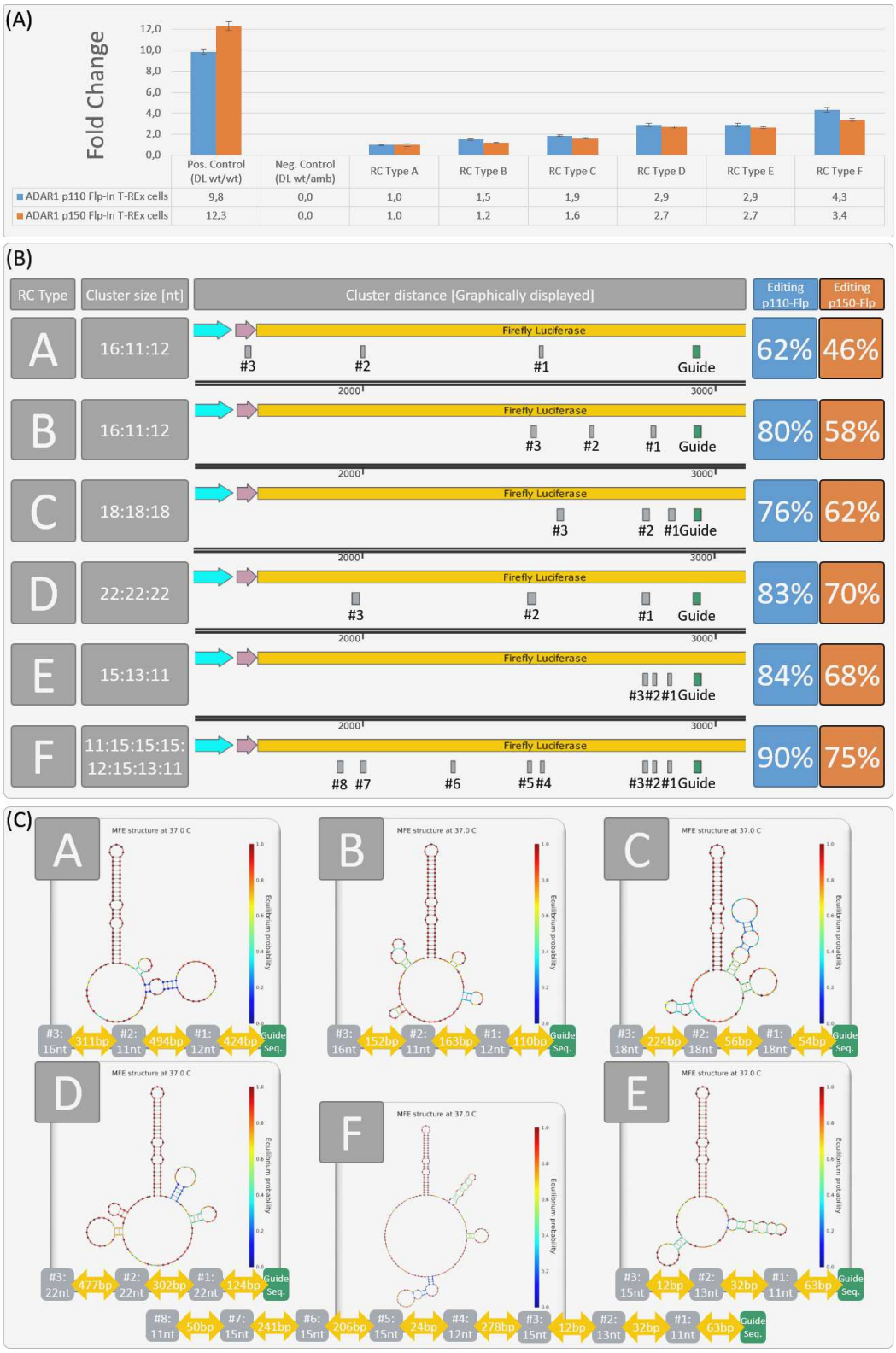


Figure 3-27: Characterization of recruitment clusters differing in size, distance, and number.

(A) Relative fold change at protein level in ADAR1 p110 or ADAR1 p150 Flp-In T-REx cells, quantified using the dual-luciferase reporter assay. (B) Editing at RNA level in ADAR1 p110 or ADAR1 p150 Flp-In T-REx cells, quantified using Sanger sequencing. The cluster size in nt is displayed on the left. 16:11:12 means cluster #3 is 16 nt long, cluster #2 is 11 nt long, and cluster #1 is 12 nt long. The distance between the clusters on the reporter plasmid (dual-luciferase reporter) is displayed graphically in the middle. (C) Secondary structure of the different recruitment-cluster-containing R/G gRNAs, as well as their exact cluster size in nt (grey), distance between the clusters (yellow), and distance to the guide sequence (green) on the reporter plasmid in bp.

Figure 3-27 continued: Characterization of recruitment clusters differing in size, distance and number.

For the dual-luciferase reporter assay readout in panel A, $4x10^4$ ADAR1 p110 Flp-In T-REx cells per well were seeded in PDL-coated 96-well plates. The cells were induced with 10 ng/ml doxycycline at seeding. 24h post seeding, at 70-90% confluence, the cells were transfected with 60 ng dual-luciferase reporter plasmid and 260 ng R/G gRNA plasmid using Lipofectamine 2000 in a 1:0.8 ratio. 72h post transfection, the dual-luciferase assay was performed with biological triplicates for each sample, which were then measured as technical duplicates. For the Sanger sequencing readout in panel B, $2x10^5$ ADAR1 p110 Flp-In T-REx cells per well were seeded in PDL-coated 24-well plates. The cells were induced with 10 ng/ml doxycycline at seeding. 24h post seeding, at 70-90% confluence, the cells were transfected with 300 ng dual-luciferase reporter plasmid and 1300 ng R/G gRNA plasmid using Lipofectamine 2000 in a 1:3 ratio. 72h post transfection, the cells were harvested. This was followed by RNA isolation with the Quiagen RNeasy Mini Kit, DNase-I digestion, RT-PCR, and Sanger sequencing. The R/G gRNA secondary structures were generated using Nupack [197].

3.12.7 Design #6, #7 and #8 – Comparing the novel *In-silico*-optimized recruitment clusters to a reference design for conventionally extended antisense parts

For the editing of endogenous ORF targets, my co-worker Tobias Merkle had created a R/G gRNA with an antisense part extended by 22 nt between his original 18 nt long antisense part and the R/G motif (Pub. 4, Figure 3A). The target adenosine was arranged at position 30 of the gRNA-mRNA duplex, which had a total size of 40 nt (Figure 3-23C). This R/G gRNA was an example for a well-working, conventionally extended R/G gRNA. The chemical modifications he incorporated into this R/G gRNA, which prevented some of the above-mentioned shortcomings of long antisense parts (e.g. gRNA-mRNA duplex off-target editing), however, could not be transferred to the encodable R/G gRNA system. Nevertheless, it was a suitable reference design to compare the efficiency of the novel RC-containing R/G gRNAs with that of a functional R/G gRNA, which had been conventionally extended. A plasmid-encoded variant of this design, called design #6 and targeting the firefly luciferase W417X amber mRNA, was then compared to the two best-performing, RC-containing R/G gRNAs created so far, design #7 (3 RCs, RC Type E) and #8 (8 RCs, RC Type F).

Although design #6 achieved good editing yields of 68% in ADAR1 p110 Flp-In T-REx cells, and 55% in ADAR1 p150 Flp-In T-REx cells, the new RC-containing R/G gRNAs clearly performed better. Design #7 reached 84% (ADAR1 p110) and 68% (ADAR1 p150), and design #8 even achieved 90% (ADAR1 p110) and 75% (ADAR1 p150) editing at RNA level. At protein level, using the mild transfection conditions that had also been previously used for the dual-luciferase assay, this difference was even more distinct.

While design #6 restored ~30-fold (ADAR1 p110) and ~11-fold (ADAR1 p150) more firefly luciferase activity than the original design #1, design #7 caused a ~90-fold (ADAR1 p110) and ~15-fold (ADAR1 p150) improvement. Design #7 was still outperformed by design #8, which restored ~135-fold (ADAR1 p110) and ~19-fold (ADAR1 p150) more firefly luciferase activity than the original design #1 (Figure 3-24). The designs #7 and #8 were especially successful in the recruitment of ADAR1 p110, which is the most abundant and constitutively expressed ADAR in the human body.

In addition to the high editing yields of the recruitment cluster R/G gRNAs, it could be shown that they caused less off-target editing within the gRNA-mRNA duplex (Figure 3-28). While the conventionally extended design #6 R/G gRNA caused off-target editing at three sites, up to 16%, the design #7 caused no off-target editing within the RC-mRNA duplex and only weak off-target editing (7%) at one position within the guide sequence. The first three RCs of design #8 were also off-target editing free. The clusters 4-8 were not analyzed. Within the guide sequence of design #8, there was one off-target site again, in this case with 16% conversion. However, this off-target editing was accompanied by 90% on-target editing, compared to only 68% for design #6, and 84% for design #7 R/G gRNAs.

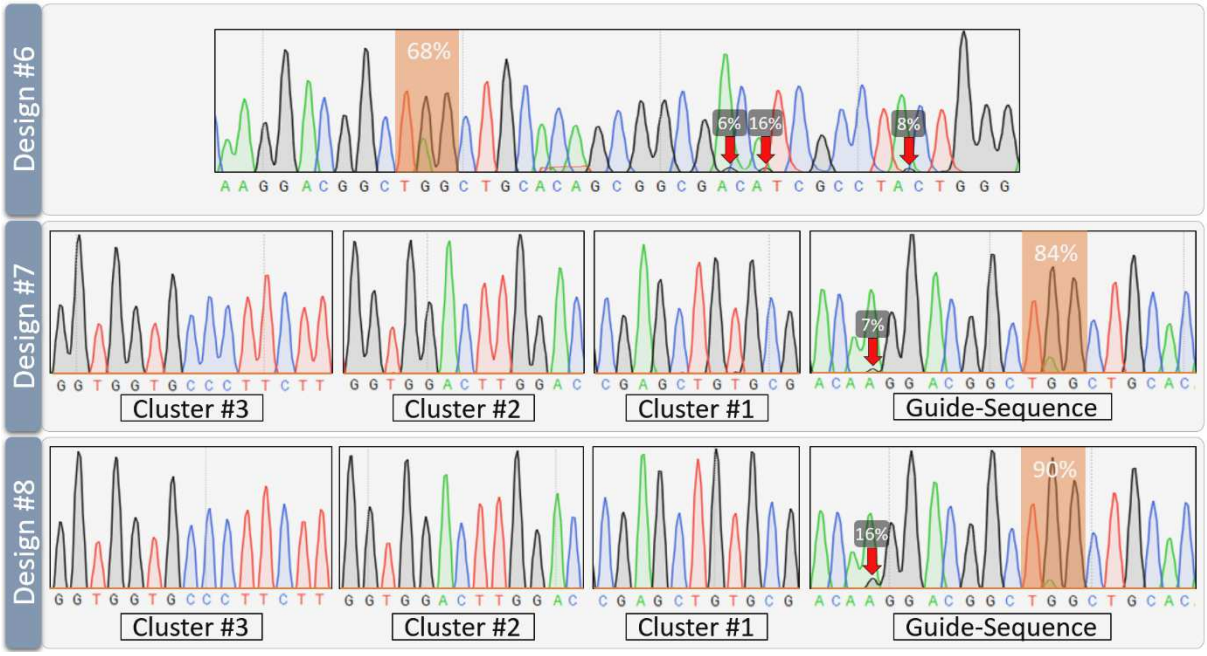


Figure 3-28: Sanger sequencing reads of the R/G gRNA designs #6, #7, and #8, comparing their off-target editing within the antisense part.

These are extended excerpts of the sequencing reads in Figure 3-24. The background was removed. The 40 nt long antisense part of R/G design #6 causes up to 16% off-target editing at three different sites within the gRNA-mRNA duplex, while reaching only 68% editing at the target site. Design #7 shows no off-target editing within the recruitment clusters. Design #8 shows also no off-target editing within the first three recruitment clusters. The clusters 4-8 were not analyzed. The sequences between the RC binding sites were also evaluated and showed no detectable off-target editing (data not shown). Both, design #7 and #8, show off-target editing within the guide sequences, but only at one site and, regarding design #7, to a very low degree. Design #7 and #8 show greatly improved on-target editing yield, compared to design #6.

3.13 The R/G gRNA designs #6, #7, and #8 allow the efficient recruitment of endogenous human ADAR1

Due to the high efficiency of the novel recruitment cluster R/G gRNAs in ADAR1 p110 Flp-In T-REx cells, designs #7 and #8 were analyzed for their recruitment of endogenous ADAR1 in HeLa cells. To assess the advantage of *in-silico*-optimized recruitment clusters over previous designs, they were compared to design #5, which is identical to design #7 and #8 but does not contain RCs, and to design #6, which contains a conventionally extended antisense part.

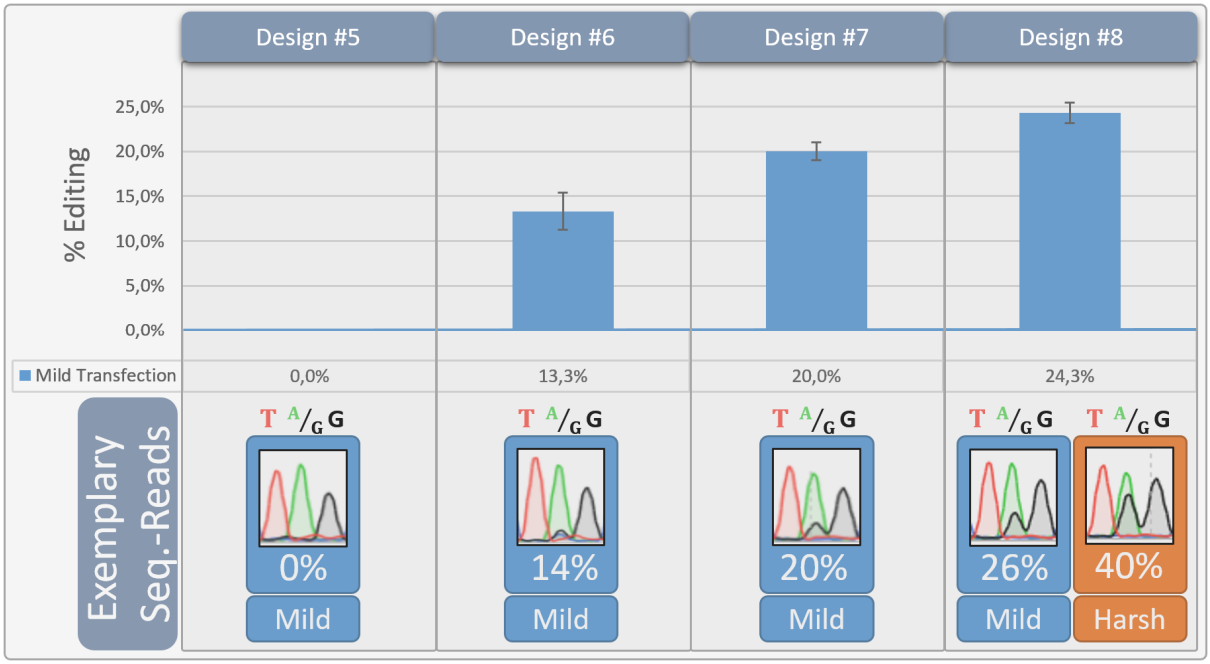


Figure 3-29: Comparative evaluation of editing yields achieved by the designs #5, #6, #7, and #8 using endogenous human ADAR1 in HeLa cells via Sanger sequencing.

The upper part of this figure shows mean editing yields at RNA level after mild transfection, quantified using Sanger sequencing. The lower part shows exemplary sequencing reads. The experiments were performed in collaboration with Dr. Wettengel. The mild transfection (blue) was performed with n=3 replicates. The harsh transfection (orange) was performed once, and only for design #8. The experiments were performed in 24-well scale. HeLa cells were seeded in 24-well scale and transfected 24h thereafter. Mild transfection was performed using $6x10^4$ cells per well transfected with 100 ng dual-luciferase reporter and 400 ng R/G gRNA plasmid, using Lipofectamine 3000 in a 1:1.5 ratio. Harsh transfection was performed using $1.2x10^5$ cells per well transfected with 200 ng dual-luciferase reporter and 800 ng R/G gRNA plasmid, using Lipofectamine 3000 in a 1:1.5 ratio. The cells were harvested 72h post transfection. This was followed by RNA isolation using the Quiagen RNeasy Mini Kit, DNase-I digestion, RT-PCR, and Sanger sequencing.

As expected, design #5 was ineffective at endogenous ADAR expression levels. The designs #6, #7, and #8, however, raised the editing yield to levels detectable by Sanger sequencing for the first time. The conventionally extended design #6 achieved ~13% correction at RNA level, but was again outperformed by the RC containing designs #7 and #8, which achieved ~20% and ~24%, respectively. By using design #8 under harsher transfection settings, up to 40% correction could be achieved (Figure 3-29).

The antisense part length seems to be a very important factor for the recruitment of endogenous ADAR1. At the moment, R/G gRNAs with conventionally extended antisense parts (e.g. design #6) are much easier to engineer, compared to R/G gRNAs containing *in-silico*-optimized recruitment clusters. The conventional extension of antisense parts, however, is prone to masking and misfolding. Meanwhile, advanced versions of the RCF should considerably simplify the engineering of RC R/G gRNAs in the future.

Interestingly, these levels of correction were achieved with plasmid-encoded R/G gRNAs (Figure 3-29). The experiments in section 3.8 demonstrated that a change from plasmid- to AdV-encoded R/G gRNAs could increase the editing yield ~35-fold, when recruiting endogenous ADAR1 in HeLa cells. An AdV-encoded design #7 or #8 R/G gRNA might therefore perform even better.

The results in Figure 3-29 marked the last, and also the most profound, breakthrough in this thesis. For the first time, it was possible to perform efficient site-directed RNA editing using only endogenous ADAR1 and a genetically-encoded R/G gRNA.

4. CONCLUSION

As indicated by the growing number of emerging strategies, the progress in the field of site-directed RNA editing (SDRE) keeps accelerating and the interest for its therapeutic utilization is growing. This thesis answered fundamental questions crucial to the further development of SDRE into a therapeutic strategy.

One prerequisite for the intended therapeutic use of SDRE is the ability to edit endogenous mRNAs. Transfection experiments with R/G gRNAs in doxycycline induced ADAR2 Flp-In T-REx cells demonstrated for the first time that SDRE in the 3' UTRs of endogenous mRNAs is possible. Editing was successful at yields of 9-38% in 12 out of 13 target sites within six different endogenous mRNAs.

Another fundamental requirement is that mRNA repair can indeed functionally restore a loss-of-function mutation in a signaling cascade connected to a human disease. A proof-of-principle experiment with a PINK1 W437X amber loss-of-function mutation in HeLa cells showed that SDRE through the R/G gRNA system could functionally rescue PINK1-Parkin-mediated mitophagy, which is linked to the etiology of Parkinson's disease.

As a means to assess the recruitment success of endogenous human ADAR enzymes, a highly sensitive dual-luciferase SDRE reporter system was established and characterized. It allowed the reliable detection of editing yields far below the Sanger sequencing detection limit. Screening of human lung, liver, CNS, and other cell lines for endogenous ADAR expression levels revealed only ADAR1 to be sufficiently expressed, which made it the prime target for recruitment experiments. The principal recruitability of ADAR1 by the R/G gRNA system could be demonstrated by using doxycycline-induced ADAR1 Flp-In T-REx cells and R/G gRNAs of design #3. Switching to an adenovirus vector substantially increased R/G gRNA expression levels, compared to previous, plasmid-based transfection, and allowed transfer into a panel of new and hard-to-transfect cell lines. After optimization, the AdV approach produced ~100-fold higher R/G gRNA expression levels in HeLa cells, and similarly high values in A549 (lung), Huh7 (liver) and SK-N-BE (CNS) cells. The AdV-encoded R/G gRNA (design #3) achieved editing yields of up to 5.3% in the mentioned cell lines, as measured by the highly sensitive dual-luciferase SDRE reporter system, recruiting only endogenous ADAR1.

Selective knock-down of endogenous ADARs in HeLa cells showed that almost exclusively the nuclear isoform ADAR1 p110 was recruited. Contrarily, the cytosolic ADAR1 p150 was not recruited, despite qPCR-based localization experiments verifying the co-localisation of AdV-encoded R/G gRNAs in the cytosol. This suggests that additional, currently unidentified factors affect the selective recruitment of ADAR enzymes by R/G gRNAs. Based on these results, the further development of novel encodable R/G gRNA designs was mainly assessed in ADAR1 p110 Flp-In T-REx cells.

Successive rounds of both rational and *in-silico* design optimization of the R/G gRNA architecture improved the ADAR1-p110-dependent repair of the dual-luciferase reporter by more than two orders of magnitude at protein level. The encodable R/G gRNA design #8 recruited sufficient endogenous ADAR1 to be detected by Sanger sequencing for the first time, and converted up to 40% of a premature 5'-UAG stop-codon into a 5'-UIG tryptophancodon, when targeting the dual-luciferase reporter in HeLa cells. Since these yields were achieved with plasmid-encoded R/G gRNAs, an AdV-encoded version can be expected to produce even higher editing, but is currently not yet available.

A major advantage of the encodable R/G gRNA system is that it no longer requires editase overexpression, which is an inherent limitation of the engineered editases in other SDRE systems. Follow-up studies could adapt the established combination of endogenous ADAR1 and AdV-encoded design #8 R/G gRNAs to edit triplets within 3' UTRs and ORFs of endogenous mRNAs. First experiments in higher animals can also be considered at this point. The results discussed herein represent a foundation on which further development of therapeutic SDRE can be built upon.

5. REFERENCES

1. Matthews, M.M., et al., *Structures of human ADAR2 bound to dsRNA reveal base-flipping mechanism and basis for site selectivity.* Nat Struct Mol Biol, 2016. **23**(5): p. 426-33.
2. Wulff, B.E. and K. Nishikura, *Substitutional A-to-I RNA editing.* Wiley Interdiscip Rev RNA, 2010. **1**(1): p. 90-101.
3. Daniel, C., et al., *Editing inducer elements increases A-to-I editing efficiency in the mammalian transcriptome.* Genome Biol, 2017. **18**(1): p. 195.
4. Farajollahi, S. and S. Maas, *Molecular diversity through RNA editing: a balancing act.* Trends Genet, 2010. **26**(5): p. 221-30.
5. Gommans, W.M., S.P. Mullen, and S. Maas, *RNA editing: a driving force for adaptive evolution?* Bioessays, 2009. **31**(10): p. 1137-45.
6. Nishikura, K., *Functions and regulation of RNA editing by ADAR deaminases.* Annu Rev Biochem, 2010. **79**: p. 321-49.
7. Brown, T. *Nucleic Acids Book*. 2005 [cited 2019 14. January]; Available from: https://www.atdbio.com/nucleic-acids-book.
8. Aphasizhev, R. and I. Aphasizheva, *Mitochondrial RNA editing in trypanosomes: small RNAs in control.* Biochimie, 2014. **100**: p. 125-31.
9. Cho, D.S., et al., *Requirement of dimerization for RNA editing activity of adenosine deaminases acting on RNA.* J Biol Chem, 2003. **278**(19): p. 17093-102.
10. Licht, K., et al., *Inosine induces context-dependent recoding and translational stalling.* Nucleic Acids Res, 2019. **47**(1): p. 3-14.
11. Rueter, S.M., T.R. Dawson, and R.B. Emeson, *Regulation of alternative splicing by RNA editing.* Nature, 1999. **399**(6731): p. 75-80.
12. Kawahara, Y., et al., *Redirection of silencing targets by adenosine-to-inosine editing of miRNAs.* Science, 2007. **315**(5815): p. 1137-40.
13. Paul, D., et al., *A-to-I editing in human miRNAs is enriched in seed sequence, influenced by sequence contexts and significantly hypoedited in glioblastoma multiforme.* Sci Rep, 2017. **7**(1): p. 2466.
14. Nishikura, K., *A-to-I editing of coding and non-coding RNAs by ADARs.* Nat Rev Mol Cell Biol, 2016. **17**(2): p. 83-96.
15. Kung, C.P., L.B. Maggi, Jr., and J.D. Weber, *The Role of RNA Editing in Cancer Development and Metabolic Disorders.* Front Endocrinol (Lausanne), 2018. **9**: p. 762.
16. Sinigaglia, K., et al., *ADAR RNA editing in innate immune response phasing, in circadian clocks and in sleep.* Biochim Biophys Acta Gene Regul Mech, 2018.
17. Wang, Q., et al., *RNA Editing, ADAR1, and the Innate Immune Response.* Genes (Basel), 2017. **8**(1).
18. Sommer, B., et al., *RNA editing in brain controls a determinant of ion flow in glutamate-gated channels.* Cell, 1991. **67**(1): p. 11-9.
19. Lomeli, H., et al., *Control of kinetic properties of AMPA receptor channels by nuclear RNA editing.* Science, 1994. **266**(5191): p. 1709-13.
20. Burnashev, N., et al., *Divalent ion permeability of AMPA receptor channels is dominated by the edited form of a single subunit.* Neuron, 1992. **8**(1): p. 189-98.
21. Colina, C., et al., *Regulation of Na+/K+ ATPase transport velocity by RNA editing.* PLoS Biol, 2010. **8**(11): p. e1000540.
22. Burns, C.M., et al., *Regulation of serotonin-2C receptor G-protein coupling by RNA editing.* Nature, 1997. **387**(6630): p. 303-8.
23. Bhalla, T., et al., *Control of human potassium channel inactivation by editing of a small mRNA hairpin.* Nat Struct Mol Biol, 2004. **11**(10): p. 950-6.
24. Zhou, Z.Y., et al., *Genome wide analyses uncover allele-specific RNA editing in human and mouse.* Nucleic Acids Res, 2018. **46**(17): p. 8888-8897.

25. Peng, Z., et al., *Comprehensive analysis of RNA-Seq data reveals extensive RNA editing in a human transcriptome.* Nat Biotechnol, 2012. **30**(3): p. 253-60.
26. Stulic, M. and M.F. Jantsch, *Spatio-temporal profiling of Filamin A RNA-editing reveals ADAR preferences and high editing levels outside neuronal tissues.* RNA Biol, 2013. **10**(10): p. 1611-7.
27. Wahlstedt, H., et al., *Large-scale mRNA sequencing determines global regulation of RNA editing during brain development.* Genome Res, 2009. **19**(6): p. 978-86.
28. Lykke-Andersen, S., S. Pinol-Roma, and J. Kjems, *Alternative splicing of the ADAR1 transcript in a region that functions either as a 5'-UTR or an ORF.* RNA, 2007. **13**(10): p. 1732-44.
29. Schwartz, T., et al., *Proteolytic dissection of Zab, the Z-DNA-binding domain of human ADAR1.* J Biol Chem, 1999. **274**(5): p. 2899-906.
30. Kawakubo, K. and C.E. Samuel, *Human RNA-specific adenosine deaminase (ADAR1) gene specifies transcripts that initiate from a constitutively active alternative promoter.* Gene, 2000. **258**(1-2): p. 165-72.
31. Barraud, P., et al., *A bimodular nuclear localization signal assembled via an extended double-stranded RNA-binding domain acts as an RNA-sensing signal for transportin 1.* Proc Natl Acad Sci U S A, 2014. **111**(18): p. E1852-61.
32. Desterro, J.M., et al., *Dynamic association of RNA-editing enzymes with the nucleolus.* J Cell Sci, 2003. **116**(Pt 9): p. 1805-18.
33. George, C.X. and C.E. Samuel, *Human RNA-specific adenosine deaminase ADAR1 transcripts possess alternative exon 1 structures that initiate from different promoters, one constitutively active and the other interferon inducible.* Proc Natl Acad Sci U S A, 1999. **96**(8): p. 4621-6.
34. Patterson, J.B., et al., *Mechanism of interferon action: double-stranded RNA-specific adenosine deaminase from human cells is inducible by alpha and gamma interferons.* Virology, 1995. **210**(2): p. 508-11.
35. Sakurai, M., et al., *ADAR1 controls apoptosis of stressed cells by inhibiting Staufen1-mediated mRNA decay.* Nat Struct Mol Biol, 2017.
36. Park, E. and L.E. Maquat, *Staufen-mediated mRNA decay.* Wiley Interdiscip Rev RNA, 2013. **4**(4): p. 423-35.
37. Di Giammartino, D.C., K. Nishida, and J.L. Manley, *Mechanisms and consequences of alternative polyadenylation.* Mol Cell, 2011. **43**(6): p. 853-66.
38. Bahn, J.H., et al., *Genomic analysis of ADAR1 binding and its involvement in multiple RNA processing pathways.* Nat Commun, 2015. **6**: p. 6355.
39. Herbert, A., et al., *The Zalpha domain from human ADAR1 binds to the Z-DNA conformer of many different sequences.* Nucleic Acids Res, 1998. **26**(15): p. 3486-93.
40. Koeris, M., et al., *Modulation of ADAR1 editing activity by Z-RNA in vitro.* Nucleic Acids Res, 2005. **33**(16): p. 5362-70.
41. Athanasiadis, A., *Zalpha-domains: at the intersection between RNA editing and innate immunity.* Semin Cell Dev Biol, 2012. **23**(3): p. 275-80.
42. Ng, S.K., et al., *Proteins that contain a functional Z-DNA-binding domain localize to cytoplasmic stress granules.* Nucleic Acids Res, 2013. **41**(21): p. 9786-99.
43. Weissbach, R. and A.D. Scadden, *Tudor-SN and ADAR1 are components of cytoplasmic stress granules.* RNA, 2012. **18**(3): p. 462-71.
44. Li, C.L., et al., *Structural and functional insights into human Tudor-SN, a key component linking RNA interference and editing.* Nucleic Acids Res, 2008. **36**(11): p. 3579-89.
45. Scadden, A.D., *The RISC subunit Tudor-SN binds to hyper-edited double-stranded RNA and promotes its cleavage.* Nat Struct Mol Biol, 2005. **12**(6): p. 489-96.
46. Morita, Y., et al., *Human endonuclease V is a ribonuclease specific for inosine-containing RNA.* Nat Commun, 2013. **4**: p. 2273.
47. Pestal, K., et al., *Isoforms of RNA-Editing Enzyme ADAR1 Independently Control Nucleic Acid Sensor MDA5-Driven Autoimmunity and Multi-organ Development.* Immunity, 2015. **43**(5): p. 933-44.

48. Swiecki, M. and M. Colonna, *Type I interferons: diversity of sources, production pathways and effects on immune responses.* Curr Opin Virol, 2011. **1**(6): p. 463-75.
49. Ahmad, S., et al., *Breaching Self-Tolerance to Alu Duplex RNA Underlies MDA5-Mediated Inflammation.* Cell, 2018. **172**(4): p. 797-810 e13.
50. Yu, Z., T. Chen, and X. Cao, *RNA editing by ADAR1 marks dsRNA as "self".* Cell Res, 2015. **25**(12): p. 1283-4.
51. Liddicoat, B.J., A.M. Chalk, and C.R. Walkley, *ADAR1, inosine and the immune sensing system: distinguishing self from non-self.* Wiley Interdiscip Rev RNA, 2016. **7**(2): p. 157-72.
52. Chung, H., et al., *Human ADAR1 Prevents Endogenous RNA from Triggering Translational Shutdown.* Cell, 2018. **172**(4): p. 811-824 e14.
53. Chen, C., H. Huang, and C.H. Wu, *Protein Bioinformatics Databases and Resources.* Methods Mol Biol, 2017. **1558**: p. 3-39.
54. Gerber, A., M.A. O'Connell, and W. Keller, *Two forms of human double-stranded RNA-specific editase 1 (hRED1) generated by the insertion of an Alu cassette.* RNA, 1997. **3**(5): p. 453-63.
55. Sansam, C.L., K.S. Wells, and R.B. Emeson, *Modulation of RNA editing by functional nucleolar sequestration of ADAR2.* Proc Natl Acad Sci U S A, 2003. **100**(24): p. 14018-23.
56. Maas, S. and W.M. Gommans, *Identification of a selective nuclear import signal in adenosine deaminases acting on RNA.* Nucleic Acids Res, 2009. **37**(17): p. 5822-9.
57. Marcucci, R., et al., *Pin1 and WWP2 regulate GluR2 Q/R site RNA editing by ADAR2 with opposing effects.* EMBO J, 2011. **30**(20): p. 4211-22.
58. Feng, Y., et al., *Altered RNA editing in mice lacking ADAR2 autoregulation.* Mol Cell Biol, 2006. **26**(2): p. 480-8.
59. LeGates, T.A. and C.M. Altimus, *Measuring circadian and acute light responses in mice using wheel running activity.* J Vis Exp, 2011(48).
60. Terajima, H., et al., *A-to-I RNA editing enzyme ADAR2 regulates light-induced circadian phase-shift.* Sci Rep, 2018. **8**(1): p. 14848.
61. Terajima, H., et al., *ADARB1 catalyzes circadian A-to-I editing and regulates RNA rhythm.* Nat Genet, 2017. **49**(1): p. 146-151.
62. Chen, C.X., et al., *A third member of the RNA-specific adenosine deaminase gene family, ADAR3, contains both single- and double-stranded RNA binding domains.* RNA, 2000. **6**(5): p. 755-67.
63. Oakes, E., et al., *Adenosine Deaminase That Acts on RNA 3 (ADAR3) Binding to Glutamate Receptor Subunit B Pre-mRNA Inhibits RNA Editing in Glioblastoma.* J Biol Chem, 2017. **292**(10): p. 4326-4335.
64. Mladenova, D., et al., *Adar3 Is Involved in Learning and Memory in Mice.* Front Neurosci, 2018. **12**: p. 243.
65. Hartner, J.C., et al., *Liver disintegration in the mouse embryo caused by deficiency in the RNA-editing enzyme ADAR1.* J Biol Chem, 2004. **279**(6): p. 4894-902.
66. Hartner, J.C., et al., *ADAR1 is essential for the maintenance of hematopoiesis and suppression of interferon signaling.* Nat Immunol, 2009. **10**(1): p. 109-15.
67. Wang, Q., et al., *Requirement of the RNA editing deaminase ADAR1 gene for embryonic erythropoiesis.* Science, 2000. **290**(5497): p. 1765-8.
68. Liddicoat, B.J., et al., *RNA editing by ADAR1 prevents MDA5 sensing of endogenous dsRNA as nonself.* Science, 2015. **349**(6252): p. 1115-20.
69. Rice, G.I., et al., *Mutations in ADAR1 cause Aicardi-Goutieres syndrome associated with a type I interferon signature.* Nat Genet, 2012. **44**(11): p. 1243-8.
70. Crow, Y.J., *Aicardi-Goutieres Syndrome*, in *GeneReviews((R))*, M.P. Adam, et al., Editors. 1993: Seattle (WA).
71. Miyamura, Y., et al., *Mutations of the RNA-specific adenosine deaminase gene (DSRAD) are involved in dyschromatosis symmetrica hereditaria.* Am J Hum Genet, 2003. **73**(3): p. 693-9.

72. Kondo, T., et al., *Six novel mutations of the ADAR1 gene in patients with dyschromatosis symmetrica hereditaria: histological observation and comparison of genotypes and clinical phenotypes.* J Dermatol, 2008. **35**(7): p. 395-406.
73. Livingston, J.H., et al., *A type I interferon signature identifies bilateral striatal necrosis due to mutations in ADAR1.* J Med Genet, 2014. **51**(2): p. 76-82.
74. Crow, Y.J., et al., *Mutations in ADAR1, IFIH1, and RNASEH2B presenting as spastic paraplegia.* Neuropediatrics, 2014. **45**(6): p. 386-93.
75. Mannion, N.M., et al., *The RNA-editing enzyme ADAR1 controls innate immune responses to RNA.* Cell Rep, 2014. **9**(4): p. 1482-94.
76. Song, C., et al., *Functions of the RNA Editing Enzyme ADAR1 and Their Relevance to Human Diseases.* Genes (Basel), 2016. **7**(12).
77. Shimokawa, T., et al., *RNA editing of the GLI1 transcription factor modulates the output of Hedgehog signaling.* RNA Biol, 2013. **10**(2): p. 321-33.
78. Song, I.H., et al., *ADAR1 expression is associated with tumour-infiltrating lymphocytes in triple-negative breast cancer.* Tumour Biol, 2017. **39**(10): p. 1010428317734816.
79. Sagredo, E.A., et al., *ADAR1-mediated RNA-editing of 3'UTRs in breast cancer.* Biol Res, 2018. **51**(1): p. 36.
80. Chen, Y., et al., *ADAR1 overexpression is associated with cervical cancer progression and angiogenesis.* Diagn Pathol, 2017. **12**(1): p. 12.
81. Shigeyasu, K., et al., *AZIN1 RNA editing confers cancer stemness and enhances oncogenic potential in colorectal cancer.* JCI Insight, 2018. **3**(12).
82. Qiao, J.J., et al., *ADAR1: a promising new biomarker for esophageal squamous cell carcinoma?* Expert Rev Anticancer Ther, 2014. **14**(8): p. 865-8.
83. Dou, N., et al., *Aberrant overexpression of ADAR1 promotes gastric cancer progression by activating mTOR/p70S6K signaling.* Oncotarget, 2016. **7**(52): p. 86161-86173.
84. Cho, C.J., et al., *Combinatory RNA-Sequencing Analyses Reveal a Dual Mode of Gene Regulation by ADAR1 in Gastric Cancer.* Dig Dis Sci, 2018. **63**(7): p. 1835-1850.
85. Beyer, U., et al., *Rare ADAR and RNASEH2B variants and a type I interferon signature in glioma and prostate carcinoma risk and tumorigenesis.* Acta Neuropathol, 2017. **134**(6): p. 905-922.
86. Dong, X., et al., *CDK13 RNA Over-Editing Mediated by ADAR1 Associates with Poor Prognosis of Hepatocellular Carcinoma Patients.* Cell Physiol Biochem, 2018. **47**(6): p. 2602-2612.
87. Zipeto, M.A., et al., *ADAR1 Activation Drives Leukemia Stem Cell Self-Renewal by Impairing Let-7 Biogenesis.* Cell Stem Cell, 2016. **19**(2): p. 177-191.
88. Anadon, C., et al., *Gene amplification-associated overexpression of the RNA editing enzyme ADAR1 enhances human lung tumorigenesis.* Oncogene, 2016. **35**(33): p. 4407-13.
89. Lazzari, E., et al., *Alu-dependent RNA editing of GLI1 promotes malignant regeneration in multiple myeloma.* Nat Commun, 2017. **8**(1): p. 1922.
90. Nemlich, Y., et al., *ADAR1-mediated regulation of melanoma invasion.* Nat Commun, 2018. **9**(1): p. 2154.
91. Wahlstedt, H. and M. Ohman, *Site-selective versus promiscuous A-to-I editing.* Wiley Interdiscip Rev RNA, 2011. **2**(6): p. 761-71.
92. Wright, A. and B. Vissel, *The essential role of AMPA receptor GluR2 subunit RNA editing in the normal and diseased brain.* Front Mol Neurosci, 2012. **5**: p. 34.
93. Polson, A.G. and B.L. Bass, *Preferential selection of adenosines for modification by double-stranded RNA adenosine deaminase.* EMBO J, 1994. **13**(23): p. 5701-11.
94. Deffit, S.N. and H.A. Hundley, *To edit or not to edit: regulation of ADAR editing specificity and efficiency.* Wiley Interdiscip Rev RNA, 2016. **7**(1): p. 113-27.
95. Bazak, L., E.Y. Levanon, and E. Eisenberg, *Genome-wide analysis of Alu editability.* Nucleic Acids Res, 2014. **42**(11): p. 6876-84.
96. Tan, M.H., et al., *Dynamic landscape and regulation of RNA editing in mammals.* Nature, 2017. **550**(7675): p. 249-254.

97. Bazak, L., et al., *A-to-I RNA editing occurs at over a hundred million genomic sites, located in a majority of human genes.* Genome Res, 2014. **24**(3): p. 365-76.
98. Kuttan, A. and B.L. Bass, *Mechanistic insights into editing-site specificity of ADARs.* Proc Natl Acad Sci U S A, 2012. **109**(48): p. E3295-304.
99. Ryter, J.M. and S.C. Schultz, *Molecular basis of double-stranded RNA-protein interactions: structure of a dsRNA-binding domain complexed with dsRNA.* EMBO J, 1998. **17**(24): p. 7505-13.
100. Schmedt, C., et al., *Functional characterization of the RNA-binding domain and motif of the double-stranded RNA-dependent protein kinase DAI (PKR).* J Mol Biol, 1995. **249**(1): p. 29-44.
101. Gan, J., et al., *Structural insight into the mechanism of double-stranded RNA processing by ribonuclease III.* Cell, 2006. **124**(2): p. 355-66.
102. Stefl, R., et al., *Structure and specific RNA binding of ADAR2 double-stranded RNA binding motifs.* Structure, 2006. **14**(2): p. 345-55.
103. Stefl, R., et al., *The solution structure of the ADAR2 dsRBM-RNA complex reveals a sequence-specific readout of the minor groove.* Cell, 2010. **143**(2): p. 225-37.
104. Masliah, G., P. Barraud, and F.H. Allain, *RNA recognition by double-stranded RNA binding domains: a matter of shape and sequence.* Cell Mol Life Sci, 2013. **70**(11): p. 1875-95.
105. Lehmann, K.A. and B.L. Bass, *The importance of internal loops within RNA substrates of ADAR1.* J Mol Biol, 1999. **291**(1): p. 1-13.
106. Wang, Y., S. Park, and P.A. Beal, *Selective Recognition of RNA Substrates by ADAR Deaminase Domains.* Biochemistry, 2018. **57**(10): p. 1640-1651.
107. Klaue, Y., et al., *Biochemical analysis and scanning force microscopy reveal productive and nonproductive ADAR2 binding to RNA substrates.* RNA, 2003. **9**(7): p. 839-46.
108. Kallman, A.M., M. Sahlin, and M. Ohman, *ADAR2 A-->I editing: site selectivity and editing efficiency are separate events.* Nucleic Acids Res, 2003. **31**(16): p. 4874-81.
109. Lehmann, K.A. and B.L. Bass, *Double-stranded RNA adenosine deaminases ADAR1 and ADAR2 have overlapping specificities.* Biochemistry, 2000. **39**(42): p. 12875-84.
110. Eggington, J.M., T. Greene, and B.L. Bass, *Predicting sites of ADAR editing in double-stranded RNA.* Nat Commun, 2011. **2**: p. 319.
111. Wong, S.K., S. Sato, and D.W. Lazinski, *Substrate recognition by ADAR1 and ADAR2.* RNA, 2001. **7**(6): p. 846-58.
112. Grosjean, H., *RNA modification: the Golden Period 1995-2015.* RNA, 2015. **21**(4): p. 625-6.
113. Xiang, J.F., et al., *N(6)-Methyladenosines Modulate A-to-I RNA Editing.* Mol Cell, 2018. **69**(1): p. 126-135 e6.
114. Roundtree, I.A., et al., *Dynamic RNA Modifications in Gene Expression Regulation.* Cell, 2017. **169**(7): p. 1187-1200.
115. Meyer, K.D., et al., *Comprehensive analysis of mRNA methylation reveals enrichment in 3' UTRs and near stop codons.* Cell, 2012. **149**(7): p. 1635-46.
116. Wei, C.M. and B. Moss, *Nucleotide sequences at the N6-methyladenosine sites of HeLa cell messenger ribonucleic acid.* Biochemistry, 1977. **16**(8): p. 1672-6.
117. Vogel, P., A. Hanswillemenke, and T. Stafforst, *Switching Protein Localization by Site-Directed RNA Editing under Control of Light.* ACS Synth Biol, 2017. **6**(9): p. 1642-1649.
118. Wettengel, J., et al., *Harnessing human ADAR2 for RNA repair - Recoding a PINK1 mutation rescues mitophagy.* Nucleic Acids Res, 2017. **45**(5): p. 2797-2808.
119. Schneider, M.F., et al., *Optimal guideRNAs for re-directing deaminase activity of hADAR1 and hADAR2 in trans.* Nucleic Acids Res, 2014. **42**(10): p. e87.
120. Vogel, P., et al., *Efficient and precise editing of endogenous transcripts with SNAP-tagged ADARs.* Nat Methods, 2018. **15**(7): p. 535-538.
121. Gong, S., et al., *NECTAR: a database of codon-centric missense variant annotations.* Nucleic Acids Res, 2014. **42**(Database issue): p. D1013-9.

122. Xu, T., et al., *Hepatic phosphorylation status of serine/threonine kinase 1, mammalian target of rapamycin signaling proteins, and growth rate in Holstein heifer calves in response to maternal supply of methionine.* J Dairy Sci, 2018. **101**(9): p. 8476-8491.
123. Rao, R.S., et al., *Convergent signaling pathways--interaction between methionine oxidation and serine/threonine/tyrosine O-phosphorylation.* Cell Stress Chaperones, 2015. **20**(1): p. 15-21.
124. Souchelnytskyi, S., et al., *Phosphorylation of Smad signaling proteins by receptor serine/threonine kinases.* Methods Mol Biol, 2001. **124**: p. 107-20.
125. Zhang, P., et al., *Serine/threonine phosphorylation in cellular signaling for alveolar macrophage phagocytic response to endotoxin.* Shock, 2000. **13**(1): p. 34-40.
126. Gaidenko, T.A., et al., *Threonine phosphorylation of modulator protein RsbR governs its ability to regulate a serine kinase in the environmental stress signaling pathway of Bacillus subtilis.* J Mol Biol, 1999. **288**(1): p. 29-39.
127. Tanti, J.F., et al., *Serine/threonine phosphorylation of insulin receptor substrate 1 modulates insulin receptor signaling.* J Biol Chem, 1994. **269**(8): p. 6051-7.
128. Singer, K.L., et al., *Relationship of serine/threonine phosphorylation/dephosphorylation signaling to glucocorticoid regulation of tight junction permeability and ZO-1 distribution in nontransformed mammary epithelial cells.* J Biol Chem, 1994. **269**(23): p. 16108-15.
129. Defaus, S., et al., *Mammalian protein glycosylation--structure versus function.* Analyst, 2014. **139**(12): p. 2944-67.
130. O'Connor, S.E. and B. Imperiali, *Modulation of protein structure and function by asparagine-linked glycosylation.* Chem Biol, 1996. **3**(10): p. 803-12.
131. Woolf, T.M., J.M. Chase, and D.T. Stinchcomb, *Toward the therapeutic editing of mutated RNA sequences.* Proc Natl Acad Sci U S A, 1995. **92**(18): p. 8298-302.
132. Montiel-Gonzalez, M.F., et al., *Correction of mutations within the cystic fibrosis transmembrane conductance regulator by site-directed RNA editing.* Proc Natl Acad Sci U S A, 2013. **110**(45): p. 18285-90.
133. Stafforst, T. and M.F. Schneider, *An RNA-deaminase conjugate selectively repairs point mutations.* Angew Chem Int Ed Engl, 2012. **51**(44): p. 11166-9.
134. Azad, M.T.A., S. Bhakta, and T. Tsukahara, *Site-directed RNA editing by adenosine deaminase acting on RNA for correction of the genetic code in gene therapy.* Gene Ther, 2017. **24**(12): p. 779-786.
135. Fukuda, M., et al., *Construction of a guide-RNA for site-directed RNA mutagenesis utilising intracellular A-to-I RNA editing.* Sci Rep, 2017. **7**: p. 41478.
136. Cox, D.B.T., et al., *RNA editing with CRISPR-Cas13.* Science, 2017.
137. Cox, D.B.T., et al., *RNA editing with CRISPR-Cas13.* Science, 2017. **358**(6366): p. 1019-1027.
138. Montiel-Gonzalez, M.F., I.C. Vallecillo-Viejo, and J.J. Rosenthal, *An efficient system for selectively altering genetic information within mRNAs.* Nucleic Acids Res, 2016. **44**(21): p. e157.
139. Vogel, P. and T. Stafforst, *Critical review on engineering deaminases for site-directed RNA editing.* Curr Opin Biotechnol, 2018. **55**: p. 74-80.
140. Krutzfeldt, J., et al., *Silencing of microRNAs in vivo with 'antagomirs'.* Nature, 2005. **438**(7068): p. 685-9.
141. Keppler, A., et al., *A general method for the covalent labeling of fusion proteins with small molecules in vivo.* Nat Biotechnol, 2003. **21**(1): p. 86-9.
142. Vogel, P., et al., *Improving site-directed RNA editing in vitro and in cell culture by chemical modification of the guideRNA.* Angew Chem Int Ed Engl, 2014. **53**(24): p. 6267-71.
143. Hanswillemenke, A., et al., *Site-Directed RNA Editing in Vivo Can Be Triggered by the Light-Driven Assembly of an Artificial Riboprotein.* J Am Chem Soc, 2015. **137**(50): p. 15875-81.
144. Austin, R.J., et al., *Designed arginine-rich RNA-binding peptides with picomolar affinity.* J Am Chem Soc, 2002. **124**(37): p. 10966-7.
145. Vallecillo-Viejo, I.C., et al., *Abundant off-target edits from site-directed RNA editing can be reduced by nuclear localization of the editing enzyme.* RNA Biol, 2018. **15**(1): p. 104-114.

146. Sinnamon, J.R., et al., *Site-directed RNA repair of endogenous Mecp2 RNA in neurons.* Proc Natl Acad Sci U S A, 2017. **114**(44): p. E9395-E9402.
147. Stefl, R. and F.H. Allain, *A novel RNA pentaloop fold involved in targeting ADAR2.* RNA, 2005. **11**(5): p. 592-7.
148. Wettengel, J. and E.K.U.t. Tübingen, *Harnessing human ADAR2 for site-directed RNA editing*. 2016, Tübingen. VII, 223 Seiten.
149. Katrekar, D., et al., *In vivo RNA editing of point mutations via RNA-guided adenosine deaminases.* Nat Methods, 2019.
150. Hagedorn, C., et al., *S/MAR Element Facilitates Episomal Long-Term Persistence of Adeno-Associated Virus Vector Genomes in Proliferating Cells.* Hum Gene Ther, 2017. **28**(12): p. 1169-1179.
151. Monteleone, L.R., et al., *A Bump-Hole Approach for Directed RNA Editing.* Cell Chem Biol, 2018.
152. Crudele, J.M. and J.S. Chamberlain, *Cas9 immunity creates challenges for CRISPR gene editing therapies.* Nat Commun, 2018. **9**(1): p. 3497.
153. Peabody, D.S., *The RNA binding site of bacteriophage MS2 coat protein.* EMBO J, 1993. **12**(2): p. 595-600.
154. Bolognesi, B. and B. Lehner, *Reaching the limit.* Elife, 2018. **7**.
155. Moriya, H., *Quantitative nature of overexpression experiments.* Mol Biol Cell, 2015. **26**(22): p. 3932-9.
156. Naso, M.F., et al., *Adeno-Associated Virus (AAV) as a Vector for Gene Therapy.* BioDrugs, 2017. **31**(4): p. 317-334.
157. Srivastava, A., *In vivo tissue-tropism of adeno-associated viral vectors.* Curr Opin Virol, 2016. **21**: p. 75-80.
158. Juliano, R.L., *The delivery of therapeutic oligonucleotides.* Nucleic Acids Res, 2016. **44**(14): p. 6518-48.
159. Tatiparti, K., et al., *siRNA Delivery Strategies: A Comprehensive Review of Recent Developments.* Nanomaterials (Basel), 2017. **7**(4).
160. Tai, W. and X. Gao, *Functional peptides for siRNA delivery.* Adv Drug Deliv Rev, 2017. **110-111**: p. 157-168.
161. Springer, A.D. and S.F. Dowdy, *GalNAc-siRNA Conjugates: Leading the Way for Delivery of RNAi Therapeutics.* Nucleic Acid Ther, 2018. **28**(3): p. 109-118.
162. Kaczmarek, J.C., P.S. Kowalski, and D.G. Anderson, *Advances in the delivery of RNA therapeutics: from concept to clinical reality.* Genome Med, 2017. **9**(1): p. 60.
163. Geary, R.S., et al., *Pharmacokinetics, biodistribution and cell uptake of antisense oligonucleotides.* Adv Drug Deliv Rev, 2015. **87**: p. 46-51.
164. Guha, T.K., A. Wai, and G. Hausner, *Programmable Genome Editing Tools and their Regulation for Efficient Genome Engineering.* Comput Struct Biotechnol J, 2017. **15**: p. 146-160.
165. Shah, S.A., et al., *Protospacer recognition motifs: mixed identities and functional diversity.* RNA Biol, 2013. **10**(5): p. 891-9.
166. Hu, J.H., et al., *Evolved Cas9 variants with broad PAM compatibility and high DNA specificity.* Nature, 2018. **556**(7699): p. 57-63.
167. Vu, L.T. and T. Tsukahara, *C-to-U editing and site-directed RNA editing for the correction of genetic mutations.* Biosci Trends, 2017. **11**(3): p. 243-253.
168. Mao, Z., et al., *Comparison of nonhomologous end joining and homologous recombination in human cells.* DNA Repair (Amst), 2008. **7**(10): p. 1765-71.
169. Heyer, W.D., K.T. Ehmsen, and J. Liu, *Regulation of homologous recombination in eukaryotes.* Annu Rev Genet, 2010. **44**: p. 113-39.
170. Peddle, C.F. and R.E. MacLaren, *The Application of CRISPR/Cas9 for the Treatment of Retinal Diseases.* Yale J Biol Med, 2017. **90**(4): p. 533-541.
171. de Buhr, H. and R.J. Lebbink, *Harnessing CRISPR to combat human viral infections.* Curr Opin Immunol, 2018. **54**: p. 123-129.

172. Reautschnig, P., P. Vogel, and T. Stafforst, *The notorious R.N.A. in the spotlight - drug or target for the treatment of disease.* RNA Biol, 2017. **14**(5): p. 651-668.
173. Chew, W.L., et al., *A multifunctional AAV-CRISPR-Cas9 and its host response.* Nat Methods, 2016. **13**(10): p. 868-74.
174. Araki, M. and T. Ishii, *Providing Appropriate Risk Information on Genome Editing for Patients.* Trends Biotechnol, 2016. **34**(2): p. 86-90.
175. *Keep off-target effects in focus.* Nat Med, 2018. **24**(8): p. 1081.
176. Eid, A., S. Alshareef, and M.M. Mahfouz, *CRISPR base editors: genome editing without double-stranded breaks.* Biochem J, 2018. **475**(11): p. 1955-1964.
177. Lam, J.K., et al., *siRNA Versus miRNA as Therapeutics for Gene Silencing.* Mol Ther Nucleic Acids, 2015. **4**: p. e252.
178. Bracken, C.P., H.S. Scott, and G.J. Goodall, *A network-biology perspective of microRNA function and dysfunction in cancer.* Nat Rev Genet, 2016. **17**(12): p. 719-732.
179. Cilek, E.E., H. Ozturk, and B. Gur Dedeoglu, *Construction of miRNA-miRNA networks revealing the complexity of miRNA-mediated mechanisms in trastuzumab treated breast cancer cell lines.* PLoS One, 2017. **12**(10): p. e0185558.
180. Christopher, A.F., et al., *MicroRNA therapeutics: Discovering novel targets and developing specific therapy.* Perspect Clin Res, 2016. **7**(2): p. 68-74.
181. Zhang, N., et al., *Engineering Artificial MicroRNAs for Multiplex Gene Silencing and Simplified Transgenic Screen.* Plant Physiol, 2018. **178**(3): p. 989-1001.
182. Hoy, S.M., *Patisiran: First Global Approval.* Drugs, 2018. **78**(15): p. 1625-1631.
183. Ahmadzada, T., G. Reid, and D.R. McKenzie, *Fundamentals of siRNA and miRNA therapeutics and a review of targeted nanoparticle delivery systems in breast cancer.* Biophys Rev, 2018. **10**(1): p. 69-86.
184. Rao, D.D., et al., *siRNA vs. shRNA: similarities and differences.* Adv Drug Deliv Rev, 2009. **61**(9): p. 746-59.
185. Carthew, R.W. and E.J. Sontheimer, *Origins and Mechanisms of miRNAs and siRNAs.* Cell, 2009. **136**(4): p. 642-55.
186. Bartlett, D.W. and M.E. Davis, *Insights into the kinetics of siRNA-mediated gene silencing from live-cell and live-animal bioluminescent imaging.* Nucleic Acids Res, 2006. **34**(1): p. 322-33.
187. Noland, C.L. and J.A. Doudna, *Multiple sensors ensure guide strand selection in human RNAi pathways.* RNA, 2013. **19**(5): p. 639-48.
188. Hohjoh, H., *Disease-causing allele-specific silencing by RNA interference.* Pharmaceuticals (Basel), 2013. **6**(4): p. 522-35.
189. Mansoori, B., S. Sandoghchian Shotorbani, and B. Baradaran, *RNA interference and its role in cancer therapy.* Adv Pharm Bull, 2014. **4**(4): p. 313-21.
190. Aguiar, S., B. van der Gaag, and F.A.B. Cortese, *RNAi mechanisms in Huntington's disease therapy: siRNA versus shRNA.* Transl Neurodegener, 2017. **6**: p. 30.
191. Ding, H., et al., *Selective silencing by RNAi of a dominant allele that causes amyotrophic lateral sclerosis.* Aging Cell, 2003. **2**(4): p. 209-17.
192. Grimm, D., *The dose can make the poison: lessons learned from adverse in vivo toxicities caused by RNAi overexpression.* Silence, 2011. **2**: p. 8.
193. Peters, S.U., et al., *The behavioral phenotype in MECP2 duplication syndrome: a comparison with idiopathic autism.* Autism Res, 2013. **6**(1): p. 42-50.
194. UniProt Consortium, T., *UniProt: the universal protein knowledgebase.* Nucleic Acids Res, 2018. **46**(5): p. 2699.
195. Zerbino, D.R., et al., *Ensembl 2018.* Nucleic Acids Res, 2018. **46**(D1): p. D754-D761.
196. Lorenz, R., et al., *ViennaRNA Package 2.0.* Algorithms Mol Biol, 2011. **6**: p. 26.
197. Zadeh, J.N., et al., *NUPACK: Analysis and design of nucleic acid systems.* J Comput Chem, 2011. **32**(1): p. 170-3.
198. Coordinators, N.R., *Database resources of the National Center for Biotechnology Information.* Nucleic Acids Res, 2018. **46**(D1): p. D8-D13.

199. Wang, Y., W. Zhu, and D.E. Levy, *Nuclear and cytoplasmic mRNA quantification by SYBR green based real-time RT-PCR.* Methods, 2006. **39**(4): p. 356-62.
200. Schmittgen, T.D. and K.J. Livak, *Analyzing real-time PCR data by the comparative C(T) method.* Nat Protoc, 2008. **3**(6): p. 1101-8.
201. Pfaffl, M.W., *Real-Time RT-PCR: Neue Ansätze zur exakten mRNA Quantifizierung.* 2004.
202. He, T.C., et al., *A simplified system for generating recombinant adenoviruses.* Proc Natl Acad Sci U S A, 1998. **95**(5): p. 2509-14.
203. Luo, J., et al., *A protocol for rapid generation of recombinant adenoviruses using the AdEasy system.* Nat Protoc, 2007. **2**(5): p. 1236-47.
204. Yamano, K., N. Matsuda, and K. Tanaka, *The ubiquitin signal and autophagy: an orchestrated dance leading to mitochondrial degradation.* EMBO Rep, 2016. **17**(3): p. 300-16.
205. Springer, W. and P.J. Kahle, *Regulation of PINK1-Parkin-mediated mitophagy.* Autophagy, 2011. **7**(3): p. 266-78.
206. Okatsu, K., et al., *PINK1 autophosphorylation upon membrane potential dissipation is essential for Parkin recruitment to damaged mitochondria.* Nat Commun, 2012. **3**: p. 1016.
207. Scornaienchi, V., et al., *Mutation analysis of the PINK1 gene in Southern Italian patients with early- and late-onset parkinsonism.* Parkinsonism Relat Disord, 2012. **18**(5): p. 651-3.
208. Darquet, A.M., et al., *A new DNA vehicle for nonviral gene delivery: supercoiled minicircle.* Gene Ther, 1997. **4**(12): p. 1341-9.
209. Fritzell, K., et al., *Sensitive ADAR editing reporter in cancer cells enables high-throughput screening of small molecule libraries.* Nucleic Acids Res, 2018.
210. Rossetti, C., et al., *RNA editing signature during myeloid leukemia cell differentiation.* Leukemia, 2017.
211. Donnelly, M.L., et al., *The cleavage activities of aphthovirus and cardiovirus 2A proteins.* J Gen Virol, 1997. **78 (Pt 1)**: p. 13-21.
212. Donnelly, M.L., et al., *Analysis of the aphthovirus 2A/2B polyprotein 'cleavage' mechanism indicates not a proteolytic reaction, but a novel translational effect: a putative ribosomal 'skip'.* J Gen Virol, 2001. **82**(Pt 5): p. 1013-25.
213. de Felipe, P., et al., *Co-translational, intraribosomal cleavage of polypeptides by the foot-and-mouth disease virus 2A peptide.* J Biol Chem, 2003. **278**(13): p. 11441-8.
214. Kim, J.H., et al., *High cleavage efficiency of a 2A peptide derived from porcine teschovirus-1 in human cell lines, zebrafish and mice.* PLoS One, 2011. **6**(4): p. e18556.
215. Kjaer, J. and G.J. Belsham, *Selection of functional 2A sequences within foot-and-mouth disease virus; requirements for the NPGP motif with a distinct codon bias.* RNA, 2018. **24**(1): p. 12-17.
216. Mizuguchi, H., et al., *IRES-dependent second gene expression is significantly lower than cap-dependent first gene expression in a bicistronic vector.* Mol Ther, 2000. **1**(4): p. 376-82.
217. Lee, C.S., et al., *Adenovirus-Mediated Gene Delivery: Potential Applications for Gene and Cell-Based Therapies in the New Era of Personalized Medicine.* Genes Dis, 2017. **4**(2): p. 43-63.
218. Dormond, E., M. Perrier, and A. Kamen, *From the first to the third generation adenoviral vector: what parameters are governing the production yield?* Biotechnol Adv, 2009. **27**(2): p. 133-44.
219. Ye, J., et al., *Primer-BLAST: a tool to design target-specific primers for polymerase chain reaction.* BMC Bioinformatics, 2012. **13**: p. 134.
220. Luan, L., et al., *Comparative Transcriptome Profiles of Human Blood in Response to the Toll-like Receptor 4 Ligands Lipopolysaccharide and Monophosphoryl Lipid A.* Sci Rep, 2017. **7**: p. 40050.
221. Ji, Q., et al., *Long non-coding RNA MALAT1 promotes tumour growth and metastasis in colorectal cancer through binding to SFPQ and releasing oncogene PTBP2 from SFPQ/PTBP2 complex.* Br J Cancer, 2014. **111**(4): p. 736-48.
222. Yang, L.H., et al., *Universal stem-loop primer method for screening and quantification of microRNA.* PLoS One, 2014. **9**(12): p. e115293.
223. Wittrup, A., et al., *Visualizing lipid-formulated siRNA release from endosomes and target gene knockdown.* Nat Biotechnol, 2015. **33**(8): p. 870-6.

224. Lee, N.S., et al., *Functional and intracellular localization properties of U6 promoter-expressed siRNAs, shRNAs, and chimeric VA1 shRNAs in mammalian cells.* RNA, 2008. **14**(9): p. 1823-33.
225. Engelke, D.R. and J.J. Rossi, *RNA interference*, in *Methods in enzymology*. 2005, Elsevier Acad. Press: Amsterdam Heidelberg San Diego. p. 1 Online-Ressource (XXXVII, 453 Seiten).
226. Paul, C.P., et al., *Localized expression of small RNA inhibitors in human cells.* Mol Ther, 2003. **7**(2): p. 237-47.
227. Paul, C.P., et al., *Effective expression of small interfering RNA in human cells.* Nat Biotechnol, 2002. **20**(5): p. 505-8.
228. Good, P.D., et al., *Expression of small, therapeutic RNAs in human cell nuclei.* Gene Ther, 1997. **4**(1): p. 45-54.
229. Hartig, R., et al., *Active nuclear import of single-stranded oligonucleotides and their complexes with non-karyophilic macromolecules.* Biol Cell, 1998. **90**(5): p. 407-26.
230. Lorenz, P., et al., *Nucleocytoplasmic shuttling: a novel in vivo property of antisense phosphorothioate oligodeoxynucleotides.* Nucleic Acids Res, 2000. **28**(2): p. 582-92.
231. Crooke, S.T., et al., *Cellular uptake and trafficking of antisense oligonucleotides.* Nat Biotechnol, 2017. **35**(3): p. 230-237.
232. Bailey, J.K., et al., *Nucleic acid binding proteins affect the subcellular distribution of phosphorothioate antisense oligonucleotides.* Nucleic Acids Res, 2017. **45**(18): p. 10649-10671.
233. Strunze, S., et al., *Kinesin-1-mediated capsid disassembly and disruption of the nuclear pore complex promote virus infection.* Cell Host Microbe, 2011. **10**(3): p. 210-23.
234. Kuwabara, T., et al., *Significantly higher activity of a cytoplasmic hammerhead ribozyme than a corresponding nuclear counterpart: engineered tRNAs with an extended 3' end can be exported efficiently and specifically to the cytoplasm in mammalian cells.* Nucleic Acids Res, 2001. **29**(13): p. 2780-8.
235. Heep, M. and E.K.U.t. Tübingen, *Optimization of the guide RNA for recruitment of ADAR1 and ADAR2.* 2016: p. 97.
236. Phelps, K.J., et al., *Recognition of duplex RNA by the deaminase domain of the RNA editing enzyme ADAR2.* Nucleic Acids Res, 2015.
237. Pearson, W.R., *An introduction to sequence similarity ("homology") searching.* Curr Protoc Bioinformatics, 2013. **Chapter 3**: p. Unit3 1.
238. Komura, R., et al., *High-throughput evaluation of T7 promoter variants using biased randomization and DNA barcoding.* PLoS One, 2018. **13**(5): p. e0196905.
239. Burnett, J.C. and J.J. Rossi, *RNA-based therapeutics: current progress and future prospects.* Chem Biol, 2012. **19**(1): p. 60-71.

6. APPENDIX

6.1 KEY FUNCTIONS OF THE RECRUITMENT CLUSTER FINDER V1.0.1 PYTHON SOURCE CODE

```
def findRangeIndices():
    blockMin = vBlockSizeMin.get()
    blockMax = vBlockSizeMax.get()
    hits = vORF.get()
    hitsSplitted = hits.split(vGuide1.get())
    hitsReplaced = re.sub(r'GA', 'bb', hitsSplitted[0], flags=re.I)
    hitsReplaced = re.sub(r'[GTC]', 'B', hitsReplaced, flags = re.I)
    my_Regex = r"B{" + re.escape(blockMin) + r",}"
    result = re.finditer(my_Regex, hitsReplaced, flags=re.I)
    resultList = []
    hitsBracket = list(hits)
    for x in result:
        resultList.insert(0, (x.start(0), x.end(0)))

    print ("\nPossible Block Range Indices of " +str(len(resultList))  + "
    Blocks: \n")
    guideIndex = len(hitsSplitted[0])
    print("Guide Sequence at " +str(guideIndex) + ", Distance to next Block: " +
    str(guideIndex - resultList[0][1]))
    print(resultListToString(resultList))
    return(hits, resultList, guideIndex)
```

Function 1: This function searches for G, C, T, GA blocks within the target sequence, which is followed by determination of start- and end-indices.

```
def recombineAllLists(lis):
    result = None
    for curr in lis:
        if(result == None):
            result = [[x] for x in curr]
        else:
            temp = []
            for x in range(0, len(result)):
                temp = temp + [result[x] + [y] for y in curr]
            result = temp
    return(result)
for currBlock in allBlocksAll:
                    bool = True
                    bool = (distToGuideSeq - currBlock[clusterNumber-1][1])>=
                    distanceMin and (distToGuideSeq - currBlock[clusterNumber-1][1
                    ])<= distanceMax
                    for i in sizeRange:
                        if(not bool):
                            break
                        bool = bool and (currBlock[i+1][0] - currBlock[i][1]) >=
                        distanceMin and  (currBlock[i+1][0]-currBlock[i][1])<=
                        distanceMax
                    if(bool):
                        counterDistanceGood += 1
                        clustStr = "> "
                        for n in currBlock:
                            clustStr = clustStr + tupleToStr(n)
                        file.write(clustStr +   "\n" + generateRcRNA(currBlock) +
                        "\n")
                    else:
                        counterDistanceBad += 1
        print("Hits:" + str(counterDistanceGood) + "   Filtered out: " + str(
        counterDistanceBad))
```

Function 2: This function recombines the clusters to cluster groups. The cluster groups are then filtered for the input criteria, cluster size and distance. All groups that fulfill the criteria (hits) are saved into a file that is later used by ViennaRNA.

```
def generateRcRNA(blocks):
    dna = vOptional.get()
    for block in blocks:
        dna = dna + block[2] + vLinker.get()
    dna = dna + vGuide2.get() + vMotif.get()
    rcRNA = Seq.Seq(dna).transcribe().reverse_complement()
    return(str(rcRNA))
```

Function 3: This functions converts blocks reverse complementarily

```
def calcOutput():
    startTime = time.time()
    output = subprocess.check_output(["ViennaRNA Package/RNAfold.exe","--noPS",
    "temp/sequences.txt"])
    root.outputList = outputToList(output)
    printCalculatedOutput()
    print("Run Time: " + str(time.time()-startTime))
    print("\nCalculation completed!")
```

Function 4: This functions runs ViennaRNA [196] to fold the R/G gRNAs. The results are saved in an output list.

```
def printCalculatedOutput():
    outputList = root.outputList
    if(len(outputList)>0):
        max = min(int(vNumberResults.get()), (len(outputList[0])))
        order = vOutputSort.get()
        if(order == sMinEnergy):
            bestSortedValues = heapq.nsmallest(max, outputList[4])
            bestSortedValues = list(sorted(set(bestSortedValues), reverse = False))
        else:
            if(order == sDotBracket):
                bestSortedValues = heapq.nlargest(max, outputList[5])
                bestSortedValues = list(sorted(set(bestSortedValues), reverse =
                True))
            else:
                bestSortedValues = outputList[5][:max]
        isShowBlocks = (vShowBlocks.get()==1)
        listbox.delete(0, Tk.END)
        result = []
        for current in bestSortedValues:
            if(order == sMinEnergy):
                indices = [i for i, x in enumerate(outputList[4]) if x == current]
            else:
                if(order == sDotBracket):
                    indices = [i for i, x in enumerate(outputList[5]) if x ==
                    current]
                else:
                    indices = range(0, max)
            for currentMin in indices:
                if(max > 0):
                    string1 = str(outputList[1][currentMin]) + "\n"
                    string2 = str(outputList[2][currentMin]) + "\n"
                    string3 = str(outputList[3][currentMin]) + "\n"
                    if(isShowBlocks):
                        seq = outputList[1][currentMin]
                        blockIndices = [int(match.group()) for match in re.finditer
                        ("[0-9]+", seq)]
                        orf = vORF.get()
                        stringSeq1 = orf[blockIndices[0]: blockIndices[1]] + "\n"
                        stringSeq2 = orf[blockIndices[2]: blockIndices[3]] + "\n"
                        stringSeq3 = orf[blockIndices[4]: blockIndices[5]] + "\n"
                    string4 = " Energy: " + str(outputList[4][currentMin]) + ",
                    Dots/Brackets Ratio: " + str(outputList[5][currentMin]) + "\n"
                    listbox.insert(Tk.END, string1)
                    listbox.insert(Tk.END, string2)
                    listbox.insert(Tk.END, string3)
                    if(isShowBlocks):
                        listbox.insert(Tk.END, stringSeq1)
                        listbox.insert(Tk.END, stringSeq2)
                        listbox.insert(Tk.END, stringSeq3)
                    listbox.insert(Tk.END, string4)
                    if(isShowBlocks):
                        result.append([string1, string2, string3, stringSeq1,
                        stringSeq2, stringSeq3, string4])
                    else:
                        result.append([string1, string2, string3, string4])
                    max -= 1
        root.exportResult = result
```

Function 5: The ViennaRNA results are sorted by the initially chosen criteria (“numerical order”, “dot/bracket ratio order”, or “minimal energy order”). Then the result table is generated.

6.2 R/G gRNA SEQUENCES

Table 6-1: R/G gRNA sequences (Plasmids)

Plasmid #	Design #	Target (Triplett)	R/G motif version Editing position Guide length + #RC RC length [nt]	Sequence of the mRNA binding site (5'→3')
pTS121	#1	Firefly W417X (5'-UAG)	R/G-V1 Pos. 6 16 nt / No RC	GCAG**CCA**GCCGTCCTT
pTS170	#2	Firefly W417X (5'-UAG)	R/G-V1 Pos. 8 16 nt / No RC	GTGCAG**CCA**GCCGTCC
pTS408	ND	β-actin 3'UTR (5'-UAG#1)	R/G-V1 Pos. 8 16 nt / No RC	ACGCAA**CCA**AGTCATA
pTS558	#3	Firefly W417X (5'-UAG)	R/G-V20 Pos. 8 16 nt / No RC	GTGCAG**CCA**GCCGTCC
pTS863	#4	Firefly W417X (5'-UAG)	R/G-V20 Pos. 8 20 nt / No RC	GTGCAG**CCA**GCCGTCCTTGT
pTS930	ND	Firefly W417X (5'-UAG)	R/G-V20 Pos. 11 21 nt / No RC	GCTGTGCAG**CCA**GCCGTCCGT
pTS931	ND	Firefly W417X (5'-UAG)	R/G-V20 Pos. 16 31 nt / No RC	TCGCCGCGGGG-CAG**CCA**GCCGTCCTGGTCGA
pTS932	ND	Firefly W417X (5'-UAG)	R/G-V20 Pos. 21 41 nt / No RC	CGAGGTCGCCGCGGGG-CAG**CCA**GCCGTCCTGGTCGAG-GAGA
pTS933	ND	Firefly W417X (5'-UAG)	R/G-V20 Pos. 26 51 nt / No RC	GGAGGCGAGGTCGCCGCGGGG-CAG**CCA**GCCGTCCTGGTCGAG-GAGAGCGGT
pTS934	ND	Firefly W417X (5'-UAG)	R/G-V20 Pos. 11 21 nt / 1x RC 15	GCTGTGCAG**CCA**GCCGTCCGT RC#1:GCCACGGACGCACAG
pTS935	ND	Firefly W417X (5'-UAG)	R/G-V20 Pos. 11 21 nt / 2x RC 15:13	GCTGTGCAG**CCA**GCCGTCCGT RC#1:GCCACGGACGCACAG RC#2:CGAAGAAGGGCAC
pTS936	ND	Firefly W417X (5'-UAG)	R/G-V20 Pos. 11 21 nt / 3x RC 15:13:15	GCTGTGCAG**CCA**GCCGTCCGT RC#1:GCCACGGACGCACAG RC#2:CGAAGAAGGGCAC RC#3:GAGCGGCGCCCCGCC
pTS946	#6	Firefly W417X (5'-UAG)	R/G-V21 Pos. 30 40 nt / No RC	CCCAG-TAGGCGATGTCGCCGCTGTG-CAG**CCA**GCCGTCCTT
pTS961	#7	Firefly W417X (5'-UAG)	R/G-V21 Pos. 8 20 nt / 3x RC 15:13:11	GTGCAG**CCA**GCCGTCCTTGT RC#1:CGCACAGCTCG RC#2:GTCCAAGTCCACC RC#3:AAGAAGGGCACCACC

Plasmid #	Design #	Target (Triplett)	R/G motif version Editing position Guide length + #RC RC length [nt]	Sequence of the mRNA binding site (5'→3')
pTS962	ND	Firefly W417X (5'-UAG)	R/G-V21 Pos. 8 20 nt + 3x RC 16:11:12	GTGCAG**CCA**GCCGTCCTTGT RC#1:AAGTCCACCACC RC#2:GCGCCCCGCCG RC#3:CACGACCCGAAAGCCG
pTS963	ND	Firefly W417X (5'-UAG)	R/G-V21 Pos. 8 20 nt + 3x RC 16:11:12	GTGCAG**CCA**GCCGTCCTTGT RC#1:CTCCTCCTCGAA RC#2:CCACAGCCACA RC#3:CTCCACGTCTCCAGCC
pTS964	ND	Firefly W417X (5'-UAG)	R/G-V21 Pos. 8 20 nt + 3x RC 18:18:18	GTGCAG**CCA**GCCGTCCTTGT RC#1:GGCCACGGACGCACAGCT RC#2:TCGAAGAAGGGCACCACC RC#3:GGGCACCAGCAGGGCAGA
pTS966	ND	Firefly W417X (5'-UAG)	R/G-V21 Pos. 8 20 nt + 3x RC 22:22:22	GTGCAG**CCA**GCCGTCCTTGT RC#1:AGCCTCGAAGAAGGG-CACCACC RC#2:CACGACCCGAAAGCCG-CAGATC RC#3:AACAGGGCACCCAACAC-GGGCA
pTS967	#5	Firefly W417X (5'-UAG)	R/G-V21 Pos. 8 20 nt / No RC	GTGCAG**CCA**GCCGTCCTTGT
pTS970	#8	Firefly W417X (5'-UAG))	R/G-V21 Pos. 8 20 nt + 8x RC 11:15:15:15:12:15:13 :11	GTGCAG**CCA**GCCGTCCTTGT RC#1:CGCACAGCTCG RC#2:GTCCAAGTCCACC RC#3:AAGAAGGGCACCACC RC#4:CTCCTCCTCGAA RC#5:GCCGCAGATCAA RC#6:CGAAGCTCTCG RC#7:CCACAGCCACACCGA RC#8:CACACCACGATCCGA

Table 6-2: R/G gRNA sequences (AdV)

Plasmid # (Shuttle-vector)	Design #	Target (Triplett)	R/G motif version Editing position Guide length + #RC RC length [nt]	Sequence of the mRNA binding site (5'→3')
No pTS# (pAd-Track)	#2	β-actin 3'UTR (5'-UAG#1)	R/G-V1 Pos. 8 16 nt / No RC	ACGCAA**CCA**AGTCATA
pTS668 (pShuttle)	#3	Firefly W417X (5'-UAG-3')	R/G-V20 CCAp8 16 nt / No RC	GTGCAG**CCA**GCCGTCC

6.3 R/G MOTIF VERSIONS

Table 6-3: R/G motif versions

R/G Motif Version	Sequence (5'→3')
V1	GTGGAATAGTATAACAATATGCTAAATGTTGTTATAGTATCCCAC
V2	GTGGAAGAGGAGAACAATATGCTAAATGTTGTTCTCGTCTCCCAC
V3	GTGGACGCGGAGACCAATATGCTAAATGTTGGTCTCGGCGCCCAC
V4	GTGGAGCGGCGCGCCGGCGCGCTAAGCGCCGGCGCGGCGCCCCAC
V20	GTGGTCGAGAAGAGGAGAACAATATGCTAAATGTTGTTCTCGTCTCCTCGACCAC
V21	GGTGTCGAGGAGACGAGAACAACATTTAGCATATTGTTCTCCTCTTCTCGACACC

6.4 VECTOR SEQUENCES

Table 6-4: Vector sequences

pTS247(pSilencer R/G gRNA cloning vector)						
1	TCGCGCGTTT	CGGTGATGAC	GGTGAAAACC	TCTGACACAT	GCAGCTCCCG	GAGACGGTCA
61	CAGCTTGTCT	GTAAGCGGAT	GCCGGGAGCA	GACAAGCCCG	TCAGGGCGCG	TCAGCGGGTG
121	TTGGCGGGTG	TCGGGGCTGG	CTTAACTATG	CGGCATCAGA	GCAGATTGTA	CTGAGAGTGC
181	ACCATATGCG	GTGTGAAATA	CCGCACAGAT	GCGTAAGGAG	AAAATACCGC	ATCAGGCGCC
241	ATTCGCCATT	CAGGCTGCGC	AACTGTTGGG	AAGGGCGATC	GGTGCGGGCC	TCTTCGCTAT
301	TACGCCAGCT	GGCGAAAGGG	GGATGTGCTG	CAAGGCGATT	AAGTTGGGTA	ACGCCAGGGT
361	TTTCCCAGTC	ACGACGTTGT	AAAACGACGG	CCAGTGCCAA	GCTTATCGAC	TACGGAAGAC
421	TAGTGGGATA	CTATAACAAC	ATTTAGCATA	TTGTTATACT	ATTCCACGGA	TCCCGCGTCC
481	TTTCCACAAG	ATATATAAAC	CCAAGAAATC	GAAATACTTT	CAAGTTACGG	TAAGCATATG
541	ATAGTCCATT	TTAAAACATA	ATTTTAAAAC	TGCAAACTAC	CCAAGAAATT	ATTACTTTCT
601	ACGTCACGTA	TTTTGTACTA	ATATCTTTGT	GTTTACAGTC	AAATTAATTC	TAATTATCTC
661	TCTAACAGCC	TTGTATCGTA	TATGCAAATA	TGAAGGAATC	ATGGGAAATA	GGCCCTCTTC
721	CTGCCCGACC	TTGGCGCGCG	CTCGGCGCGC	GGTCACGCTC	CGTCACGTGG	TGCGTTTTGC
781	CTGCGCGTCT	TTCCACTGGG	GAATTCATGC	TTCTCCTCCC	TTTAGTGAGG	GTAATTCTCT
841	CTCTCTCCCT	ATAGTGAGTC	GTATTAATTC	CTTCTCTTCT	ATAGTGTCAC	CTAAATCGTT
901	GCAATTCGTA	ATCATGTCAT	AGCTGTTTCC	TGTGTGAAAT	TGTTATCCGC	TCACAATTCC
961	ACACAACATA	CGAGCCGGAA	GCATAAAGTG	TAAAGCCTGG	GGTGCCTAAT	GAGTGAGCTA
1021	ACTCACATTA	ATTGCGTTGC	GCTCACTGCC	CGCTTTCCAG	TCGGGAAACC	TGTCGTGCCA
1081	GCTGCATTAA	TGAATCGGCC	AACGCGCGGG	GAGAGGCGGT	TTGCGTATTG	GGCGCTCTTC
1141	CGCTTCCTCG	CTCACTGACT	CGCTGCGCTC	GGTCGTTCGG	CTGCGGCGAG	CGGTATCAGC
1201	TCACTCAAAG	GCGGTAATAC	GGTTATCCAC	AGAATCAGGG	GATAACGCAG	GAAAGAACAT
1261	GTGAGCAAAA	GGCCAGCAAA	AGGCCAGGAA	CCGTAAAAAG	GCCGCGTTGC	TGGCGTTTTT
1321	CCATAGGCTC	CGCCCCCCTG	ACGAGCATCA	CAAAAATCGA	CGCTCAAGTC	AGAGGTGGCG
1381	AAACCCGACA	GGACTATAAA	GATACCAGGC	GTTTCCCCCT	GGAAGCTCCC	TCGTGCGCTC
1441	TCCTGTTCCG	ACCCTGCCGC	TTACCGGATA	CCTGTCCGCC	TTTCTCCCTT	CGGGAAGCGT
1501	GGCGCTTTCT	CATAGCTCAC	GCTGTAGGTA	TCTCAGTTCG	GTGTAGGTCG	TTCGCTCCAA
1561	GCTGGGCTGT	GTGCACGAAC	CCCCCGTTCA	GCCCGACCGC	TGCGCCTTAT	CCGGTAACTA
1621	TCGTCTTGAG	TCCAACCCGG	TAAGACACGA	CTTATCGCCA	CTGGCAGCAG	CCACTGGTAA
1681	CAGGATTAGC	AGAGCGAGGT	ATGTAGGCGG	TGCTACAGAG	TTCTTGAAGT	GGTGGCCTAA
1741	CTACGGCTAC	ACTAGAAGAA	CAGTATTTGG	TATCTGCGCT	CTGCTGAAGC	CAGTTACCTT
1801	CGGAAAAAGA	GTTGGTAGCT	CTTGATCCGG	CAAAAAAACC	ACCGCTGGTA	GCGGTGGTTT
1861	TTTTGTTTGC	AAGCAGCAGA	TTACGCGCAG	AAAAAAAGGA	TCTCAAGAAG	ATCCTTTGAT
1921	CTTTTCTACG	GGGTCTGACG	CTCAGTGGAA	CGAAAACTCA	CGTTAAGGGA	TTTTGGTCAT
1981	GAGATTATCA	AAAAGGATCT	TCACCTAGAT	CCTTTTAAAT	TAAAAATGAA	GTTTTAAATC
2041	AATCTAAAGT	ATATATGAGT	AAACTTGGTC	TGACAGTTAC	CAATGCTTAA	TCAGTGAGGC

```
2101    ACCTATCTCA GCGATCTGTC TATTTCGTTC ATCCATAGTT GCCTGACTCC CCGTCGTGTA
2161    GATAACTACG ATACGGGAGG GCTTACCATC TGGCCCCAGT GCTGCAATGA TACCGCGAGA
2221    CCCACGCTCA CCGGCTCCAG ATTTATCAGC AATAAACCAG CCAGCCGGAA GGGCCGAGCG
2281    CAGAAGTGGT CCTGCAACTT TATCCGCCTC CATCCAGTCT ATTAATTGTT GCCGGGAAGC
2341    TAGAGTAAGT AGTTCGCCAG TTAATAGTTT GCGCAACGTT GTTGCCATTG CTACAGGCAT
2401    CGTGGTGTCA CGCTCGTCGT TTGGTATGGC TTCATTCAGC TCCGGTTCCC AACGATCAAG
2461    GCGAGTTACA TGATCCCCCA TGTTGTGCAA AAAAGCGGTT AGCTCCTTCG GTCCTCCGAT
2521    CGTTGTCAGA AGTAAGTTGG CCGCAGTGTT ATCACTCATG GTTATGGCAG CACTGCATAA
2581    TTCTCTTACT GTCATGCCAT CCGTAAGATG CTTTTCTGTG ACTGGTGAGT ACTCAACCAA
2641    GTCATTCTGA GAATAGTGTA TGCGGCGACC GAGTTGCTCT TGCCCGGCGT CAATACGGGA
2701    TAATACCGCG CCACATAGCA GAACTTTAAA AGTGCTCATC ATTGGAAAAC GTTCTTCGGG
2761    GCGAAAACTC TCAAGGATCT TACCGCTGTT GAGATCCAGT TCGATGTAAC CCACTCGTGC
2821    ACCCAACTGA TCTTCAGCAT CTTTTACTTT CACCAGCGTT TCTGGGTGAG CAAAAACAGG
2881    AAGGCAAAAT GCCGCAAAAA AGGGAATAAG GGCGACACGG AAATGTTGAA TACTCATACT
2941    CTTCCTTTTT CAATATTCAG GCTTACCTTT ATGCTTCCGG CTCGTATGTT GTGTGGAATT
3001    GTGAGCGGAT AACAATTTCA CACAGGAAAC AGCTATGACC ATGATTACGC CAAGCTCTAG
3061    CTAGAGGTCG ACGGTATACA GACATGATAA GATACATTGA TGAGTTTGGA CAAACCACAA
3121    CTAGAATGCA GTGAAAAAAA TGCTTTATTT GTGAAATTTG TGATGCTATT GCTTTATTTG
3181    TAACCATTAT AAGCTGCAAT AAACAAGTTG GGGTGGGCGA AGAACTCCAG CATGAGATCC
3241    CCGCGCTGGA GGATCATCCA GCCGGCGTCC CGGAAAACGA TTCCGAAGCC CAACCTTTCA
3301    TAGAAGGCGG CGGTGGAATC GAAATCTCGT AGCACGTGCT ATTCCTTTGC CCTCGGACGA
3361    GTGCTGGGGC GTCGGTTTCC ACTATCGGCG AGTACTTCTA CACAGCCATC GGTCCAGACG
3421    GCCGCGCTTC TGCGGGCGAT TTGTGTACGC CCGACAGTCC CGGCTCCGGA TCGGACGATT
3481    GCGTCGCATC GACCCTGCGC CCAAGCTGCA TCATCGAAAT TGCCGTCAAC CAAGCTCTGA
3541    TAGAGTTGGT CAAGACCAAT GCGGAGCATA TACGCCCGGA GCCGCGGCGA TCCTGCAAGC
3601    TCCGGATGCC TCCGCTCGAA GTAGCGCGTC TGCTGCTCCA TACAAGCCAA CCACGGCCTC
3661    CAGAAGAAGA TGTTGGCGAC CTCGTATTGG GAATCCCCGA ACATCGCCTC GCTCCAGTCA
3721    ATGACCGCTG TTATGCGGCC ATTGTCCGTC AGGACATTGT TGGAGCCGAA ATCCGCGTGC
3781    ACGAGGTGCC GGACTTCGGG GCAGTCCTCG GCCCAAAGCA TCAGCTCATC GAGAGCCTGC
3841    GCGACGGACG CACTGACGGT GTCGTCCATC ACAGTTTGCC AGTGATACAC ATGGGGATCA
3901    GCAATCGCGC ATATGAAATC ACGCCATGTA GTGTATTGAC CGATTCCTTG CGGTCCGAAT
3961    GGGCCGAACC CGCTCGTCTG GCTAAGATCG GCCGCAGCGA TCGCATCCAT GGCCTCCGCG
4021    ACCGGCTGCA GAACAGCGGG CAGTTCGGTT TCAGGCAGGT CTTGCAACGT GACACCCTGT
4081    GCACGGCGGG AGATGCAATA GGTCAGGCTC TCGCTGAATT CCCCAATGTC AAGCACTTCC
4141    GGAATCGGGA GCGCGGCCGA TGCAAAGTGC CGATAAACAT AACGATCTTT GTAGAAACCA
4201    TCGGCGCAGC TATTTACCCG CAGGACATAT CCACGCCCTC CTACATCGAA GCTGAAAGCA
4261    CGAGATTCTT CGCCCTCCGA GAGCTGCATC AGGTCGGAGA CGCTGTCGAA CTTTTCGATC
4321    AGAAACTTCT CGACAGACGT CGCGGTGAGT TCAGGCTTTT TCATCACGTG CTGATCAGAT
4381    CCGAAAATGG ATATACAAGC TCCCGGGAGC TTTTTGCAAA AGCCTAGGCC TCCAAAAAAG
4441    CCTCCCCACT ACTTCTGGAA TAGCTCAGAG GCAGAGGCGG CCTCGGCCTC TGCATAAATA
4501    AAAAAAATTA GTCAGCCATG GGGCGGAGAA TGGGCGGAAC TGGGCGGAGT TAGGGGCGGG
4561    ATGGGCGGAG TTAGGGGCGG GACTATGGTT GCTGACTAAT TGAGATGCAT GCTTTGCATA
4621    CTTCTGCCTG CTGGGGAGCC TGGGGACTTT CCACACCTGG TTGCTAACTA ATTGAGATGC
4681    ATGCTTTGCA TACTTCTGCC TGCTGGAATA TTATTGAAGC ATTTATCAGG GTTATTGTCT
4741    CATGAGCGGA TACATATTTG AATGTATTTA GAAAAATAAA CAAATAGGGG TTCCGCGCAC
4801    ATTTCCCCGA AAAGTGCCAC CTGACGTCTA AGAAACCATT ATTATCATGA CATTAACCTA
4861    TAAAAATAGG CGTATCACGA GGCCCTTTCG TC
```

pTS296(pEdit1.2 containing ADAR2)

```
1       GCTAGCGCGA TGTACGGGCC AGATATACGC GTTGACATTG ATTATTGACT AGTTATTAAT
61      AGTAATCAAT TACGGGGTCA TTAGTTCATA GCCCATATAT GGAGTTCCGC GTTACATAAC
121     TTACGGTAAA TGGCCCGCCT GGCTGACCGC CCAACGACCC CCGCCCATTG ACGTCAATAA
181     TGACGTATGT TCCCATAGTA ACGCCAATAG GGACTTTCCA TTGACGTCAA TGGGTGGAGT
241     ATTTACGGTA AACTGCCCAC TTGGCAGTAC ATCAAGTGTA TCATATGCCA AGTACGCCCC
301     CTATTGACGT CAATGACGGT AAATGGCCCG CCTGGCATTA TGCCCAGTAC ATGACCTTAT
361     GGGACTTTCC TACTTGGCAG TACATCTACG TATTAGTCAT CGCTATTACC ATGGTGATGC
421     GGTTTTGGCA GTACATCAAT GGGCGTGGAT AGCGGTTTGA CTCACGGGGA TTTCCAAGTC
481     TCCACCCCAT TGACGTCAAT GGGAGTTTGT TTTGGCACCA AAATCAACGG GACTTTCCAA
541     AATGTCGTAA CAACTCCGCC CCATTGACGC AAATGGGCGG TAGGCGTGTA CGGTGGGAGG
601     TCTATATAAG CAGAGCTCTC TGGCTAACTA GAGAACCCAC TGCTTACTGG CTTATCGAAA
661     TTAATACGAC TCACTATAGG GAGACCCAAG CTTGGTACCG AGCTCGGATC CACCATGGAT
721     ATAGAAGATG AAGAAAACAT GAGTTCCAGC AGCACTGATG TGAAGGAAAA CCGCAATCTG
781     GACAACGTGT CCCCCAAGGA TGGCAGCACA CCTGGGCCTG GCGAGGGCTC TCAGCTCTCC
```

```
841     AATGGGGGTG GTGGTGGCCC CGGCAGAAAG CGGCCCCTGG AGGAGGGCAG CAATGGCCAC
901     TCCAAGTACC GCCTGAAGAA AAGGAGGAAA ACACCAGGGC CCGTCCTCCC CAAGAACGCC
961     CTGATGCAGC TGAATGAGAT CAAGCCTGGT TTGCAGTACA CACTCCTGTC CCAGACTGGG
1021    CCCGTGCACG CGCCTTTGTT TGTCATGTCT GTGGAGGTGA ATGGCCAGGT TTTTGAGGGC
1081    TCTGGTCCCA CAAAGAAAAA GGCAAAACTC CATGCTGCTG AGAAGGCCTT GAGGTCTTTC
1141    GTTCAGTTTC CTAATGCCTC TGAGGCCCAC CTGGCCATGG GGAGGACCCT GTCTGTCAAC
1201    ACGGACTTCA CATCTGACCA GGCCGACTTC CCTGACACGC TCTTCAATGG TTTTGAAACT
1261    CCTGACAAGG CGGAGCCTCC CTTTTACGTG GGCTCCAATG GGGATGACTC CTTCAGTTCC
1321    AGCGGGGACC TCAGCTTGTC TGCTTCCCCG GTGCCTGCCA GCCTAGCCCA GCCTCCTCTC
1381    CCTGCCTTAC CACCATTCCC ACCCCCGAGT GGGAAGAATC CCGTGATGAT CTTGAACGAA
1441    CTGCGCCCAG GACTCAAGTA TGACTTCCTC TCCGAGAGCG GGGAGAGCCA TGCCAAGAGC
1501    TTCGTCATGT CTGTGGTCGT GGATGGTCAG TTCTTTGAAG GCTCGGGGAG AAACAAGAAA
1561    CTTGCCAAGG CCCGGGCTGC GCAGTCTGCC CTGGCCGCCA TTTTTAACTT GCACTTGGAT
1621    CAGACGCCAT CTCGCCAGCC TATTCCCAGT GAGGGTCTCC AGCTGCATTT ACCGCAGGTT
1681    TTAGCTGACG CTGTCTCACG CCTGGTCCTG GGTAAGTTTG GTGACCTGAC CGACAACTTC
1741    TCCTCCCCTC ACGCTCGCAG AAAAGTGCTG GCTGGAGTCG TCATGACAAC AGGCACAGAT
1801    GTTAAAGATG CCAAGGTGAT AAGTGTTTCT ACAGGAACAA AATGTATTAA TGGTGAATAC
1861    ATGAGTGATC GTGGCCTTGC ATTAAATGAC TGCCATGCAG AAATAATATC TCGGAGATCC
1921    TTGCTCAGAT TTCTTTATAC ACAACTTGAG CTTTACTTAA ATAACAAAGA TGATCAAAAA
1981    AGATCCATCT TTCAGAAATC AGAGCGAGGG GGGTTTAGGC TGAAGGAGAA TGTCCAGTTT
2041    CATCTGTACA TCAGCACCTC TCCCTGTGGA GATGCCAGAA TCTTCTCACC ACATGAGCCA
2101    ATCCTGGAAG AACCAGCAGA TAGACACCCA AATCGTAAAG CAAGAGGACA GCTACGGACC
2161    AAAATAGAGT CTGGTGAGGG GACGATTCCA GTGCGCTCCA ATGCGAGCAT CCAAACGTGG
2221    GACGGGGTGC TGCAAGGGGA GCGGCTGCTC ACCATGTCCT GCAGTGACAA GATTGCACGC
2281    TGGAACGTGG TGGGCATCCA GGGTTCCCTG CTCAGCATTT TCGTGGAGCC CATTTACTTC
2341    TCGAGCATCA TCCTGGGCAG CCTTTACCAC GGGGACCACC TTTCCAGGGC CATGTACCAG
2401    CGGATCTCCA ACATAGAGGA CCTGCCACCT CTCTACACCC TCAACAAGCC TTTGCTCAGT
2461    GGCATCAGCA ATGCAGAAGC ACGGCAGCCA GGGAAGGCCC CCAACTTCAG TGTCAACTGG
2521    ACGGTAGGCG ACTCCGCTAT TGAGGTCATC AACGCCACGA CTGGGAAGGA TGAGCTGGGC
2581    CGCGCGTCCC GCCTGTGTAA GCACGCGTTG TACTGTCGCT GGATGCGTGT GCACGGCAAG
2641    GTTCCCTCCC ACTTACTACG CTCCAAGATT ACCAAACCCA ACGTGTACCA TGAGTCCAAG
2701    CTGGCGGCAA AGGAGTACCA GGCCGCCAAG GCGCGTCTGT TCACAGCCTT CATCAAGGCG
2761    GGGCTGGGGG CCTGGGTGGA GAAGCCCACC GAGCAGGACC AGTTCTCACT CACGCCCTCT
2821    AGAGGGCCCT ATTCTATAGT GTCACCTAAA TGCGGGGGAG GCTCACATCA TCACCATCAC
2881    CATTAGGAAT TCCGACTGTG CCTTCTAGTT GCCAGCCATC TGTTGTTTGC CCCTCCCCCG
2941    TGCCTTCCTT GACCCTGGAA GGTGCCACTC CCACTGTCCT TTCCTAATAA AATGAGGAAA
3001    TTGCATCGCA TTGTCTGAGT AGGTGTCATT CTATTCTGGG GGGTGGGGTG GGGCAGGACA
3061    GCAAGGGGGA GGATTGGGAA GACAATAGCA GGCATGCTGG GGATGCGGTG GGCTCTATGG
3121    GGTACCCCGC CCATCCCGCC CCTAACTCCG CCCAGTTCCG CCCATTCTCC GCCCCATGGC
3181    TGACTAATTT TTTTTATTTA TGCAGAGGCC GAGGCCGCCT CTGCCTCTGA GCTATTCACC
3241    GGTGGCCGCG TTGCTGGCGT TTTTCCATAG GCTCCGCCCC CCTGACGAGC ATCACAAAAA
3301    TCGACGCTCA AGTCAGAGGT GGCGAAACCC GACAGGACTA TAAAGATACC AGGCGTTTCC
3361    CCCTGGAAGC TCCCTCGTGC GCTCTCCTGT TCCGACCCTG CCGCTTACCG GATACCTGTC
3421    CGCCTTTCTC CCTTCGGGAA GCGTGGCGCT TTCTCATAGC TCACGCTGTA GGTATCTCAG
3481    TTCGGTGTAG GTCGTTCGCT CCAAGCTGGG CTGTGTGCAC GAACCCCCCG TTCAGCCCGA
3541    CCGCTGCGCC TTATCCGGTA ACTATCGTCT TGAGTCCAAC CCGGTAAGAC ACGACTTATC
3601    GCCACTGGCA GCAGCCACTG GTAACAGGAT TAGCAGAGCG AGGTATGTAG GCGGTGCTAC
3661    AGAGTTCTTG AAGTGGTGGC CTAACTACGG CTACACTAGA AGAACAGTAT TTGGTATCTG
3721    CGCTCTGCTG AAGCCAGTTA CCTTCGGAAA AAGAGTTGGT AGCTCTTGAT CCGGCAAACA
3781    AACCACCGCT GGTAGCGGTG GTTTTTTTGT TTGCAAGCAG CAGATTACGC GCAGAAAAAA
3841    AGGATCTCAA GAAGATCCTT TGATCTTTTC TACGGGGTCT GACGCTCAGT GGAACGAAAA
3901    CTCACGTTAA GGGATTTTGG TCATGAGATT ATCAAAAAGG ATCTTCACCT AGATCCTTTT
3961    AAATTAAAAA TGAAGTTTTA AATCAATCTA AAGTATATAT GAGTAAACTT GGTCTGACAG
4021    TTACCAATGC TTAATCAGTG AGGCACCTAT CTCAGCGATC TGTCTATTTC GTTCATCCAT
4081    AGTTGCCTGA CTCCCCGTCG TGTAGATAAC TACGATACGG GAGGGCTTAC CATCTGGCCC
4141    CAGTGCTGCA ATGATACCGC GAGACCCACG CTCACCGGCT CCAGATTTAT CAGCAATAAA
4201    CCAGCCAGCC GGAAGGGCCG AGCGCAGAAG TGGTCCTGCA ACTTTATCCG CCTCCATCCA
4261    GTCTATTAAT TGTTGCCGGG AAGCTAGAGT AAGTAGTTCG CCAGTTAATA GTTTGCGCAA
4321    CGTTGTTGCC ATTGCTACAG GCATCGTGGT GTCACGCTCG TCGTTTGGTA TGGCTTCATT
4381    CAGCTCCGGT TCCCAACGAT CAAGGCGAGT TACATGATCC CCCATGTTGT GCAAAAAAGC
4441    GGTTAGCTCC TTCGGTCCTC CGATCGTTGT CAGAAGTAAG TTGGCCGCAG TGTTATCACT
4501    CATGGTTATG GCAGCACTGC ATAATTCTCT TACTGTCATG CCATCCGTAA GATGCTTTTC
```

```
4561    TGTGACTGGT GAGTACTCAA CCAAGTCATT CTGAGAATAG TGTATGCGGC GACCGAGTTG
4621    CTCTTGCCCG GCGTCAATAC GGGATAATAC CGCGCCACAT AGCAGAACTT TAAAAGTGCT
4681    CATCATTGGA AAACGTTCTT CGGGGCGAAA ACTCTCAAGG ATCTTACCGC TGTTGAGATC
4741    CAGTTCGATG TAACCCACTC GTGCACCCAA CTGATCTTCA GCATCTTTTA CTTTCACCAG
4801    CGTTTCTGGG TGAGCAAAAA CAGGAAGGCA AAATGCCGCA AAAAAGGGAA TAAGGGCGAC
4861    ACGGAAATGT TGAATACTCA TACTCTTCCT TTTTCAATAT TATTGAAGCA TTTATCAGGG
4921    TTATTGTCTC ATGAGCGGAT ACATATTTGA ACGGCCGGAT TTCTTGTCAC GCTTTGATGC
4981    ATCGTTGGTG GTTGATGGAT ATCTGACGAA CATGTGCACA GTGGTACAAG GTATTCCTGG
5041    TTG
```

pTS363 (pEdit1.2 containing ADAR2 and 5x PINK1 R/G gRNAs)

```
1       GCTAGCGCGA TGTACGGGCC AGATATACGC GTTGACATTG ATTATTGACT AGTTATTAAT
61      AGTAATCAAT TACGGGGTCA TTAGTTCATA GCCCATATAT GGAGTTCCGC GTTACATAAC
121     TTACGGTAAA TGGCCCGCCT GGCTGACCGC CCAACGACCC CCGCCCATTG ACGTCAATAA
181     TGACGTATGT TCCCATAGTA ACGCCAATAG GGACTTTCCA TTGACGTCAA TGGGTGGAGT
241     ATTTACGGTA AACTGCCCAC TTGGCAGTAC ATCAAGTGTA TCATATGCCA AGTACGCCCC
301     CTATTGACGT CAATGACGGT AAATGGCCCG CCTGGCATTA TGCCCAGTAC ATGACCTTAT
361     GGGACTTTCC TACTTGGCAG TACATCTACG TATTAGTCAT CGCTATTACC ATGGTGATGC
421     GGTTTTGGCA GTACATCAAT GGGCGTGGAT AGCGGTTTGA CTCACGGGGA TTTCCAAGTC
481     TCCACCCCAT TGACGTCAAT GGGAGTTTGT TTTGGCACCA AAATCAACGG GACTTTCCAA
541     AATGTCGTAA CAACTCCGCC CCATTGACGC AAATGGGCGG TAGGCGTGTA CGGTGGGAGG
601     TCTATATAAG CAGAGCTCTC TGGCTAACTA GAGAACCCAC TGCTTACTGG CTTATCGAAA
661     TTAATACGAC TCACTATAGG GAGACCCAAG CTTGGTACCG AGCTCGGATC CACCATGGAT
721     ATAGAAGATG AAGAAAACAT GAGTTCCAGC AGCACTGATG TGAAGGAAAA CCGCAATCTG
781     GACAACGTGT CCCCCAAGGA TGGCAGCACA CCTGGGCCTG GCGAGGGCTC TCAGCTCTCC
841     AATGGGGGTG GTGGTGGCCC CGGCAGAAAG CGGCCCCTGG AGGAGGGCAG CAATGGCCAC
901     TCCAAGTACC GCCTGAAGAA AAGGAGGAAA ACACCAGGGC CCGTCCTCCC CAAGAACGCC
961     CTGATGCAGC TGAATGAGAT CAAGCCTGGT TTGCAGTACA CACTCCTGTC CCAGACTGGG
1021    CCCGTGCACG CGCCTTTGTT TGTCATGTCT GTGGAGGTGA ATGGCCAGGT TTTTGAGGGC
1081    TCTGGTCCCA CAAAGAAAAA GGCAAAACTC CATGCTGCTG AGAAGGCCTT GAGGTCTTTC
1141    GTTCAGTTTC CTAATGCCTC TGAGGCCCAC CTGGCCATGG GGAGGACCCT GTCTGTCAAC
1201    ACGGACTTCA CATCTGACCA GGCCGACTTC CCTGACACGC TCTTCAATGG TTTTGAAACT
1261    CCTGACAAGG CGGAGCCTCC CTTTTACGTG GGCTCCAATG GGGATGACTC CTTCAGTTCC
1321    AGCGGGGACC TCAGCTTGTC TGCTTCCCCG GTGCCTGCCA GCCTAGCCCA GCCTCCTCTC
1381    CCTGCCTTAC CACCATTCCC ACCCCCGAGT GGGAAGAATC CCGTGATGAT CTTGAACGAA
1441    CTGCGCCCAG GACTCAAGTA TGACTTCCTC TCCGAGAGCG GGGAGAGCCA TGCCAAGAGC
1501    TTCGTCATGT CTGTGGTCGT GGATGGTCAG TTCTTTGAAG GCTCGGGGAG AAACAAGAAA
1561    CTTGCCAAGG CCCGGGCTGC GCAGTCTGCC CTGGCCGCCA TTTTTAACTT GCACTTGGAT
1621    CAGACGCCAT CTCGCCAGCC TATTCCCAGT GAGGGTCTCC AGCTGCATTT ACCGCAGGTT
1681    TTAGCTGACG CTGTCTCACG CCTGGTCCTG GGTAAGTTTG GTGACCTGAC CGACAACTTC
1741    TCCTCCCCTC ACGCTCGCAG AAAAGTGCTG GCTGGAGTCG TCATGACAAC AGGCACAGAT
1801    GTTAAAGATG CCAAGGTGAT AAGTGTTTCT ACAGGAACAA AATGTATTAA TGGTGAATAC
1861    ATGAGTGATC GTGGCCTTGC ATTAAATGAC TGCCATGCAG AAATAATATC TCGGAGATCC
1921    TTGCTCAGAT TTCTTTATAC ACAACTTGAG CTTTACTTAA ATAACAAAGA TGATCAAAAA
1981    AGATCCATCT TTCAGAAATC AGAGCGAGGG GGGTTTAGGC TGAAGGAGAA TGTCCAGTTT
2041    CATCTGTACA TCAGCACCTC TCCCTGTGGA GATGCCAGAA TCTTCTCACC ACATGAGCCA
2101    ATCCTGGAAG AACCAGCAGA TAGACACCCA AATCGTAAAG CAAGAGGACA GCTACGGACC
2161    AAAATAGAGT CTGGTGAGGG GACGATTCCA GTGCGCTCCA ATGCGAGCAT CCAAACGTGG
2221    GACGGGGTGC TGCAAGGGGA GCGGCTGCTC ACCATGTCCT GCAGTGACAA GATTGCACGC
2281    TGGAACGTGG TGGGCATCCA GGGTTCCCTG CTCAGCATTT TCGTGGAGCC CATTTACTTC
2341    TCGAGCATCA TCCTGGGCAG CCTTTACCAC GGGGACCACC TTTCCAGGGC CATGTACCAG
2401    CGGATCTCCA ACATAGAGGA CCTGCCACCT CTCTACACCC TCAACAAGCC TTTGCTCAGT
2461    GGCATCAGCA ATGCAGAAGC ACGGCAGCCA GGGAAGGCCC CCAACTTCAG TGTCAACTGG
2521    ACGGTAGGCG ACTCCGCTAT TGAGGTCATC AACGCCACGA CTGGGAAGGA TGAGCTGGGC
2581    CGCGCGTCCC GCCTGTGTAA GCACGCGTTG TACTGTCGCT GGATGCGTGT GCACGGCAAG
2641    GTTCCCTCCC ACTTACTACG CTCCAAGATT ACCAAACCCA ACGTGTACCA TGAGTCCAAG
2701    CTGGCGGCAA AGGAGTACCA GGCCGCCAAG GCGCGTCTGT TCACAGCCTT CATCAAGGCG
2761    GGGCTGGGGG CCTGGGTGGA GAAGCCCACC GAGCAGGACC AGTTCTCACT CACGCCCTCT
2821    AGAGGGCCCT ATTCTATAGT GTCACCTAAA TGCGGGGGAG GCTCACATCA TCACCATCAC
2881    CATTAGGAAT TCCGACTGTG CCTTCTAGTT GCCAGCCATC TGTTGTTTGC CCCTCCCCCG
2941    TGCCTTCCTT GACCCTGGAA GGTGCCACTC CCACTGTCCT TTCCTAATAA AATGAGGAAA
3001    TTGCATCGCA TTGTCTGAGT AGGTGTCATT CTATTCTGGG GGGTGGGGTG GGGCAGGACA
3061    GCAAGGGGGA GGATTGGGAA GACAATAGCA GGCATGCTGG GGATGCGGTG GGCTCTATGG
```

```
3121     GGTACCCCGC CCATCCCGCC CCTAACTCCG CCCAGTTCCG CCCATTCTCC GCCCCATGGC
3181     TGACTAATTT TTTTTATTTA TGCAGAGGCC GAGGCCGCCT CTGCCTCTGA GCTATTCACC
3241     GGTGGCCGCG TTGCTGGCGT TTTTCCATAG GCTCCGCCCC CCTGACGAGC ATCACAAAAA
3301     TCGACGCTCA AGTCAGAGGT GGCGAAACCC GACAGGACTA TAAAGATACC AGGCGTTTCC
3361     CCCTGGAAGC TCCCTCGTGC GCTCTCCTGT TCCGACCCTG CCGCTTACCG GATACCTGTC
3421     CGCCTTTCTC CCTTCGGGAA GCGTGGCGCT TTCTCATAGC TCACGCTGTA GGTATCTCAG
3481     TTCGGTGTAG GTCGTTCGCT CCAAGCTGGG CTGTGTGCAC GAACCCCCCG TTCAGCCCGA
3541     CCGCTGCGCC TTATCCGGTA ACTATCGTCT TGAGTCCAAC CCGGTAAGAC ACGACTTATC
3601     GCCACTGGCA GCAGCCACTG GTAACAGGAT TAGCAGAGCG AGGTATGTAG GCGGTGCTAC
3661     AGAGTTCTTG AAGTGGTGGC CTAACTACGG CTACACTAGA AGAACAGTAT TTGGTATCTG
3721     CGCTCTGCTG AAGCCAGTTA CCTTCGGAAA AAGAGTTGGT AGCTCTTGAT CCGGCAAACA
3781     AACCACCGCT GGTAGCGGTG GTTTTTTTGT TTGCAAGCAG CAGATTACGC GCAGAAAAAA
3841     AGGATCTCAA GAAGATCCTT TGATCTTTTC TACGGGGTCT GACGCTCAGT GGAACGAAAA
3901     CTCACGTTAA GGGATTTTGG TCATGAGATT ATCAAAAAGG ATCTTCACCT AGATCCTTTT
3961     AAATTAAAAA TGAAGTTTTA AATCAATCTA AAGTATATAT GAGTAAACTT GGTCTGACAG
4021     TTACCAATGC TTAATCAGTG AGGCACCTAT CTCAGCGATC TGTCTATTTC GTTCATCCAT
4081     AGTTGCCTGA CTCCCCGTCG TGTAGATAAC TACGATACGG GAGGGCTTAC CATCTGGCCC
4141     CAGTGCTGCA ATGATACCGC GAGACCCACG CTCACCGGCT CCAGATTTAT CAGCAATAAA
4201     CCAGCCAGCC GGAAGGGCCG AGCGCAGAAG TGGTCCTGCA ACTTTATCCG CCTCCATCCA
4261     GTCTATTAAT TGTTGCCGGG AAGCTAGAGT AAGTAGTTCG CCAGTTAATA GTTTGCGCAA
4321     CGTTGTTGCC ATTGCTACAG GCATCGTGGT GTCACGCTCG TCGTTTGGTA TGGCTTCATT
4381     CAGCTCCGGT TCCCAACGAT CAAGGCGAGT TACATGATCC CCCATGTTGT GCAAAAAAGC
4441     GGTTAGCTCC TTCGGTCCTC CGATCGTTGT CAGAAGTAAG TTGGCCGCAG TGTTATCACT
4501     CATGGTTATG GCAGCACTGC ATAATTCTCT TACTGTCATG CCATCCGTAA GATGCTTTTC
4561     TGTGACTGGT GAGTACTCAA CCAAGTCATT CTGAGAATAG TGTATGCGGC GACCGAGTTG
4621     CTCTTGCCCG GCGTCAATAC GGGATAATAC CGCGCCACAT AGCAGAACTT TAAAAGTGCT
4681     CATCATTGGA AAACGTTCTT CGGGGCGAAA ACTCTCAAGG ATCTTACCGC TGTTGAGATC
4741     CAGTTCGATG TAACCCACTC GTGCACCCAA CTGATCTTCA GCATCTTTTA CTTTCACCAG
4801     CGTTTCTGGG TGAGCAAAAA CAGGAAGGCA AAATGCCGCA AAAAAGGGAA TAAGGGCGAC
4861     ACGGAAATGT TGAATACTCA TACTCTTCCT TTTTCAATAT TATTGAAGCA TTTATCAGGG
4921     TTATTGTCTC ATGAGCGGAT ACATATTTGA ACGGCCGGAT TTCTTGTCAC GCTTTGATGC
4981     ATCGTTGGTG GTTGATGGAT GACGTTGTAA AACGACGGCC AGTGCCAAGC TCCCAAAAAA
5041     GGGCCCTGAA GAGGGCCCTG ATGCCTGGGC AGTGGTGGGA TACTATAACA ACATTTAGCA
5101     TATTGTTATA CTATTCCACG GATCCCGCGT CCTTTCCACA AGATATATAA ACCCAAGAAA
5161     TCGAAATACT TTCAAGTTAC GGTAAGCATA TGATAGTCCA TTTTAAAACA TAATTTTAAA
5221     ACTGCAAACT ACCCAAGAAA TTATTACTTT CTACGTCACG TATTTTGTAC TAATATCTTT
5281     GTGTTTACAG TCAAATTAAT TCTAATTATC TCTCTAACAG CCTTGTATCG TATATGCAAA
5341     TATGAAGGAA TCATGGGAAA TAGGCCCTCT TCCTGCCCGA CCTTGGCGCG CGCTCGGCGC
5401     GCGGTCACGC TCCGTCACGT GGTGCGTTTT GCCTGCGCGT CTTTCCACTG GGCATTACCT
5461     GTAGGTCACC TTTCGGACAC GGTCTCCCTG GTGACGTTGT AAAACGACGG CCAGTGCCAA
5521     GCTCCCAAAA AAGGGCCCTG AAGAGGGCCC TGATGCCTGG GCAGTGGTGG GATACTATAA
5581     CAACATTTAG CATATTGTTA TACTATTCCA CGGATCCCGC GTCCTTTCCA CAAGATATAT
5641     AAACCCAAGA AATCGAAATA CTTTCAAGTT ACGGTAAGCA TATGATAGTC CATTTTAAAA
5701     CATAATTTTA AAACTGCAAA CTACCCAAGA AATTATTACT TTCTACGTCA CGTATTTTGT
5761     ACTAATATCT TTGTGTTTAC AGTCAAATTA ATTCTAATTA TCTCTCTAAC AGCCTTGTAT
5821     CGTATATGCA AATATGAAGG AATCATGGGA AATAGGCCCT CTTCCTGCCC GACCTTGGCG
5881     CGCGCTCGGC GCGCGGTCAC GCTCCGTCAC GTGGTGCGTT TTGCCTGCGC GTCTTTCCTC
5941     TGGGACCCTC ACTCTTTGCC CCGCTATACC TACCCACCCT CATAGACGTT GTAAAACGAC
6001     GGCCAGTGCC AAGCTCCCAA AAAAGGGCCC TGAAGAGGGC CCTGATGCCT GGGCAGTGGT
6061     GGGATACTAT AACAACATTT AGCATATTGT TATACTATTC CACGGATCCC GCGTCCTTTC
6121     CACAAGATAT ATAAACCCAA GAAATCGAAA TACTTTCAAG TTACGGTAAG CATATGATAG
6181     TCCATTTTAA AACATAATTT TAAAACTGCA AACTACCCAA GAAATTATTA CTTTCTACGT
6241     CACGTATTTT GTACTAATAT CTTTGTGTTT ACAGTCAAAT TAATTCTAAT TATCTCTCTA
6301     ACAGCCTTGT ATCGTATATG CAAATATGAA GGAATCATGG GAAATAGGCC CTCTTCCTGC
6361     CCGACCTTGG CGCGCGCTCG GCGCGCGGTC ACGCTCCGTC ACGTGGTGCG TTTTGCCTGC
6421     GCGTCTTTCC ACTGGGCCGG TACATCTAGT AAGCAGTGAT CAGACGGCAG TATGTGGAGT
6481     TGTAAAACGA CGGCCAGTGC CAAGCTCCCA AAAAAGGGCC CTGAAGAGGG CCCTGATGCC
6541     TGGGCAGTGG TGGGATACTA TAACAACATT TAGCATATTG TTATACTATT CCACGGATCC
6601     CGCGTCCTTT CCACAAGATA TATAAACCCA AGAAATCGAA ATACTTTCAA GTTACGGTAA
6661     GCATATGATA GTCCATTTTA AAACATAATT TTAAAACTGC AAACTACCCA AGAAATTATT
6721     ACTTTCTACG TCACGTATTT TGTACTAATA TCTTTGTGTT TACAGTCAAA TTAATTCTAA
6781     TTATCTCTCT AACAGCCTTG TATCGTATAT GCAAATATGA AGGAATCATG GGAAATAGGC
```

```
6841     CCTCTTCCTG CCCGACCTTG GCGCGCGCTC GGCGCGCGGT CACGCTCCGT CACGTGGTGC
6901     GTTTTGCCTG CGCGTCTTTC CACTGGGTCG CACTACGGAA TATATTTGCG CTACTGCCTG
6961     CATAATGGGA CGTTGTAAAA CGACGGCCAG TGCCAAGCTC CCAAAAAAGG GCCCTGAAGA
7021     GGGCCCTGAT GCCTGGGCAG TGGTGGGATA CTATAACAAC ATTTAGCATA TTGTTATACT
7081     ATTCCACGGA TCCCGCGTCC TTTCCACAAG ATATATAAAC CCAAGAAATC GAAATACTTT
7141     CAAGTTACGG TAAGCATATG ATAGTCCATT TTAAAACATA ATTTTAAAAC TGCAAACTAC
7201     CCAAGAAATT ATTACTTTCT ACGTCACGTA TTTTGTACTA ATATCTTTGT GTTTACAGTC
7261     AAATTAATTC TAATTATCTC TCTAACAGCC TTGTATCGTA TATGCAAATA TGAAGGAATC
7321     ATGGGAAATA GGCCCTCTTC CTGCCCGACC TTGGCGCGCG CTCGGCGCGC GGTCACGCTC
7381     CGTCACGTGG TGCGTTTTGC CTGCGCGTCT TTCCACTGGG ATCTGACGAA CATGTGCACA
7441     GTGGTACAAG GTATTCCTGG TTG
```

pTS554 (pShuttle-CMV containing dual-luciferase wt/wt)

```
1        NNNTTAATTA ANNNTCCCTT CCAGCTCTCT GCCCCTTTTG GATTGAAGCC AATATGATAA
61       TGAGGGGGTG GAGTTTGTGA CGTGGCGCGG GGCGTGGGAA CGGGGCGGGT GACGTAGTAG
121      TGTGGCGGAA GTGTGATGTT GCAAGTGTGG CGGAACACAT GTAAGCGACG GATGTGGCAA
181      AAGTGACGTT TTTGGTGTGC GCCGGTGTAC ACAGGAAGTG ACAATTTTCG CGCGGTTTTA
241      GGCGGATGTT GTAGTAAATT TGGGCGTAAC CGAGTAAGAT TTGGCCATTT TCGCGGGAAA
301      ACTGAATAAG AGGAAGTGAA ATCTGAATAA TTTTGTGTTA CTCATAGCGC GTAANNNNTA
361      ATAGTAATCA ATTACGGGGT CATTAGTTCA TAGCCCATAT ATGGAGTTCC GCGTTACATA
421      ACTTACGGTA AATGGCCCGC CTGGCTGACC GCCCAACGAC CCCCGCCCAT TGACGTCAAT
481      AATGACGTAT GTTCCCATAG TAACGCCAAT AGGGACTTTC CATTGACGTC AATGGGTGGA
541      GTATTTACGG TAAACTGCCC ACTTGGCAGT ACATCAAGTG TATCATATGC CAAGTACGCC
601      CCCTATTGAC GTCAATGACG GTAAATGGCC CGCCTGGCAT TATGCCCAGT ACATGACCTT
661      ATGGGACTTT CCTACTTGGC AGTACATCTA CGTATTAGTC ATCGCTATTA CCATGGTGAT
721      GCGGTTTTGG CAGTACATCA ATGGGCGTGG ATAGCGGTTT GACTCACGGG GATTTCCAAG
781      TCTCCACCCC ATTGACGTCA ATGGGAGTTT GTTTTGGCAC CAAAATCAAC GGGACTTTCC
841      AAAATGTCGT AACAACTCCG CCCCATTGAC GCAAATGGGC GGTAGGCGTG TACGGTGGGA
901      GGTCTATATA AGCAGAGCTG GTTTAGTGAA CCGTCAGATC CGCTAGAGAT CTGGTACCGT
961      CGACGCGGCC GCTCGAGCCT AAGCTTCTAG ATAAGATATG ACTTCGAAAG TTTATGATCC
1021     AGAACAAAGG AAACGGATGA TAACTGGTCC GCAGTGGTGG GCCAGATGTA AACAAATGAA
1081     TGTTCTTGAT TCATTTATTA ATTATTATGA TTCAGAAAAA CATGCAGAAA ATGCTGTTAT
1141     TTTTTTACAT GGTAACGCGG CCTCTTCTTA TTTATGGCGA CATGTTGTGC CACATATTGA
1201     GCCAGTAGCG CGGTGTATTA TACCAGACCT TATTGGTATG GGCAAATCAG GCAAATCTGG
1261     TAATGGTTCT TATAGGTTAC TTGATCATTA CAAATATCTT ACTGCATGGT TTGAACTTCT
1321     TAATTTACCA AAGAAGATCA TTTTTGTCGG CCATGATTGG GGTGCTTGTT TGGCATTTCA
1381     TTATAGCTAT GAGCATCAAG ATAAGATCAA AGCAATAGTT CACGCTGAAA GTGTAGTAGA
1441     TGTGATTGAA TCATGGGATG AATGGCCTGA TATTGAAGAA GATATTGCGT TGATCAAATC
1501     TGAAGAAGGA GAAAAAATGG TTTTGGAGAA TAACTTCTTC GTGGAAACCA TGTTGCCATC
1561     AAAAATCATG AGAAAGTTAG AACCAGAAGA ATTTGCAGCA TATCTTGAAC CATTCAAAGA
1621     GAAAGGTGAA GTTCGTCGTC CAACATTATC ATGGCCTCGT GAAATCCCGT TAGTAAAAGG
1681     TGGTAAACCT GACGTTGTAC AAATTGTTAG GAATTATAAT GCTTATCTAC GTGCAAGTGA
1741     TGATTTACCA AAAATGTTTA TTGAATCGGA CCCAGGATTC TTTTCCAATG CTATTGTTGA
1801     AGGTGCCAAG AAGTTTCCTA ATACTGAATT TGTCAAAGTA AAAGGTCTTC ATTTTTCGCA
1861     AGAAGATGCA CCTGATGAAA TGGGAAAATA TATCAAATCG TTCGTTGAGC GAGTTCTCAA
1921     AAATGAACAA GGAAGCGGAG CTACTAACTT CAGCCTGCTG AAGCAGGCTG GAGACGTGGA
1981     GGAGAACCCT GGACCTATGG AAGATGCCAA AAACATTAAG AAGGGCCCAG CGCCATTCTA
2041     CCCACTCGAA GACGGGACCG CCGGCGAGCA GCTGCACAAA GCCATGAAGC GCTACGCCCT
2101     GGTGCCCGGC ACCATCGCCT TTACCGACGC ACATATCGAG GTGGACATTA CCTACGCCGA
2161     GTACTTCGAG ATGAGCGTTC GGCTGGCAGA AGCTATGAAG CGCTATGGGC TGAATACAAA
2221     CCATCGGATC GTGGTGTGCA GCGAGAATAG CTTGCAGTTC TTCATGCCCG TGTTGGGTGC
2281     CCTGTTCATC GGTGTGGCTG TGGCCCCAGC TAACGACATC TACAACGAGC GCGAGCTGCT
2341     GAACAGCATG GGCATCAGCC AGCCCACCGT CGTATTCGTG AGCAAGAAAG GGCTGCAAAA
2401     GATCCTCAAC GTGCAAAAGA AGCTACCGAT CATACAAAAG ATCATCATCA TGGATAGCAA
2461     GACCGACTAC CAGGGCTTCC AAAGCATGTA CACCTTCGTG ACTTCCCATT TGCCACCCGG
2521     CTTCAACGAG TACGACTTCG TGCCCGAGAG CTTCGACCGG GACAAAACCA TCGCCCTGAT
2581     CATGAACAGT AGTGGCAGTA CCGGATTGCC CAAGGGCGTA GCCCTACCGC ACCGCACCGC
2641     TTGTGTCCGA TTCAGTCATG CCCGCGACCC CATCTTCGGC AACCAGATCA TCCCCGACAC
2701     CGCTATCCTC AGCGTGGTGC CATTTCACCA CGGCTTCGGC ATGTTCACCA CGCTGGGCTA
2761     CTTGATCTGC GGCTTTCGGG TCGTGCTCAT GTACCGCTTC GAGGAGGAGC TATTCTTGCG
2821     CAGCTTGCAA GACTATAAGA TTCAATCTGC CCTGCTGGTG CCCACACTAT TTAGCTTCTT
2881     CGCTAAGAGC ACTCTCATCG ACAAGTACGA CCTAAGCAAC TTGCACGAGA TCGCCAGCGG
2941     CGGGGCGCCG CTCAGCAAGG AGGTAGGTGA GGCCGTGGCC AAACGCTTCC ACCTACCAGG
```

```
3001     CATCCGCCAG GGCTACGGCC TGACAGAAAC AACCAGCGCC ATTCTGATCA CCCCCGAAGG
3061     GGACGACAAG CCTGGCGCAG TAGGCAAGGT GGTGCCCTTC TTCGAGGCTA AGGTGGTGGA
3121     CTTGGACACC GGTAAGACAC TGGGTGTGAA CCAGCGCGGC GAGCTGTGCG TCCGTGGCCC
3181     CATGATCATG AGCGGCTACG TTAACAACCC CGAGGCTACA AACGCTCTCA TCGACAAGGA
3241     CGGCTGGCTG CACAGCGGCG ACATCGCCTA CTGGGACGAG GACGAGCACT TCTTCATCGT
3301     GGACCGGCTG AAGAGCCTGA TCAAATACAA GGGCTACCAG GTAGCCCCAG CCGAACTGGA
3361     GAGCATCCTG CTGCAACACC CCAACATCTT CGACGCCGGG GTCGCCGGCC TGCCCGACGA
3421     CGATGCCGGC GAGCTGCCCG CCGCAGTCGT CGTGCTGGAA CACGGTAAAA CCATGACCGA
3481     GAAGGAGATC GTGGACTATG TGGCCAGCCA GGTCACAACC GCCAAGAAGC TGCGCGGTGG
3541     TGTTGTGTTC GTGGACGAGG TGCCTAAAGG ACTGACCGGC AAGTTGGACG CCCGCAAGAT
3601     CCGCGAGATT CTCATTAAGG CCAAGAAGGG CGGCAAGATC GCCGTGTAAG AATTCATCCG
3661     ATCCACCGGA TCTAGATAAC TGATCATAAT CAGCCATACC ACATTTGTAG AGGTTTTACT
3721     TGCTTTAAAA AACCTCCCAC ACCTCCCCCT GAACCTGAAA CATAAAATGA ATGCAATTGT
3781     TGTTGTTAAC TTGTTTATTG CAGCTTATAA TGGTTACAAA TAAAGCAATA GCATCACAAA
3841     TTTCACAAAT AAAGCATTTT TTTCACTGCA TTCTAGTTGT GGTTTGTCCA AACTCATCAA
3901     TGTATCTTAA CGCNNNNTAA GGGTGGGAAA GAATATATAA GGTGGGGGTC TTATGTAGTT
3961     TTGTATCTGT TTTGCAGCAG CCGCCGCCGC CATGAGCACC AACTCGTTTG ATGGAAGCAT
4021     TGTGAGCTCA TATTTGACAA CGCGCATGCC CCCATGGGCC GGGGTGCGTC AGAATGTGAT
4081     GGGCTCCAGC ATTGATGGTC GCCCCGTCCT GCCCGCAAAC TCTACTACCT TGACCTACGA
4141     GACCGTGTCT GGAACGCCGT TGGAGACTGC AGCCTCCGCC GCCGCTTCAG CCGCTGCAGC
4201     CACCGCCCGC GGGATTGTGA CTGACTTTGC TTTCCTGAGC CCGCTTGCAA GCAGTGCAGC
4261     TTCCCGTTCA TCCGCCCGCG ATGACAAGTT GACGGCTCTT TTGGCACAAT TGGATTCTTT
4321     GACCCGGGAA CTTAATGTCG TTTCTCAGCA GCTGTTGGAT CTGCGCCAGC AGGTTTCTGC
4381     CCTGAAGGCT TCCTCCCCTC CCAATGCGGT TTAAAACATA AATAAAAAAC CAGACTCTGT
4441     TTGGATTTGG ATCAAGCAAG TGTCTTGCTG TCTTTATTTA GGGGTTTTGC GCGCGCGGTA
4501     GGCCCGGGAC CAGCGGTCTC GGTCGTTGAG GGTCCTGTGT ATTTTTTCCA GGACGTGGTA
4561     AAGGTGACTC TGGATGTTCA GATACATGGG CATAAGCCCG TCTCTGGGGT GGAGGTAGCA
4621     CCACTGCAGA GCTTCATGCT GCGGGGTGGT GTTGTAGATG ATCCAGTCGT AGCAGGAGCG
4681     CTGGGCGTGG TGCCTAAAAA TGTCTTTCAG TAGCAAGCTG ATTGCCAGGG GCAGGCCCTT
4741     GGTGTAAGTG TTTACAAAGC GGTTAAGCTG GGATGGGTGC ATACGTGGGG ATATGAGATG
4801     CATCTTGGAC TGTATTTTTA GGTTGGCTAT GTTCCCAGCC ATATCCCTCC GGGGATTCAT
4861     GTTGTGCAGA ACCACCAGCA CAGTGTATCC GGTGCACTTG GGAAATTTGT CATGTAGCTT
4921     AGAAGGAAAT GCGTGGAAGA ACTTGGAGAC GCCCTTGTGA CCTCCAAGAT TTTCCATGCA
4981     TTCGTCCATA ATGATGGCAA TGGGCCCACG GGCGGCGGCC TGGGCGAAGA TATTTCTGGG
5041     ATCACTAACG TCATAGTTGT GTTCCAGGAT GAGATCGTCA TAGGCCATTT TTACAAAGCG
5101     CGGGCGGAGG GTGCCAGACT GCGGTATAAT GGTTCCATCC GGCCCAGGGG CGTAGTTACC
5161     CTCACAGATT TGCATTTCCC ACGCTTTGAG TTCAGATGGG GGGATCATGT CTACCTGCGG
5221     GGCGATGAAG AAAACGGTTT CCGGGGTAGG GGAGATCAGC TGGGAAGAAA GCAGGTTCCT
5281     GAGCAGCTGC GACTTACCGC AGCCGGTGGG CCCGTAAATC ACACCTATTA CCGGGTGCAA
5341     CTGGTAGTTA AGAGAGCTGC AGCTGCCGTC ATCCCTGAGC AGGGGGGCCA CTTCGTTAAG
5401     CATGTCCCTG ACTCGCATGT TTTCCCTGAC CAAATCCGCC AGAAGGCGCT CGCCGCCCAG
5461     CGATAGCAGT TCTTGCAAGG AAGCAAAGTT TTTCAACGGT TTGAGACCGT CCGCCGTAGG
5521     CATGCTTTTG AGCGTTTGAC CAAGCAGTTC CAGGCGGTCC CACAGCTCGG TCACCTGCTC
5581     TACGGCATCT CGATCCAGCA TATCTCCTCG TTTCGCGGGT TGGGGCGGCT TTCGCTGTAC
5641     GGCAGTAGTC GGTGCTCGTC CAGACGGGCC AGGGTCATGT CTTTCCACGG GCGCAGGGTC
5701     CTCGTCAGCG TAGTCTGGGT CACGGTGAAG GGGTGCGCTC CGGGCTGCGC GCTGGCCAGG
5761     GTGCGCTTGA GGCTGGTCCT GCTGGTGCTG AAGCGCTGCC GGTCTTCGCC CTGCGCGTCG
5821     GCCAGGTAGC ATTTGACCAT GGTGTCATAG TCCAGCCCCT CCGCGGCGTG GCCCTTGGCG
5881     CGCAGCTTGC CCTTGGAGGA GGCGCCGCAC GAGGGGCAGT GCAGACTTTT GAGGGCGTAG
5941     AGCTTGGGCG CGAGAAATAC CGATTCCGGG GAGTAGGCAT CCGCGCCGCA GGCCCCGCAG
6001     ACGGTCTCGC ATTCCACGAG CCAGGTGAGC TCTGGCCGTT CGGGGTCAAA AACCAGGTTT
6061     CCCCCATGCT TTTTGATGCG TTTCTTACCT CTGGTTTCCA TGAGCCGGTG TCCACGCTCG
6121     GTGACGAAAA GGCTGTCCGT GTCCCCGTAT ACAGACTNNN GTTTAAACGA ATTCNNNATA
6181     TAAAATGCAA GGTGCTGCTC AAAAAATCAG GCAAAGCCTC GCGCAAAAAA GAAAGCACAT
6241     CGTAGTCATG CTCATGCAGA TAAAGGCAGG TAAGCTCCGG AACCACCACA GAAAAAGACA
6301     CCATTTTTCT CTCAAACATG TCTGCGGGTT TCTGCATAAA CACAAAATAA AATAACAAAA
6361     AAACATTTAA ACATTAGAAG CCTGTCTTAC AACAGGAAAA ACAACCCTTA TAAGCATAAG
6421     ACGGACTACG GCCATGCCGG CGTGACCGTA AAAAAACTGG TCACCGTGAT TAAAAAGCAC
6481     CACCGACAGC TCCTCGGTCA TGTCCGGAGT CATAATGTAA GACTCGGTAA ACACATCAGG
6541     TTGATTCATC GGTCAGTGCT AAAAAGCGAC CGAAATAGCC CGGGGGAATA CATACCCGCA
6601     GGCGTAGAGA CAACATTACA GCCCCCATAG GAGGTATAAC AAAATTAATA GGAGAGAAAA
6661     ACACATAAAC ACCTGAAAAA CCCTCCTGCC TAGGCAAAAT AGCACCCTCC CGCTCCAGAA
```

```
6721     CAACATACAG CGCTTCACAG CGGCAGCCTA ACAGTCAGCC TTACCAGTAA AAAAGAAAAC
6781     CTATTAAAAA AACACCACTC GACACGGCAC CAGCTCAATC AGTCACAGTG TAAAAAAGGG
6841     CCAAGTGCAG AGCGAGTATA TATAGGACTA AAAAATGACG TAACGGTTAA AGTCCACAAA
6901     AAACACCCAG AAAACCGCAC GCGAACCTAC GCCCAGAAAC GAAAGCCAAA AAACCCACAA
6961     CTTCCTCAAA TCGTCACTTC CGTTTTCCCA CGTTACGTAA CTTCCCATTT TAAGAAAACT
7021     ACAATTCCCA ACACATACAA GTTACTCCGC CCTAAAACCT ACGTCACCCG CCCCGTTCCC
7081     ACGCCCCGCG CCACGTCACA AACTCCACCC CCTCATTATC ATATTGGCTT CAATCCAAAA
7141     TAAGGTATAT TATTGATGAT NNNTTAATTA AGGATCCNNN CGGTGTGAAA TACCGCACAG
7201     ATGCGTAAGG AGAAAATACC GCATCAGGCG CTCTTCCGCT TCCTCGCTCA CTGACTCGCT
7261     GCGCTCGGTC GTTCGGCTGC GGCGAGCGGT ATCAGCTCAC TCAAAGGCGG TAATACGGTT
7321     ATCCACAGAA TCAGGGGATA ACGCAGGAAA GAACATGTGA GCAAAAGGCC AGCAAAAGGC
7381     CAGGAACCGT AAAAAGGCCG CGTTGCTGGC GTTTTTCCAT AGGCTCCGCC CCCCTGACGA
7441     GCATCACAAA AATCGACGCT CAAGTCAGAG GTGGCGAAAC CCGACAGGAC TATAAAGATA
7501     CCAGGCGTTT CCCCCTGGAA GCTCCCTCGT GCGCTCTCCT GTTCCGACCC TGCCGCTTAC
7561     CGGATACCTG TCCGCCTTTC TCCCTTCGGG AAGCGTGGCG CTTTCTCATA GCTCACGCTG
7621     TAGGTATCTC AGTTCGGTGT AGGTCGTTCG CTCCAAGCTG GGCTGTGTGC ACGAACCCCC
7681     CGTTCAGCCC GACCGCTGCG CCTTATCCGG TAACTATCGT CTTGAGTCCA ACCCGGTAAG
7741     ACACGACTTA TCGCCACTGG CAGCAGCCAC TGGTAACAGG ATTAGCAGAG CGAGGTATGT
7801     AGGCGGTGCT ACAGAGTTCT TGAAGTGGTG GCCTAACTAC GGCTACACTA GAAGGACAGT
7861     ATTTGGTATC TGCGCTCTGC TGAAGCCAGT TACCTTCGGA AAAAGAGTTG GTAGCTCTTG
7921     ATCCGGCAAA CAAACCACCG CTGGTAGCGG TGGTTTTTTT GTTTGCAAGC AGCAGATTAC
7981     GCGCAGAAAA AAAGGATCTC AAGAAGATCC TTTGATCTTT TCTACGGGGT CTGACGCTCA
8041     GTGGAACGAA AACTCACGTT AAGGGATTTT GGTCATGAGA TTATCAAAAA GGATCTTCAC
8101     CTAGATCCTT TTAAATTAAA AATGAAGTTT TAAATCAATC TAAAGTATAT ATGAGTAAAC
8161     TTGGTCTGAC AGTTACCAAT GCTTAATCAG TGAGGCACCT ATCTCAGCGA TCTGTCTATT
8221     TCGTTCATCC ATAGTTGCCT GACTCCCCGT CGTGTAGATA ACTACGATAC GGGAGGGCTT
8281     ACCATCTGGC CCCAGTGCTG CAATGATACC GCGAGACCCA CGCTCACCGG CTCCAGATTT
8341     ATCAGCAATA AACCAGCCAG CCGGAAGGGC CGAGCGCAGA AGTGGTCCTG CAACTTTATC
8401     CGCCTCCATC CAGTCTATTA ATTGTTGCCG GGAAGCTAGA GTAAGTAGTT CGCCAGTTAA
8461     TAGTTTGCGC AACGTTGTTG NNNNAAAAAG GATCTTCACC TAGATCCTTT TCACGTAGAA
8521     AGCCAGTCCG CAGAAACGGT GCTGACCCCG GATGAATGTC AGCTACTGGG CTATCTGGAC
8581     AAGGGAAAAC GCAAGCGCAA AGAGAAAGCA GGTAGCTTGC AGTGGGCTTA CATGGCGATA
8641     GCTAGACTGG GCGGTTTTAT GGACAGCAAG CGAACCGGAA TTGCCAGCTG GGGCGCCCTC
8701     TGGTAAGGTT GGGAAGCCCT GCAAAGTAAA CTGGATGGCT TTCTCGCCGC CAAGGATCTG
8761     ATGGCGCAGG GGATCAAGCT CTGATCAAGA GACAGGATGA GGATCGTTTC GCATGATTGA
8821     ACAAGATGGA TTGCACGCAG GTTCTCCGGC CGCTTGGGTG GAGAGGCTAT TCGGCTATGA
8881     CTGGGCACAA CAGACAATCG GCTGCTCTGA TGCCGCCGTG TTCCGGCTGT CAGCGCAGGG
8941     GCGCCCGGTT CTTTTTGTCA AGACCGACCT GTCCGGTGCC CTGAATGAAC TGCAAGACGA
9001     GGCAGCGCGG CTATCGTGGC TGGCCACGAC GGGCGTTCCT TGCGCAGCTG TGCTCGACGT
9061     TGTCACTGAA GCGGGAAGGG ACTGGCTGCT ATTGGGCGAA GTGCCGGGGC AGGATCTCCT
9121     GTCATCTCAC CTTGCTCCTG CCGAGAAAGT ATCCATCATG GCTGATGCAA TGCGGCGGCT
9181     GCATACGCTT GATCCGGCTA CCTGCCCATT CGACCACCAA GCGAAACATC GCATCGAGCG
9241     AGCACGTACT CGGATGGAAG CCGGTCTTGT CGATCAGGAT GATCTGGACG AAGAGCATCA
9301     GGGGCTCGCG CCAGCCGAAC TGTTCGCCAG GCTCAAGGCG AGCATGCCCG ACGGCGAGGA
9361     TCTCGTCGTG ACCCATGGCG ATGCCTGCTT GCCGAATATC ATGGTGGAAA ATGGCCGCTT
9421     TTCTGGATTC ATCGACTGTG GCCGGCTGGG TGTGGCGGAC CGCTATCAGG ACATAGCGTT
9481     GGCTACCCGT GATATTGCTG AAGAGCTTGG CGGCGAATGG GCTGACCGCT TCCTCGTGCT
9541     TTACGGTATC GCCGCTCCCG ATTCGCAGCG CATCGCCTTC TATCGCCTTC TTGACGAGTT
9601     CTTCTGAATT TTGTTAAAAT TTTTGTTAAA TCAGCTCATT TTTTAACCAA TAGGCCGAAA
9661     TCGGCAACAT CCCTTATAAA TCAAAAGAAT AGACCGCGAT AGGGTTGAGT GTTGTTCCAG
9721     TTTGGAACAA GAGTCCACTA TTAAAGAACG TGGACTCCAA CGTCAAAGGG CGAAAAACCG
9781     TCTATCAGGG CGATGGCCCA CTACGTGAAC CATCACCCAA ATCAAGTTTT TTGCGGTCGA
9841     GGTGCCGTAA AGCTCTAAAT CGGAACCCTA AAGGGAGCCC CCGATTTAGA GCTTGACGGG
9901     GAAAGCCGGC GAACGTGGCG AGAAAGGAAG GGAAGAAAGC GAAAGGAGCG GGCGCTAGGG
9961     CGCTGGCAAG TGTAGCGGTC ACGCTGCGCG TAACCACCAC ACCCGCGCGC TTAATGCGCC
10021    GNNNNNNNNN NNNNNNNNNN NN
```

pTS555(pShuttle-CMV containing dual-luciferase wt/amb)

```
1        NNNTTAATTA ANNNTCCCTT CCAGCTCTCT GCCCCTTTTG GATTGAAGCC AATATGATAA
61       TGAGGGGGTG GAGTTTGTGA CGTGGCGCGG GGCGTGGGAA CGGGGCGGGT GACGTAGTAG
121      TGTGGCGGAA GTGTGATGTT GCAAGTGTGG CGGAACACAT GTAAGCGACG GATGTGGCAA
181      AAGTGACGTT TTTGGTGTGC GCCGGTGTAC ACAGGAAGTG ACAATTTTCG CGCGGTTTTA
241      GGCGGATGTT GTAGTAAATT TGGGCGTAAC CGAGTAAGAT TTGGCCATTT TCGCGGGAAA
```

```
301     ACTGAATAAG AGGAAGTGAA ATCTGAATAA TTTTGTGTTA CTCATAGCGC GTAANNNNTA
361     ATAGTAATCA ATTACGGGGT CATTAGTTCA TAGCCCATAT ATGGAGTTCC GCGTTACATA
421     ACTTACGGTA AATGGCCCGC CTGGCTGACC GCCCAACGAC CCCCGCCCAT TGACGTCAAT
481     AATGACGTAT GTTCCCATAG TAACGCCAAT AGGGACTTTC CATTGACGTC AATGGGTGGA
541     GTATTTACGG TAAACTGCCC ACTTGGCAGT ACATCAAGTG TATCATATGC CAAGTACGCC
601     CCCTATTGAC GTCAATGACG GTAAATGGCC CGCCTGGCAT TATGCCCAGT ACATGACCTT
661     ATGGGACTTT CCTACTTGGC AGTACATCTA CGTATTAGTC ATCGCTATTA CCATGGTGAT
721     GCGGTTTTGG CAGTACATCA ATGGGCGTGG ATAGCGGTTT GACTCACGGG GATTTCCAAG
781     TCTCCACCCC ATTGACGTCA ATGGGAGTTT GTTTTGGCAC CAAAATCAAC GGGACTTTCC
841     AAAATGTCGT AACAACTCCG CCCCATTGAC GCAAATGGGC GGTAGGCGTG TACGGTGGGA
901     GGTCTATATA AGCAGAGCTG GTTTAGTGAA CCGTCAGATC CGCTAGAGAT CTGGTACCGT
961     CGACGCGGCC GCTCGAGCCT AAGCTTCTAG ATAAGATATG ACTTCGAAAG TTTATGATCC
1021    AGAACAAAGG AAACGGATGA TAACTGGTCC GCAGTGGTGG GCCAGATGTA AACAAATGAA
1081    TGTTCTTGAT TCATTTATTA ATTATTATGA TTCAGAAAAA CATGCAGAAA ATGCTGTTAT
1141    TTTTTTACAT GGTAACGCGG CCTCTTCTTA TTTATGGCGA CATGTTGTGC CACATATTGA
1201    GCCAGTAGCG CGGTGTATTA TACCAGACCT TATTGGTATG GGCAAATCAG GCAAATCTGG
1261    TAATGGTTCT TATAGGTTAC TTGATCATTA CAAATATCTT ACTGCATGGT TTGAACTTCT
1321    TAATTTACCA AAGAAGATCA TTTTTGTCGG CCATGATTGG GGTGCTTGTT TGGCATTTCA
1381    TTATAGCTAT GAGCATCAAG ATAAGATCAA AGCAATAGTT CACGCTGAAA GTGTAGTAGA
1441    TGTGATTGAA TCATGGGATG AATGGCCTGA TATTGAAGAA GATATTGCGT TGATCAAATC
1501    TGAAGAAGGA GAAAAAATGG TTTTGGAGAA TAACTTCTTC GTGGAAACCA TGTTGCCATC
1561    AAAAATCATG AGAAAGTTAG AACCAGAAGA ATTTGCAGCA TATCTTGAAC CATTCAAAGA
1621    GAAAGGTGAA GTTCGTCGTC CAACATTATC ATGGCCTCGT GAAATCCCGT TAGTAAAAGG
1681    TGGTAAACCT GACGTTGTAC AAATTGTTAG GAATTATAAT GCTTATCTAC GTGCAAGTGA
1741    TGATTTACCA AAAATGTTTA TTGAATCGGA CCCAGGATTC TTTTCCAATG CTATTGTTGA
1801    AGGTGCCAAG AAGTTTCCTA ATACTGAATT TGTCAAAGTA AAAGGTCTTC ATTTTTCGCA
1861    AGAAGATGCA CCTGATGAAA TGGGAAAATA TATCAAATCG TTCGTTGAGC GAGTTCTCAA
1921    AAATGAACAA GGAAGCGGAG CTACTAACTT CAGCCTGCTG AAGCAGGCTG GAGACGTGGA
1981    GGAGAACCCT GGACCTATGG AAGATGCCAA AAACATTAAG AAGGGCCCAG CGCCATTCTA
2041    CCCACTCGAA GACGGGACCG CCGGCGAGCA GCTGCACAAA GCCATGAAGC GCTACGCCCT
2101    GGTGCCCGGC ACCATCGCCT TTACCGACGC ACATATCGAG GTGGACATTA CCTACGCCGA
2161    GTACTTCGAG ATGAGCGTTC GGCTGGCAGA AGCTATGAAG CGCTATGGGC TGAATACAAA
2221    CCATCGGATC GTGGTGTGCA GCGAGAATAG CTTGCAGTTC TTCATGCCCG TGTTGGGTGC
2281    CCTGTTCATC GGTGTGGCTG TGGCCCCAGC TAACGACATC TACAACGAGC GCGAGCTGCT
2341    GAACAGCATG GGCATCAGCC AGCCCACCGT CGTATTCGTG AGCAAGAAAG GGCTGCAAAA
2401    GATCCTCAAC GTGCAAAAGA AGCTACCGAT CATACAAAAG ATCATCATCA TGGATAGCAA
2461    GACCGACTAC CAGGGCTTCC AAAGCATGTA CACCTTCGTG ACTTCCCATT TGCCACCCGG
2521    CTTCAACGAG TACGACTTCG TGCCCGAGAG CTTCGACCGG GACAAAACCA TCGCCCTGAT
2581    CATGAACAGT AGTGGCAGTA CCGGATTGCC CAAGGGCGTA GCCCTACCGC ACCGCACCGC
2641    TTGTGTCCGA TTCAGTCATG CCCGCGACCC CATCTTCGGC AACCAGATCA TCCCCGACAC
2701    CGCTATCCTC AGCGTGGTGC CATTTCACCA CGGCTTCGGC ATGTTCACCA CGCTGGGCTA
2761    CTTGATCTGC GGCTTTCGGG TCGTGCTCAT GTACCGCTTC GAGGAGGAGC TATTCTTGCG
2821    CAGCTTGCAA GACTATAAGA TTCAATCTGC CCTGCTGGTG CCCACACTAT TTAGCTTCTT
2881    CGCTAAGAGC ACTCTCATCG ACAAGTACGA CCTAAGCAAC TTGCACGAGA TCGCCAGCGG
2941    CGGGGCGCCG CTCAGCAAGG AGGTAGGTGA GGCCGTGGCC AAACGCTTCC ACCTACCAGG
3001    CATCCGCCAG GGCTACGGCC TGACAGAAAC AACCAGCGCC ATTCTGATCA CCCCCGAAGG
3061    GGACGACAAG CCTGGCGCAG TAGGCAAGGT GGTGCCCTTC TTCGAGGCTA AGGTGGTGGA
3121    CTTGGACACC GGTAAGACAC TGGGTGTGAA CCAGCGCGGC GAGCTGTGCG TCCGTGGCCC
3181    CATGATCATG AGCGGCTACG TTAACAACCC CGAGGCTACA AACGCTCTCA TCGACAAGGA
3241    CGGCTAGCTG CACAGCGGCG ACATCGCCTA CTGGGACGAG GACGAGCACT TCTTCATCGT
3301    GGACCGGCTG AAGAGCCTGA TCAAATACAA GGGCTACCAG GTAGCCCCAG CCGAACTGGA
3361    GAGCATCCTG CTGCAACACC CCAACATCTT CGACGCCGGG GTCGCCGGCC TGCCCGACGA
3421    CGATGCCGGC GAGCTGCCCG CCGCAGTCGT CGTGCTGGAA CACGGTAAAA CCATGACCGA
3481    GAAGGAGATC GTGGACTATG TGGCCAGCCA GGTCACAACC GCCAAGAAGC TGCGCGGTGG
3541    TGTTGTGTTC GTGGACGAGG TGCCTAAAGG ACTGACCGGC AAGTTGGACG CCCGCAAGAT
3601    CCGCGAGATT CTCATTAAGG CCAAGAAGGG CGGCAAGATC GCCGTGTAAG AATTCATCCG
3661    ATCCACCGGA TCTAGATAAC TGATCATAAT CAGCCATACC ACATTTGTAG AGGTTTTACT
3721    TGCTTTAAAA AACCTCCCAC ACCTCCCCCT GAACCTGAAA CATAAAATGA ATGCAATTGT
3781    TGTTGTTAAC TTGTTTATTG CAGCTTATAA TGGTTACAAA TAAAGCAATA GCATCACAAA
3841    TTTCACAAAT AAAGCATTTT TTTCACTGCA TTCTAGTTGT GGTTTGTCCA AACTCATCAA
3901    TGTATCTTAA CGCNNNNTAA GGGTGGGAAA GAATATATAA GGTGGGGGTC TTATGTAGTT
3961    TTGTATCTGT TTTGCAGCAG CCGCCGCCGC CATGAGCACC AACTCGTTTG ATGGAAGCAT
```

```
4021     TGTGAGCTCA TATTTGACAA CGCGCATGCC CCCATGGGCC GGGGTGCGTC AGAATGTGAT
4081     GGGCTCCAGC ATTGATGGTC GCCCCGTCCT GCCCGCAAAC TCTACTACCT TGACCTACGA
4141     GACCGTGTCT GGAACGCCGT TGGAGACTGC AGCCTCCGCC GCCGCTTCAG CCGCTGCAGC
4201     CACCGCCCGC GGGATTGTGA CTGACTTTGC TTTCCTGAGC CCGCTTGCAA GCAGTGCAGC
4261     TTCCCGTTCA TCCGCCCGCG ATGACAAGTT GACGGCTCTT TTGGCACAAT TGGATTCTTT
4321     GACCCGGGAA CTTAATGTCG TTTCTCAGCA GCTGTTGGAT CTGCGCCAGC AGGTTTCTGC
4381     CCTGAAGGCT TCCTCCCCTC CCAATGCGGT TTAAAACATA AATAAAAAAC CAGACTCTGT
4441     TTGGATTTGG ATCAAGCAAG TGTCTTGCTG TCTTTATTTA GGGGTTTTGC GCGCGCGGTA
4501     GGCCCGGGAC CAGCGGTCTC GGTCGTTGAG GGTCCTGTGT ATTTTTTCCA GGACGTGGTA
4561     AAGGTGACTC TGGATGTTCA GATACATGGG CATAAGCCCG TCTCTGGGGT GGAGGTAGCA
4621     CCACTGCAGA GCTTCATGCT GCGGGGTGGT GTTGTAGATG ATCCAGTCGT AGCAGGAGCG
4681     CTGGGCGTGG TGCCTAAAAA TGTCTTTCAG TAGCAAGCTG ATTGCCAGGG GCAGGCCCTT
4741     GGTGTAAGTG TTTACAAAGC GGTTAAGCTG GGATGGGTGC ATACGTGGGG ATATGAGATG
4801     CATCTTGGAC TGTATTTTTA GGTTGGCTAT GTTCCCAGCC ATATCCCTCC GGGGATTCAT
4861     GTTGTGCAGA ACCACCAGCA CAGTGTATCC GGTGCACTTG GGAAATTTGT CATGTAGCTT
4921     AGAAGGAAAT GCGTGGAAGA ACTTGGAGAC GCCCTTGTGA CCTCCAAGAT TTTCCATGCA
4981     TTCGTCCATA ATGATGGCAA TGGGCCCACG GGCGGCGGCC TGGGCGAAGA TATTTCTGGG
5041     ATCACTAACG TCATAGTTGT GTTCCAGGAT GAGATCGTCA TAGGCCATTT TTACAAAGCG
5101     CGGGCGGAGG GTGCCAGACT GCGGTATAAT GGTTCCATCC GGCCCAGGGG CGTAGTTACC
5161     CTCACAGATT TGCATTTCCC ACGCTTTGAG TTCAGATGGG GGGATCATGT CTACCTGCGG
5221     GGCGATGAAG AAAACGGTTT CCGGGGTAGG GGAGATCAGC TGGGAAGAAA GCAGGTTCCT
5281     GAGCAGCTGC GACTTACCGC AGCCGGTGGG CCCGTAAATC ACACCTATTA CCGGGTGCAA
5341     CTGGTAGTTA AGAGAGCTGC AGCTGCCGTC ATCCCTGAGC AGGGGGGCCA CTTCGTTAAG
5401     CATGTCCCTG ACTCGCATGT TTTCCCTGAC CAAATCCGCC AGAAGGCGCT CGCCGCCCAG
5461     CGATAGCAGT TCTTGCAAGG AAGCAAAGTT TTTCAACGGT TTGAGACCGT CCGCCGTAGG
5521     CATGCTTTTG AGCGTTTGAC CAAGCAGTTC CAGGCGGTCC CACAGCTCGG TCACCTGCTC
5581     TACGGCATCT CGATCCAGCA TATCTCCTCG TTTCGCGGGT TGGGGCGGCT TTCGCTGTAC
5641     GGCAGTAGTC GGTGCTCGTC CAGACGGGCC AGGGTCATGT CTTTCCACGG GCGCAGGGTC
5701     CTCGTCAGCG TAGTCTGGGT CACGGTGAAG GGGTGCGCTC CGGGCTGCGC GCTGGCCAGG
5761     GTGCGCTTGA GGCTGGTCCT GCTGGTGCTG AAGCGCTGCC GGTCTTCGCC CTGCGCGTCG
5821     GCCAGGTAGC ATTTGACCAT GGTGTCATAG TCCAGCCCCT CCGCGGCGTG GCCCTTGGCG
5881     CGCAGCTTGC CCTTGGAGGA GGCGCCGCAC GAGGGGCAGT GCAGACTTTT GAGGGCGTAG
5941     AGCTTGGGCG CGAGAAATAC CGATTCCGGG GAGTAGGCAT CCGCGCCGCA GGCCCCGCAG
6001     ACGGTCTCGC ATTCCACGAG CCAGGTGAGC TCTGGCCGTT CGGGGTCAAA AACCAGGTTT
6061     CCCCCATGCT TTTTGATGCG TTTCTTACCT CTGGTTTCCA TGAGCCGGTG TCCACGCTCG
6121     GTGACGAAAA GGCTGTCCGT GTCCCCGTAT ACAGACTNNN GTTTAAACGA ATTCNNNATA
6181     TAAAATGCAA GGTGCTGCTC AAAAAATCAG GCAAAGCCTC GCGCAAAAAA GAAAGCACAT
6241     CGTAGTCATG CTCATGCAGA TAAAGGCAGG TAAGCTCCGG AACCACCACA GAAAAAGACA
6301     CCATTTTTCT CTCAAACATG TCTGCGGGTT TCTGCATAAA CACAAAATAA AATAACAAAA
6361     AAACATTTAA ACATTAGAAG CCTGTCTTAC AACAGGAAAA ACAACCCTTA TAAGCATAAG
6421     ACGGACTACG GCCATGCCGG CGTGACCGTA AAAAAACTGG TCACCGTGAT TAAAAAGCAC
6481     CACCGACAGC TCCTCGGTCA TGTCCGGAGT CATAATGTAA GACTCGGTAA ACACATCAGG
6541     TTGATTCATC GGTCAGTGCT AAAAAGCGAC CGAAATAGCC CGGGGGAATA CATACCCGCA
6601     GGCGTAGAGA CAACATTACA GCCCCCATAG GAGGTATAAC AAAATTAATA GGAGAGAAAA
6661     ACACATAAAC ACCTGAAAAA CCCTCCTGCC TAGGCAAAAT AGCACCCTCC CGCTCCAGAA
6721     CAACATACAG CGCTTCACAG CGGCAGCCTA ACAGTCAGCC TTACCAGTAA AAAAGAAAAC
6781     CTATTAAAAA AACACCACTC GACACGGCAC CAGCTCAATC AGTCACAGTG TAAAAAAGGG
6841     CCAAGTGCAG AGCGAGTATA TATAGGACTA AAAAATGACG TAACGGTTAA AGTCCACAAA
6901     AAACACCCAG AAAACCGCAC GCGAACCTAC GCCCAGAAAC GAAAGCCAAA AAACCCACAA
6961     CTTCCTCAAA TCGTCACTTC CGTTTTCCCA CGTTACGTAA CTTCCCATTT TAAGAAAACT
7021     ACAATTCCCA ACACATACAA GTTACTCCGC CCTAAAACCT ACGTCACCCG CCCCGTTCCC
7081     ACGCCCCGCG CCACGTCACA AACTCCACCC CCTCATTATC ATATTGGCTT CAATCCAAAA
7141     TAAGGTATAT TATTGATGAT NNNTTAATTA AGGATCCNNN CGGTGTGAAA TACCGCACAG
7201     ATGCGTAAGG AGAAAATACC GCATCAGGCG CTCTTCCGCT TCCTCGCTCA CTGACTCGCT
7261     GCGCTCGGTC GTTCGGCTGC GGCGAGCGGT ATCAGCTCAC TCAAAGGCGG TAATACGGTT
7321     ATCCACAGAA TCAGGGGATA ACGCAGGAAA GAACATGTGA GCAAAAGGCC AGCAAAAGGC
7381     CAGGAACCGT AAAAAGGCCG CGTTGCTGGC GTTTTTCCAT AGGCTCCGCC CCCCTGACGA
7441     GCATCACAAA AATCGACGCT CAAGTCAGAG GTGGCGAAAC CCGACAGGAC TATAAAGATA
7501     CCAGGCGTTT CCCCCTGGAA GCTCCCTCGT GCGCTCTCCT GTTCCGACCC TGCCGCTTAC
7561     CGGATACCTG TCCGCCTTTC TCCCTTCGGG AAGCGTGGCG CTTTCTCATA GCTCACGCTG
7621     TAGGTATCTC AGTTCGGTGT AGGTCGTTCG CTCCAAGCTG GGCTGTGTGC ACGAACCCCC
7681     CGTTCAGCCC GACCGCTGCG CCTTATCCGG TAACTATCGT CTTGAGTCCA ACCCGGTAAG
```

```
7741     ACACGACTTA TCGCCACTGG CAGCAGCCAC TGGTAACAGG ATTAGCAGAG CGAGGTATGT
7801     AGGCGGTGCT ACAGAGTTCT TGAAGTGGTG GCCTAACTAC GGCTACACTA GAAGGACAGT
7861     ATTTGGTATC TGCGCTCTGC TGAAGCCAGT TACCTTCGGA AAAAGAGTTG GTAGCTCTTG
7921     ATCCGGCAAA CAAACCACCG CTGGTAGCGG TGGTTTTTTT GTTTGCAAGC AGCAGATTAC
7981     GCGCAGAAAA AAAGGATCTC AAGAAGATCC TTTGATCTTT TCTACGGGGT CTGACGCTCA
8041     GTGGAACGAA AACTCACGTT AAGGGATTTT GGTCATGAGA TTATCAAAAA GGATCTTCAC
8101     CTAGATCCTT TTAAATTAAA AATGAAGTTT TAAATCAATC TAAAGTATAT ATGAGTAAAC
8161     TTGGTCTGAC AGTTACCAAT GCTTAATCAG TGAGGCACCT ATCTCAGCGA TCTGTCTATT
8221     TCGTTCATCC ATAGTTGCCT GACTCCCCGT CGTGTAGATA ACTACGATAC GGGAGGGCTT
8281     ACCATCTGGC CCCAGTGCTG CAATGATACC GCGAGACCCA CGCTCACCGG CTCCAGATTT
8341     ATCAGCAATA AACCAGCCAG CCGGAAGGGC CGAGCGCAGA AGTGGTCCTG CAACTTTATC
8401     CGCCTCCATC CAGTCTATTA ATTGTTGCCG GGAAGCTAGA GTAAGTAGTT CGCCAGTTAA
8461     TAGTTTGCGC AACGTTGTTG NNNNAAAAAG GATCTTCACC TAGATCCTTT TCACGTAGAA
8521     AGCCAGTCCG CAGAAACGGT GCTGACCCCG GATGAATGTC AGCTACTGGG CTATCTGGAC
8581     AAGGGAAAAC GCAAGCGCAA AGAGAAAGCA GGTAGCTTGC AGTGGGCTTA CATGGCGATA
8641     GCTAGACTGG GCGGTTTTAT GGACAGCAAG CGAACCGGAA TTGCCAGCTG GGGCGCCCTC
8701     TGGTAAGGTT GGGAAGCCCT GCAAAGTAAA CTGGATGGCT TTCTCGCCGC CAAGGATCTG
8761     ATGGCGCAGG GGATCAAGCT CTGATCAAGA GACAGGATGA GGATCGTTTC GCATGATTGA
8821     ACAAGATGGA TTGCACGCAG GTTCTCCGGC CGCTTGGGTG GAGAGGCTAT TCGGCTATGA
8881     CTGGGCACAA CAGACAATCG GCTGCTCTGA TGCCGCCGTG TTCCGGCTGT CAGCGCAGGG
8941     GCGCCCGGTT CTTTTTGTCA AGACCGACCT GTCCGGTGCC CTGAATGAAC TGCAAGACGA
9001     GGCAGCGCGG CTATCGTGGC TGGCCACGAC GGGCGTTCCT TGCGCAGCTG TGCTCGACGT
9061     TGTCACTGAA GCGGGAAGGG ACTGGCTGCT ATTGGGCGAA GTGCCGGGGC AGGATCTCCT
9121     GTCATCTCAC CTTGCTCCTG CCGAGAAAGT ATCCATCATG GCTGATGCAA TGCGGCGGCT
9181     GCATACGCTT GATCCGGCTA CCTGCCCATT CGACCACCAA GCGAAACATC GCATCGAGCG
9241     AGCACGTACT CGGATGGAAG CCGGTCTTGT CGATCAGGAT GATCTGGACG AAGAGCATCA
9301     GGGGCTCGCG CCAGCCGAAC TGTTCGCCAG GCTCAAGGCG AGCATGCCCG ACGGCGAGGA
9361     TCTCGTCGTG ACCCATGGCG ATGCCTGCTT GCCGAATATC ATGGTGGAAA ATGGCCGCTT
9421     TTCTGGATTC ATCGACTGTG GCCGGCTGGG TGTGGCGGAC CGCTATCAGG ACATAGCGTT
9481     GGCTACCCGT GATATTGCTG AAGAGCTTGG CGGCGAATGG GCTGACCGCT TCCTCGTGCT
9541     TTACGGTATC GCCGCTCCCG ATTCGCAGCG CATCGCCTTC TATCGCCTTC TTGACGAGTT
9601     CTTCTGAATT TTGTTAAAAT TTTTGTTAAA TCAGCTCATT TTTTAACCAA TAGGCCGAAA
9661     TCGGCAACAT CCCTTATAAA TCAAAAGAAT AGACCGCGAT AGGGTTGAGT GTTGTTCCAG
9721     TTTGGAACAA GAGTCCACTA TTAAAGAACG TGGACTCCAA CGTCAAAGGG CGAAAAACCG
9781     TCTATCAGGG CGATGGCCCA CTACGTGAAC CATCACCCAA ATCAAGTTTT TTGCGGTCGA
9841     GGTGCCGTAA AGCTCTAAAT CGGAACCCTA AAGGGAGCCC CCGATTTAGA GCTTGACGGG
9901     GAAAGCCGGC GAACGTGGCG AGAAAGGAAG GGAAGAAAGC GAAAGGAGCG GGCGCTAGGG
9961     CGCTGGCAAG TGTAGCGGTC ACGCTGCGCG TAACCACCAC ACCCGCGCGC TTAATGCGCC
10021    GNNNNNNNNN NNNNNNNNNN NN
```

6.5 Conference & Retreat contributions

Talks

Reautschnig, P., Wettengel J., & Stafforst T.

Harnessing human ADAR2 for RNA repair- Recoding a PINK1 mutation rescues mitophagy.

Oligonucleotide Therapeutics Society Meeting, Montreal, Canada, September 25-28, 2016

Reautschnig, P., Wettengel J., & Stafforst T.

Harnessing human ADAR2 for Site-Directed RNA Editing

IFIB Retreat 2016, Freudenstadt, Germany, April 22-24, 2016

Poster presentations

Reautschnig, P., Wettengel J., & Stafforst T.

Towards a Deeper Understanding of Site-Directed RNA Editing

IFIB Retreat 2018, Hechingen, Germany, July 26-27, 2018

Reautschnig, P., Wettengel J., Geisler S., Kahle P. & Stafforst T.

Harnessing human ADAR2 for RNA repair – Recoding a PINK1 mutation rescues mitophagy.

1st Symposium on Nucleic Acid Modifications, Mainz, Germany, September 4-6, 2017.

Reautschnig, P., Wettengel J., Geisler S., Kahle P. & Stafforst T.

Harnessing human ADAR2 for RNA repair - Recoding a PINK1 mutation rescues mitophagy.

IFIB Retreat 2017, Freudenstadt, Germany, April 5-7, 2017

Reautschnig, P., Wettengel J., Flad S. & Stafforst, T.

Engineering of guideRNAs to harness ADAR for site-directed RNA editing

IFIB Retreat 2016, Freudenstadt, Germany, April 22-24, 2016

Reautschnig, P., Wettengel J., Flad S. & Stafforst, T.

Engineering of guideRNAs to harness ADAR for site-directed RNA editing.

67. Mosbacher Kolloquium, Mosbach, Germany, March 31-April 2, 2016.

6.6 Publications

6.6.1 **Publication 1:** Reautschnig P.*, Vogel P.*, Stafforst T. The notorious R.N.A. in the spotlight - drug or target for the treatment of disease. RNA Biol, 2017. 14(5): p. 651-668. * equal contribution

RNA BIOLOGY
2017, VOL. 14, NO. 5, 651–668
http://dx.doi.org/10.1080/15476286.2016.1208323

Taylor & Francis
Taylor & Francis Group

REVIEW

OPEN ACCESS

The notorious R.N.A. in the spotlight - drug or target for the treatment of disease

Philipp Reautschnig*, Paul Vogel*, and Thorsten Stafforst

Interfaculty Institute of Biochemistry, University of Tübingen Auf der Morgenstelle, Tübingen, Germany

ABSTRACT

mRNA is an attractive drug target for therapeutic interventions. In this review we highlight the current state, clinical trials, and developments in antisense therapy, including the classical approaches like RNaseH-dependent oligomers, splice-switching oligomers, aptamers, and therapeutic RNA interference. Furthermore, we provide an overview on emerging concepts for using RNA in therapeutic settings including protein replacement by in-vitro-transcribed mRNAs, mRNA as vaccines and anti-allergic drugs. Finally, we give a brief outlook on early-stage RNA repair approaches that apply endogenous or engineered proteins in combination with short RNAs or chemically stabilized oligomers for the re-programming of point mutations, RNA modifications, and frame shift mutations directly on the endogenous mRNA.

Abbreviations: ASO, Antisense oligonucleotide; CD, Cluster of differentiation; CFTR, Cystic fibrosis transmembrane conductance regulator; CRISPR/Cas9, Clustered regularly interspaced short palindromic repeats/CRISPR-associated 9; FDA, US Food and Drug Administration; GalNAc, N-acetyl galactosamine; IVT-mRNA, In-vitro transcribed mRNA; MHC, Major histocompatibility complex; miRNA, microRNA; MOE, 2′-O-methoxyethyl; mRNA, messenger RNA; ψ, pseudouridine; PS, Phosphothioate; RNAi, RNA interference; siRNA, Short interfering RNA; SSO, Splice-switching oligonucleotide; SMN2, Survival of motor neuron 2; TALEN, Transcription activator-like effector nuclease; TLR, Toll-like receptor; $T_H1/2$ cell, Type 1/2 T helper cell; T_R1, Type 1 regulatory T cell; VEGF, Vascular endothelial growth factor; VEGFR-1, Vascular endothelial growth factor receptor 1; ZFN, Zinc finger nuclease

ARTICLE HISTORY
Received 8 April 2016
Revised 1 June 2016
Accepted 27 June 2016

KEYWORDS
Antisense oligonucleotide; chemically modified oligonucleotides; genetic disease; RNA vaccines; RNA repair; RNA interference; site-directed RNA editing; splice-switching oligonucleotide; therapeutic aptamer; therapeutic mRNA

Introduction

During the last 15 y the diverse roles of RNA in regular but also pathological cellular processes became increasingly clear. RNA is not only a short-lived messenger and part of the translational machinery but RNA contributes significantly to the regulation and diversification of the genetic information. There is now increasing insight into the mechanistic role of defective RNA processing, including (alternative) splicing, modification, translation, and decay for the etiology of various diseases.[1-4] However, not only mis-regulation and defective processing cause disease, but even RNA species themselves can initiate disease processes independent of their protein-coding function. Nucleotide repeat diseases are typical examples.[5] To employ this new mechanistic knowledge and to translate it into therapy requires drugs that reliably target nucleic acids in a sequence-specific manner. However, there are only few small molecule drugs that target nucleic acids and those are limited in their capacity of sequence addressing. In contrast, oligonucleotide analogs provide a basis for the rational design of highly sequence-specific drugs to target virtually any cellular nucleic acid in a specific manner.[6] Classical drugs like small molecules target enzymes and receptors to block or alter their specific functions. In contrast, the interference at the nucleic acid level would allow to manipulate the transcriptome and the proteome itself. This is not limited to the simple up- or down-regulation of target gene expression. Most appealing is the possibility of actively creating new transcript and protein isoforms with altered properties and functions, for instance by re-programming a protein-coding stretch, or by altering splice sites, modification patterns, polyadenylation states, miRNA binding sites, etc.[7] Affecting the cell by targeting its nucleic acids clearly enlarges the scope of currently available therapeutic interventions including the causal treatment of some genetic diseases.

However, already short oligonucleotides have unfavorable pharmacological properties. They are hydrophilic, polyanionic macromolecules that can hardly overcome cellular membranes, are unstable against RNases, and suffer from rapid renal clearance.[8] This leads to short half-life and low bioavailability. Furthermore, adverse toxic effects may appear that include immune-reactions and off-target binding to non-targeted cellular nucleic acids. Together, oligonucleotide drugs are often characterized by low efficacy and high toxicity which strongly limits their clinical application.[6] During the last decades, medicinal chemists have put enormous effort into the development of new chemistries that improve lifetime, delivery, potency, and efficacy of the drugs while reducing their toxicity and immunogenicity. These new chemistries are now approaching clinical trials and will hopefully pave the way for

CONTACT Thorsten Stafforst thorsten.stafforst@uni-tuebingen.de
*These authors equally contributed to this work.

the broad clinical application of oligonucleotide drugs. An overview on recent developments in oligonucleotide medicinal chemistry can be found elsewhere.[6,7]

In principle, interference with the genetic information could be achieved permanently at the DNA- or transiently at the RNA-level. In this review we will focus on the RNA-level. Even though novel approaches for genome engineering are currently keenly explored,[9] we believe that it would be foolish to carelessly discard the RNA alternative. With respect to ethical issues and safety aspects, the transient and thus reversible nature of RNA manipulation could turn out as a blessing in disguise. Both, the therapeutic effects and the potential adverse effects, are likely to be tunable. Furthermore, manipulations are conceivable that are inaccessible or difficult to realize on the genome level per se. This includes amino acid changes or transcript level changes that would kill a cell if they are permanently enforced. Potentially lethal interventions on kinases, apoptosis factors, transcription or translation factors could be realized on the RNA-level suddenly, transiently or partially to obtain a therapeutic effect, for instance. Manipulation at the RNA-level might also be much more efficient compared to HDR-dependent genomic knock-in, which remained persistently inefficient in vivo, in particular in postmitotic tissues like the brain.[9] For many genetic diseases, which are caused by loss-of-function mutations, a patient would benefit more from a drug that can restore a small fraction (like 5%) of functional gene product in a large fraction of a the tissue than from a drug that can restore full gene function (100%) but only in a small fraction of the tissue. A typical example is cystic fibrosis.[10]

In this review we will first update on recent developments in the classical approaches, like RNaseH-dependent decay, chemically stabilized oligonucleotides that target mRNAs to induce splice-switching, aptamers, and the knock-down via RNAi (Fig. 1). After painful years of repeated relapse one seems to

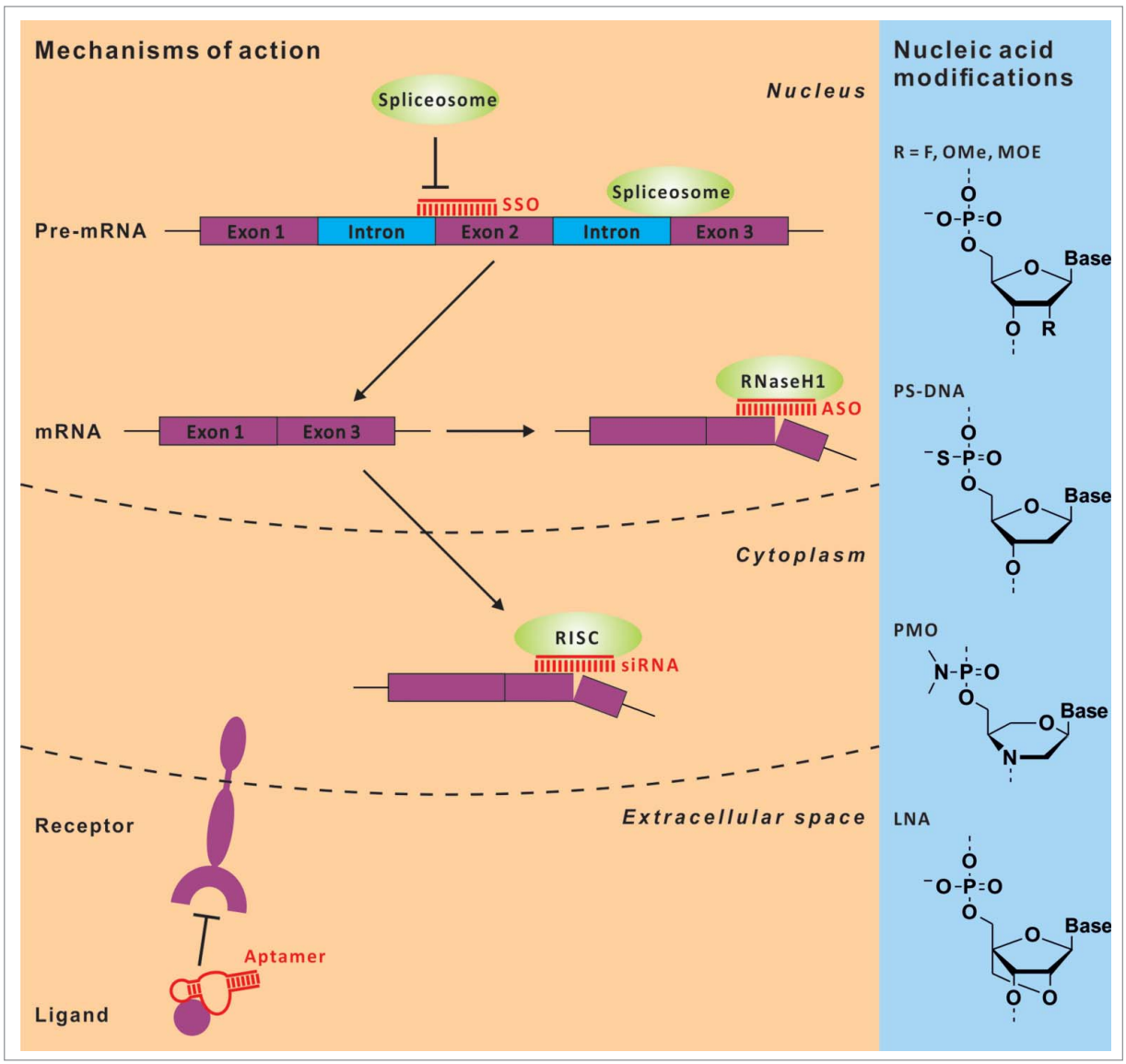


Figure 1. Chemically stabilized, short oligonucleotides can employ various mechanisms for their therapeutic effects ranging from blocking ligand – receptor binding, RNA degradation via RISC or RNaseH(1) recruitment, and alteration of splicing. The classical modes of action are shown on the left panel, a small section of typically used chemical backbone modifications are depicted on the right.

have learned the lessons and have now substantially improved the effectiveness of such drugs. For instance, in 2015 therapeutic RNAi was demonstrated in a relevant monkey model by subcutaneous administration of a chemically stabilized siRNA that partially knocks down antithrombin in the monkey's liver.[11] The problem of delivery and toxicity seems to be solved, at least for simple oligonucleotide drugs and for some organs, and allows therapeutic intervention with an affordable amount of the drug under compliant administration routes. Consequently, the number of promising clinical phase II and III studies has increased during the last few years (see Table 1).

Every new discovery in RNA function and regulation offers a starting point to develop novel therapies. After its discovery in 1998 we now find numerous drug candidates in clinical studies that apply the RNAi mechanism (Table 1).[12] In the second part of this review we highlight emerging concepts that are still in the pre-clinical or very early clinical exploration stage but that have the potential to become medicines of the future. This includes therapeutic mRNAs, mRNAs as vaccine, and RNA repair approaches. The latter apply endogenous or engineered enzymes to repair, re-program, or modify a target RNA at a specific site in order to provoke a therapeutically relevant effect (Fig. 2).

Update on established approaches

RNaseH-dependent antisense oligonucleotides

Oligonucleotides working through an RNaseH-dependent cleavage mechanism are the oldest class of antisense oligonucleotides (ASO). They are extensively explored and represent the largest class of nucleic acid analog drugs in clinical trials. RNaseH-dependent ASOs are short DNA oligomers targeting mRNA. Once the DNA-oligo/mRNA heteroduplex is formed, human RNaseH1 binds to it and catalyzes RNA cleavage under release of the intact DNA oligomer.[13]

Medicinal chemists have undertaken great efforts to improve ASO design regarding nuclease resistance, circulation half-life, target affinity (potency), and tissue specificity. The first ASOs tested in clinical trials, also referred to as 1st generation ASOs, have been modified by oxygen-to-sulfur substitutions in the phosphate backbone. ASOs with such a phosphothioate (PS) backbone show enhanced nuclease resistance and prolonged plasma half-life due to non-specific binding to plasma proteins preventing them from rapid renal filtration. However, numerous toxicities were also associated with that type of modification.[6] In 1998, fomivirsen was the first FDA-approved ASO and was applied for the treatment of human cytomegalovirus-induced retinitis in HIV patients.[14-16] Marketed as Vitravene, the 21 nt PS-oligonucleotide was administered by intravitreal injection to target the immediate early region 2 of the viral mRNA. Since the approval of fomivirsen, several ASOs belonging to the 1st generation are under clinical review. For instance, targeting the mRNA of intercellular adhesion molecule 1 and the insulin receptor substrate 1 are advanced in the treatment of pouchitis[17,18] and vascular disorders in the eye,[19-22] respectively. The RNaseH-mediated degradation of Akt-1 mRNA to impede tumor proliferation[23] is currently tested for clinical application.[24-26]

Due to the early success with 1st generation ASO, further medicinal chemistry was explored to improve half-life and potency of the drugs in order to reduce the administered dose, the application frequency, the costs, and to minimize adverse effects.[27] This resulted in the 2nd generation ASOs, also referred to as gapmers. A typical gapmer is a 20 nt oligonucleotide comprising a PS backbone and 5 flanking 2′O-methoxyethyl (MOE) groups at both termini. Due to the unmodified internal DNA gap, such ASOs remain good substrates for RNaseH, whereas the terminal MOE modifications increase nuclease resistance and enhances the binding of the ASO to the target mRNA.[28] 2nd generation ASOs entered clinical trials for various therapeutic applications. The most prominent representative of the 2nd generation is the MOE gapmer mipomersen as the second FDA-approved RNaseH-dependent ASO. The compound targets apolipoprotein B-100 mRNA and is subcutaneously administered to treat familiar hypercholesterolemia. The genetic disorder is caused by the loss of low-density lipoprotein (LDL) receptor function leading to high LDL cholesterol plasma concentration and early cardiovascular disease. Phase III trials had demonstrated an efficient decrease of LDL cholesterol by lowering ApoB-100 amount in patients obtaining mipomersen.[29-31] The treatment obviously profited from the general pharmacokinetics of systemically administered ASOs which preferably accumulate in the liver where ApoB-100 synthesis takes place.[8] Recently, an RNase-dependent ASO[32] has reached clinical phase III to reduce transthyretin expression in patients suffering from familial amyloid polyneuropathy.[33-35] Chemotherapy combined with RNaseH-mediated degradation of clusterin mRNA is a potential therapeutic option in the treatment of prostate[36-38] and lung cancer.[39,40]

Generation 2.5 ASO are derived from the traditional gapmer design. For this, the MOE modifications are replaced by 2′,4′-constrained ethyl (cEt) bridges in the flanking nucleotides. It was found that cEt-modified oligonucleotides provide the same superior target affinity, but increased nuclease resistance as compared to locked nucleic acid (LNA)-containing oligonucleotides.[41] One of the generation 2.5 ASOs targets the mRNA of signal transducer and activator of transcription 3[42] and is currently tested for the treatment of various cancer types.[43-46]

Most recently, a new chemistry has been developed that strongly increases the liver-specific uptake of oligonucleotide drugs, including ASO and siRNA therapeutics. For this, ASOs[47] and siRNAs[48] are conjugated with triantennary N-acetyl galactosamine ($GalNAc_3$). $GalNAc_3$ mediates liver-specific uptake through the asialoglycoprotein receptor (ASGPR) that is exclusively expressed on hepatocytes. Marketed as ligand-conjugated antisense (LICA) technology (Ionis Pharmaceuticals), it could be shown that the conjugation increases the potency of MOE gapmers up to 10-fold for inhibiting the expression of hepatic genes in mice.[49] When using a $GalNAc_3$-conjugated cEt gapmer, the RNaseH-mediated mRNA degradation was enhanced around 60-fold as compared to the corresponding 2nd generation MOE ASO. Additionally, Ionis Pharmaceuticals announced that its LICA drug targeting apolipoprotein(a) was 30-fold more potent in a phase I study than the unconjugated MOE gapmer.[50,51]

Table 1. Overview on the most recent and advanced clinical trials in the corresponding fields not claiming completeness. For RNAi, failed early trials are also listed that have been terminated in 2009 or before.

Table 1: Clinical Trials using RNA-Therapeutics*

RNaseH-dependent Antisense Oligonucleotides

Compound	Application Route	Target	Indication			
Fomivirsen (Vitravene®, ISIS 2922)	IVT	IE2	CMV retinitis	Ionis Pharmaceuticals /Ciba Vision	FDA Approved	14-16
Mipomersen (KYNAMRO®, ISIS 301012)	SC	ApoB-100	HoFH	Ionis Pharmaceuticals / Genzyme	FDA Approved	29-31
Custirsen (OGX-011)	IV	Clusterin	Prostate cancer	Ionis Phamaceuticals / OncoGenex Technologies	III - Completed	NCT01188187 [36]
Custirsen (OGX-011)	IV	Clusterin	Prostate cancer	Ionis Phamaceuticals / OncoGenex Technologies	III - Active	NCT01578655 [37]
Custirsen (OGX-011)	IV	Clusterin	Lung cancer	Ionis Phamaceuticals / OncoGenex Technologies	III - Recruiting	NCT01630733 [40]
Aganirsen (GS-101)	Eye drops	IRS-1	Corneal neovascularization in graft patients	Gene Signal	III - Completed	EudraCT 2008-005388-33 [21]
Aganirsen (GS-101)	Eye drops	IRS-1	Ischemic central retinal vein occlusion	Gene Signal	II/III - Recruiting	EudraCT 2014-000239-18 [22]
Alicaforsen (ISIS 2302)	Enema	ICAM-1	Pouchitis	Ionis Pharmaceuticals / Atlantic Healthcare	III - Recruiting	NCT02525523 [18]
IONIS-TTRRx (ISIS 420915)	SC	TTR	FAP, TRR Amyloidosis	Ionis Pharmaceuticals / GlaxoSmithKline	II/III - Active	NCT01737398 [33]
IONIS-TTRRx (ISIS 420915)	SC	TTR	FAP, TRR Amyloidosis	Ionis Pharmaceuticals / GlaxoSmithKline	III - Enrolling by invitation	NCT02175004 [34]
RX-0210 (Archexin®)	IV	Akt-1	Pancreatic cancer	Rexahn Pharmaceuticals	II - Completed	NCT01028495 [24]
RX-0210 (Archexin®)	IV	Akt-1	Renal cell cancer	Rexahn Pharmaceuticals	IB/II - Recruiting	NCT02089334 [25]
IONIS-STAT3-2.5Rx (AZD9150, ISIS 481464)	IV	STAT3	Gastrointestinal or ovarian cancer	Ionis Pharmaceuticals / AstraZeneca	II - Recruiting	NCT02417753 [43]
IONIS-STAT3-2.5Rx (AZD9150, ISIS 481464)	IV	STAT3	Metastatic Squamous Cell Carcinoma of the Head and Neck	Ionis Pharmaceuticals / AstraZeneca	IB/II - Recruiting	NCT02499328 [44]
IONIS-STAT3-2.5Rx (AZD9150, ISIS 481464)	IV	STAT3	Lymphoma	Ionis Pharmaceuticals / AstraZeneca	I/II - Active	NCT01563302 [45]
IONIS-STAT3-2.5Rx (AZD9150, ISIS 481464)	IV	STAT3	Lymphoma	Ionis Pharmaceuticals / AstraZeneca	IB - Not yet recruiting	NCT02549651 [46]
IONIS-APO(a)-LRx (ISIS 681257)	SC	ApoA	Elevated Lipoprotein(a)	Ionis Pharmaceutical / Akcea Thaerapeutics	I - Completed	NCT02414594 [51]

Splice-switching Oligonucleotides

Compound	Application Route	Target	Indication	Company / Initiator	Phase / Status	Trial ID / Reference
PRO051/GSK2402968 (drisapersen/ Kyndrisa®)	SC	Dystrophin exon 51 (skipping)	DMD	BioMarin Pharmaceuticals (developed by Prosensa Therapeutics/GlaxoSmithKline)	III - Recruiting	NCT01803412 [56]
PRO051/GSK2402968 (drisapersen/ Kyndrisa®)	SC	Dystrophin exon 51 (skipping)	DMD	BioMarin Pharmaceuticals (developed by Prosensa Therapeutics/GlaxoSmithKline)	IIIB - Enrolling by invitation	NCT02636686 [57]
PRO044/BMN 044	SC/IV	Dystrophin exon 44 (skipping)	DMD	BioMarin Pharmaceuticals (developed by Prosensa Therapeutics)	II - Enrolling by invitation	NCT02329769 [245]
PRO045/BMN 045	SC	Dystrophin exon 45 (skipping)	DMD	BioMarin Pharmaceuticals (developed by Prosensa Therapeutics)	IIB - Active	NCT01826474 [246]
PRO053/BMN 053	SC/IV	Dystrophin exon 53 (skipping)	DMD	BioMarin Pharmaceuticals (developed by Prosensa Therapeutics)	I/II - Active	NCT01957059 [247]

AVI-4685 (eteplirsen)	IV	Dystrophin exon 51 (skipping)	DMD	Sarepta Therapeutics	III - Recruiting	NCT02255552 [59]
SRP-4045	IV	Dystrophin exon 45 (skipping)	DMD	Sarepta Therapeutics	III - Not yet recruiting	NCT02500381 [248]
SRP-4053	IV	Dystrophin exon 53 (skipping)	DMD	Sarepta Therapeutics	III - Not yet recruiting	NCT02500381 [248]
DS-5141b	SC	Dystrophin exon 45 (skipping)	DMD	Daiichi Sankyo / Orphan Disease Treatment Instiute	I/II - Recruiting	NCT02667483 [65]
NS-065/NCNP-01	IV	Dystrophin exon 53 (skipping)	DMD	Nippon Shinyaku / National Center of Neurology and Psychiatry	II	not yet registered[62]
IONIS-SMNRx /ISIS 396443/ASO-10-27 (nusinersen)	IT	SNM2 exon 7 (retention)	SMA	Ionis Pharmaceuticals / Biogen	III - Recruiting	NCT02193074 [70]
IONIS-SMNRx /ISIS396443/ASO-10-27 (nusinersen)	IT	SNM2 exon 7 (retention)	SMA	Ionis Pharmaceuticals / Biogen	III - Active	NCT02292537 [71]

Aptamers

Compound	Application Route	Target	Indication	Company / Initiator	Phase / Status	Trial ID / Reference
Pegaptanib sodium (Macugen®)	IVT	VEGF	AMD	Evetech Pharmaceuticals, Pfizer	FDA - Approved	[75,76]
RB006 (pegnivacogin) as component of REG1	IV	Coagolation factor IXa	CAD	Regado Bioscience	III - Terminated	NCT01848106 [78]
E10030 (Fovista®)	IVT	PDGF	AMD	Ophthotech	III - Recruiting	NCT01944839 [80] NCT01940887 [81]
ARC1905 (Zimura®)	IVT	complement component C5	AMD	Ophthotech	II/III - Recruiting	NCT02686658 [83]
NOX-A12 (olaptesed pegol)	IV	CXCL12/SDF-1	MM	Noxxon Pharma	IIA - Completed	NCT01521533 [88]
NOX-A12 (olaptesed pegol)	IV	CXCL12/SDF-1	CLL	Noxxon Pharma	IIA - Active	NCT01486797 [89]
NOX-E36 (emapticap pegol)	IV	CCL2/MCP-1	DM2, Albuminuria	Noxxon Pharma	IIA - Completed	NCT01547897 [91]
NOX-H94 (lexaptepid pegol)	IV	Hepcidin	Anemia	Noxxon Pharma	IIA - Completed	NCT01691040 [92]

RNA Interference

Compound	Application Route	Target	Indication	Company / Initiator	Phase / Status	Trial ID / Reference
Bevasiranib	IVT	VEGF	Macular Degeneration	OPKO Health	III - Terminated	NCT00499590 [101]
AGN 211745	IVT	VEGFR-1	CNV, AMD	Allergan	II - Terminated	NCT00395057 [104]
I5NP	IV	p53	AKI	Quark Pharmaceuticals	I - Terminated	NCT00683553 [100]
PF-04523655	IVT	DDIT4	Diabetic Retinopathy, Diabetes Complications	Quark Pharmaceuticals	II - Terminated	NCT00701181 [102]
PRO-040201	IV	ApoB	Hypercholesterolemia	Arbutus Biopharma	I - Terminated	NCT00927459 [103]
ND-L02-s0201	IV	HSP47	Hepatic Fibrosis	Nitto Denko	I - Ongoing	NCT02227459 [114]
ALN-AS1	SC	ALAS1	AIP	Alnylam Pharmaceuticals	I - Recruiting	NCT02452372 [116]
pbi-shRNA STMN1 LP	IT	STMN1	Advanced / Metastatic Cancer, Solid Tumors	Gradalis	I - Ongoing	NCT01505153 [119]
Bamosiran	Eye drops	ADRB2	OAG, Ocular Hypertension	Sylentis	II - Completed	NCT02250612 [121]
ARB-001467	IV	Three sites of the HBV genome	Chronic Hepatitis B	Arbutus Biopharma	II - Recruiting	NCT02631096 [115]
ALN-TTRsc/ revusiran	SC	TTR	FAC	Alnylam Pharmaceuticals	III - Recruiting	NCT02319005 [117]
ALN-AT3	SC	AT	Hemophilia	Alnylam Pharmaceuticals	I - Recruiting	NCT02035605 [118]
siRNA-transfected PBMCs (APN401)	IV	Cbl-b	Several types of cancer	Comprehensive Cancer Center of Wake Forest University	I - Recruiting	NCT02166255 [120]
QPI-1007	IVT	CASP2	NAION	Quark Pharmaceuticals	II/III - Recruiting	NCT02341560 [122]
MRX34	IV	miRNA34a	Several types of cancer	Mirna Therapeutics	I - Recruiting	NCT01829971 [128]
TargoiRs	IV	miRNA16	Malignant Pleural Mesothelioma NSCLC	University of Sydney	I - Recruiting	NCT02369198 [127]
AZD4076	SC	miRNA103/107	NASH	AstraZeneca	I - Recruiting	NCT02612662 [249]
Miravirsen	SC	miRNA122	Hepatitis C	Santaris Pharma A/S	II - Ongoing	NCT02031133 [125]

(*continued*)

Table 1. (*Continued*)

Table 1: Clinical Trials using RNA-Therapeutics*

Compound	Application Route	Target	Indication	Company / Initiator	Phase / Status	Trial ID / Reference
			mRNA Therapy			
CV9104	ID	—	Prostate Cancer	CureVac	II - Ongoing	NCT01817738 [208]
CV9202	ID	—	Stage IV NSCLC	CureVac	I - Ongoing	NCT01915524 [206]
CV7201	IM	—	Rabies	CureVac	I - Ongoing	NCT02241135 [204]
CV9104	ID	—	Prostate Carcinoma	CureVac	II - Ongoing	NCT02140138 [209]
IVAC_W_bre1_uID IVAC_W_bre1_uID/IVAC_M_uID	Not available	—	TNBC	BioNtech	I - Not yet Recruiting	NCT02316457 [250]
IVAC MUTANOME RBL001/RBL002	IN	—	Melanoma	BioNtech	I - Ongoing	NCT02035956 [210]
Lipo-MERIT	IV	—	Melanoma	BioNtech	I - Recruiting	NCT02410733 [172]
AGS-003	ID	—	mRCC	Argos Therapeutics	III - Ongoing	NCT01582672 [196]
Dendritic cell vaccine (plus temozolomide chemotherapy)	ID	—	GBM	University Hospital, Antwerp	II - Recruiting	NCT02649582 [200]
Dendritic cell vaccine (plus chemotherapy)	IV	—	MPM	University Hospital, Antwerp	II - Recruiting	NCT02649829 [201]
pp65 Dendritic cell vaccine	SC	—	GBM	University of Florida	II - Not yet Recruiting	NCT02465268 [202]
ZFN Modified CD4+ T Cells	Infusion	CCR5	HIV	University of Pennsylvania	I - Recruiting	NCT02388594 [251]

*Abbreviations of targets, indications and administration routes: ADRB2: Adrenergic receptor β 2, AIP: Acute intermittent porphyria, AKI: Acute kidney Injury, ALAS: ALA-synthase, AMD: Age-related macular degeneration, ApoA/B: Apolipoprotein A/B, ApoB-100: Apolipoprotein B-100, AT: Antithrombin, CAD: Coronary artery disease, CASP2: Caspase 2, Cbl-b: Casitas B-lineage lymphoma proto-oncogene B, CCL2/MCP-1: CC-chemokine ligand 2/monocyte chemotactic protein 1, CMV: Cytomegalovirus, CNV: Choroid neovascularization, CLL: Chronic lymphocytic leukemia, CXCL-12: C-X-C motif chemokine ligand 12, DDIT4: DNA damage-inducible transcript 4, DMD: Duchenne muscular dystrophy, FAP: Familial amyloid polyneuropathy, FAC: Familial amyloidotic cardiomyopathy, GBM: Glioblastoma multiforme, HSP47: Heat shock protein 47, HoFH: Homozygous familial hypercholesterolemia, HIV: Human immunodeficiency virus, ICAM-1: Intercellular adhesion molecule 1, ID: intra-dermal, IE2: Immediate early 2 protein, IM: intra-muscular, IN: intra-nodal, IRS-1: Insulin receptor substrate 1, IT: intra-tumoral, IV: intra-venous, IVT: intra-vitreal, MPM: Malignant pleural mesothelioma, mRCC: Metastatic renal cell carcinoma, MM: Multiple myeloma, Akt-1: Murine thymoma viral oncogene homolog 1, NASH: Non-alcoholic steatohepatitis, NAION: Nonarteritic anterior ischemic optic neuropathy, NSCLC: Non-small cell lung cancer, OAG: Open angle glaucoma, PDGF: Platelet-derived growth factor, SDF-1: Stromal cell-derived factor 1, STAT3: Signal transducer and activator of transcription 3, SMA: Spinal muscular atrophy, STMN1: Stathmin-1, SC: Sub-Cutaneous, TTR: Transthyretin, DM2: Type 2 diabetes mellitus, VEGF: Vascular endothelial growth factor, VEGFR-1: Vascular endothelial growth factor receptor 1.

Splice-switching oligonucleotides

Pre-mRNA is matured during a complex nuclear process called splicing that removes the introns (non-coding sequences) and joins the exons (coding sequences). By applying alternative splice sites and by occasional inclusion or exclusion of exons and introns, multiple protein variants are derived from one gene (alternative splicing). Several diseases are related to aberrant RNA-splicing leading to non-functional proteins, and great efforts have been undertaken to develop antisense oligonucleotides, referred to as splice-switching oligonucleotides (SSOs) that manipulate splicing. Therapeutic SSOs promoting exon skipping and exon retention for the treatment for Duchenne muscular dystrophy (DMD) and spinal muscular atrophy (SMA) are currently evaluated in clinical trials.[7]

Dystrophin, the protein encoded by the DMD gene, is crucial for the integrity of muscle tissue.[52] In rare cases, newborn males harbor a defect dystrophin gene on their X chromosome. The patients suffer from successive muscle wasting resulting in a premature death due to respiratory or cardiac failure. In most cases, the loss-of-protein-function results from exonic out-of-frame deletions. In many cases the reading frame can be restored by skipping the aberrant exon by addressing a SSO to an internal exonic splicing enhancer.[53] The resulting truncated dystrophin protein retains partial function and gives the less severe Becker muscular dystrophy phenotype.[54] Several SSOs have been developed that are clinically evaluated for the skipping of exons 44, 45, 51, and 53, including drisapersen and eteplirsen (Table 1). Recently, both companies submitted new drug applications for their lead compounds drisapersen[55-57] and eteplirsen,[58,59], both amenable to exon 51 skipping. In case of drisapersen, the FDA rejected the application due to major concerns about the efficacy and safety of the drug.[60] The high dosage required led to severe adverse effects including renal and vascular injury. To improve efficacy and safety other SSO chemistries might be more successful. Whereas drisapersen is a 20 nt 2′O-methoxy phosphorothioate RNA analog, eteplirsen is a 30 nt phosphorodiamidate oligomer, a so-called morpholino. The final decision on the efficacy and safety evaluation by the FDA is still pending for eteplirsen. Additionally, a new, morpholino-based SSO for exon 53 skipping is currently under clinical evaluation (NS-065/NCNP-01).[61,62] For the future, we can hope in new chemistries. A SSO that relies on 2′O,4′C-ethylene-bridged nucleosides (ENA oligonucleotides)[63] which mediate nuclease resistance and improved binding affinity to RNA has now entered a clinical phase I/II trial for the treatment of DMD (DS-4151b).[64,65]

Spinal muscular atrophy (SMA) is a rare genetic disorder caused by survival of motor neuron 1 (SMN1) gene mutations.[66] Infant patients affected by this disease suffer from the loss of motor neurons and associated muscle wasting. However, there is a therapeutic approach by activating the SMN2 gene, which is almost identical to SMN1, but a single mutation in a splicing enhancer strongly prevents the inclusion of exon 7 resulting in an unstable protein unable to replace the lost SMN1 function.[67] In a mouse model, a highly potent 2′O-methoxyethyl PS SSO for exon 7 retention in SMN2 was identified (IONIS-SMN_{Rx}).[68] The drug is injected in the spinal cord ensuring the direct delivery to the affected motor neurons without the need to cross the blood-brain barrier. After promising clinical phase II results regarding efficacy and safety of the drug candidate,[69] two phase III trials were recently initiated for evaluating IONIS-SMN_{Rx}.[70,71]

Although the SSO design remains challenging, several new therapeutic applications were successfully validated in preclinical studies.[72] Possible drug approvals of eteplirsen or IONIS-SMN_{Rx} in the near future could eventually proof the feasibility of the splice-modulating antisense oligonucleotide approach.

Aptamers

Aptamers are 20 – 100 nt long oligomers that adopt complex three dimensional structures that allow them to interact potently and specifically with various proteins typically achieving nM- to pM binding affinities.[73] They are readily obtained in an iterative laboratory evolution procedure called SELEX (systematic evolution of ligands by exponential enrichment).[74] Currently, aptamers are mainly targeting extracellular structures such as plasma proteins and cell surface receptors thus avoiding the problem of intracellular delivery. Hence, aptamers are comparable in many aspects to antibodies, however, aptamers are much smaller, can penetrate tissues deeper, are chemically synthesized to highest purity and homogeneity and differ in their toxicity and immunogenicity profile. To improve their plasma life-time and to adjust their toxicity, aptamers are typically chemically stabilized (2′-OMe, 2′-F, 3′ inverted dT) and PEGylated.

In 2004, the first (and until today the only) aptamer, Macugen, was approved by the FDA for clinical therapy of AMD (age-related macular degeneration). The 27-nt chemically stabilized RNA oligomer is directed against the vascular endothelial growth factor (isoform 165) and blocks VEGF-receptor-induced neovascularization.[75,76] After achieving its highest sales in 2010, it has now almost entirely been displaced by antibodies (Ranibizumab and Bevacizumab, for instance) which can bind additional VEGF isoforms besides VEGF-165 and thus benefit for their poorer specificity compared to the aptamer. After this early breakthrough with Macugen, numerous aptamers have been explored in clinical settings. However, some programs suffered very unfortunate setbacks at late clinical trial states, like the aptamer-containing anticoagulation system REG1 which was terminated in 2014 in a phase III study due to unexpected toxicity / immunogenicity issues (Table 1).[77,78]

Currently, several aptamers for the local treatment of eye diseases are in late clinic trials (II and III), for instance the aptamers Fovista[79-81] and Zimura,[82,83] which target PDFG (it is a growth factor) and C5, respectively. In combination with VEGF inhibitors they might find application in the treatment of AMD in the near future. To overcome the prevalent problems with toxicity and immunogenicity, NOXXON Pharma develops so-called Spiegelmer therapeutics.[84] These drugs apply stereochemically inverted nucleotides based on L-ribose instead of the natural D-ribose, can be evolved via SELEX, and are suggested to be resistant against nucleases[85] and invisible for the immune system.[86] Currently, 3 Spiegelmer aptamers[86-92] are in clinical phase II studies (Table 1).

Therapeutic RNAi

RNA interference (RNAi) is a mechanism of posttranscriptional gene regulation that was discovered in 1998.[12] RNAi can

interfere with gene expression in various ways including the degradation of a specific mRNA target via endonucleolytic cleavage, or via recruitment of deadenylation / decapping enzymes, but it can also positively affect the stability and translation of a specific mRNA. The mechanistic details that lead to the respective responses are still under exploration. In principle, a dsRNA that is introduced into the cytoplasm is processed by the RNase dicer into ~22 bp RNA duplexes and loaded onto the endonuclease Argonaut-2 (Ago-2). Ago-2 slices the passenger strand of the RNA-duplex and applies the remaining guide strand for sequence-specific mRNA-targeting.[93] While short interfering RNAs (siRNAs) are fully complementary to their target mRNA and promote cleavage (knock-down), micro RNAs (miRNAs) contain bulges and loops that prohibit slicing by Ago-2, but alter the stability and translational activity of the target.[94]

Allowing the selective knock-down of genes in cell culture and animal-models, RNAi quickly became a valuable tool in basic biology.[95-97] In parallel a race started to exploit the RNAi mechanism for therapeutic purposes and several big pharma companies, like Merck, Roche, and Pfizer made large investments that resulted in the first clinical trials in 2004, already 6 y after the discovery of RNAi.[98,99] However, in the aftermath those early trials mostly failed due to strong innate immune reactions and/or lack of patients´ benefit, and in the consequence big pharma left RNAi again.[100-104] In the 18 y since its discovery the field of therapeutic RNAi went from enthusiastic interest over despondence and back again, resulting in a reassessment of the technological obstacles and more realistic expectations for clinical trials. This has been accompanied by commentary elsewhere.[105,106]

However, after recent successes in clinical trials, showing the efficacy of RNAi therapeutics to reduce transthyretin[107] and PCSK9[108] in patients, the interest in RNAi is currently growing and even big pharma including Sanofi and Roche started to invest again.[98] The initial drawbacks in clinical trials were mostly related to the low efficacy of the drugs, off-target issues and immune-related toxicity.[109] Off-target effects include immune-reactions induced by the siRNA/miRNA precursors, and up- and downregulation of non-target mRNAs due to saturation of the RNAi machinery and off-target binding of the siRNA.[110] There is now increasing success in tackling all those issues. Current innovations include chemical modification / sequence optimization of siRNAs and its precursors, and new solutions to the delivery problem. The latter include various forms of (lipid) nanoparticles and bioconjugates. The details of this progress are comprehensively reviewed elsewhere.[110-113] Briefly, clinical trials seem more successful when they are confined to readily accessible organs like the liver, cancer, and immune-privileged areas like the eye.[114-122] Whereas the eye is a good target for naked siRNAs, treatment of the liver benefited from lipid-based nanoparticles and the above-mentioned $GalNAc_3$ conjugates.[116] In particular the $GalNAc_3$ approach has significantly improved the efficacy of siRNA-conjugates, allowing now the weekly administration of liver-targeting siRNA via subcutaneous injection in non-human primates to knock-down antithrombin to clinically relevant levels.[11] Notable in this approach is that it allows to knockdown an essential protein (like antithrombin) in a tunable and reversible manner, whereas the permanent knock-out of antithrombin (for instance at the DNA-level) is lethal.[11] Overall, more than 20 siRNA drugs in various formulations are in clinical trials now (up to phase III, Table 1).[123] RNAi-therapy clearly has the potential to tackle currently undruggable diseases and to appear in the clinics soon.

The therapeutic use of the miRNA-related mechanism (not applying the slicing activity of Ago2) is still in its infancy. Attractive is the possibility of manipulating larger networks of genes simultaneously in both, a negative and positive manner.[124] This might become interesting for the treatment of complex diseases like cancer. On the other hand, endogenous miRNAs are involved in many cellular processes and their manipulation could also be disease-relevant. The knockdown of miRNA 122 with antisense oligonucleotides was shown to interfere with hepatitis C virus progression and is currently in phase II clinical studies.[125] As the hepatitis virus seems to require the endogenous miRNA for its functioning the knockdown of this host-specific factor is particularly promising as the virus cannot adapt easily by evolution.[126] Other miRNAs that are linked to cancer like miRNA 16 and 34a are also targeted with ASOs and are currently in clinical trials phase I.[127,128]

Emerging concepts for therapy

Therapeutic mRNA

For a long time it has been believed that only short, chemically stabilized oligonucleotides are suitable as drugs. However, long (protein-encoding) mRNAs haven recently proven their enormous therapeutic potential. Protein replacement experiments were first performed in the early 1990ties with naked mRNA in mice and rats.[129,130] Even though replacement experiments were successful to some degree, there have been massive problems related to the well-known RNA-dependent immune-stimulation through interferon-I (IFN-I) and a generally low translation efficiency.[131,132]

However, during the last 15 years, our mechanistic understanding of the immune-stimulatory effect of RNA has substantially improved. This was due to the discovery of RNA sensors including the Toll-like receptors (TLR) 3, 7, 8, Melanoma differentiation-associated protein 5 (MDA-5), Retinoic acid inducible gene I (RIG-I), as well as various RNA helicases.[133] Besides the activation of the innate immune response under release of the respective signaling molecules we have also learned how these RNA-sensing events are directly linked to the general repression of mRNA translation in the affected cells. Among others, general translation repression is mediated by phosphorylation of translation initiation factor 2α via protein kinase R activation.[134,135] In the worst case, IFN-I activates 2′–5′-adenylate synthase and RNaseL and leads to apoptosis.[136]

RNA replacement strategies aim to achieve high translation levels under minimal immune stimulation. Both can be achieved by designing mRNAs that evade RNA-sensing. The following strategies turned out as particularly successful.

a) Chemically modified pyrimide nucleotides like pseudouridine (ψ), 2-thiouridine (s2U), and 5-methylcytidine (m5C) are incorporated into mRNAs during in-vitro-transcription to minimize recognition by RNA

sensors.[137] Substitution of uridine by pseudouridine was shown to diminish recognition by TLR-3, -7, -8, and RIG-I.[137,138] To fine-tune effects on translation efficiency, nucleotide analogs are often mixed with their natural counterparts. The extent to which these modifications may induce mistranslation is yet unknown.[139]

b) Rigorous purification of the mRNA product from unincorporated nucleoside triphosphates, small abortive transcripts, remaining DNA templates, and in particular dsRNA via HPLC (High performance liquid chromatography) was shown to dramatically reduce immunogenicity of the transcripts and can increase the translation 10- to 1000-fold.[140,141]

c) Synthetic cap analog structures like ARCA (anti-reverse-cap-analog) can further decrease immune response and improve translation. In contrast to older cap analogs, ARCA is always incorporated in correct orientation.[142,143] A new ARCA variant contains a phosphothioate that resists enzymatic decapping and can increase the half-life of the mRNA.[144]

d) Computational sequence design allows to reduce the number of particularly immune-stimulatory nucleotides and combinations (like UW, with W = A or U). [145-147] Furthermore, transcript stability can be optimized by the introduction of 3′-UTRs (or some elements) taken from other mammalian or viral genes as well as addition of Poly(A)-tails.[148-153]

The RNA replacement strategy is particularly advantageous when a transient, burst-like expression of a protein is desired. Typical examples for the latter are the epigenetic re-programming (induced pluripotency), wound healing, and genome editing. In this sense, in-vitro transcribed mRNA (IVT-mRNAs) has been used to deliver a) human bone morphogenetic protein 2 (hBMP-2) to support bone regeneration in rats; to deliver b) the transcription factor mix that induces pluripotency; and to deliver c) vascular endothelial growth factor-A (VEGF-A) into a mouse model for myocardial infarction resulting in an improved heart function and enhanced survival.[154-160] Furthermore, IVT-mRNAs have been successful in the delivery of surfactant protein B in deficient mice, and in the delivery of murine erythropoietin to increase the hematocrit.[138,161]

IVT-mRNA could turn out as a valuable tool for genome editing. Genome editing holds great promise for the treatment of various diseases by a permanent repair of a gene via a site-directed knock-in or knockout.[162] However, the respective nucleases that induce the required double-strand DNA breaks including ZFNs, Talens, and CRISPR/Cas, should not be persistently expressed as this would dramatically increase the chance of off-target genome editing.[9] Consequently, its delivery as an mRNA is beneficial compared to a DNA vector and also circumvents the typical safety risks of viral and non-viral DNA-based methods like genomic insertion and antivector immunogenicity. Encoding of genome editing tools via IVT-mRNAs has already been widely used to generate transgenic animals.[163-169] In a proof-of-concept study, gene function was restored via homology-directed promotor exchange in a surfactant-B-deficient mouse model by in-vivo-delivery of the ZFN in form of an IVT-mRNA. However, this required the additional delivery of the repair template (with the promotor) in form of an AAV6 (Adeno-associated-virus serotype 6).[170] Successful promoter exchange was demonstrated and resulted in a prolonged life of the treated mice. IVT-mRNA encoded Talen have been used successfully to disrupt the CCR5 (CC chemokine receptor type 5) gene via non-homologous-end-joining in the T-cell line PM1. As the loss of CCR5 function confers resistance toward R5-tropic HIV-1 infection, side-directed nucleases are promising to target this infectious disease.[171] An initial clinical phase I study is currently starting.[172] As IVT-mRNA is a young field, this study represents the first clinical study that uses IVT-mRNAs, but more are likely to follow soon.

mRNA can have many advantages over DNA vectors to deliver therapeutic proteins. Besides its transient nature, we want note that mRNA is very well and quickly translated in postmitotic cells that are difficult to transfect with DNA vectors. mRNA also works independent of a promotor, but this can potentially limit its application if tissue-specificity is required. However, we know from various studies that there is a large number of regulatory elements, typically in the 3′-UTR, including miRNA binding sites, stabilizing and destabilizing elements that could allow to manipulate the expression of an IVT-mRNA in a tissue-specific manner in the future.[94,173]

Oligonucleotides for vaccination and desensitization

As indicated above, very successful strategies haven't been developed to evade the RNA-sensing event and to trick the innate immune system. However, inducing a specific immune response can be highly desired. Thus the recent knowledge on the immune stimulation by RNA can be used for the latter. Currently, the classical vaccination is based on the delivery of inactivated or living viruses, virus-like particles, or antigenic peptides. While the antigenic peptides require additional vaccination adjuvants like alum salts, the other entities contain sufficient pathogen-associated-molecular-patterns (PAMPs) in form of proteins, nucleic-acids, and lipopolysaccharides. These PAMPs are detected by pattern-recognition-receptors (including the above-mentioned RNA sensors) and induce the release of type-I interferons, pro-inflammatory cytokines, and chemokines. This is reviewed in-depth elsewhere.[174,175] Short peptide fragments are then presented to the immune system via MHC-complexes on dendritic cells and other antigen presenting cells.[176] This process finally induces a humoral as well as cellular immune response of the adaptive immune system.

The presented antigens are mainly protein-derived peptides. This opens the intriguing possibility to deliver antigens for MHC-presentation encoded as IVT-mRNAs under simultaneous induction of the necessary innate and adaptive immune stimulation as the IVT-mRNA itself can function as PAMP. By doing so, it is well conceivable to create specific immune responses not only against viruses and bacteria, but also against cancer cells or for allergy treatment.[177-181] The design of such mRNA-based vaccines would be highly rational, fast, cheap, and could be done in a personalized manner, for instance against the specific transcriptome of a patient-specific cancer.[182] IVT-mRNA vaccines would be faster available as the generation of virus-particles (and similar entities) would be circumvented. Lyophilized mRNA vaccines can be stored at 37 °C for several weeks.[183] This allows the transport of vaccines into

regions that cannot provide an uninterrupted cold chain. The safety-profile could also be better compared to DNA-based methods (insertion mutagenesis, low efficiency) or virus-like entities (therapy-induced virus-specific humoral immune response).[184-186] Again, also for vaccination, the transient nature of RNA expression is beneficial, as a low-level, long-term expression of an antigen might induce tolerance.[187]

Two major IVT-mRNA-based vaccination strategies are currently explored: the ex-vivo and the in-vivo approach. The first, which was earlier developed, is based on the ex-vivo pulsing of allogenic (= patient-derived) dendritic cells with antigen-encoding mRNA, which allows the redirection of the adaptive immune system to target cancer or virus-infected cells. The feasibility and safety of this method was proven in pre- and clinical trials focused on HIV and various cancer types. However, personalized ex-vivo therapies require time-consuming and expensive individualized manufacturing processes which currently limit their broad clinical application.[188-195] Nevertheless, further clinical trials up to phase III are currently running.[196-202]

Even though cumbersome, the ex-vivo strategy allows to optimize and control mRNA transfection and immune stimulation more carefully. The in-vivo approach, however, is potentially more simple and elegant, but encounters additional problems. Whereas all IVT-mRNA strategies require stable and highly translatable transcripts, the in-vivo strategy requires additionally the immune-stimulatory effect that counteracts translation. It was found that complexation of IVT-mRNA with protamine enhances immunogenicity via TLR-7 activation and simultaneously improves stability, however, with the downside of low antigen expression.[203] Anyway, a combination of protamine-complexed IVT-mRNA together with naked IVT-mRNA of the same sequence turned out to satisfy both needs at the same time: high translation efficiency and immune stimulation. Those self-adjuvanting mRNAs are currently in phase I and II clinical trials against prostate cancer, late stage lung cancer, and rabies; pre-clinical trials against influenza have been performed.[183,186,204-209] We wish to mention that also other approaches that apply naked or formulated IVT-mRNAs are in clinical trials, for instance for targeting other cancer entities.[172,210] Furthermore, non-coding RNA can also be used as a vaccination adjuvant replacing the classical alum salts as adjuvant of protein- or peptide-based vaccines.[211]

Currently, IVT-mRNA are expensive therapies. On one hand, the GMP (Good manufacturing practice) production of IVT-mRNA in large scale is not yet fully established, but CureVac has announced significant progress here.[212] On the other hand the potency of IVT-mRNA could be further improved by assisted delivery via lipid-nanoparticels, polymeric nanoparticles, gold nanoparticles, among others, as reviewed elsewhere.[213] Furthermore, there are promising attempts to develop self-replicating RNA-vaccines that apply viral RNA-dependent RNA-polymerases (from α-virus) to produce the RNA vaccine from a dilute IVT-mRNA template.[214-216] However, there are safety concerns related to the control of the replication process and the tolerance against the viral RNA-polymerase, but the strategy is still in the pre-clinical exploration phase.[217]

Finally, mRNA vaccines could also be used in allergy treatment to desensitize the immune system against a specific antigen. Desensitization against type-I allergies is typically accomplished through repeated intra-dermal, intra-nodal, or sub-lingual application of allergens. Whereas a strong Immunglobuline E and $CD8^+$ T-cell responses is intended during vaccination, desensitization aims to change the T_H1/T_R1 to T_H2 cell ratio toward T_H1/T_R1 to fine-tune the immune response and to induce tolerance.[178] Application of low-dose IVT-mRNA could be used for that purpose, and there is pre-clinical data that prove efficacy and suggest a long-term protective effect.[218] One can expect first clinical trials to start within the next few years. Applying mRNA as an anti-allergic vaccine has several advantages compared to the classical allergen extract (like standardized cat extract) or DNA-based vaccines.[219,220] IVT-mRNA is

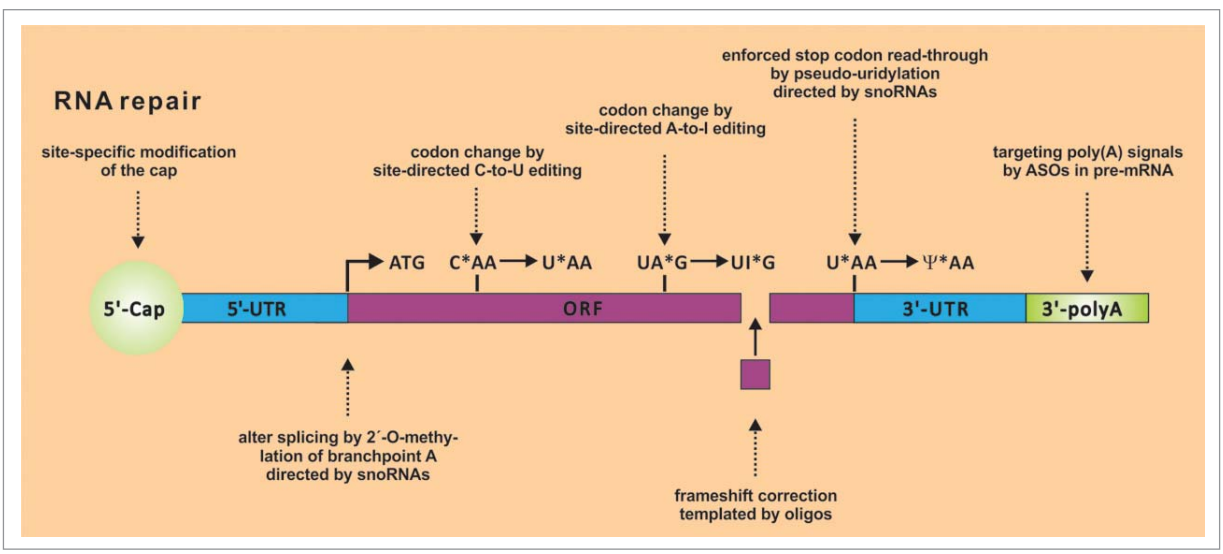


Figure 2. Overview on selected enzymatic processes that could be harnessed to restore gene function by repairing or re-programming mRNA site-specifically. Site-directed A-to-I editing, 2´-O-methylation, pseudouridylation, and frameshift correction via expression or administration of short guideRNAs has already been demonstrated. Many other processes are conceivable and currently under exploration.

obtained in a defined and highly pure state thus avoiding unintended antigens that can be included in allergen extracts.[221,222] DNA-based allergy treatment on the other hand suffers from the above mentioned safety concerns and thus harbors disproportional risk in the context of a preventative therapy.

RNA repair

Besides the manipulation of splicing, most interventions on the RNA-level aim to destroy or block their endogenous targets. Strategies to restore the function of an RNA that is corrupted by missense, nonsense or frameshift mutation, or by defective processing are rare. In case of loss-of-function mutations, the administration of a therapeutic mRNA to replace the non-functional variant might solve the problem, as discussed above. However, this is only feasible with a small number of therapeutic mRNAs that can be translated under low control of translation level and tissues specificity. Indeed, many transcripts are tightly regulated with respect to their dose and tissue specificity and come as a mixture of various isoforms due to alternative promotor usage, alternative splicing, alternative polyadenylation and alternative posttranscriptional modification. Such transcript variants may differ in their function, localization, stability, etc. To address this variety in an mRNA replacement strategy seems impractical. A better alternative would be the repair of the endogenously expressed but defective RNA transcript, a strategy, we call RNA repair.

Very recently, we and others have engineered artificial RNA-guided editing machineries that allow to re-program genetic information at the RNA level.[223-225] For this, adenosine-to-inosine (A-to-I) RNA editing enzymes[226,227] are directed toward specific sites on selected transcripts and allow for the precise posttranscriptional manipulation of the genetic information. The manipulation results from the fact that inosine is biochemically interpreted as guanosine. Thus, formal A-to-G conversions become accessible, in a highly site-specific manner. The specificity comes from the guide-RNA that addresses the editing enzymes and can be readily programmed in rational way, simply by applying Watson-Crick pairing rules.[228] Even though only A-to-G mutations are accessible the scope of manipulations is large. Twelve out of the 20 canonical amino acids can be manipulated, comprising almost all of the polar ones which are essential for protein function.[223] Furthermore, START and STOP codon, splice elements, polyadenylation signals, and viral RNA are potential targets.[226,227] We and others have shown that such strategies work inside mammalian cell culture[229] and even in a simple organism[230] and allow the repair of disease-relevant genes, like the CFTR mRNA.[225]

Other people have recently shown the possibility of re-directing snoRNA-guided RNA modification machineries, like the 2′-O-methylation[231] and the pseudouridylation machinery.[139] The first modification allows interference with splicing, the second allows the read-through of premature STOP codons. Mammalian cells harbor a plethora of RNA modifying and processing enzymes. There is no need to restrict ourselves to the usage of nucleases, like RISC, RNaseH, and RNaseP.[232] Just to give a few examples, there are RNA editing and modifying enzymes inside the cell that can change nucleotides (A-to-I, C-to-U[233], U-to-ψ, A-to-m6A,[234] and many more for the tRNAs[235]), that add the cap[236] and the poly(A)-tail,[237] RNAs can be precisely processed, for instance by the CCA-adding enzymes,[238,239] TUTases,[240] etc.[241] Thus, even complex repair processes are conceivable, including the repair of insertion and deletion mutations at the RNA-level. In this respect, we want to recall a largely overseen work from 2004, done by Paul Zamecnik, the pioneer of antisense therapy, in his early nineties shortly before he passed away. He demonstrated the possibility of repairing the terrible Δ508 deletion mutation in the CFTR gene, the main cause of cystic fibrosis, simply by administration of 2 chemically stabilized RNA oligomers.[242] In cell culture, the efficiency of mRNA repair was sufficient to restore the chloride channel function. Unfortunately, he was unable to elucidate the mechanism, but he could clearly demonstrate the repair to take place at the mRNA. Such a complex repair requires a concerted nuclease, ligase (and polymerase) activity at a specific site on an mRNA molecule. In summary, it seems that numerous endogenous enzymes stand ready inside the cell for RNA repair processes. We just have to learn how to make use of them.[243,244] If successful, one can establish novel platforms for therapeutic intervention.

Conclusions

While splice-switching oligomers and aptamers are still struggling on their ways to the clinic, major progress has been made for RNaseH-dependent ASOs and for therapeutic RNAi with chemically stabilized siRNAs. This is due to the development of new chemistries that improve efficacy and delivery of the drugs to some specific organs. An impressive example is the development of the $GalNAc_3$ conjugation that clearly improves liver targeting and might allow for the administration of siRNA and ASO by subcutaneous administration in the future. However, overcoming problems with delivery and efficacy remains elusive for many organs and will require massive basic research in the future.

Among the emerging approaches, the usage of in-vitro-transcribed mRNA for protein replacement and vaccination has made impressive progress. This was mainly due to the tailored suppression or harnessing of the RNA-induced immune response by chemical modification and formulation. The approach has the potential to find wide application in the clinics whenever a transient, burst-like expression is advantageous. The RNA repair approach is still in its infancy, but we believe that the harnessing of artificial and in particular endogenous RNA repair proteins might enable new therapies, complementing the above-mentioned classical RNA-based and the approaching genome editing methods, and being superior to the latter with respect to safety and ethical issues.

Overall, the progress during last years is impressive. The increasing number of clinical trials for various approaches makes us feel optimistic that numerous nucleic-acid-based drugs will soon find their ways to the patients to enable novel therapies.

Disclosure of potential conflicts of interest

No potential conflicts of interest were disclosed.

Funding

We gratefully acknowledge support from the University of Tübingen and the Deutsche Forschungsgemeinschaft (STA 1053/3–2, STA 1053/4–1). This work has received funding from the European Research Council (ERC) under the European Union's Horizon 2020 research and innovation program (grant agreement No 647328, RNArepair).

References

1. Klungland A, Dahl JA. Dynamic RNA modifications in disease. Curr Opin Genet Dev 2014; 26:47-52; PMID:25005745; http://dx.doi.org/10.1016/j.gde.2014.05.006
2. Schoenberg DR, Maquat LE. Regulation of cytoplasmic mRNA decay. Nat Rev Genet 2012; 13:246-59; PMID:22392217; http://dx.doi.org/10.1038/nrg3254
3. Chabot B, Shkreta L. Defective control of pre-messenger RNA splicing in human disease. J Cell Biol 2016; 212:13-27; PMID:26728853; http://dx.doi.org/10.1083/jcb.201510032
4. Scheper GC, van der Knaap MS, Proud CG. Translation matters: protein synthesis defects in inherited disease. Nat Rev Genet 2007; 8:711-23; PMID:17680008; http://dx.doi.org/10.1038/nrg2142
5. Nalavade R, Griesche N, Ryan DP, Hildebrand S, Krauss S. Mechanisms of RNA-induced toxicity in CAG repeat disorders. Cell Death Dis 2013; 4:e752; PMID:23907466; http://dx.doi.org/10.1038/cddis.2013.276
6. Bennett CF, Swayze EE. RNA targeting therapeutics: molecular mechanisms of antisense oligonucleotides as a therapeutic platform. Annu Rev Pharmacol Toxicol 2010; 50:259-93; PMID:20055705; http://dx.doi.org/10.1146/annurev.pharmtox.010909.105654
7. Kole R, Krainer AR, Altman S. RNA therapeutics: beyond RNA interference and antisense oligonucleotides. Nat Rev Drug Discov 2012; 11:125-40; PMID:22262036; http:dx.doi.org/10.1038/nrd3625
8. Geary RS. Antisense oligonucleotide pharmacokinetics and metabolism. Expert Opin Drug Metab Toxicol 2009; 5:381-91; PMID:19379126; http://dx.doi.org/10.1517/17425250902877680
9. Cox DB, Platt RJ, Zhang F. Therapeutic genome editing: prospects and challenges. Nat Med 2015; 21:121-31; PMID:25654603; http://dx.doi.org/10.1038/nm.3793
10. Myra Stern and Eric WFW Alton. Use of Liposomes in the Treatment of Cystic Fibrosis. Gene Therapy in Lung Disease CRC Press, 2002; 383–396; http://dx.doi.org/10.3109/9780203908822-18
11. Sehgal A, Barros S, Ivanciu L, Cooley B, Qin J, Racie T, Hettinger J, Carioto M, Jiang Y, Brodsky J, et al. An RNAi therapeutic targeting antithrombin to rebalance the coagulation system and promote hemostasis in hemophilia. Nat Med 2015; 21:492-7; PMID:25849132; http://dx.doi.org/10.1038/nm.3847
12. Fire A, Xu S, Montgomery MK, Kostas SA, Driver SE, Mello CC. Potent and specific genetic interference by double-stranded RNA in Caenorhabditis elegans. Nature 1998; 391:806-11; PMID:9486653; http://dx.doi.org/10.1038/35888
13. Wu H, Lima WF, Zhang H, Fan A, Sun H, Crooke ST. Determination of the role of the human RNase H1 in the pharmacology of DNA-like antisense drugs. J Biol Chem 2004; 279:17181-9; PMID:14960586; http://dx.doi.org/10.1074/jbc.M311683200
14. Vitravene Study Group. A randomized controlled clinical trial of intravitreous fomivirsen for treatment of newly diagnosed peripheral cytomegalovirus retinitis in patients with AIDS. Am J Ophthalmol 2002; 133:467-74; PMID:11931780; http://dx.doi.org/10.1016/S0002-9394(02)01327-2
15. Vitravene Study Group. Randomized dose-comparison studies of intravitreous fomivirsen for treatment of cytomegalovirus retinitis that has reactivated or is persistently active despite other therapies in patients with AIDS. Am J Ophthalmol 2002; 133:475-83; PMID:11931781; http://dx.doi.org/10.1016/S0002-9394(02)01326-0
16. Vitravene Study Group. Safety of intravitreous fomivirsen for treatment of cytomegalovirus retinitis in patients with AIDS. Am J Ophthalmol 2002; 133:484-98; PMID:11931782; http://dx.doi.org/10.1016/S0002-9394(02)01332-6
17. Greuter T, Biedermann L, Rogler G, Sauter B, Seibold F. Alicaforsen, an antisense inhibitor of ICAM-1, as treatment for chronic refractory pouchitis after proctocolectomy: A case series. United European Gastroenterol J 2016; 4:97-104; PMID:26966529; http://dx.doi.org/10.1177/2050640615593681
18. Randomized Study of Topical Alicaforsen Enema in Antibiotic Refractory Pouchitis. (2015). Clinical Phase Trial III
19. Cloutier F, Lawrence M, Goody R, Lamoureux S, Al-Mahmood S, Colin S, Ferry A, Conduzorgues JP, Hadri A, Cursiefen C, et al. Anti-angiogenic activity of aganirsen in nonhuman primate and rodent models of retinal neovascular disease after topical administration. Invest Ophthalmol Vis Sci 2012; 53:1195-203; PMID:22323484; http://dx.doi.org/10.1167/iovs.11-9064
20. Cursiefen C, Viaud E, Bock F, Geudelin B, Ferry A, Kadlecová P, Lévy M, Al Mahmood S, Colin S, Thorin E, et al. Aganirsen antisense oligonucleotide eye drops inhibit keratitis-induced corneal neovascularization and reduce need for transplantation: the I-CAN study. Ophthalmology 2014; 121:1683-92; PMID:24811963; http://dx.doi.org/10.1016/j.ophtha.2014.03.038
21. A multicentre double-blind randomized study to investigate the efficacy and tolerability of GS-101 eye drops, an antisense oligonucleotide, versus placebo on inhibition of corneal neovascularization, a major risk factor of corneal graft rejection: The I-GRAFT study. (2009). Clinical Trial Phase III
22. The STRONG Study. (http://strong-nvg.com/the-study/). Clinical Trial Phase II/III
23. Yoon H, Kim DJ, Ahn EH, Gellert GC, Shay JW, Ahn CH, Lee YB. Antitumor activity of a novel antisense oligonucleotide against Akt1. J Cell Biochem 2009; 108:832-8; PMID:19693774; http://dx.doi.org/10.1002/jcb.22311
24. A Safety and Efficacy Study of RX-0201 Plus Gemcitabine in Metastatic Pancreatic Cancer. (2009). Clinical Trial Phase II
25. Dose-Finding, Safety and Efficacy Study of RX-0201 Plus Everolimus in Metastatic Renal Cell Cancer. (2014). Clinical Trial Phase IB/II
26. Tagawa ST, Chatta GS, Mazhari R, Benaim E. Archexin, a novel AKT-1-specific inhibitor for the treatment of metastatic renal cancer: Preliminary phase I data. ASCO Annual Meeting Proceedings 2016; (suppl 2S; abstr 550);
27. MacLeod AR. Antisense therapies for cancer: bridging the pharmacogenomic divide. Drug Discovery Today: Therapeutic Strategies 2013; 10:e157-e63
28. Teplova M, Minasov G, Tereshko V, Inamati GB, Cook PD, Manoharan M, Egli M. Crystal structure and improved antisense properties of 2′-O-(2-methoxyethyl)-RNA. Nat Struct Mol Biol 1999; 6:535-9; PMID:10360355; http://dx.doi.org/10.1038/9304.
29. Raal FJ, Santos RD, Blom DJ, Marais AD, Charng M-J, Cromwell WC, Lachmann RH, Gaudet D, Tan JL, Chasan-Taber S, et al. Mipomersen, an apolipoprotein B synthesis inhibitor, for lowering of LDL cholesterol concentrations in patients with homozygous familial hypercholesterolaemia: a randomised, double-blind, placebo-controlled trial. The Lancet 2010; 375:998-1006; http://dx.doi.org/10.1016/S0140-6736(10)60284-X
30. Stein EA, Dufour R, Gagne C, Gaudet D, East C, Donovan JM, Chin W, Tribble DL, McGowan M. Apolipoprotein B synthesis inhibition with mipomersen in heterozygous familial hypercholesterolemia: results of a randomized, double-blind, placebo controlled trial to assess efficacy and safety as add-on therapy in patients with coronary artery disease. Circulation 2012; 126:2283-92; PMID:23060426; http://dx.doi.org/10.1161/CIRCULATIONAHA.112.104125
31. Santos RD, Duell PB, East C, Guyton JR, Moriarty PM, Chin W, Mittleman RS. Long-term efficacy and safety of mipomersen in patients with familial hypercholesterolaemia: 2-year interim results of an open-label extension. Eur Heart J 2015; 36:566-75; PMID:24366918; http://dx.doi.org/10.1093/eurheartj/eht549
32. Ackermann EJ, Guo S, Booten S, Alvarado L, Benson M, Hughes S, Monia BP. Clinical development of an antisense therapy for the treatment of transthyretin-associated polyneuropathy. Amyloid 2012; 19:43-4; PMID:22494066; http://dx.doi.org/10.3109/13506129.2012.673140
33. Efficacy and Safety of IONIS-TTR Rx in Familial Amyloid Polyneuropathy. (2012). Clinical Phase trial II/III

34. Open-Label Extension Assessing Long Term Safety and Efficacy of IONIS-TTR Rx in Familial Amyloid Polyneuropathy (FAP). (2012). Clinical Phase Trial III
35. Ionis Pharmaceuticals Press Release. (http://ir.isispharm.com/phoenix.zhtml?c=222170&p=irol-newsArticle&ID2105651). 2015
36. Comparison of Docetaxel/Prednisone to Docetaxel/Prednisone in Combination With OGX-011 in Men With Prostate Cancer (SYNERGY). (2010). Clinical Trial Phase III
37. Comparison of Cabazitaxel/Prednisone Alone or in Combination With Custirsen for 2nd Line Chemotherapy in Prostate Cancer (AFFINITY). (2012). Clinical Trial Phase III
38. OncoGenex Press Release. (http://ir.oncogenex.com/releasedetail.cfm?ReleaseID=933276). 2015
39. Laskin JJ, Nicholas G, Lee C, Gitlitz B, Vincent M, Cormier Y, Stephenson J, Ung Y, Sanborn R, Pressnail B, et al. Phase I/II trial of custirsen (OGX-011), an inhibitor of clusterin, in combination with a gemcitabine and platinum regimen in patients with previously untreated advanced non-small cell lung cancer. J Thorac Oncol 2012; 7:579-86; PMID:22198426; http://dx.doi.org/10.1097/JTO.0b013e31823f459c
40. A Multinational, Randomized, Open-Label Study of Custirsen In Patients With Advanced or Metastatic (Stage IV) Non-Small Cell Lung Cancer. (2012). Clinical Trial Phase III
41. Seth PP, Vasquez G, Allerson CA, Berdeja A, Gaus H, Kinberger GA, Prakash TP, Migawa MT, Bhat B, Swayze EE. Synthesis and biophysical evaluation of 2′, 4′-constrained 2′ O-methoxyethyl and 2′, 4′-constrained 2′ O-ethyl nucleic acid analogues. J Org Chem 2010; 75:1569-81; PMID:20136157; http://dx.doi.org/10.1021/jo902560f
42. Hong D, Kurzrock R, Kim Y, Woessner R, Younes A, Nemunaitis J, Fowler N, Zhou T, Schmidt J, Jo M, et al. AZD9150, a next-generation antisense oligonucleotide inhibitor of STAT3 with early evidence of clinical activity in lymphoma and lung cancer. Sci Translat Med 2015; 7:314ra185-314ra185; http://dx.doi.org/10.1126/scitranslmed.aac5272
43. AZD9150, a STAT3 Antisense Oligonucleotide, in People With Malignant Ascites. (2015). Clinical Trial Phase II
44. Study to Assess MEDI4736 With Either AZD9150 or AZD5069 in Relapsed Metastatic Squamous Cell Carcinoma of Head & Neck. (2015). Clinical Trial Phase IB/II
45. Phase 1/2, Open-label, Dose-escalation Study of IONIS-STAT3Rx, Administered to Patients With Advanced Cancers. (2012). Clinical Trial Phase I/II
46. MEDI4736 Alone and in Combination With Tremelimumab or AZD9150 in Adult Subjects With Diffuse Large B-cell Lymphoma (D4190C00023). (2015). Clinical Trial Phase IB
47. Prakash PT, Seth PP, Swayze EE, Graham MJ. Compositions and methods for modulating apolipoprotein (a) expression (US. Patent No. Nine,181,550). 2015
48. Nair JK, Willoughby JL, Chan A, Charisse K, Alam MR, Wang Q, Hoekstra M, Kandasamy P, Kel'in AV, Milstein S, et al. Multivalent N-acetylgalactosamine-conjugated siRNA localizes in hepatocytes and elicits robust RNAi-mediated gene silencing. J Am Chem Soc 2014; 136:16958-61; PMID:25434769; http://dx.doi.org/10.1021/ja505986a
49. Prakash TP, Graham MJ, Yu J, Carty R, Low A, Chappell A, Schmidt K, Zhao C, Aghajan M, Murray HF, et al. Targeted delivery of antisense oligonucleotides to hepatocytes using triantennary N-acetyl galactosamine improves potency 10-fold in mice. Nucleic Acids Res 2014; 42:8796-807; PMID:24992960; http://dx.doi.org/10.1093/nar/gku531
50. Ionis Pharmaceuticals Press Release. (http://ir.isispharm.com/phoenix.zhtml?c=222170&p=irol-newsArticle&ID=2110180) 2015
51. Safety, Tolerability, Pharmacokinetics, and Pharmacodynamics of IONIS APO(a)-LRx in Healthy Volunteers With Elevated Lipoprotein(a). (2015). Clinical Trial Phase I
52. Hoffman EP, Brown RH, Kunkel LM. Dystrophin: the protein product of the Duchenne muscular dystrophy locus. Cell 1987; 51:919-28; PMID:3319190; http://dx.doi.org/10.1016/0092-8674(87)90579-4
53. Aartsma−Rus A, Fokkema I, Verschuuren J, Ginjaar I, van Deutekom J, van Ommen GJ, den Dunnen JT. Theoretic applicability of antisense−mediated exon skipping for Duchenne muscular dystrophy mutations. Hum Mutat 2009; 30:293-9; PMID:19156838; http://dx.doi.org/10.1002/humu.20918
54. Koenig M, Beggs A, Moyer M, Scherpf S, Heindrich K, Bettecken T, Meng G, Müller CR, Lindlöf M, Kaariainen H, et al. The molecular basis for Duchenne versus Becker muscular dystrophy: correlation of severity with type of deletion. Am J Hum Genet 1989; 45:498; PMID:2491009
55. Voit T, Topaloglu H, Straub V, Muntoni F, Deconinck N, Campion G, De Kimpe SJ, Eagle M, Guglieri M, Hood S, et al. Safety and efficacy of drisapersen for the treatment of Duchenne muscular dystrophy (DEMAND II): an exploratory, randomised, placebo-controlled phase 2 study. Lancet Neurol 2014; 13:987-96; PMID:25209738; http://dx.doi.org/10.1016/S1474-4422(14)70195-4
56. A Study of the Safety, Tolerability & Efficacy of Long-term Administration of Drisapersen in US & Canadian Subjects. (2013). Clinical Trial Phase III
57. Extension Study of Drisapersen in DMD Subjects. (2015). Clinical Trial Phase IIIB
58. Mendell JR, Rodino−Klapac LR, Sahenk Z, Roush K, Bird L, Lowes LP, Alfano L, Gomez AM, Lewis S, Kota J, et al. Eteplirsen for the treatment of Duchenne muscular dystrophy. Ann Neurol 2013; 74:637-47; PMID:23907995; http://dx.doi.org/10.1002/ana.23982
59. Confirmatory Study of Eteplirsen in DMD Patients (PROMOVI). (2014). Clinical Trial Phase III
60. Drisapersen FDA Report. (http://www.fda.gov/downloads/AdvisoryCommittees/CommitteesMeetingMaterials/Drugs/PeripheralandCentralNervousSystemDrugsAdvisoryCommittee/UCM473737.pdf). 2015
61. Komaki H, Nagata T, Saito T, Masuda S, Takeshita E, Tachimori H, et al. GP 251-Exon 53 skipping of the dystrophin gene in patients with Duchenne muscular dystrophy by systemic administration of NS-065/NCNP-01: A phase 1, dose escalation, first-in-human study. Neuromuscular Disorders 2015; 25:S261-S2; http://dx.doi.org/10.1016/j.nmd.2015.06.276
62. Nippon Shinyaku Press Release. (https://www.nippon-shinyaku.co.jp/english/news/?id=2920). 2016
63. Koizumi M. 2′-O, 4′-C-Ethylene-bridged nucleic acids (ENA) as next-generation antisense and antigene agents. Biol Pharm Bull 2004; 27:453-6; PMID:15056846; http://dx.doi.org/10.1248/bpb.27.453
64. Daiichi Sankyo Press Release. (http://www.daiichisankyo.com/media_investors/media_relations/press_releases/detail/006412/160225_635_E.pdf). 2016
65. Study of DS-5141b in Patients With Duchenne Muscular Dystrophy. (2016). Clinical Trial Phase I/II
66. Lefebvre S, Bürglen L, Reboullet S, Clermont O, Burlet P, Viollet L, Benichou B, Cruaud C, Millasseau P, Zeviani M, et al. Identification and characterization of a spinal muscular atrophy-determining gene. Cell 1995; 80:155-65; PMID:7813012; http://dx.doi.org/10.1016/0092-8674(95)90460-3
67. Lorson CL, Hahnen E, Androphy EJ, Wirth B. A single nucleotide in the SMN gene regulates splicing and is responsible for spinal muscular atrophy. Proc Natl Acad Sci U S A 1999; 96:6307-11; PMID:103395583; http://dx.doi.org/10.1073/pnas.96.11.6307
68. Hua Y, Vickers TA, Okunola HL, Bennett CF, Krainer AR. Antisense masking of an hnRNP A1/A2 intronic splicing silencer corrects SMN2 splicing in transgenic mice. Am J Hum Genet 2008; 82:834-48; PMID:18371932; http://dx.doi.org/10.1016/j.ajhg.2008.01.014
69. Ionis Pharmaceuticals. (http://ir.ionispharma.com/phoenix.zhtml?c=222170&p=irol-newsArticle&ID=2097778). 2015
70. A Study to Assess the Efficacy and Safety of IONIS-SMN Rx in Infants With Spinal Muscular Atrophy. (2014). Clinical Trial Phase III
71. A Study to Assess the Efficacy and Safety of IONIS-SMN Rx in Patients With Later-onset Spinal Muscular Atrophy. (2014). Clinical Phase Trial III
72. Disterer P, Kryczka A, Liu Y, Badi YE, Wong JJ, Owen JS, Khoo B. Development of therapeutic splice-switching oligonucleotides. Hum Gene Ther 2014; 25:587-98; PMID:24826963; http://dx.doi.org/10.1089/hum.2013.234
73. Keefe AD, Pai S, Ellington A. Aptamers as therapeutics. Nat Rev Drug Discov 2010; 9:537-50; PMID:20592747; http://dx.doi.org/10.1038/nrd3141

74. Bouchard PR, Hutabarat RM, Thompson KM. Discovery and Development of Therapeutic Aptamers. Annu Rev Pharmacol Toxicol 2010; 50:237-57; PMID:20055704; http://dx.doi.org/10.1146/annurev.pharmtox.010909.105547
75. Ruckman J, Green LS, Beeson J, Waugh S, Gillette WL, Henninger DD, Claesson-Welsh L, Janjić N. 2′-Fluoropyrimidine RNA-based aptamers to the 165-amino acid form of vascular endothelial growth factor (VEGF165). Inhibition of receptor binding and VEGF-induced vascular permeability through interactions requiring the exon 7-encoded domain. J Biol Chem 1998; 273:20556-67; PMID:9685413; http://dx.doi.org/10.1074/jbc.273.32.20556
76. Ng EWM, Shima DT, Calias P, Cunningham ET, Guyer DR, Adamis AP. Pegaptanib, a targeted anti-VEGF aptamer for ocular vascular disease. Nat Rev Drug Discov 2006; 5:123-32; PMID:16518379; http://dx.doi.org/10.1038/nrd1955
77. Lincoff AM, Mehran R, Povsic TJ, Zelenkofske SL, Huang Z, Armstrong PW, Steg PG, Bode C, Cohen MG, Buller C, et al. Effect of the REG1 anticoagulation system versus bivalirudin on outcomes after percutaneous coronary intervention (REGULATE-PCI): a randomised clinical trial. Lancet 2016; 387:349-56; PMID:26547100; http://dx.doi.org/10.1016/S0140-6736(15)00515-2
78. A Study To Determine the Efficacy and Safety of REG1 Compared to Bivalirudin in Patients Undergoing PCI (Regulate). (2013). Clinical Trial Phase III (Terminated)
79. Dugel PU. Anti-PDGF combination therapy in neovascular age-related macular degeneration: results of a phase 2b study. Retina Today 2013; 8:65-71
80. A Phase 3 Safety and Efficacy Study of Fovista® (E10030) Intravitreous Administration in Combination With Lucentis® Compared to Lucentis® Monotherapy. (2013). Clinical Trial Phase III
81. A Phase 3 Safety and Efficacy Study of Fovista® (E10030) Intravitreous Administration in Combination With Either Avastin® or Eylea® Compared to Avastin® or Eylea® Monotherapy. (2013). Clinical Trial Phase III
82. Monés J. Complement factor 5 inhibition in age-related macular degeneration. Retina Today 2010; 5:52-5
83. A Phase 2/3 Trial to Assess the Safety and Efficacy of Intravitreous Administration of Zimura® (Anti-C5 Aptamer) in Subjects With Geographic Atrophy Secondary to Dry Age-Related Macular Degeneration. (2016). Clinical Trial Phase II/III
84. Vater A, Klussmann S. Turning mirror-image oligonucleotides into drugs: the evolution of Spiegelmer® therapeutics. Drug Discovery Today 2015; 20:147-55; PMID:25236655; http://dx.doi.org/10.1016/j.drudis.2014.09.004
85. Ashley GW. Modeling, synthesis, and hybridization properties of (L)-ribonucleic acid. J Am Chem Soc 1992; 114:9731-6; http://dx.doi.org/10.1021/ja00051a001
86. van Eijk LT, John AS, Schwoebel F, Summo L, Vauléon S, Zöllner S, Laarakkers CM, Kox M, van der Hoeven JG, Swinkels DW, et al. Effect of the antihepcidin Spiegelmer lexaptepid on inflammation-induced decrease in serum iron in humans. Blood 2014; 124:2643-6; PMID:25163699; http://dx.doi.org/10.1182/blood-2014-03-559484
87. Roccaro AM, Sacco A, Purschke WG, Moschetta M, Buchner K, Maasch C, Zboralski D, Zöllner S, Vonhoff S, Mishima Y, et al. SDF-1 inhibition targets the bone marrow niche for cancer therapy. Cell Rep 2014; 9:118-28; PMID:25263552; http://dx.doi.org/10.1016/j.celrep.2014.08.042
88. NOX-A12 in Combination With Bortezomib and Dexamethasone in Relapsed Multiple Myeloma. (2012). Clinical Trial Phase IIA
89. NOX-A12 in Combination With Bendamustine and Rituximab in Relapsed Chronic Lymphocytic Leukemia (CLL). (2011). Clinical Trial Phase IIA
90. Oberthür D, Achenbach J, Gabdulkhakov A, Buchner K, Maasch C, Falke S, Rehders D, Klussmann S, Betzel C. Crystal structure of a mirror-image L-RNA aptamer (Spiegelmer) in complex with the natural L-protein target CCL2. Nat Commun 2015; 6:6923; PMID:25901662; http://dx.doi.org/10.1038/ncomms7923
91. NOX-E36 in Patients With Type 2 Diabetes Mellitus and Albuminuria. (2012). Clinical Trial Phase IIA
92. Efficacy of NOX-H94 on Anemia of Chronic Disease in Patients With Cancer. (2012). Clinical Trial Phase IIA
93. MacRae IJ, Ma E, Zhou M, Robinson CV, Doudna JA. In vitro reconstitution of the human RISC-loading complex. Proc Natl Acad Sci U S A 2008; 105:512-7; PMID:18178619; http://dx.doi.org/10.1073/pnas.0710869105
94. Valinezhad Orang A, Safaralizadeh R, Kazemzadeh-Bavili M. Mechanisms of miRNA-Mediated Gene Regulation from Common Downregulation to mRNA-Specific Upregulation. Int J Genomics 2014; 2014:970607; PMID:25180174
95. Elbashir SM, Harborth J, Lendeckel W, Yalcin A, Weber K, Tuschl T. Duplexes of 21-nucleotide RNAs mediate RNA interference in cultured mammalian cells. Nature 2001; 411:494-8; PMID:11373684; http://dx.doi.org/10.1038/35078107
96. Mellitzer G, Hallonet M, Chen L, Ang SL. Spatial and temporal 'knock down' of gene expression by electroporation of double-stranded RNA and morpholinos into early postimplantation mouse embryos. Mech Dev 2002; 118:57-63; PMID:12351170; http://dx.doi.org/10.1016/S0925-4773(02)00191-0
97. Calegari F, Haubensak W, Yang D, Huttner WB, Buchholz F. Tissue-specific RNA interference in postimplantation mouse embryos with endoribonuclease-prepared short interfering RNA. Proc Natl Acad Sci U S A 2002; 99:14236-40; PMID:12391321; http://dx.doi.org/10.1073/pnas.192559699
98. Zeliadt N. Big pharma shows signs of renewed interest in RNAi drugs. Nat Med 2014; 20:109; PMID:24504395; http://dx.doi.org/10.1038/nm0214-109
99. Open Label Study for the Evaluation of Tolerability of Five Dose Levels of Cand5. (2004). Clinical Trial Phase I
100. A Dose Escalation and Safety Study of I5NP to Prevent Acute Kidney Injury (AKI) in Patients at High Risk of AKI Undergoing Major Cardiovascular Surgery (QRK.004). (2008). Clinical Trial Phase I (Terminated)
101. Safety & Efficacy Study Evaluating the Combination of Bevasiranib & Lucentis Therapy in Wet AMD (COBALT). (2007). Clinical Trial Phase III (Terminated)
102. Study Evaluating Efficacy and Safety of PF-04523655 Versus Laser in Subjects With Diabetic Macular Edema (DEGAS). (2008). Clinical Trial Phase II (Terminated)
103. Study to Evaluate the Safety, Tolerability, Pharmacokinetics (PK), and Pharmacodynamics (PD) of Liposomal siRNA in Subjects With High Cholesterol. (2009). Clinical Trial Phase I (Terminated)
104. A Study Using Intravitreal Injections of a Small Interfering RNA in Patients With Age-Related Macular Degeneration. (2006). Clinical Trial Phase II (Terminated)
105. Conde J, Artzi N. Are RNAi and miRNA therapeutics truly dead? Trends Biotechnol 2015; 33:141-4; PMID:25595555; http://dx.doi.org/10.1016/j.tibtech.2014.12.005
106. Krieg AM. Is RNAi dead? Mol Ther 2011; 19:1001-2; PMID:21629254; http://dx.doi.org/10.1038/mt.2011.94
107. Coelho T, Adams D, Silva A, Lozeron P, Hawkins PN, Mant T, Perez J, Chiesa J, Warrington S, Tranter E, et al. Safety and efficacy of RNAi therapy for transthyretin amyloidosis. N Engl J Med 2013; 369:819-29; PMID:23984729; http://dx.doi.org/10.1056/NEJMoa1208760
108. Fitzgerald K, Frank-Kamenetsky M, Shulga-Morskaya S, Liebow A, Bettencourt BR, Sutherland JE, Hutabarat RM, Clausen VA, Karsten V, Cehelsky J, et al. Effect of an RNA interference drug on the synthesis of proprotein convertase subtilisin/kexin type 9 (PCSK9) and the concentration of serum LDL cholesterol in healthy volunteers: a randomised, single-blind, placebo-controlled, phase 1 trial. Lancet 2014; 383:60-8; PMID:24094767; http://dx.doi.org/10.1016/S0140-6736(13)61914-5
109. van de Water FM, Boerman OC, Wouterse AC, Peters JG, Russel FG, Masereeuw R. Intravenously administered short interfering RNA accumulates in the kidney and selectively suppresses gene function

in renal proximal tubules. Drug Metab Dispos 2006; 34:1393-7; PMID:16714375; http://dx.doi.org/10.1124/dmd.106.009555
110. Alagia A, Eritja R. siRNA and RNAi optimization. Wiley interdisciplinary reviews RNA, 2016 7; 316-329; PMID:26840434; http://dx.doi.org/10.1002/wrna.1337
111. Lorenzer C, Dirin M, Winkler AM, Baumann V, Winkler J. Going beyond the liver: progress and challenges of targeted delivery of siRNA therapeutics. J Control Release 2015; 203:1-15; PMID:25660205; http://dx.doi.org/10.1016/j.jconrel.2015.02.003
112. Jeong EH, Kim H, Jang B, Cho H, Ryu J, Kim B, Park Y, Kim J, Lee JB, Lee H. Technological development of structural DNA/RNA-based RNAi systems and their applications. Adv Drug Deliv Rev 2016; 107:29-43; PMID:26494399; http://dx.doi.org/10.1016/j.addr.2015.10.008
113. Bramsen JB, Kjems J. Engineering small interfering RNAs by strategic chemical modification. Methods Mol Biol 2013; 942:87-109; PMID:23027047; http://dx.doi.org/10.1007/978-1-62703-119-6_5
114. Phase 1b/2, Open Label, Repeat Dose, Dose Escalation Study of ND-L02-s0201 Injection in Subjects With Moderate to Extensive Fibrosis (METAVIR F3-4). (2014). Clinical Trial Phase I
115. Study of ARB-001467 in Subjects With Chronic HBV Infection Receiving Nucleos(t)Ide Analogue Therapy. (2015). Clinical Trial Phase II
116. A Phase 1 Study of ALN-AS1 in Patients With Acute Intermittent Porphyria (AIP). (2015). Clinical Trial Phase I
117. ENDEAVOUR: Phase 3 Multicenter Study of Revusiran (ALN-TTRSC) in Patients With Transthyretin (TTR) Mediated Familial Amyloidotic Cardiomyopathy (FAC). (2014). Clinical Trial Phase III
118. A Phase 1 Study of an Investigational Drug, ALN-AT3SC, in Healthy Volunteers and Hemophilia A or B Patients. (2014). Clinical Trial Phase I
119. Phase I Intratumoral Pbi-shRNA STMN1 LP in Advanced and/or Metastatic Cancer (STMN1-LP). (2012). Clinical Trial Phase I
120. APN401 in Treating Patients With Melanoma, Kidney Cancer, Pancreatic Cancer, or Other Solid Tumors That Are Metastatic or Cannot Be Removed By Surgery. (2014). Clinical Trial Phase I
121. SYL040012, Treatment for Open Angle Glaucoma (SYLTAG). (2014). Clinical Trial Phase II
122. Phase 2/3, Randomized, Double-Masked, Sham-Controlled Trial of QPI-1007 in Subjects With Acute Nonarteritic Anterior Ischemic Optic Neuropathy (NAION). (2015). Clinical Trial Phase II/III
123. Bobbin ML, Rossi JJ. RNA Interference (RNAi)-Based Therapeutics: Delivering on the Promise? Annu Rev Pharmacol Toxicol 2016; 56:103-22; PMID:26738473; http://dx.doi.org/10.1146/annurev-pharmtox-010715-103633
124. Lam JK, Chow MY, Zhang Y, Leung SW. siRNA Versus miRNA as Therapeutics for Gene Silencing. Mol Ther Nucleic Acids 2015; 4: e252; PMID:26372022; http://dx.doi.org/10.1038/mtna.2015.23
125. Long-term Extension to Miravirsen Study in Null Responder to Pegylated Interferon Alpha Plus Ribavirin Subjects With Chronic Hepatitis C. (2014). Clinical Trial Phase II
126. Lee CH, Kim JH, Lee SW. The role of microRNAs in hepatitis C virus replication and related liver diseases. J Microbiol 2014; 52:445-51; PMID:24871972; http://dx.doi.org/10.1007/s12275-014-4267-x
127. MesomiR 1: A Phase I Study of TargomiRs as 2nd or 3rd Line Treatment for Patients With Recurrent MPM and NSCLC. (2015). Clinical Trial Phase I
128. A Multicenter Phase I Study of MRX34, MicroRNA miR-RX34 Liposomal Injection. (2013). Clinical Trial Phase I
129. Wolff JA, Malone RW, Williams P, Chong W, Acsadi G, Jani A, Felgner PL. Direct gene transfer into mouse muscle in vivo. Science 1990; 247:1465-8; PMID:1690918; http://dx.doi.org/10.1126/science.1690918
130. Jirikowski GF, Sanna PP, Maciejewski-Lenoir D, Bloom FE. Reversal of diabetes insipidus in Brattleboro rats: intrahypothalamic injection of vasopressin mRNA. Science 1992; 255:996-8; PMID:1546298; http://dx.doi.org/10.1126/science.1546298
131. Isaacs A, Cox RA, Rotem Z. Foreign nucleic acids as the stimulus to make interferon. Lancet 1963; 2:113-6; PMID:13956740; http://dx.doi.org/10.1016/S0140-6736(63)92585-6
132. Weissman D. mRNA transcript therapy. Expert Rev Vaccines 2015; 14:265-81; PMID:25359562; http://dx.doi.org/10.1586/14760584.2015.973859
133. Alexopoulou L, Holt AC, Medzhitov R, Flavell RA. Recognition of double-stranded RNA and activation of NF-kappaB by Toll-like receptor 3. Nature 2001; 413:732-8; PMID:11607032; http://dx.doi.org/10.1038/35099560
134. Anderson BR, Muramatsu H, Nallagatla SR, Bevilacqua PC, Sansing LH, Weissman D, Karikó K. Incorporation of pseudouridine into mRNA enhances translation by diminishing PKR activation. Nucleic Acids Res 2010; 38:5884-92; PMID:20457754; http://dx.doi.org/10.1093/nar/gkq347
135. Kariko K, Muramatsu H, Welsh FA, Ludwig J, Kato H, Akira S, Weissman D. Incorporation of pseudouridine into mRNA yields superior nonimmunogenic vector with increased translational capacity and biological stability. Mol Ther 2008; 16:1833-40; PMID:18797453; http://dx.doi.org/10.1038/mt.2008.200
136. Goubau D, Deddouche S, Reis e Sousa C. Cytosolic sensing of viruses. Immunity 2013; 38:855-69; PMID:23706667; http://dx.doi.org/10.1016/j.immuni.2013.05.007
137. Kariko K, Buckstein M, Ni H, Weissman D. Suppression of RNA recognition by Toll-like receptors: the impact of nucleoside modification and the evolutionary origin of RNA. Immunity 2005; 23:165-75; PMID:16111635; http://dx.doi.org/10.1016/j.immuni.2005.06.008
138. Kormann MS, Hasenpusch G, Aneja MK, Nica G, Flemmer AW, Herber-Jonat S, Huppmann M, Mays LE, Illenyi M, Schams A, et al. Expression of therapeutic proteins after delivery of chemically modified mRNA in mice. Nat Biotechnol 2011; 29:154-7; PMID:21217696; http://dx.doi.org/10.1038/nbt.1733
139. Karijolich J, Yu YT. Converting nonsense codons into sense codons by targeted pseudouridylation. Nature 2011; 474:395-8; PMID:21677757; http://dx.doi.org/10.1038/nature10165
140. Summer H, Gramer R, Droge P. Denaturing urea polyacrylamide gel electrophoresis (Urea PAGE). J Vis Exp 2009; PMID:19865070; http://dx.doi.org/10.3791/1485 (2009)
141. Kariko K, Muramatsu H, Ludwig J, Weissman D. Generating the optimal mRNA for therapy: HPLC purification eliminates immune activation and improves translation of nucleoside-modified, protein-encoding mRNA. Nucleic Acids Res 2011; 39:e142; PMID:21890902; http://dx.doi.org/10.1093/nar/gkr695
142. Pasquinelli AE, Dahlberg JE, Lund E. Reverse 5' caps in RNAs made in vitro by phage RNA polymerases. RNA 1995; 1:957-67; PMID:8548660
143. Stepinski J, Waddell C, Stolarski R, Darzynkiewicz E, Rhoads RE. Synthesis and properties of mRNAs containing the novel "anti-reverse" cap analogs 7-methyl(3'-O-methyl)GpppG and 7-methyl (3'-deoxy)GpppG. RNA 2001; 7:1486-95; PMID:11680853; http://dx.doi.org/10.1017.S1355838201014078
144. Grudzien-Nogalska E, Jemielity J, Kowalska J, Darzynkiewicz E, Rhoads RE. Phosphorothioate cap analogs stabilize mRNA and increase translational efficiency in mammalian cells. RNA 2007; 13:1745-55; PMID:17720878; http://dx.doi.org/10.1261/rna.701307
145. Fath S, Bauer AP, Liss M, Spriestersbach A, Maertens B, Hahn P, Ludwig C, Schäfer F, Graf M, Wagner R. Multiparameter RNA and codon optimization: a standardized tool to assess and enhance autologous mammalian gene expression. PLoS one 2011; 6:e17596; PMID:21408612; http://dx.doi.org/10.1371/journal.pone.0017596
146. Al-Saif M, Khabar KS. UU/UA dinucleotide frequency reduction in coding regions results in increased mRNA stability and protein expression. Mol Ther 2012; 20:954-9; PMID:22434136; http://dx.doi.org/10.1038/mt.2012.29
147. Weissman D, Kariko K. mRNA: Fulfilling the Promise of Gene Therapy. Mol Ther 2015; 23:1416-7; PMID:26321183; http://dx.doi.org/10.1038/mt.2015.138
148. Zubiaga AM, Belasco JG, Greenberg ME. The nonamer UUAUUUAUU is the key AU-rich sequence motif that mediates mRNA degradation. Mol Cell Biol 1995; 15:2219-30; PMID:7891716; http://dx.doi.org/10.1128/MCB.15.4.2219

149. Peixeiro I, Silva AL, Romao L. Control of human β-globin mRNA stability and its impact on β-thalassemia phenotype. Haematologica 2011; 96:905-13; PMID:21357703; http://dx.doi.org/10.3324/haematol.2010.039206
150. Waggoner SA, Liebhaber SA. Regulation of α-globin mRNA stability. Exp Biol Med (Maywood) 2003; 228:387-95; PMID:12671183
151. Withers JB, Beemon KL. The structure and function of the rous sarcoma virus RNA stability element. J Cell Biochem 2011; 112:3085-92; PMID:21769913; http://dx.doi.org/10.1002/jcb.23272
152. Gallie DR. The cap and poly(A) tail function synergistically to regulate mRNA translational efficiency. Genes Dev 1991; 5:2108-16; PMID:1682219; http://dx.doi.org/10.1101/gad.5.11.2108
153. Peng J, Murray EL, Schoenberg DR. In vivo and in vitro analysis of poly(A) length effects on mRNA translation. Methods Mol Biol 2008; 419:215-30; PMID:18369986; http://dx.doi.org/10.1007/978-1-59745-033-1_15
154. Balmayor ER, Geiger JP, Aneja MK, Berezhanskyy T, Utzinger M, Mykhaylyk O, Rudolph C, Plank C. Chemically modified RNA induces osteogenesis of stem cells and human tissue explants as well as accelerates bone healing in rats. Biomaterials 2016; 87:131-46; PMID:26923361; http://dx.doi.org/10.1016/j.biomaterials.2016.02.018
155. Elangovan S, Khorsand B, Do AV, Hong L, Dewerth A, Kormann M, Ross RD, Sumner DR, Allamargot C, Salem AK. Chemically modified RNA activated matrices enhance bone regeneration. J Control Release 2015; 218:22-8; PMID:26415855; http://dx.doi.org/10.1016/j.jconrel.2015.09.050
156. Plews JR, Li J, Jones M, Moore HD, Mason C, Andrews PW, Na J. Activation of pluripotency genes in human fibroblast cells by a novel mRNA based approach. PloS one 2010; 5:e14397; PMID:21209933; http://dx.doi.org/10.1371/journal.pone.0014397
157. Warren L, Manos PD, Ahfeldt T, Loh YH, Li H, Lau F, Ebina W, Mandal PK, Smith ZD, Meissner A, et al. Highly efficient reprogramming to pluripotency and directed differentiation of human cells with synthetic modified mRNA. Cell Stem Cell 2010; 7:618-30; PMID:20888316; http://dx.doi.org/10.1016/j.stem.2010.08.012
158. Mandal PK, Rossi DJ. Reprogramming human fibroblasts to pluripotency using modified mRNA. Nat Protoc 2013; 8:568-82; PMID:23429718; http://dx.doi.org/10.1038/nprot.2013.019
159. Preskey D, Allison TF, Jones M, Mamchaoui K, Unger C. Synthetically modified mRNA for efficient and fast human iPS cell generation and direct transdifferentiation to myoblasts. Biochem Biophys Res Commun 2015; PMID:26449459; http://dx.doi.org/10.1016/j.bbrc.2015.09.102
160. Simeonov KP, Uppal H. Direct reprogramming of human fibroblasts to hepatocyte-like cells by synthetic modified mRNAs. PloS One 2014; 9:e100134; PMID:24963715; http://dx.doi.org/10.1371/journal.pone.0100134
161. Kariko K, Muramatsu H, Keller JM, Weissman D. Increased erythropoiesis in mice injected with submicrogram quantities of pseudouridine-containing mRNA encoding erythropoietin. Mol Ther 2012; 20:948-53; PMID:22334017; http://dx.doi.org/10.1038/mt.2012.7
162. Yin H, Xue W, Chen S, Bogorad RL, Benedetti E, Grompe M, Koteliansky V, Sharp PA, Jacks T, Anderson DG. Genome editing with Cas9 in adult mice corrects a disease mutation and phenotype. Nat Biotechnol 2014; 32:551-3; PMID:24681508; http://dx.doi.org/10.1038/nbt.2884
163. Wang X, Yu H, Lei A, Zhou J, Zeng W, Zhu H, Niu Y, Shi B, Cai B, Liu J, et al. Generation of gene-modified goats targeting MSTN and FGF5 via zygote injection of CRISPR/Cas9 system. Sci Rep 2015; 5:13878; PMID:26354037; http://dx.doi.org/10.1038/srep13878
164. Kaneko T, Mashimo T. Simple Genome Editing of Rodent Intact Embryos by Electroporation. PloS one 2015; 10:e0142755; PMID:26556280; http://dx.doi.org/10.1371/journal.pone.0142755
165. Wang T, Hong Y. Direct gene disruption by TALENs in medaka embryos. Gene 2014; 543:28-33; PMID:24713411; http://dx.doi.org/10.1016/j.gene.2014.04.013
166. Yang D, Zhang J, Xu J, Zhu T, Fan Y, Fan J, Chen YE. Production of apolipoprotein C-III knockout rabbits using zinc finger nucleases. J Vis Exp 2013; 81; e50957; PMID:24301055; http://dx.doi.org/10.3791/50957
167. Flisikowska T, Thorey IS, Offner S, Ros F, Lifke V, Zeitler B, Rottmann O, Vincent A, Zhang L, Jenkins S, et al. Efficient immunoglobulin gene disruption and targeted replacement in rabbit using zinc finger nucleases. PloS One 2011; 6:e21045; PMID:21695153; http://dx.doi.org/10.1371/journal.pone.0021045
168. Williams DA, Thrasher AJ. Concise review: lessons learned from clinical trials of gene therapy in monogenic immunodeficiency diseases. Stem Cells Transl Med 2014; 3:636-42; PMID:24682287; http://dx.doi.org/10.5966/sctm.2013-0206
169. Schirmbeck R, Reimann J, Kochanek S, Kreppel F. The immunogenicity of adenovirus vectors limits the multispecificity of CD8 T-cell responses to vector-encoded transgenic antigens. Mol Ther 2008; 16:1609-16; PMID:18612271; http://dx.doi.org/10.1038/mt.2008.141
170. Mahiny AJ, Dewerth A, Mays LE, Alkhaled M, Mothes B, Malaeksefat E, Loretz B, Rottenberger J, Brosch DM, Reautschnig P, et al. In vivo genome editing using nuclease-encoding mRNA corrects SP-B deficiency. Nat Biotechnol 2015; 33:584-6; PMID:25985262; http://dx.doi.or/10.1038/nbt.3241
171. Allers K, Hutter G, Hofmann J, Loddenkemper C, Rieger K, Thiel E, Schneider T. Evidence for the cure of HIV infection by CCR5Delta32/Delta32 stem cell transplantation. Blood 2011; 117:2791-9; PMID:21148083; http://dx.doi.org/10.1182/blood-2010-09-309591
172. Evaluation the Safety and Tolerability of i.v. Administration of a Cancer Vaccine in Patients With Advanced Melanoma (Lipo-MERIT). (2015). Clinical Trial Phase I
173. Wissink EM, Fogarty EA, Grimson A. High-throughput discovery of post-transcriptional cis-regulatory elements. BMC Genomics 2016; 17:177; PMID:26941072; http://dx.doi.org/10.1186/s12864-016-2479-7
174. Desmet CJ, Ishii KJ. Nucleic acid sensing at the interface between innate and adaptive immunity in vaccination. Nat Rev Immunol 2012; 12:479-91; PMID:22728526; http://dx.doi.org/10.1038/nri3247
175. Brubaker SW, Bonham KS, Zanoni I, Kagan JC. Innate immune pattern recognition: a cell biological perspective. Annu Rev Immunol 2015; 33:257-90; PMID:25581309; http://dx.doi.org/10.1146/annurev-immunol-032414-112240
176. Clark GJ, Angel N, Kato M, Lopez JA, MacDonald K, Vuckovic S, Hart DN. The role of dendritic cells in the innate immune system. Microbes Infect 2000; 2:257-72; PMID:10758402; http://dx.doi.org/10.1016/S1286-4579(00)00302-6
177. Jacobson JM, Routy JP, Welles S, DeBenedette M, Tcherepanova I, Angel JB, Asmuth DM, Stein DK, Baril JG, McKellar M, et al. Dendritic Cell Immunotherapy for HIV-1 Infection Using Autologous HIV-1 RNA: A Randomized, Double-Blind, Placebo-Controlled Clinical Trial. J Acquir Immune Defic Syndr 2016; 72:31-8; PMID:26751016; http://dx.doi.org/10.1097/QAI.0000000000000926
178. Weiss R, Scheiblhofer S, Thalhamer J. Allergens are not pathogens: why immunization against allergy differs from vaccination against infectious diseases. Hum Vaccin Immunother 2014; 10:703-7; PMID:24280693; http://dx.doi.org/10.4161/hv.27183
179. Weiss R, Scheiblhofer S, Roesler E, Ferreira F, Thalhamer J. Prophylactic mRNA vaccination against allergy. Curr Opin Allergy Clin Immunol 2010; 10:567-74; PMID:20856111; http://dx.doi.org/10.1097/ACI.0b013e32833fd5b6
180. Fotin-Mleczek M, Duchardt KM, Lorenz C, Pfeiffer R, Ojkic-Zrna S, Probst J, Kallen KJ. Messenger RNA-based vaccines with dual activity induce balanced TLR-7 dependent adaptive immune responses and provide antitumor activity. J Immunother 2011; 34:1-15; PMID:21150709; http://dx.doi.org/10.1097/CJI.0b013e3181f7dbe8
181. Fotin-Mleczek M, Zanzinger K, Heidenreich R, Lorenz C, Kowalczyk A, Kallen KJ, et al. mRNA-based vaccines synergize with radiation therapy to eradicate established tumors. Radiat Oncol 2014; 9:180; PMID:25127546; http://dx.doi.org/10.1186/1748-717X-9-180
182. Schmidt MA, Goodwin TJ. Personalized medicine in human space flight: using Omics based analyses to develop individualized countermeasures that enhance astronaut safety and performance. Metabolomics 2013; 9:1134-56; PMID:24273472; http://dx.doi.org/10.1007/s11306-013-0556-3

183. Petsch B, Schnee M, Vogel AB, Lange E, Hoffmann B, Voss D, Schlake T, Thess A, Kallen KJ, Stitz L, et al. Protective efficacy of in vitro synthesized, specific mRNA vaccines against influenza A virus infection. Nat Biotechnol 2012; 30:1210-6; PMID:23159882; http://dx.doi.org/10.1038/nbt.2436
184. Pascolo S. Vaccination with messenger RNA. Methods Mol Med 2006; 127:23-40; PMID:16988444; htpp://dx.doi.org/10.1385/1-59745-168-1:23
185. Mortimer I, Tam P, MacLachlan I, Graham RW, Saravolac EG, Joshi PB. Cationic lipid-mediated transfection of cells in culture requires mitotic activity. Gene Ther 1999; 6:403-11; PMID:10435090; http://dx.doi.org/10.1038/sj.gt.3300837
186. Wong SS, Webby RJ. An mRNA vaccine for influenza. Nat Biotechnol 2012; 30:1202-4; PMID:23222788; http://dx.doi.org/10.1038/nbt.2439
187. Ichino M, Mor G, Conover J, Weiss WR, Takeno M, Ishii KJ, Klinman DM. Factors associated with the development of neonatal tolerance after the administration of a plasmid DNA vaccine. J Immunol 1999; 162:3814-8; PMID:10201898
188. Ponsaerts P, Van Tendeloo VF, Berneman ZN. Cancer immunotherapy using RNA-loaded dendritic cells. Clin Exp Immunol 2003; 134:378-84; PMID:14632740; http://dx.doi.org/10.1046/j.1365-2249.2003.02286.x
189. Vik-Mo EO, Nyakas M, Mikkelsen BV, Moe MC, Due-Tonnesen P, Suso EM, Sæbøe-Larssen S, Sandberg C, Brinchmann JE, Helseth E, et al. Therapeutic vaccination against autologous cancer stem cells with mRNA-transfected dendritic cells in patients with glioblastoma. Cancer Immunol Immunother 2013; 62:1499-509; PMID:23817721; http://dx.doi.org/10.1007/s00262-013-1453-3
190. Van Gulck E, Vlieghe E, Vekemans M, Van Tendeloo VF, Van De Velde A, Smits E, Anguille S, Cools N, Goossens H, Mertens L, et al. mRNA-based dendritic cell vaccination induces potent antiviral T-cell responses in HIV-1-infected patients. AIDS 2012; 26:F1-12; PMID:22156965; http://dx.doi.org/10.1097/QAD.0b013e32834f33e8
191. Aarntzen EH, Schreibelt G, Bol K, Lesterhuis WJ, Croockewit AJ, de Wilt JH, van Rossum MM, Blokx WA, Jacobs JF, Duiveman-de Boer T, et al. Vaccination with mRNA-electroporated dendritic cells induces robust tumor antigen-specific CD4+ and CD8+ T cells responses in stage III and IV melanoma patients. Clin Cancer Res 2012; 18:5460-70; PMID:22896657; http://dx.doi.org/10.1158/1078-0432.CCR-11-3368
192. Suso EM, Dueland S, Rasmussen AM, Vetrhus T, Aamdal S, Kvalheim G, Gaudernack G. hTERT mRNA dendritic cell vaccination: complete response in a pancreatic cancer patient associated with response against several hTERT epitopes. Cancer Immunol Immunother 2011; 60:809-18; PMID:21365467; http://dx.doi.org/10.1007/s00262-011-0991-9
193. Van Tendeloo VF, Van de Velde A, Van Driessche A, Cools N, Anguille S, Ladell K, Gostick E, Vermeulen K, Pieters K, Nijs G, et al. Induction of complete and molecular remissions in acute myeloid leukemia by Wilms' tumor 1 antigen-targeted dendritic cell vaccination. Proc Natl Acad Sci U S A 2010; 107:13824-9; PMID:20631300; http://dx.doi.org/10.1073/pnas.1008051107
194. Routy JP, Boulassel MR, Yassine-Diab B, Nicolette C, Healey D, Jain R, Landry C, Yegorov O, Tcherepanova I, Monesmith T, et al. Immunologic activity and safety of autologous HIV RNA-electroporated dendritic cells in HIV-1 infected patients receiving antiretroviral therapy. Clin Immunol 2010; 134:140-7; PMID:19889582; http://dx.doi.org/10.1016/j.clim.2009.09.009
195. Coosemans A, Wolfl M, Berneman ZN, Van Tendeloo V, Vergote I, Amant F, Van Gool SW. Immunological response after therapeutic vaccination with WT1 mRNA-loaded dendritic cells in end-stage endometrial carcinoma. Anticancer Res 2010; 30:3709-14; PMID:20944158
196. Phase 3 Trial of Autologous Dendritic Cell Immunotherapy (AGS-003) Plus Standard Treatment of Advanced Renal Cell Carcinoma (RCC) (ADAPT). (2012). Clinical Trial Phase III
197. Wilgenhof S, Corthals J, Heirman C, van Baren N, Lucas S, Kvistborg P, Thielemans K, Neyns B. Phase II Study of Autologous Monocyte-Derived mRNA Electroporated Dendritic Cells (TriMixDC-MEL) Plus Ipilimumab in Patients With Pretreated Advanced Melanoma. J Clin Oncol 2016; 34:1330-8; PMID:26926680; http://dx.doi.org/10.1200/JCO.2015.63.4121
198. Bol KF, Figdor CG, Aarntzen EH, Welzen ME, van Rossum MM, Blokx WA, van de Rakt MW, Scharenborg NM, de Boer AJ, Pots JM, et al. Intranodal vaccination with mRNA-optimized dendritic cells in metastatic melanoma patients. Oncoimmunology 2015; 4:e1019197; PMID:26405571; http://dx.doi.org/10.1080/2162402X.2015.1019197
199. MiHA-loaded PD-L-silenced DC Vaccination After Allogeneic SCT (PSCT19). (2016). Clinical Trial Phase II
200. Adjuvant Dendritic Cell-immunotherapy Plus Temozolomide in Glioblastoma Patients (ADDIT-GLIO). (2016). Clinical Trial Phase II
201. Autologous Dendritic Cell Vaccination in Mesothelioma (MESO-DEC). (2016). Clinical Trial Phase II
202. Vaccine Therapy for the Treatment of Newly Diagnosed Glioblastoma Multiforme (ATTAC-II). (2016). Clinical Trial Phase II
203. Weide B, Pascolo S, Scheel B, Derhovanessian E, Pflugfelder A, Eigentler TK, Pawelec G, Hoerr I, Rammensee HG, Garbe C. Direct injection of protamine-protected mRNA: results of a phase 1/2 vaccination trial in metastatic melanoma patients. J Immunother 2009; 32:498-507; PMID:19609242; http://dx.doi.org/10.1097/CJI.0b013e3181a00068
204. RNActive® Rabies Vaccine (CV7201) in Healthy Adults. (2014). Clinical Trial Phase I
205. Hekele A, Bertholet S, Archer J, Gibson DG, Palladino G, Brito LA, Otten GR, Brazzoli M, Buccato S, Bonci A, et al. Rapidly produced SAM((R)) vaccine against H7N9 influenza is immunogenic in mice. Emerg Microbes Infect 2013; 2:e52; PMID:26038486; http://dx.doi.org/10.1038/emi.2013.54
206. Trial of RNActive®-Derived Cancer Vaccine and Local Radiation in in Stage IV Non Small Cell Lung Cancer (NSCLC). (2013). Clinical Trial Phase I
207. Safety and Efficacy Trial of a RNActive®-Derived Prostate Cancer Vaccine in Hormone Refractory Disease. (2009). Clinical Trial Phase II
208. Trial of RNActive®-Derived Prostate Cancer Vaccine in Metastatic Castrate-refractory Prostate Cancer. (2013). Clinical Trial Phase II
209. An Open Label Randomised Trial of RNActive® Cancer Vaccine in High Risk and Intermediate Risk Patients With Prostate Cancer. (2014). Clinical Trial Phase II
210. IVAC MUTANOME Phase I Clinical Trial. (2014). Clinical Trial Phase I
211. Heidenreich R, Jasny E, Kowalczyk A, Lutz J, Probst J, Baumhof P, Scheel B, Voss S, Kallen KJ, Fotin-Mleczek M. A novel RNA-based adjuvant combines strong immunostimulatory capacities with a favorable safety profile. Int J Cancer 2015; 137:372-84; PMID:25530186; http://dx.doi.org/10.1002/ijc.29402
212. CureVac Press Release. (http://www.curevac.com/fileadmin/curevac.de/media/Content/Newsroom/20151221_CureVac_Press_release_JP_Morgan.pdf). 2015
213. Islam MA, Reesor EK, Xu Y, Zope HR, Zetter BR, Shi J. Biomaterials for mRNA delivery. Biomater Sci 2015; 3:1519-33; PMID:26280625; http://dx.doi.org/10.1039/C5BM00198F
214. Geall AJ, Verma A, Otten GR, Shaw CA, Hekele A, Banerjee K, Cu Y, Beard CW, Brito LA, Krucker T, et al. Nonviral delivery of self-amplifying RNA vaccines. Proc Natl Acad Sci U S A 2012; 109:14604-9; PMID:22908294; http://dx.doi.org/10.1073/pnas.1209367109
215. Rodriguez-Gascon A, del Pozo-Rodriguez A, Solinis MA. Development of nucleic acid vaccines: use of self-amplifying RNA in lipid nanoparticles. Int J Nanomedicine 2014; 9:1833-43; PMID:24748793; http://dx.doi.org/10.2147/IJN.S39810
216. Xu J, Luft JC, Yi X, Tian S, Owens G, Wang J, Johnson A, Berglund P, Smith J, Napier ME, et al. RNA replicon delivery via lipid-complexed PRINT protein particles. Mol Pharm 2013; 10:3366-74; PMID:23924216; http://dx.doi.org/10.1021/mp400190z
217. Pascolo S. The messenger's great message for vaccination. Expert Rev Vaccines 2015; 14:153-6; PMID:25586101; http://dx.doi.org/10.1586/14760584.2015.1000871
218. Hattinger E, Scheiblhofer S, Roesler E, Thalhamer T, Thalhamer J, Weiss R. Prophylactic mRNA Vaccination against Allergy Confers Long-Term Memory Responses and Persistent Protection in Mice. J

Immunol Res 2015; 2015:797421; PMID:26557723; http://dx.doi.org/10.1155/2015/797421
219. Nelson HS, Oppenheimer J, Vatsia GA, Buchmeier A. A double-blind, placebo-controlled evaluation of sublingual immunotherapy with standardized cat extract. J Allergy Clin Immunol 1993; 92:229-36; PMID:8349933; http://dx.doi.org/10.1016/0091-6749(93)90166-D
220. Roesler E, Weiss R, Weinberger EE, Fruehwirth A, Stoecklinger A, Mostbock S, Ferreira F, Thalhamer J, Scheiblhofer S. Immunize and disappear-safety-optimized mRNA vaccination with a panel of 29 allergens. J Allergy Clin Immunol 2009; 124:1070-7 e1-11; PMID:19665781; http://dx.doi.org/10.1016/j.jaci.2009.06.036
221. Ball T, Sperr WR, Valent P, Lidholm J, Spitzauer S, Ebner C, Kraft D, Valenta R. Induction of antibody responses to new B cell epitopes indicates vaccination character of allergen immunotherapy. Eur J Immunol 1999; 29:2026-36; PMID:10382766; http://dx.doi.org/10.1002/(SICI)1521-4141(199906)29:06%3c2026::AID-IMMU2026%3e3.0.CO;2-2
222. Van Ree R, Van Leeuwen WA, Dieges PH, Van Wijk RG, De Jong N, Brewczyski PZ, Kroon AM, Schilte PP, Tan KY, Simon-Licht IF, et al. Measurement of IgE antibodies against purified grass pollen allergens (Lol p 1, 2, 3 and 5) during immunotherapy. Clin Exp Allergy 1997; 27:68-74; PMID:9117883; http://dx.doi.org/10.1046/j.1365-2222.1997.d01-416.x
223. Stafforst T, Schneider MF. An RNA-deaminase conjugate selectively repairs point mutations. Angewandte Chemie 2012; 51:11166-9; PMID:23038402; http://dx.doi.org/10.1002/anie.201206489
224. Schneider MF, Wettengel J, Hoffmann PC, Stafforst T. Optimal guideRNAs for re-directing deaminase activity of hADAR1 and hADAR2 in trans. Nucleic Acids Res 2014; 42:e87; PMID:24744243; http://dx.doi.org/10.1093/nar/gku272
225. Montiel-Gonzalez MF, Vallecillo-Viejo I, Yudowski GA, Rosenthal JJ. Correction of mutations within the cystic fibrosis transmembrane conductance regulator by site-directed RNA editing. Proc Natl Acad Sci U S A 2013; 110:18285-90; PMID:24108353; http://dx.doi.org/10.1073/pnas.1306243110
226. Bass BL. RNA editing by adenosine deaminases that act on RNA. Annu Rev Biochem 2002; 71:817-46; PMID:12045112; http://dx.doi.org/10.1146/annurev.biochem.71.110601.135501
227. Nishikura K. Functions and regulation of RNA editing by ADAR deaminases. Annu Rev Biochem 2010; 79:321-49; PMID:20192758; http://dx.doi.org/10.1146/annurev-biochem-060208-105251
228. Vogel P, Stafforst T. Site-directed RNA editing with antagomir deaminases–a tool to study protein and RNA function. ChemMedChem 2014; 9:2021-5; PMID:24954543; http://dx.doi.org/10.1002/cmdc.201402139
229. Vogel P, Schneider MF, Wettengel J, Stafforst T. Improving site-directed RNA editing in vitro and in cell culture by chemical modification of the guideRNA. Angewandte Chemie 2014; 53:6267-71; PMID:24890431; http://dx.doi.org/10.1002/anie.201402634
230. Hanswillemenke A, Kuzdere T, Vogel P, Jékely G, Stafforst T. Site-Directed RNA Editing in Vivo Can Be Triggered by the Light-Driven Assembly of an Artificial Riboprotein. J Am Chem Soc 2015; 137:15875-81; PMID:26594902; http://dx.doi.org/10.1021/jacs.5b10216
231. Zhao X, Yu Y-T. Targeted pre-mRNA modification for gene silencing and regulation. Nat Methods 2008; 5:95-100; PMID:18066073; http://dx.doi.org/10.1038/nmeth1142
232. Kole R, Krainer AR, Altman S. RNA therapeutics: beyond RNA interference and antisense oligonucleotides. Nat Rev Drug Discov 2012; 11:125-40; PMID:22262036; http://dx.doi.org/10.1038/nrd3625
233. Blanc V, Davidson NO. C-to-U RNA editing: mechanisms leading to genetic diversity. J Biol Chem 2003; 278:1395-8; PMID:12446660; http://dx.doi.org/10.1074/jbc.R200024200
234. Liu N, Pan T. N6-methyladenosine-encoded epitranscriptomics. Nat Struct Mol Biol 2016; 23:98-102; PMID:26840897; http://dx.doi.org/10.1038/nsmb.3162
235. Machnicka MA, Milanowska K, Oglou OO, Purta E, Kurkowska M, Olchowik A, Januszewski W, Kalinowski S, Dunin-Horkawicz S, Rother KM, et al. MODOMICS: a database of RNA modification pathways—2013 update. Nucleic Acids Res 2013; 41:D262-D7; PMID:23118484; http://dx.doi.org/10.1093/nar/gks1007
236. Muttach F, Rentmeister A. A Biocatalytic Cascade for Versatile One−Pot Modification of mRNA Starting from Methionine Analogues. Angewandte Chemie International Edition 2016; 55:1917-20; http://dx.doi.org/10.1002/anie.201507577
237. Vickers TA, Wyatt JR, Burckin T, Bennett CF, Freier SM. Fully modified 2′ MOE oligonucleotides redirect polyadenylation. Nucleic Acids Res 2001; 29:1293-9; PMID:11238995; http://dx.doi.org/10.1093/nar/29.6.1293
238. Xiong Y, Steitz TA. Mechanism of transfer RNA maturation by CCA-adding enzyme without using an oligonucleotide template. Nature 2004; 430:640-5; PMID:15295590; http://dx.doi.org/10.1038/nature02711
239. Cho HD, Verlinde CL, Weiner AM. Reengineering CCA-adding enzymes to function as (U, G)-or dCdCdA-adding enzymes or poly (C, A) and poly (U, G) polymerases. Proc Natl Acad Sci U S A 2007; 104:54-9; PMID:17179213; http://dx.doi.org/10.1073/pnas.0606961104
240. Trippe R, Guschina E, Hossbach M, Urlaub H, Lührmann R, Benecke B-J. Identification, cloning, and functional analysis of the human U6 snRNA-specific terminal uridylyl transferase. RNA 2006; 12:1494-504; PMID:16790842; http://dx.doi.org/10.1261/rna.87706
241. Martin G, Keller W. RNA-specific ribonucleotidyl transferases. RNA 2007; 13:1834-49; PMID:17872511; http://dx.doi.org/10.1261/rna.652807
242. Zamecnik PC, Raychowdhury MK, Tabatadze DR, Cantiello HF. Reversal of cystic fibrosis phenotype in a cultured Δ508 cystic fibrosis transmembrane conductance regulator cell line by oligonucleotide insertion. Proc Natl Acad Sci U S A 2004; 101:8150-5; PMID:15148387; http://dx.doi.org/10.1073/pnas.0401933101
243. Exploratory Study to Evaluate QR-010 in Subjects With Cystic Fibrosis ΔF508 CFTR Mutation. (2015). Clinical Trial Phase I
244. Dose Escalation Study of QR-010 in Homozygous ΔF508 Cystic Fibrosis Patients. (2015). Clinical Trial Phase IB
245. Open Label, Extension Study of PRO044 in Duchenne Muscular Dystrophy (DMD). (2014). Clinical Trial Phase II
246. Phase IIb Study of PRO045 in Subjects With Duchenne Muscular Dystrophy. (2013). Clinical Trial Phase 2B
247. A Phase I/II Study of PRO053 in Subjects With Duchenne Muscular Dystrophy (DMD). (2013). Clinical Phase Trial I/II
248. Study of SRP-4045 and SRP-4053 in DMD Patients (ESSENCE). (2015). Clinical Trial Phase III
249. A Study to Assess the Safety and Tolerability of Single Doses of AZD4076 in Healthy Male Subjects. (2015). Clinical Trial Phase I
250. RNA-Immunotherapy of IVAC_W_bre1_uID and IVAC_M_uID (TNBC-MERIT). (2014). Clinical Trial Phase I
251. A Phase I Study of T-Cells Genetically Modified at the CCR5 Gene by Zinc Finger Nucleases SB-728mR in HIV-Infected Patients. (2015). Clinical Trial Phase I

6.6.2 **PUBLICATION 2:** WETTENGEL J.*, REAUTSCHNIG P.*, GEISLER S., KAHLE P.J., STAFFORST T. HARNESSING HUMAN ADAR2 FOR RNA REPAIR - RECODING A PINK1 MUTATION RESCUES MITOPHAGY. NUCLEIC ACIDS RES, 2017. 45(5): P. 2797-2808. * EQUAL CONTRIBUTION

Published online 7 October 2016 *Nucleic Acids Research, 2017, Vol. 45, No. 5* ***2797–2808***
doi: 10.1093/nar/gkw911

Harnessing human ADAR2 for RNA repair – Recoding a PINK1 mutation rescues mitophagy

Jacqueline Wettengel[1,†], Philipp Reautschnig[1,†], Sven Geisler[2,3], Philipp J. Kahle[2,3] and Thorsten Stafforst[1,*]

[1]Interfaculty Institute of Biochemistry, University of Tübingen, Auf der Morgenstelle 15, 72076 Tübingen, Germany, [2]Department for Neurodegenerative Diseases, Hertie Institute for Clinical Brain Research, University of Tübingen, Otfried-Müller-Strasse 27, 72076 Tübingen, Germany and [3]German Center for Neurodegenerative Diseases, Otfried-Müller-Strasse 23, 72076 Tübingen, Germany

Received June 23, 2016; Revised September 27, 2016; Accepted September 30, 2016

ABSTRACT

Site-directed A-to-I RNA editing is a technology for re-programming genetic information at the RNA-level. We describe here the first design of genetically encodable guideRNAs that enable the re-addressing of human ADAR2 toward specific sites in user-defined mRNA targets. Up to 65% editing yield has been achieved in cell culture for the recoding of a premature Stop codon (UAG) into tryptophan (UIG). In the targeted gene, editing was very specific. We applied the technology to recode a recessive loss-of-function mutation in PINK1 (W437X) in HeLa cells and showed functional rescue of PINK1/Parkin-mediated mitophagy, which is linked to the etiology of Parkinson's disease. In contrast to other editing strategies, this approach requires no artificial protein. Our novel guideRNAs may allow for the development of a platform technology that requires only the administration or expression of a guideRNA to recode genetic information, with high potential for application in biology and medicine.

INTRODUCTION

RNA editing alters genetic information at the RNA-level by insertion, deletion or modification of nucleotides (1). The catalytic deamination of adenosine (A) gives inosine (I) that is biochemically read as guanosine. In consequence, A-to-I RNA editing alters the function of RNAs in various ways. Amino acids are substituted, miRNA recognition (2,3) and splicing (4) are altered. In the human transcriptome, the classic example for amino acid substitution is the editing of the glutamate receptor GluR2 transcript at two sites, the R/G and the Q/R site, with the latter one being essential for nervous system function (5,6). A-to-I editing is carried out by two ADARs (adenosine deaminases acting on RNA), ADAR1 and ADAR2 in different isoforms (7). An alteration of RNA editing is linked to various neurological diseases including behavioral disorders, epilepsy and the Prader–Willi syndrome (8–11). Knock-out of ADAR2 in mice leads to an early dead of the newborn due to seizures. Interestingly, this phenotype can be rescued by genomic insertion of an R at the Q/R site in GluR2 (12). Mutations in ADAR1 are linked to the Aicardi–Goutieres syndrome, (13) an autoimmune disease and others, including dyschromatosis (14). Knock-out of ADAR1 function in mice results in early embryonic fatality, (15,16) which can be rescued by a simultaneous knock-out of dsRNA sensing via MDA5 (17,18). Both, hyper (19)- and hypoediting (20) have been associated with cancer (21–23). Whereas ADAR1 is ubiquitously expressed in various tissues, ADAR2 is mainly expressed in neurons (24). Both enzymes are promiscuously recruited to hundreds of thousands of double-stranded RNA structures by their N-terminal dsRNA binding domains (dsRBD) (25,26). Thus, editing is relatively unspecific, happens massively in Alu repeats, and on dsRNA structures of >30 bp (27). However, some RNA substrates, containing bulges and loops, are edited in a highly specific manner (7). In particular, the precise and efficient editing at the R/G site of the GluR2 transcript results from a defined positioning of ADAR2 by the interaction of its two dsRBDs with the exon/intron border of the transcript (28).

Re-directing RNA editing to user-defined targets allows altering genetic information in a highly rational way (29). Various applications in basic biology and medicine are conceivable. Even though limited to A-to-I substitution, the scope is large. It includes the targeting of most polar amino acids (Gln, Arg, His, Tyr, Ser, Thr and others), which play essential roles in enzyme catalysis, signaling and posttranslational modification. Furthermore, stop, start, splicing signals and miRNA recognition sites can be manipulated. Thus, site-directing RNA editing at specific sites on user-

*To whom correspondence should be addressed. Tel: +49 7071 2975376; Fax: +49 7071 295070; Email: thorsten.stafforst@uni-tuebingen.de
†These authors contributed equally to this work as the first authors.

defined targets has an immense potential for the manipulation of protein function, RNA processing, and could be used to attenuate disease phenotypes (29). Such a strategy would complement current genome editing approaches (30) in several aspects. The transient and thus reversible nature of RNA manipulation could be beneficial with respect to ethical issues and safety aspects. Both, therapeutic and potential adverse effects are likely to be tunable and reversible. Furthermore, manipulations are conceivable that are inaccessible at the genome level. This includes amino acid or transcript level changes that would kill a cell if they are permanently enforced. Potentially lethal interventions on kinases, apoptosis factors, coagulation factors, (31) transcription or translation factors could be realized on the RNA-level suddenly, transiently or partially to obtain a therapeutic effect. Manipulation at the RNA-level might also be much more efficient compared to HDR-dependent gene correction, which remained persistently inefficient *in vivo*, (30) in particular in postmitotic tissues like the brain. For many genetic diseases caused by recessive loss-of-function mutations, a drug that can restore a small fraction (like 5%) of functional gene product in a large fraction of a tissue is superior to a drug that can restore full gene function (100%) but only in a small fraction of the tissue. A typical example is cystic fibrosis (32).

Site-directed RNA editing is not a new concept. The first trial goes back to the pioneering work of Tod Woolf and colleagues who could demonstrate already in 1995 to elicit RNA editing in a reporter mRNA inside Xenopus eggs when the mRNA was hybridized with a 52 nt long, unstructured guideRNA prior to microinjection (33,29). However, the main issue of low efficiency and in particular off-site editing in the guideRNA/mRNA could not be solved at that time. In 2012 and 2013, we (34,35) and the lab of Joshua Rosenthal (36) have reanimated the concept of site-directed RNA editing by independent engineering of artificial editing enzymes that address the catalytic activity in a highly rational way with the help of external guideRNAs. Such strategies work inside mammalian cell culture and even in a simple organism (37) and allow the repair of disease-relevant genes, like the CFTR (36) mRNA. Chemical modification of the guideRNA was shown to improve specificity (35). Even though feasible and expandable for multiplexing approaches, both established strategies require the expression of an engineered deaminase. With respect to this limitation, we were wondering if it may become possible to harness the endogenous human ADARs for site-directed RNA editing, again with external guideRNAs. As ADARs are well expressed in neurons, (24) such a strategy could enable the attenuation of (neuron-related) disease phenotypes related to loss-of-function mutations simply by administration or ectopic expression of a small guideRNA.

Neurodegenerative diseases are a global challenge of tomorrow (38). Their enormous costs in healthcare threaten the welfare system. Parkinson´s disease (PD) affects the central nervous system, destroys motion control, is often accompanied by neuropsychiatric disorders and characterized by a slow progression (39). The disease results from a loss of dopaminergic neurons in the substantia nigra and is typically accompanied by the formation of Lewy bodies. Several genes are linked to inherited forms of PD including numerous mutations in α-synuclein, Parkin, PINK1 or LRRK2. However, some forms of hereditary early-onset forms of PD are linked to single mutations in one gene, like W437X in PINK1 (40). Studying such mutations has proven valuable for the elucidation of the mechanism underlying specific forms of PD. Recent research shows that PINK1 and Parkin work together in a mitochondria quality control pathway, where damaged depolarized mitochondria are eliminated by the process of autophagy, termed mitophagy. The PINK1 kinase-function is required in an initial and essential step of mitophagy, which is the recruitment of cytosolic Parkin to the damaged mitochondria and the formation of perinuclear clusters (41). A single G-to-A nucleotide exchange in PINK1 has been described that changes Trp437 into a premature Stop codon and truncates PINK1's C-terminus by 145 amino acids including the functionally important kinase domain (40). This results in the impairment of the Parkin-dependent perinuclear clustering, clearance of damaged mitochondria and is linked to early-onset PD.

Here, we describe the rationale for the design of guideRNAs to harness human ADAR2 for site-directed RNA editing. We demonstrate the feasibility of the approach by the repair of a neuron-related disease-causing point mutation and show functional rescue of a mitophagy phenotype.

MATERIALS AND METHODS

Protein production

Wild-type human ADAR2 (with a C-terminal His_6-tag) was produced from yeast (YVH10), purified by nickel and heparin affinity chromatography similar as described before (42). For details see Supplementary Material.

R/G-guideRNA synthesis

R/G-guideRNAs were produced by T7 *in vitro* transcription. The guideRNAs were cleaved during transcription from a cis-acting hammerhead ribozyme (TA TTCCACCT GA TGAGTTTTTA CGAAACGTTC CCGTGAGGGA ACGTC*GTGGAATA, the guideRNA starts after the asterisk that marks the cleavage site). The guideRNAs were purified by urea (7.5M) PAGE (8%, 1xTBE), isolated by the crush soak method and precipitated with ethanol.

In vitro editing

Editing assays were performed with purified mRNAs, guideRNAs and ADAR2-His_6 protein. mRNA (0.5 or 25 nM), guideRNA (5 or 125 nM) and ADAR2 (180 or 350 nM) were incubated in reverse transcription buffer (75 mM KCl, 25 mM Tris-HCl, 2 mM DTT, pH 8.3) and the indicated amounts of magnesium (1.5 or 3 mM), and spermidine (0, 0.5 or 2 mM). Editing reactions were cycled three times between 37°C (30 min) and 30°C (30 min) and were stopped by addition of a sense oligomer that displaces the guideRNA from the mRNA. After reverse transcription and Taq-PCR, DNA was analyzed by Sanger sequencing. The editing yields were estimated by the relative areas of the guanosine versus adenosine traces.

Cellular editing

(**a**) under transient ADAR2 expression: ADAR2 (lacking the His-tag) was subcloned into the pcDNA3.1 vector as described in the Supplementary Material. The R/G guideRNAs were subcloned into the pSilencer2.1-U6hygro vector under control of U6 promoter and terminator. The W58X eGFP gene was cloned into the pcDNA3.1 vector as described before (29). 293T cells were cultivated with Dulbecco's modified Eagle's medium (DMEM), 10% fetal bovine serum (FBS), 1% P/S, 37°C, 5% CO_2. Cells (1.75×10^5/well) were seeded into 24-well plates and were 24 h later transfected with the indicated amounts of plasmids using Lipofectamine 2000. Plasmid amounts have typically been 300 ng ADAR2, 300 ng W58X GFP, 1300–1600 ng guideRNA per well. For details see Supplementary Figures S8–S15. After the indicated time after transfection (typically 48 h) editing was evaluated by fluorescence microscopy (Zeiss AxioObserver.Z1) and RNA sequencing. For the latter, total RNA from the cells (Qiagen RNeasy Mini Kit) was DNaseI-digested, followed by reverse transcription, Taq-PCR amplification and Sanger sequencing. (**b**) under genomic ADAR2 expression: the 293 Flp-In T-REx system (Life Technologies) was used for stable integration of a single copy of ADAR2 at a genomic FRT-site in the cells. Briefly, 4×10^6 cells were seeded on a 10 cm dish. After one day, 1 μg of ADAR2 in a pcDNA5 vector under control of the tet-on CMV promoter, and 9 μg of pOG44 expressing the Flp recombinase were transfected with Lipofectamine 2000 (30 μl). One day later, the medium was changed for at least two weeks to selection medium (DMEM, 10% FBS, 100 μg/ml hygromycin B, 15 μg blasticidin S). Cells were kept in selection medium prior to the editing experiment which was then done in the absence of antibiotics. Editing: 3×10^5/well were seeded in poly-D-lysin-coated 24-well plates. Twenty four hours later, transfection was performed with GFP (300 ng) and R/G-guideRNA (1300 ng) using Lipofectamine 2000. GFP phenotype was analyzed by fluorescence microscopy. RNA was isolated and sequenced as described above 72 h post transfection. The sequences of all guideRNAs are given in Supplementary Table S1.

Mitophagy assay

HeLa cells (PINK1 wt or KO) were cultured under standard conditions (DMEM + 10% FBS, 37°C, 5% CO_2). The mitophagy assay was performed in 24-well dishes. Each well contained a cover-slip coated with poly-D-lysine (Sigma Aldrich). The cells were seeded at 2.5×10^4/well. After 24 h, cells were transfected with the indicated plasmids using FuGene6 (Promega) following the manufacturer's instructions. If not indicated plasmid amounts/well were 300 ng for EGFP-Parkin, 300 ng for PINK W437amber and 300 ng editing vector. In control experiment Figure 4, d) 1300 ng of a guideRNA plasmid based on pSilencer lacking ADAR2 was co-transfected instead of the editing vector. In control experiment Figure 4, e), 200 ng of an editing vector lacking any guideRNAs but containing ADAR2 was co-transfected instead of the original editing vector. Treatment with 10 μM CCCP (in DMEM + 10% FBS) was either performed 46 h after transfection for 2 h or 24 h after transfection for 24 h. To visualize the mitochondria with a membrane potential sensitive dye, like MitoTracker Red CMXRos, a CCCP wash out was performed. For this, the depolarizing agent CCCP was washed out by changing the media twice every 15 min. Then the cells were incubated with 100 nM MitoTracker Red CMXROS (Invitrogen, M7512) in DMEM for 30 min at 37°C prior fixation or harvesting. Always 48 h after transfection, cells were washed with phosphate buffered saline (PBS) and then either fixated (4% paraformaldehyde/PBS) for immunocytochemical staining (A) or harvested for RNA isolation (B). (A) For immunocytochemical staining, the cells were washed 3 times with PBS and permeabilized with 1% Triton-X-100 in PBS for 5 min. After additional 3 washing steps, the cells were blocked in 10% FCS in PBS for 1 h at RT and incubated with following antibodies for 2 h at RT: mouse anti-ADAR2 (Santa Cruz, sc-73409) and rabbit anti-PINK1 (Novus Biological, BC-100-494). Subsequently, the cells were incubated with secondary antibodies: goat anti-rabbit or anti-mouse Alexa Fluor 568, 647 or 350. The nuclei were counterstained with Hoechst33342 in PBS. The cover-slips were mounted on glass-slides using Dako fluorescent mounting medium. The cells were analyzed using an Axio-Imager equipped with ApoTome (Zeiss). (B) RNA was isolated from the cells (Qiagen RNeasy Mini Kit). This was followed by DNaseI digest, reverse-transcription, amplification and Sanger sequencing. For further details see the Supplementary Material.

RESULTS AND DISCUSSION

Design of a trans-acting guideRNA

At the R/G-site of the natural GluR2 transcript, a cis-located intronic sequence folds back to the exon under formation of a bulged stem loop structure that recruits ADAR2 via its two dsRBDs (Figure 1A). A trans-acting guideRNA is conceivable that contains a part of the natural cis-acting R/G-motif in combination with an mRNA-binding platform to bind an arbitrary mRNA in order to recruit ADAR2 for site-directed editing. We designed a trans-acting guideRNA based on available structural information on the binding complex of dsRBD1 and 2 with the R/G-hairpin structure (28). We decided to cut the native R/G-site between the two guanosines five and six nucleotides downstream of the editing site (Figure 1A). This position appeared most suitable to harbor the protruding mRNA under minimal interference with dsRBD2 recognition and the deaminase domain. The latter assumption was confirmed by a recent crystal structure of the ADAR2 deaminase domain with an RNA model substrate (43).

Human ADAR2 gives highly efficient editing *in vitro*

First, we studied the new editing strategy in the polymerase chain reaction (PCR) tube with purified guideRNAs and ADAR2 protein. A reporter mRNA (cyan fluorescent protein) that contains a single G-to-A point mutation generating a nonsense stop signal ($Trp^{66} \rightarrow$amber) served as substrate (Figure 1), as described before (34,42). Wild-type human ADAR2 was expressed and purified from yeast (YVH10), as described before (42). To minimize charge repulsion and crowding at the exit of the mRNA, we generated guideRNAs with strictly homogenous 5′-r(GUGG)

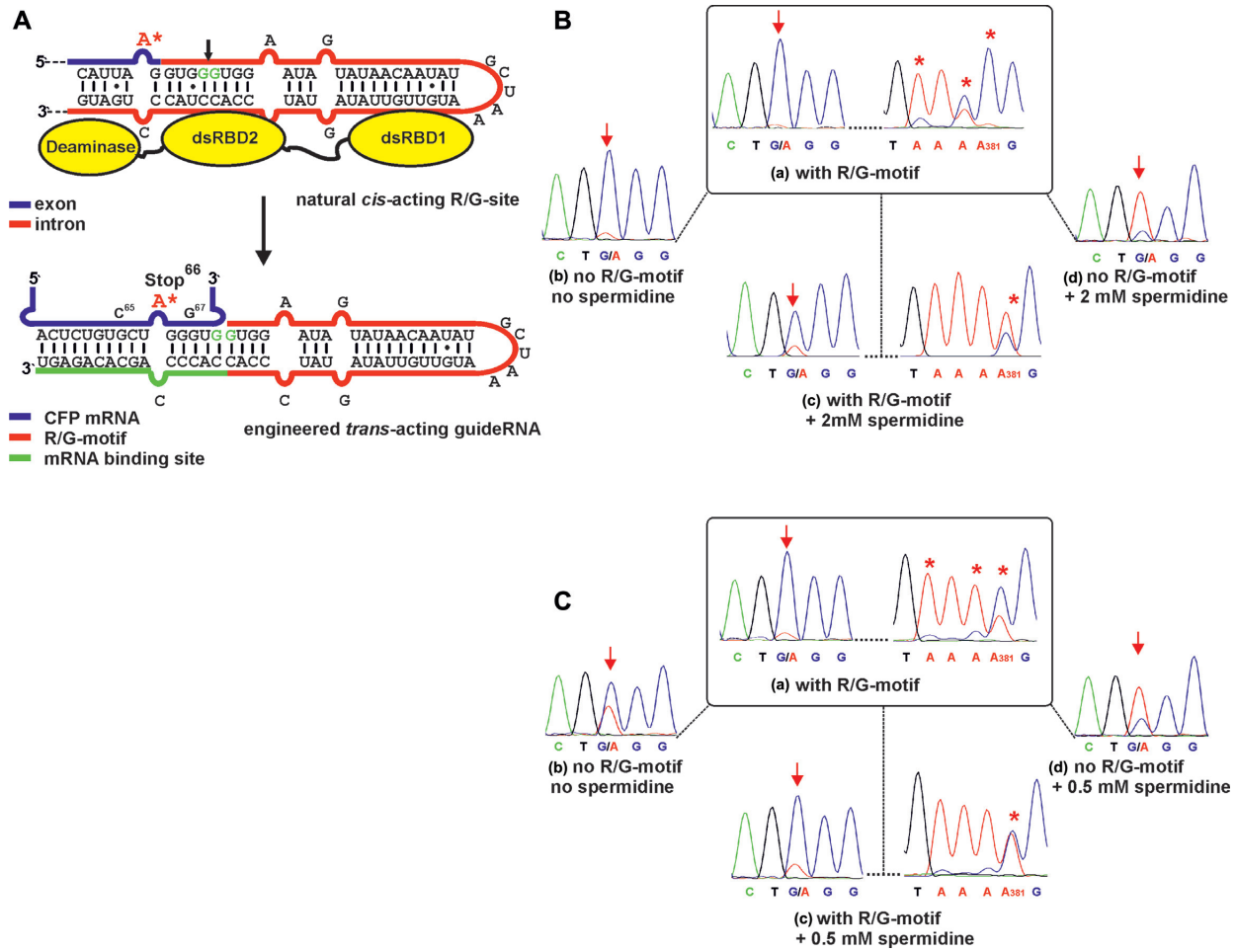


Figure 1. (**A**) Engineering the natural cis-acting R/G-site into a trans-acting guideRNA that steers wild-type hADAR2 to a reporter mRNA (CFP) to repair the Stop codon 66 (UA*G) to tryptophan. (**B** and **C**) *In vitro* RNA editing experiments in absence and presence of spermidine and of the R/G-motif in the guideRNA. Targeted adenosine is marked with an arrow; off-site editing around adenosine 381 is marked with *. The sequence of the control guideRNA lacking the R/G-motif was 5′-NNN GAACACCCC*AGCACAGA. (**B**) Editing under high concentrations ([ADAR2] = 350 nM, [guideRNA] = 125 nM, [mRNA] = 25 nM, [Mg] = 3 mM). (**C**) Editing under low concentrations ([ADAR2] = 180 nM, [guideRNA] = 5 nM, [mRNA] = 0.5 nM, [Mg] = 1.5 mM).

ends by *in vitro* transcription from a hammerhead cassette (44). Initially, we used a guideRNA with a 16 nt reverse complementary mRNA binding site that puts the targeted adenosine in a mismatch with cytosine. The targeted adenosine was kept at a distance of five intervening base pairs to the 5′-terminus of the R/G-helix and was mismatched with cytosine (Figure 1A).

First, we tested the system at high concentrations of all components ([mRNA] = 25 nM, [guideRNA] = 125 nM, [ADAR2] = 350 nM) in presence of 3 mM magnesium. Full A-to-I conversion at the targeted site (A200) was achieved (Figure 1B, (a)). However, in contrast to our engineered SNAP-ADAR deaminase, (34,42) wild-type ADAR2 was dramatically more reactive and gave massive off-target editing, with full conversion at A381 (Figure 1B, (a)), 50–70% yield at A295, A380, A476 and minor editing at various sites (Supplementary Figure S2). This off-site editing was typically guideRNA-independent. Furthermore, the presence of the R/G-motif was not strictly necessary under these conditions as demonstrated by sufficient editing when applying a 17 nt short, single-stranded guideRNA lacking the R/G-motif (Figure 1B, (b)). Apparently, short RNA duplexes are already sufficiently recognized by ADAR2 under these lax conditions. Consequently, we increased the stringency by adding spermidine (2 mM, Figure 1B, (c and d)). The additive diminished over-editing at all sites to background except adenosine 381 that retained some residual activity (Figure 1B, (c)); Supplementary Figure S3). In presence of spermidine, the guideRNA lacking the R/G-motif was clearly inferior compared to the complete R/G-guideRNA (Figure 1B, (d) versus (c)). We also tested the reverse case and demonstrated that the R/G-guideRNA does not recruit the engineered SNAP-ADAR2 deaminase that lacks both dsRNA binding domains (Supplementary Figure S5). Thus, the recognition between R/G-motif and the dsRBDs is required for editing.

Then we studied editing at low concentrations of the components (Figure 1C). For this, we halved the ADAR2 con-

centration to 180 nM and strongly decreased the concentrations of the guideRNA (5 nM) and of the mRNA (0.5 nM). Furthermore, the magnesium concentration was decreased to physiological levels (1.5 mM). Notably, the editing reaction run much more smoothly, and significant off-target editing occurred only at adenosine 381 (50% yield, Figure 1C, (a)). The addition of spermidine was only possible to 0.5 mM before losing editing yield at the target, and hence, off-site editing was only slightly reduced (Figure 1C, (c)), Supplementary Figure S4). However, 0.5 mM spermidine improved the necessity for the presence of the R/G-motif in the guideRNA clearly (Figure 1C, (b) versus (d)). Overall, off-target editing is controllable to some extent. Complete inhibition, however, appears difficult in the PCR tube.

Editing is selective and efficient in cell culture

In contrast to our alternative SNAP-tag strategy (29,34,35,37), this new strategy allows the ectopic expression of all components including the guideRNA inside the cell. To demonstrate this, we made use of the U6 promotor that enables expression of uncapped small RNAs with homogenous 5′-GUGG-termini by RNA polymerase III (45–47). Initially, we kept the successful guideRNA architecture from the *in vitro* experiment, and put the targeted adenosine 6 nt away from the R/G-motif, and kept the adenosine mismatched with cytosine (Figure 2A). Initially, the mRNA binding template was 16 nt in length. GuideRNAs were delivered on pSilencer U6 hygro plasmids (Life Technologies). ADAR2 and the fluorogenic reporter substrate were provided on pcDNA3.1 vectors under control of the CMV promoter. Due to the better brightness we changed the editing substrate from CFP to eGFP (W58amber).

After co-transfection of all three plasmids (ADAR2, GFP reporter, guideRNA) into 293T cells the fluorescence phenotype was analyzed via fluorescence microscopy (Figure 2B). Initial optimization showed that a stoichiometry of five guideRNA to one ADAR2 plasmid was optimal. Transfection with wt GFP served as positive control (Figure 2B, (a)). The fluorescence of the Stop58 eGFP transcript was only restored in presence of both components: ADAR2 and a guideRNA reverse complementary to the target site (Figure 2B, (e)). Applying no guideRNA (Figure 2B, (b)) or a guideRNA that binds to the transcript, but 24 nt downstream the target site [not shown], showed no repair activity. This is in accordance with the strict requirement for a dsRNA secondary structure at the editing site. Editing was also absent when a catalytically inactive ADAR2 variant (E396A) or no ADAR2 was used (Figure 2B, (c and d)). We extracted total RNA from the cells, reverse transcribed and amplified the eGFP transcript with specific primers to estimate the editing yield by Sanger sequencing. This revealed a single and specific A-to-G conversion at the targeted codon with approximately 25% yield after 24 h (Supplementary Figure S8) and up to 40% yield after 48 h (Figure 2B, (e)). The Sanger sequencing trace covered the whole coding sequence of the reporter and revealed a highly specific editing. Only at adenosine 381, which was already prone to off-target editing *in vitro* (Figure 1B and C), we found some minor (<5%) guideRNA-independent off-target editing (full trace see Supplementary Figure S6). However, by adjusting the amount of co-transfected ADAR2 (see next paragraph) this specific off-target editing event was completely abolished.

Further control experiments demonstrated the benefit of the R/G-motif in our guideRNAs. First, when applying our R/G-guideRNA together with SNAP-ADAR2 (lacking the dsRBD that recognize the R/G-motif) instead of human ADAR2 (Supplementary Figure S9), no editing took place. Second, chemically stabilized, single-stranded guideRNAs of 19 nt or 21 nt length that put the targeted adenosine almost centrally into an A:C mismatch but lacks the R/G-motif elicited some minor editing but was clearly inferior to the R/G-guideRNA in recruiting human ADAR2 to the reporter gene (Supplementary Figure S9).

Editing is further improved by optimizing the guideRNA

To further improve cellular editing we pursued several strategies. First, a hairpin, derived from the BoxB-motif (48,49), was included at the 3′-terminus of the R/G-guideRNA to stabilize it. However, the effect was relatively small. As the hairpin did not show any drawback for the editing yields but facilitated cloning, the hairpin was included in all guideRNA architectures that followed. We then varied the length of the mRNA binding template from 18 to 29 nt with several intermediates (Figure 2C, and Supplementary Figure S10). When using up to 20 nt, the editing yield stayed virtually unchanged around 40%. However, when increasing the mRNA binding site to 25 and 29 nt, the editing yield significantly dropped (Supplementary Figure S10). Furthermore, the risk for off-target editing in the mRNA/guideRNA duplex increased with increasing the guideRNA's length (Figure 2C, Supplementary Figure S10), reminiscent to the situation in the first editing trials by Tod Woolf (33). We then systematically varied the distance between the editing site and the 5′-terminus of the guideRNA (Figure 2D, Supplementary Figure S11). Notably, we found a bell-shaped distribution of the editing yield. Whereas almost no editing was observed with two intervening bases, the yield increased step-wise and reached an optimum with 6 to 7 intervening bases and decreased slowly for ≥8 nucleotides. Thus, changing the guideRNA architecture from the original 10+1+5 nt design to 8+1+7 nt not only shifts the targeted bases further into the middle of the mRNA/gRNA-duplex but also improves the editing yield from around 40% up to 50% (Figure 2B versus 2D and E). The same bell-shaped trend with a maximum around 7 intervening nucleotides was repeatedly found for the repair of the R407Q missense mutation (5′-CAG codon) in PINK1 (Supplementary Figure S15). Under transient transfection, editing increased until 48 h (50%), stayed constant until 72 h and then started to decline slowly (40% after 96 h, Supplementary Figure S13). Finally, the amount of ADAR2 in the transfection was optimized to reduce off-target editing. Indeed, when reducing the amount of transfected ADAR2 to a range of 25-200 ng, the off-target editing had disappeared without interfering target editing (Figure 2E, Supplementary Figures S7 and S14).

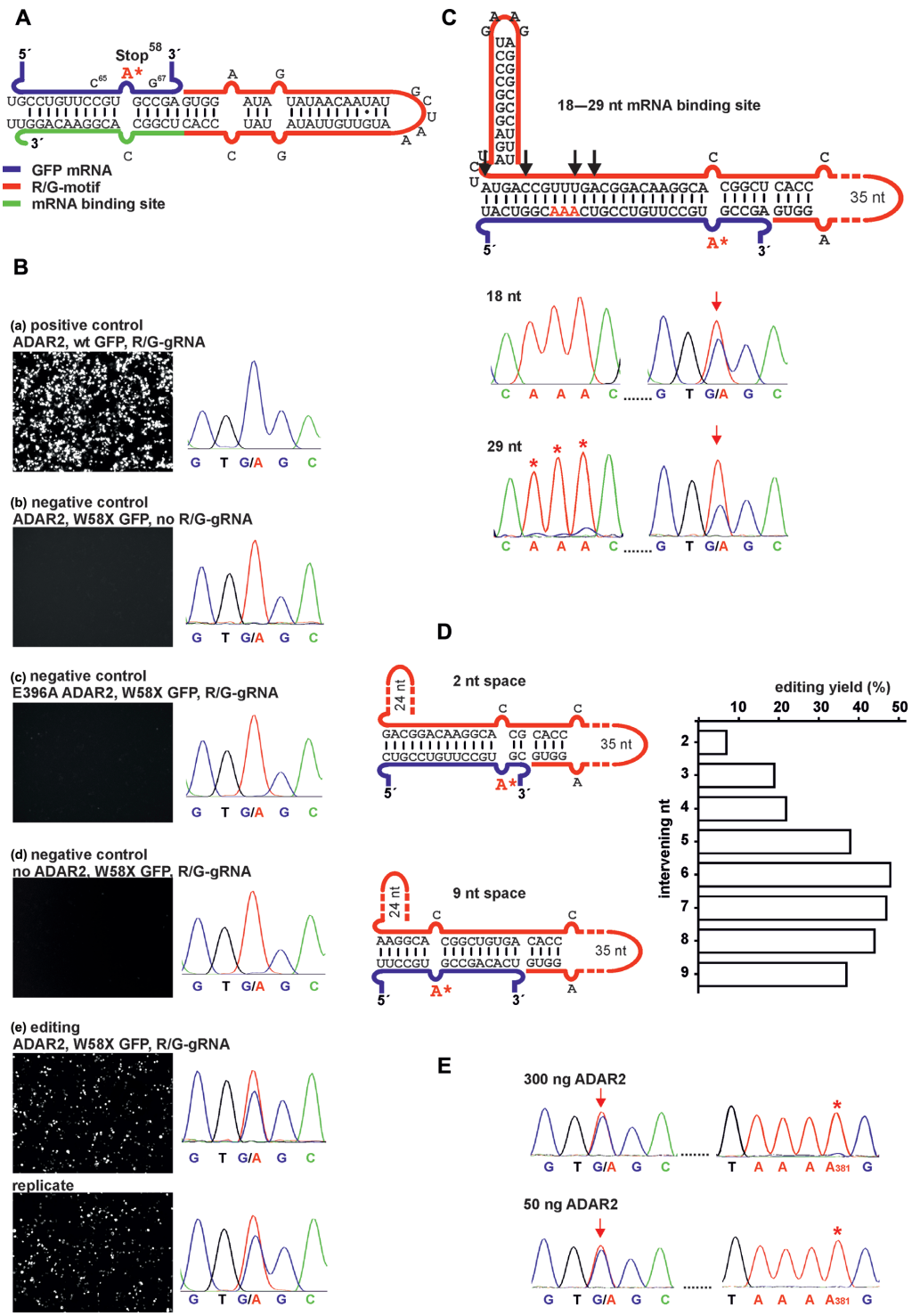


Figure 2. Editing under transient expression in 293T cells. (**A**) Scheme of the mRNA/guideRNA duplex of the editing experiment shown in B. (**B**) Fluorescence imaging of the GFP channel (50 ms exposure time for editing and positive control, 150 ms for negative controls) and RNA sequencing results for cellular editing, 48 h post co-transfection of R/G-guideRNA, ADAR2 and the W58X GFP reporter. E396A ADAR2 is a catalytically inactive mutant. For the editing experiment a duplicate is shown. (**C–E**) Optimization of cellular editing. (**C**) Inclusion of a 3′-terminal hairpin and variation of the length of the mRNA binding platform from 18 to 29 nt, black arrows indicate the different lengths used. For additional traces and fluorescence imaging see Supplementary Figure S10. Red arrows indicate the targeted base, red* indicate off-target editing in the sequencing traces. (**D**) Editing dependency when varying the number of nucleotides (2–9) intervening the targeted adenosine and the R/G-motif. (**E**) Editing versus off-target editing using the optimal guideRNA architecture (7 intervening bases, 16 nt mRNA template, 3′-hairpin) depending on the amount of transfected ADAR2. If not indicated differently, plasmid amounts were 300 ng ADAR2, 300 ng W58X GFP, 1300–1600 ng guideRNA per well in 24-well-plates. For details see Supplementary Figures S8–S15.

Editing works efficiently under genomic expression of ADAR2

Unfortunately, editing failed in 293T cells without overexpression of ADAR2 from a plasmid (Figure 2B, (d)). By RT-qPCR we could show that neither ADAR2 nor its typical endogenous target GluR2 is expressed in 293T cells. The plasmid-borne ADAR2 was on average expressed to the ∼11-fold mRNA-level of the reference gene β-actin (Supplementary Figure S16). To better mimic endogenous ADAR2-levels, we created a homogenous 293T cell line containing a single genomic copy of human ADAR2 under control of the CMV tet-on promoter. Under full induction with doxycycline, the mRNA-level of ADAR2 was induced to ∼50% of that of the β-actin gene (Supplementary Figure S16). Into these cells, we transfected the W58X GFP reporter and the optimized guideRNA (7 nt intervening, 3′-terminal hairpin) delivered on plasmids, as before. Depending on the amount of co-transfected guideRNA (400–1300 ng) editing yields of 45–65% were achieved (Figure 3A, Supplementary Figure S17). No off-target editing was observed in the entire ORF of the reporter gene including adenosine 381 (Figure 3A and Supplementary Figure S18). This is in accordance with the experiments above showing that decreasing ADAR2 expression reduces off-target editing. Notably, the editing yields under lower, but homogenous expression of ADAR2 were significantly better than under transient expression. The advantage of the R/G-motif seems even more pronounced under genomic compared to strong transient ADAR2 expression. The transfection of ADAR2-expressing 293 cells with chemically stabilized, single-stranded guideRNAs of 19 nt and 21 nt length did not elicit detectable editing (Supplementary Figure S9), even though some minor editing was found under transient ADAR2 expression.

Editing of endogenous transcripts is possible by transfection of the guideRNA only

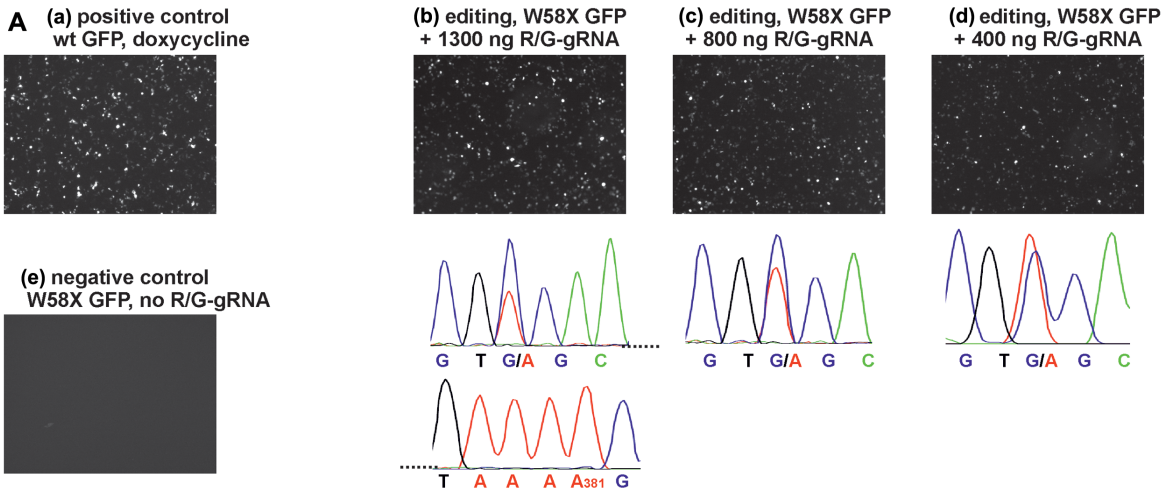
We then tested the editing of endogenous transcripts, either by co-transfection of the guideRNA with ADAR2 in 293T cells, or by sole transfection of the guideRNA into cells that express ADAR2 under control of doxycycline. Thirteen potentially editable 5′-UAG triplets in six different genes (β-actin, GAPDH, GPI, GUSB, VCP, RAB7A) were selected. The genes were chosen to cover a range from highly expressed (like β-actin) to lowly expressed (like GUSB). The editing sites were selected not to interfere with gene function, thus they have been located mainly in the 3′-UTRs. To be able to compare the editing yields, only 5′-UAG triplets were selected. The guideRNAs were equally designed for all 13 sites following the rules developed above. Twelve out of the 13 sites were editable with yields ranging from 10% to 35% (Figure 3B). With the exception of the VCP and RAB7A transcript, the editing yields seem to benefit from the 20-fold higher ADAR2-level under transient compared to genomic ADAR2 expression. In contrast, the editing yields did not depend on the transcript-levels. In β-actin, for instance, two sites were edited well, whereas one site was not edited at all. On the other hand GUSB, which is roughly 100-fold less expressed than β-actin, was edited to a similar extent as β-actin and with the same trends for transient versus genomic ADAR2 expression. Thus, other factors may determine the editing success like the accessibility of the edited site.

Guided RNA editing repairs the PINK1 W437amber mutation and rescues mitophagy

To demonstrate the practical application we aimed to repair a disease-causing point mutation under rescue of the disease-relevant phenotype. The PINK1 W437Stop mutation was selected as it is linked to an inheritable monogenetic form of Parkinson's disease (40) and is characterized by a well-known cellular phenotype (loss of mitophagy). Assuming that all three Stop codons at position 437 in PINK1 are disease-relevant, we chose the amber codon, which is the best editable (42). Following the rules developed above, a guideRNA was designed that puts the amber Stop codon at position 437 into an A:C mismatch in the middle of a 16 nt mRNA/guideRNA duplex (Figure 3D). The guideRNA was further stabilized by a hairpin at the 3′-terminus. Co-transfection of plasmids encoding PINK1 W437amber, ADAR2 and the guideRNA in 293T cells resulted in an editing yield of 35% (Figure 3D).

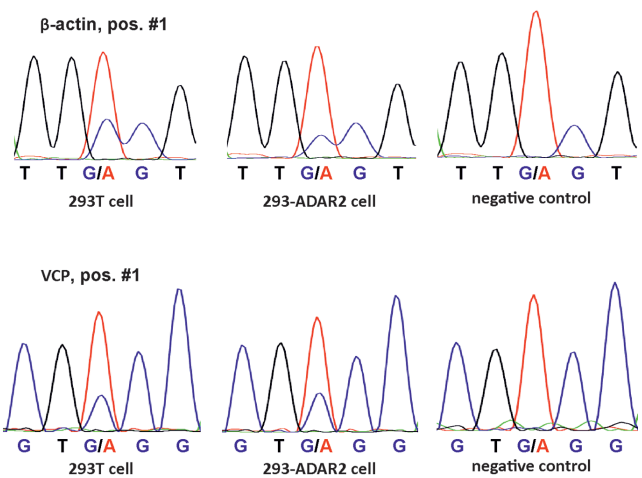
The established PINK1 functional assay in HeLa cells follows a complex protocol (41). First, endogenous wild-type PINK1 is knocked down by RNAi. Then mutated PINK1 and eGFP-tagged Parkin are co-transfected on plasmids, and the mitochondrial membrane potential of the cells is depolarized by CCCP treatment. The CCCP-induced PINK1-dependent perinuclear clustering of Parkin is typically visible after 2 h of CCCP treatment. After long-term treatment (24 h CCCP) mitochondria are cleared from the cytoplasm (mitophagy). To facilitate the protocol and reduce the number of transfections, we simplified the assay. First, a stable PINK1 knock out (KO) HeLa cell line was created using CRISPR-Cas9 (50) technology (for details see Supplementary Figure S20). Second, an editing vector was created that contains ADAR2-His_6 together with five copies of the PINK1-guideRNA, each under control of a U6-promotor (Figure 3C). This allowed us to apply a triple (PINK1, Parkin, editing vector) instead of a quadruple (PINK1, Parkin, ADAR2, guideRNA) co-transfection. Furthermore, the editing vector ensures that guideRNA and ADAR2 are taken-up in a defined stoichiometry and allows to deduce guideRNA expression from ADAR2 staining. The size of the vector was reduced to 7.4 kbp by removing all unnecessary parts of the backbone. A similar construct that targets the W58X GFP reporter was functional in 293T cells, however, suffering from slightly reduced editing yields (35% instead of 50%).

As shown in Figure 4A/B, (a), wild-type HeLa cells show the typical perinuclear clustering of Parkin after 2 h CCCP treatment in 89 ± 3% of the Parkin-positive cells. This clustering phenotype has been shown repeatedly (41,51) to require functional PINK1. As expected, Parkin-clustering disappeared in the PINK1-KO HeLa cell line, even when W437amber PINK1 was overexpressed from a plasmid (Figure 4A/B, (b)), clearly showing that W437amber is unable to rescue the phenotype. However, Parkin-clustering was fully rescued by co-transfection of PINK1 W437X



B editing of endogenous transcripts

gene	293T cell + ADAR2 + guideRNA	293T-ADAR2 cell + doxycyclin + guideRNA	negative control (no guideRNA)
β-actin			
position #1	27%, 24%	17%, 14%	0%, 0%
position #2	0%, 0%	0%, 0%	0%, 0%
position #3	24%, 23 %	14%	0%, 0%
GAPDH			
position #1	21%, 17%	10%, 10%	0%, 0%
position #2	21%, 19%	12%, 10%	0%, 0%
GPI			
position #1	15%, 16%	12%, 11%	0%, 0%
GUSB			
position #1	11%, 9%	10%, 9%	0%, 0%
position #2	26%, 19%	20%, 16%	0%, 0%
VCP			
position #1	23%	25%, 21%	0%
position #2	16%	15%	0%
position #3	12%, 11%	13%	0%
RAB7A			
position #1	32%, 30%	38%	0%, 0%
position #2	27%, 28%	33%	0%, 0%



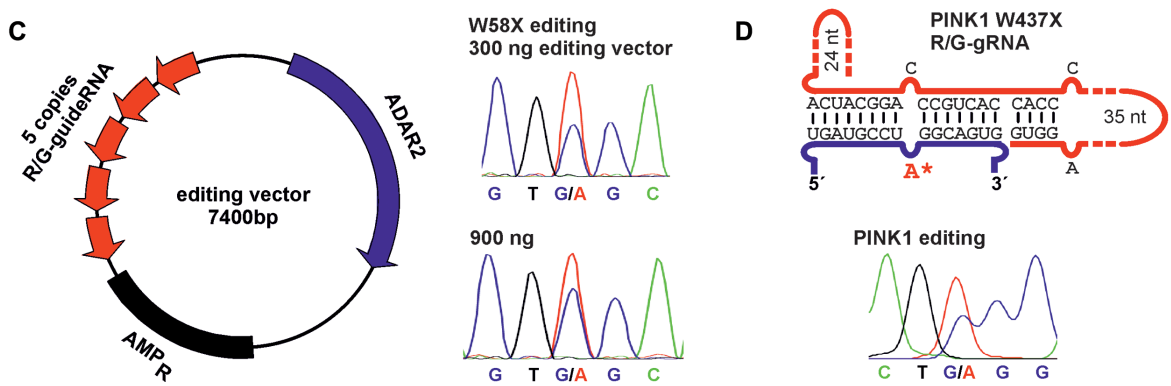


Figure 3. (**A**) Editing with genomically integrated ADAR2 in 293T cells under administration of the R/G-guideRNA and the W58X reporter only. (**B**) Editing of endogenous transcripts. 293T cells were transfected with 300 ng ADAR2 and 1300 ng R/G-gRNA, 293T-ADAR2 cells were induced with 10 ng/ml doxycycline and transfected with 1300 ng guideRNA. Most editing experiments are reported in duplicates. All sequencing traces are given in Supplementary Figure S19. (**C**) Scheme of the editing vector that combines ADAR2 and 5 identical copies of the guideRNA on a minimal plasmid. Editing of W58X GFP in 239T cells under transient expression shows the functioning. (**D**) Scheme of the R/G-guideRNA against PINK1 W437amber and editing in 293T cells under transient transfection.

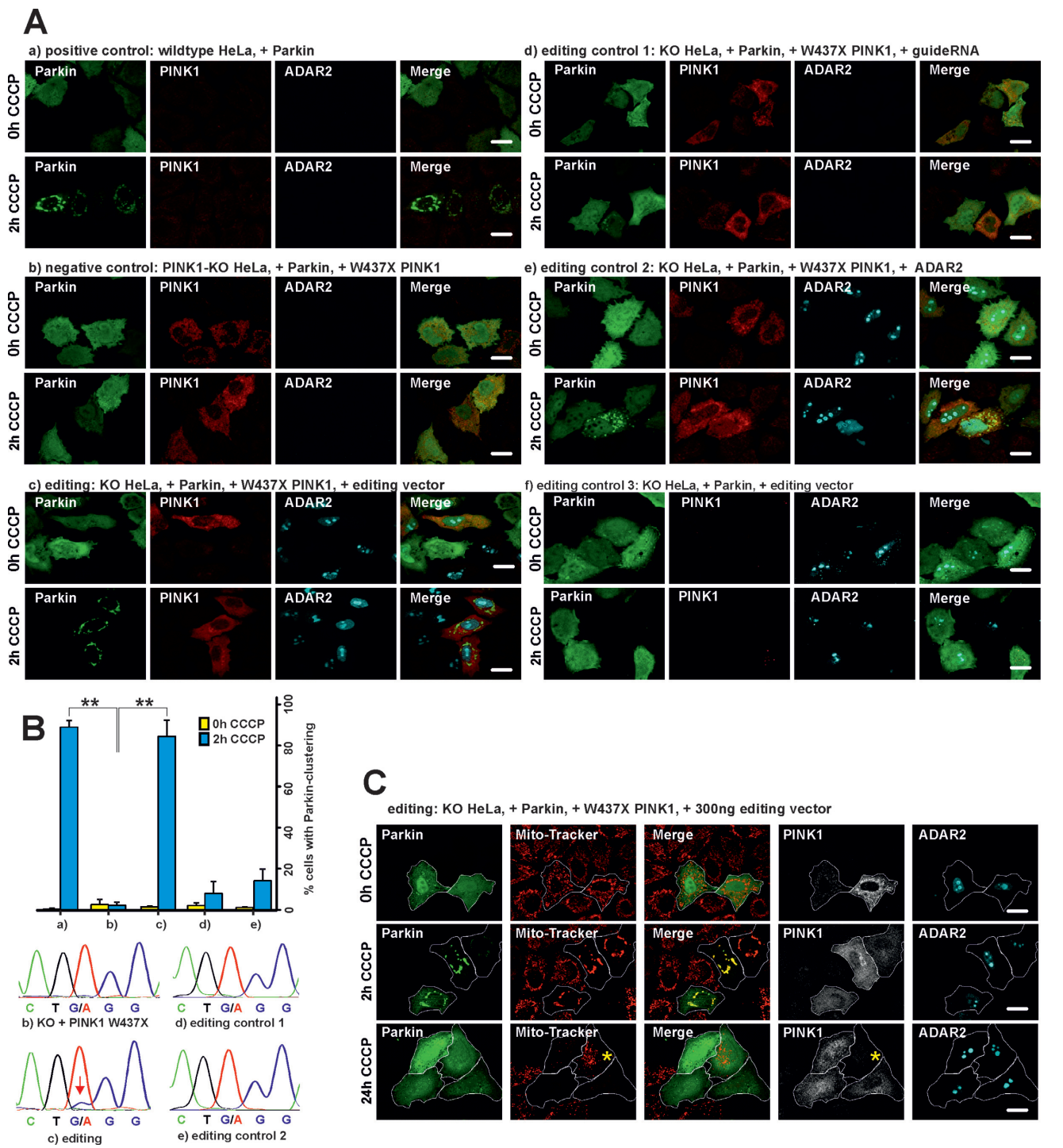


Figure 4. Assaying PINK1 function. (**A**) Parkin-clustering during CCCP treatment. HeLa wild type or PINK1-KO cells were transfected with the respective plasmids and treated either with CCCP or DMSO (=CCCP 0 h) for 2 h or 24 h as indicated. Afterwards, PINK1 and ADAR2 were stained with antibodies, Parkin was tagged with EGFP for subsequent fluorescence microscopy. The scale bars indicate 20 μm. (**B**) Analysis of the editing experiments (a) to (e) shown in panel A. Top: quantification of cells with and without CCCP treatment for Parkin-clustering. The error bars give the standard deviation for N = 3 independent replications (always around 100 cells were counted). The asterisk(s) denote statistical significance when compared to sample (b) (2 h CCCP), with $*P \leq 0.05$; $**P \leq 0.005$; $***P \leq 0.0005$. Bottom: Sanger sequencing of the PINK1 transcript after RT-PCR. Only sample (c) showed detectable editing as indicated by a red arrow. (a) to (f) in panel A and B correspond to the same experiments. (**C**) Restoration of the mitophagy phenotype by editing. After long-term CCCP treatment (24 h), triple positive cells (expressing Parkin, PINK1 and the editing vector) show clearance of the mitochondria that were visualized by Mito-Tracker (Red CMXRos) staining after CCCP washout (30 min). The Mito-Tracker and GFP channel have been merged to visualize the mitochondrial Parkin localization after 2 h CCCP treatment. For better visualization of mitophagy, the Parkin-positive cells have been encircled. The yellow asterisk (24 h CCCP treatment) marks an example of a cell that lacks PINK1 expression and hence did not perform mitophagy.

together with the editing vector in 85±8% of the cells positive for PINK1, Parkin and ADAR2 expression. Co-transfection of PINK1amber/Parkin with the guideRNA or with ADAR2 alone (Figure 4A/B, (d and e)) was unable to rescue the clustering phenotype, even though some clustering has been found occasionally in a low number of the cells (≤15%). Consequently, the presence of both components of the editing system, the guideRNA and ADAR2, are required to rescue the phenotype to the level seen for the wild-type control. To asses editing efficiency, total RNA was extracted from cells, reverse transcribed with PINK1-specific primers and was Sanger-sequenced after PCR amplification. Only in presence of both components, guideRNA and ADAR2, editing of the PINK1 W437amber codon was detectable (ca. 10% yield). The formation of full-length PINK1 after editing was further confirmed by Western blot (Supplementary Figure S21). To ensure that the rescue of Parkin-clustering was due to the repair of the PINK1 W437amber codon, we co-transfected HeLa KO cells with Parkin and the editing vector, but in absence of the editing target PINK1 W437X. As expected no rescue was detectable (Figure 4A, (f)). Obviously, restoring the loss-of-function phenotype was only mediated by site-directed RNA editing of the amber Stop codon of the transfected PINK1 W437X construct. Finally, we provide data for the PINK1/Parkin-mediated mitophagy, which is restored after editing of PINK1 W437amber (Figure 4A). For this, the mitochondria of all cells were stained after 0 h, 2 h and 24 h of CCCP with Mito-Tracker (Red CMXRos). As before, after 2 h of CCCP treatment Parkin did clearly co-localize with the perinuclear mitochondria. After 24 h of CCCP treatment, mitochondrial staining has disappeared in PINK1/Parkin/ADAR2 triple-positive cells but not in cells that lack the expression of one or several of the transfected components (Figure 4C, Supplementary Figure S23). This confirms that the two stages of mitophagy, (41,51) the Parkin-clustering (after 2 h of depolarization) and the clearance of depolarized mitochondria (after 24 h of depolarization), are restored upon editing of PINK1 W437amber.

CONCLUSION

We demonstrate here the first strategy to harness wild-type human ADAR2 to stimulate site-selective RNA editing at arbitrary mRNAs. The strategy relies on the ectopic expression of short, structured guideRNAs that base-pair to specific, user-defined mRNAs thereby mimicking the intronic R/G-motif of the glutamate receptor transcript to recruit wild-type human ADAR2 to stimulate A-to-I conversion. In the PCR tube, we demonstrated the functionality of the tool and the dependency on the structured R/G-motif. Selective and efficient editing was achieved even though the tendency for off-site editing required optimization of the reaction conditions and was not entirely controllable. As the guideRNAs can be expressed from an RNA polymerase III promoter, the strategy is fully genetically encodable. We have demonstrated the specific and efficient (up to 40%) editing of a fluorescent reporter gene in 293T cells. Editing benefited from the presence of the structured R/G-motif in the guideRNA, and the transfection of single-stranded, chemically stabilized guideRNAs (with a binding site for the target mRNA up to 21 nt) were not as efficient as our guideRNAs with a 16 nt mRNA binding site plus the additional R/G-motif. In contrast to *in vitro* editing, no off-target editing was observed in the reporter gene within cells once the amount of ADAR2 plasmid was optimized. The lower amount of ADAR2 and the RNP-landscape inside the living cell possibly protects the transcript from over-editing without suppressing editing at the targeted adenosine. Editing was further improved to a yield of ∼50% by optimizing the guideRNA architecture. In particular the placement of the targeted adenosine with respect to the R/G-motif was critical. The optimal architecture turned out to deviate slightly from the natural R/G-site of GluR2. The insertion of additional nucleotides seems reasonable to accommodate the exit of the mRNA's 3′-part (28,43). Site-directed RNA editing worked efficiently in a created 293T cell line that expresses human ADAR2 from the genome under the control of doxycycline. Even though ADAR2 mRNA-levels were >20-fold reduced compared to transient expression conditions, improved editing yields (up to 65%) were achieved. Furthermore, off-site editing in the reporter was abolished and the amount of R/G-guideRNA could be reduced. Finally, we demonstrated with a set of 13 sites in six housekeeping genes that ADAR2 can be re-directed for the editing of endogenously expressed transcripts. Importantly, this was also successful by ectopic expression of the guideRNA alone into the engineered 293 cells that express ADAR2 moderately from a single genomic copy. However, as the guideRNAs have not been optimized for these targets, editing yields stayed in a range of 10–35%. Further optimization of the guideRNAs may focus on the R/G-motif, the mRNA binding-site, and chemical modifications (29). In this work, we have mostly focused on the editing of 5′-UAG triplets. We expect many other triplets to be well editable, (42) an example for 5′-CAG is shown in the Supporting Information (Supplementary Figure S15). However, one can expect different triplets to be unequally well edited corresponding to the well-known preferences of the deaminase domain (52).

To demonstrate the usefulness of the editing approach in biological applications, we generated an editing vector that contains one copy of human ADAR2 and five copies of the R/G-guideRNA. This vector allowed us to define the stoichiometry of ADAR2 and the R/G-guideRNA in the transfected cell and to strongly reduce the amount of plasmid. We showed the functioning of the vector for the repair of the W58X GFP reporter. The respective editing vector against PINK1 W437X was functional to rescue the PINK1-mediated perinuclear clustering of Parkin and the mitophagy phenotype in PINK1-KO HeLa cells under transient transfection of the editing target. We could clearly demonstrate the rescue to require the presence of the entire editing machinery, the guideRNA and ADAR2, as well as the presence of the editing target PINK1 W437X. Editing was specific and no other A-to-I editing event was detectable in the Sanger sequencing trace, which covered ca. 600 nt (140 adenosines) of the PINK1 transcript (Supplementary Figure S22). However, to assess how the guideRNA-dependent harnessing of ADAR2 affects the editing homeostasis at natural editing sites, transcriptome-wide deep RNA-sequencing would be required, preferentially in

an *in vivo* situation where endogenous ADAR2 function is detectable and required. The editing yield of 10% for PINK1 W437X can surely be further improved, for instance by optimizing the co-transfection conditions in HeLa cells. The apparent editing yield over the entire cell culture suffers from the expression of PINK1 in the absence of the editing components. Thus, we expect an editing yield higher than 10% in those cells that actually express the editing machinery. Anyway, the rescue of this recessive loss-of-function mutation in PINK1 may indeed only require a relatively moderate repair yield.

Together, we have shown that harnessing human ADAR2 for site-directed RNA editing allows recoding mRNAs to levels high enough to manipulate disease-relevant cellular phenotypes. Other endogenous RNA-processing enzymes, including RNaseH and the RNA-induced silencing complex (RNA interference), have been shown to be re-addressable toward new targets by ectopic expression or administration of short (chemically stabilized) guideRNAs (53,54). The structured R/G-guideRNAs introduced here are capable of recruiting human ADAR2. It remains yet unclear if they are able to recruit endogenously expressed ADAR2 for site-directed RNA editing. Certainly, they are a good starting point for the development of improved guideRNA architectures in the future. While RNAi and RNaseH recruitment are limited to the up/down-regulation of target transcripts, RNA editing allows for the active recoding and hence enables an entirely novel point of attack. Currently, there is increasing success in the tailoring of oligonucleotide drugs with respect to their efficacy, stability, toxicology and delivery (55). Major breakthrough seems to have happened in the therapeutic RNA interference field currently (31,56,57). We feel confident that our work sets the stage for the development of a novel therapeutic platform that is based on the harnessing of human ADARs for the repair of disease-relevant genes by either ectopic expression or administration of chemically stabilized (29) guideRNAs.

SUPPLEMENTARY DATA

Supplementary Data are available at NAR Online.

ACKNOWLEDGEMENTS

The CRISPR/Cas9 plasmid (pSpCas9(PINK1-KO)-2A-Puro) was generously provided by Christine Bus (University of Tübingen, Germany).

FUNDING

University of Tübingen, the Hertie Foundation, the German Center for Neurodegenerative Disease and the Deutsche Forschungsgemeinschaft [STA 1053/3-2 to T.S., GE 2844/1-1 to S.G.); European Research Council (ERC) under the European Union's Horizon 2020 research and innovation program [grant no. 647328]. Funding for open access charge: DFG.

Conflict of interest statement. The Authors have filed a patent on the technology described in this article.

REFERENCES

1. Nishikura,K. (2010) Functions and Regulation of RNA Editing by ADAR Deaminases. *Annu. Rev. Biochem.*, **79**, 321–349.
2. Kawahara,Y., Zinshteyn,B., Sethupathy,P., Iizasa,H., Hatzigeorgiou,A.G. and Nishikura,K. (2007) Redirection of silencing targets by adenosine-to-inosine editing of miRNAs. *Science*, **315**, 1137–1140.
3. Yang,W., Chendrimada,T.P., Wang,Q., Higuchi,M., Seeburg,P.H., Shiekhattar,R. and Nishikura,K. (2006) Modulation of microRNA processing and expression through RNA editing by ADAR deaminases. *Nat. Struct. Mol. Biol.*, **13**, 13–21.
4. Rueter,S.M., Dawson,T.R. and Emeson,R.B. (1999) Regulation of alternative splicing by RNA editing. *Nature*, **399**, 75–80.
5. Rueter,S.M., Burns,C.M., Coode,S.A., Mookherjee,P. and Emeson,R.B. (1995) Glutamate receptor RNA editing in vitro by enzymatic conversion of adenosine to inosine. *Science*, **267**, 1491–1494.
6. Melcher,T., Maas,S., Herb,A., Sprengel,R., Seeburg,P.H. and Higuchi,M. (1996) A mammalian RNA editing enzyme, *Nature*, **379**, 460–464.
7. Bass,B.L. (2002) RNA editing by Adenosine Deaminases that act on RNA. *Annu. Rev. Biochem.*, **71**, 817–846.
8. Morabito,M.V., Abbas,A.I., Hood,J.L., Kesterson,R.A., Jacobs,M.M., Kump,D.S., Hachey,D.L., Roth,B.L. and Emeson,R.B. (2010) Mice with altered serotonin 2C receptor RNA editing display characteristics of Prader-Willi syndrome. *Neurobiol. Dis.*, **39**, 169–180.
9. Maas,S., Kawahara,Y., Tamburro,K.M. and Nishikura,K. (2006) A-to-I RNA editing and human disease. *RNA Biol.*, **3**, 1–9.
10. Slotkin,W. and Nishikura,K. (2013) Adenosine-to-inosine RNA editing and human disease. *Genome Med.*, **5**, 105–118.
11. Silberberg,G., Lundin,D., Navon,R. and Öhman,M. (2012) Deregulation of the A-to-I RNA editing mechanism in psychiatric disorders. *Hum. Mol. Genet.*, **21**, 311–321.
12. Higuchi,M., Maas,S., Single,F.N., Hartner,J., Rozov,A., Burnashev,N., Feldmeyer,D., Sprengel,R. and Seeburg,P.H. (2000) Point mutation in an AMPA receptor gene rescues lethality in mice deficient in the RNA-editing enzyme ADAR2. *Nature*, **406**, 78–81.
13. Rice,G.I., Kasher,P.R., Forte,G.M., Mannion,N.M., Greenwood,S.M., Szynkiewicz,M., Dickerson,J.E., Bhaskar,S.S., Zampini,M., Briggs,T.A. *et al.* (2012) Mutations in ADAR1 cause Aicardi-Goutières syndrome associated with a type I interferon signature. *Nat. Genet.*, **44**, 1243–1248.
14. Zhang,X.J., He,P.P., Li,M., He,C.D., Yan,K.L., Cui,Y., Yang,S., Zhang,K.Y., Gao,M., Chen,J.J. *et al.* (2004) Seven novel mutations of the ADAR gene in Chinese families and sporadic patients with dyschromatosis symmetrica hereditaria (DSH). *Hum. Mutat.*, **23**, 629–630.
15. Hartner,J.C., Schmittwolf,C., Kispert,A., Müller,A.M., Higuchi,M. and Seeburg,P.H. (2004) Liver disintegration in the mouse embryo caused by deficiency in the RNA-editing enzyme ADAR1. *J. Biol. Chem.*, **279**, 4894–4902.
16. Wang,Q., Miyakoda,M., Yang,W., Khillan,J., Stachura,D.L., Weiss,M.J. and Nishikura,K. (2004) Stress-induced apoptosis associated with null mutation of ADAR1 RNA editing deaminase gene. *J. Biol. Chem.*, **279**, 4952–4961.
17. Liddicoat,B.J., Piskol,R., Chalk,A.M., Ramaswami,G., Higuchi,M., Hartner,J.C., Li,J.B., Seeburg,P.H. and Walkley,C.R. (2015) RNA editing by ADAR1 prevents MDA5 sensing of endogenous dsRNA as nonself. *Science*, **349**, 1115–1120.
18. Mannion,N.M., Greenwood,S.M., Young,R., Cox,S., Brindle,J., Read,D., Nellaker,C., Vesely,C., Ponting,C.P., McLaughlin,P.J. *et al.* (2014) The RNA-editing enzyme ADAR1 controls innate immune responses to RNA. *Cell Rep.*, **9**, 1482–1494.
19. Chen,L., Li,Y., Lin,C.H., Chan,T.H., Chow,R.K., Song,Y., Liu,M., Yuan,Y.F., Fu,L., Kong,K.L. *et al.* (2013) Recoding RNA editing of *antizyme inhibitor 1* predisposes to hepatocellular carcinoma. *Nat. Med.*, **19**, 209–216.
20. Shimokawa,T., Rahman,M.F., Tostar,U., Sonkoly,E., Ståhle,M., Pivarcsi,A., Palaniswamy,R. and Zaphiropoulos,P.G. (2013) RNA editing of the GLI1 transcription factor modulates the output of Hedgehog signaling. *RNA Biol.*, **10**, 321–333.

21. Gallo,A. (2013) RNA editing enters the limelight in cancer. *Nat. Med.*, **19**, 130–131.
22. Paz-Yaacov,N., Bazak,L., Buchumenski,I., Porath,H.T., Danan-Gotthold,M., Knisbacher,B.A., Eisenberg,E. and Levanon,E.Y. (2015) Elevated RNA Editing Activity Is a Major Contributor to Transcriptomic Diversity in Tumors. *Cell Reports*, **13**, 267–276.
23. Han,L., Diao,L., Yu,S., Xu,X., Li,J., Zhang,R., Yang,Y., Werner,H.M., Eterovic,A.K., Yuan,Y., Li,J. *et al.* (2015) The genomic landscape and clinical relevance of A-to-I RNA editing in human cancers. *Cancer Cell*, **28**, 515–528.
24. Picardi,E., Manzari,C., Mastropasqua,F., Aiello,I., D'Erchia,A.M. and Pesole,G. (2015) Profiling RNA editing in human tissues: Towards the inosinome Atlas. *Sci. Rep.*, **5**, 14941–14956.
25. Peng,Z., Cheng,Y., Tan,B.C., Kang,L., Tian,Z., Zhu,Y., Zhang,W., Liang,Y., Hu,X., Tan,X. *et al.* (2012) Comprehensive analysis of RNA-Seq data reveals extensive RNA editing in a human transcriptome. *Nat. Biotech.*, **30**, 253–260.
26. Bazak,L., Haviv,A., Barak,M., Jacob-Hirsch,J., Deng,P., Zhang,R., Isaacs,F.J., Rechavi,G., Li,J.B., Eisenberg,E. *et al.* (2014) A-to-I RNA editing occurs at over a hundred million genomic sites, located in a majority of human genes. *Genome Res.*, **24**, 365–376.
27. Ramaswami,G., Lin,W., Piskol,R., Tan,M.H., Davis,C. and Li,J.B. (2012) Accurate identification of human *Alu* and non-*Alu* RNA editing sites. *Nat. Meth.*, **9**, 579–581.
28. Stefl,R., Oberstrass,F.C., Hood,J.L., Jourdan,M., Zimmermann,M., Skrisovska,L., Maris,C., Peng,L., Hofr,C., Emeson,R.B. *et al.* (2010) The solution structure of the ADAR2 dsRBM-RNA complex reveals a sequence-specific readout of the minor groove. *Cell*, **143**, 225–237.
29. Vogel,P. and Stafforst,T. (2014) Site-directed RNA editing with Antagomir Deaminases — A tool to study protein and RNA function. *ChemMedChem*, **9**, 2021–2025.
30. Cox,D.B., Platt,R.J. and Zhang,F. (2015) Therapeutic genome editing: prospects and challenges. *Nat. Med.*, **21**, 121–131.
31. Sehgal,A., Barros,S., Ivanciu,L., Cooley,B., Qin,J., Racie,T., Hettinger,J., Carioto,M., Jiang,Y., Brodsky,J. *et al.* (2015) An RNAi therapeutic targeting antithrombin to rebalance the coagulation system and promote hemostasis in hemophilia. *Nat. Med.*, **21**, 491–497.
32. Stern,M. and Alton,E.W.F.W. (2002) Use of liposomes in the treatment of cystic fibrosis. In: Albeda,SM (ed). *Gene Therapy in Lung Disease: Lung Biology in Health and Disease*. CRC Press, NY, Vol. **169**, pp. 383–396.
33. Woolf,T.M., Chase,J.M. and Stinchcomb,D.T. (1995) Towards the therapeutic editing of mutated RNA sequences. *Proc. Natl. Acad. Sci. U.S.A.*, **92**, 8298–8302.
34. Stafforst,T. and Schneider,M.F. (2012) An RNA–Deaminase conjugate selectively repairs point mutations. *Angew. Chem. Int. Ed. Engl.*, **51**, 11166–11169.
35. Vogel,P., Schneider,M.F., Wettengel,J. and Stafforst,T. (2014) Improving site-directed RNA editing in vitro and in cell culture by chemical modification of the guideRNA. *Angew. Chem. Int. Ed. Engl.*, **53**, 6267–6271.
36. Montiel-Gonzalez,M. F., Guillermo,I., Yudowski,A. and Rosenthal,J. J. C. (2013) Correction of mutations within the cystic fibrosis transmembrane conductance regulator by site-directed RNA editing. *Proc. Natl. Acad. Sci. U.S.A.*, **110**, 18285–18290.
37. Hanswillemenke,A., Kuzdere,T., Vogel,P., Jékely,G. and Stafforst,T. (2015) Site-directed RNA editing in vivo can be triggered by the light-driven assembly of an artificial riboprotein. *J. Am. Chem. Soc.*, **137**, 15875–15881.
38. World Health Organization (WHO). (2006) *Neurological disorders: Public health challenges*, Geneva, ISBN 9241563362.
39. Farrer,M.J. (2006) Genetics of Parkinson disease: paradigm shifts and future prospects. *Nat. Rev. Gen.*, **7**, 306–318.
40. Valente,E.M., Abou-Sleiman,P.M., Caputo,V., Muqit,M.M., Harvey,K., Gispert,S., Ali,Z., Del Turco,D., Bentivoglio,A.R., Healy,D.G. *et al.* (2004) Hereditary early-onset Parkinson's disease caused by mutations in PINK1. *Science*, **304**, 1158–1160.
41. Geisler,S., Holmström,K.M., Skujat,D., Fiesel,F.C., Rothfuss,O.C., Kahle,P.J. and Springer,W. (2010) PINK1/Parkin-mediated mitophagy is dependent on VDAC1 and p62/SQSTM1. *Nat. Cell Biol.*, **12**, 119–131.
42. Schneider,M.F., Wettengel,J., Hoffmann,P.C. and Stafforst,T. (2014) Optimal guideRNAs for re-directing deaminase activity of hADAR1 and hADAR2 in trans. *Nucleic Acids Res.*, **42**, e87.
43. Matthews,M.M., Thomas,J.M., Zheng,Y., Tran,K., Phelps,K.J., Scott,A.I., Havel,J., Fisher,A.J. and Beal,P.A. (2016) Structures of human ADAR2 bound to dsRNA reveal base-flipping mechanism and basis for site selectivity. *Nat. Struct. Mol. Biol.*, **23**, 426–433.
44. Persson,T., Hartmann,R. K. and Eckstein,F. (2002) Selection of hammerhead ribozyme variants with low Mg2+ requirement: importance of stem-loop II. *ChemBioChem*, **3**, 1066–1071.
45. Miyagishi,M. and Taira,K. (2002) U6 promoter-driven siRNAs with four uridine 3′ overhangs efficiently suppress targeted gene expression in mammalian cells. *Nat. Biotech.*, **20**, 497–500.
46. Lee,N. S., Dohjima,T., Bauer,G., Li,H., Li,M.J., Ehsani,A., Salvaterra,P. and Rossi,J.(2002) Expression of small interfering RNAs targeted against HIV-1 rev transcripts in human cells. *Nat. Biotech.*, **20**, 500–505.
47. Paul,C. P., Good,P. D., Winer,I. and Engelke,D. R. (2002) Effective expression of small interfering RNA in human cells. *Nat. Biotech.*, **20**, 505–508.
48. Daigle,N. and Ellenberg,J. (2007) LambdaN-GFP: an RNA reporter system for live-cell imaging. *Nat. Methods*, **4**, 633–636.
49. Keryer-Bibens,C., Barreau,C. and Osborne,H. B. (2008) Tethering of proteins to RNAs by bacteriophage proteins. *Biol. Cell*, **100**, 125–138.
50. Jinek,M., Chylinski,K., Fonfara,I., Hauer,M., Doudna,J.A. and Charpentier,E. (2012) A programmable dual-RNA-guided DNA endonuclease in adaptive bacterial immunity. *Science*, **337**, 816–821.
51. Vives-Bauza,C., Zhou,C., Huang,Y., Cui,M., de Vries,R.L., Kim,J., May,J., Tocilescu,M.A., Liu,W., Ko,H.S. *et al.* (2010) PINK1-dependent recruitment of Parkin to mitochondria in mitophagy. *Proc. Natl. Acad. Sci. U.S.A.*, **107**, 378–383.
52. Eggington,J.M., Greene,T. and Bass,B.L. (2011) Predicting sites of ADAR editing in double-stranded RNA. *Nat. Commun.*, **2**, doi:10.1038/ncomms1324.
53. Bennett,C.F. and Swayze,E.E. (2010) RNA targeting therapeutics: molecular mechanisms of antisense oligonucleotides as a therapeutic platform. *Annu. Rev. Pharmacol. Toxicol.*, **50**, 259–293.
54. Kole,R., Krainer,A.R. and Altman,S. (2012) RNA therapeutics: Beyond RNA interference and antisense oligonucleotides. *Nat. Rev. Drug. Discov.*, **11**, 125–140.
55. Reautschnig,P., Vogel,P. and Stafforst,T. (2016) The notorious R.N.A. in the spotlight – drug or target for the treatment of disease. *RNA Biol.*, doi:10.1080/15476286.2016.1208323.
56. Coelho,T., Adams,D., Silva,A., Lozeron,P., Hawkins,P.N., Mant,T., Perez,J., Chiesa,J., Warrington,S., Tranter,E. *et al.* (2013) Safety and efficacy of RNAi therapy for transthyretin amyloidosis. *N. Engl. J. Med.*, **369**, 819–829.
57. Fitzgerald,K., Frank-Kamenetsky,M., Shulga-Morskaya,S., Liebow,A., Bettencourt,B.R., Sutherland,J.E., Hutabarat,R.M., Clausen,V.A., Karsten,V., Cehelsky,J. *et al.* (2014) Effect of an RNA interference drug on the synthesis of proprotein convertase subtilisin/kexin type 9 (PCSK9) and the concentration of serum LDL cholesterol in healthy volunteers: A randomised, single-blind, placebo-controlled, phase 1 trial. *Lancet.*, **383**, 60–68.

Supporting Information

Harnessing human ADAR2 for RNA repair – Recoding a PINK1 mutation rescues mitophagy

Jacqueline Wettengel,[1,a] Philipp Reautschnig,[1,a] Sven Geisler,[2,3] Philipp J. Kahle[2,3] and Thorsten Stafforst[1]*

[1] Interfaculty Institute of Biochemistry, University of Tübingen

Auf der Morgenstelle 15; 72076 Tübingen (Germany)

[2] Department for Neurodegenerative Diseases, Hertie Institute for Clinical Brain Research, University of Tübingen

Otfried-Müller-Strasse 27; 72076 Tübingen (Germany)

[3] German Center for Neurodegenerative Diseases,

Otfried-Müller-Strasse 23; 72076 Tübingen (Germany)

wt hADAR2 protein production from yeast

Gene and protein sequence of the produced wt hADAR2-His6, the expression was controlled by pGal promotor and cyc1 terminator, similar as described before for SNAP-ADAR(1-3). A His_6-tag was directly cloned after ADAR2.

```
                10        20        30        40        50        60
1         GAGAAAAAACCCCGGATTCTAGAATGGATATAGAAGATGAAGAAAACATGAGTTCCAGCA
1                         Xba1   M  D  I  E  D  E  E  N  M  S  S  S

                70        80        90       100       110       120
61        GCACTGATGTGAAGGAAAACCGCAATCTGGACAACGTGTCCCCCAAGGATGGCAGCACAC
20        S  T  D  V  K  E  N  R  N  L  D  N  V  S  P  K  D  G  S  T

               130       140       150       160       170       180
121       CTGGGCCTGGCGAGGGCTCTCAGCTCTCCAATGGGGGTGGTGGTGGCCCCGGCAGAAAGC
40        P  G  P  G  E  G  S  Q  L  S  N  G  G  G  G  G  P  G  R  K

               190       200       210       220       230       240
181       GGCCCCTGGAGGAGGGCAGCAATGGCCACTCCAAGTACCGCCTGAAGAAAAGGAGGAAAA
60        R  P  L  E  E  G  S  N  G  H  S  K  Y  R  L  K  K  R  R  K

               250       260       270       280       290       300
241       CACCAGGGCCCGTCCTCCCCAAGAACGCCCTGATGCAGCTGAATGAGATCAAGCCTGGTT
80        T  P  G  P  V  L  P  K  N  A  L  M  Q  L  N  E  I  K  P  G

               310       320       330       340       350       360
301       TGCAGTACACACTCCTGTCCCAGACTGGGCCCGTGCACGCGCCTTTGTTTGTCATGTCTG
100       L  Q  Y  T  L  L  S  Q  T  G  P  V  H  A  P  L  F  V  M  S

               370       380       390       400       410       420
361       TGGAGGTGAATGGCCAGGTTTTTGAGGGCTCTGGTCCCACAAAGAAAAAGGCAAAACTCC
120       V  E  V  N  G  Q  V  F  E  G  S  G  P  T  K  K  K  A  K  L

               430       440       450       460       470       480
421       ATGCTGCTGAGAAGGCCTTGAGGTCTTTCGTTCAGTTTCCTAATGCCTCTGAGGCCCACC
140       H  A  A  E  K  A  L  R  S  F  V  Q  F  P  N  A  S  E  A  H
```

```
                490       500       510       520       530       540
481       TGGCCATGGGGAGGACCCTGTCTGTCAACACGGACTTCACATCTGACCAGGCCGACTTCC
160       L  A  M  G  R  T  L  S  V  N  T  D  F  T  S  D  Q  A  D  F

                550       560       570       580       590       600
541       CTGACACGCTCTTCAATGGTTTTGAAACTCCTGACAAGGCGGAGCCTCCCTTTTACGTGG
180       P  D  T  L  F  N  G  F  E  T  P  D  K  A  E  P  P  F  Y  V

                610       620       630       640       650       660
601       GCTCCAATGGGGATGACTCCTTCAGTTCCAGCGGGGACCTCAGCTTGTCTGCTTCCCCGG
200       G  S  N  G  D  D  S  F  S  S  S  G  D  L  S  L  S  A  S  P

                670       680       690       700       710       720
661       TGCCTGCCAGCCTAGCCCAGCCTCCTCTCCCTGCCTTACCACCATTCCCACCCCCGAGTG
220       V  P  A  S  L  A  Q  P  P  L  P  A  L  P  P  F  P  P  P  S

                730       740       750       760       770       780
721       GGAAGAATCCCGTGATGATCTTGAACGAACTGCGCCCAGGACTCAAGTATGACTTCCTCT
240       G  K  N  P  V  M  I  L  N  E  L  R  P  G  L  K  Y  D  F  L

                790       800       810       820       830       840
781       CCGAGAGCGGGGAGAGCCATGCCAAGAGCTTCGTCATGTCTGTGGTCGTGGATGGTCAGT
260       S  E  S  G  E  S  H  A  K  S  F  V  M  S  V  V  V  D  G  Q

                850       860       870       880       890       900
841       TCTTTGAAGGCTCGGGGAGAAACAAGAAGCTTGCCAAGGCCCGGGCTGCGCAGTCTGCCC
280       F  F  E  G  S  G  R  N  K  K  L  A  K  A  R  A  A  Q  S  A

                910       920       930       940       950       960
901       TGGCCGCCATTTTTAACTTGCACTTGGATCAGACGCCATCTCGCCAGCCTATTCCCAGTG
300       L  A  A  I  F  N  L  H  L  D  Q  T  P  S  R  Q  P  I  P  S

                970       980       990      1000      1010      1020
961       AGGGTCTTCAGCTGCATTTACCGCAGGTTTTAGCTGACGCTGTCTCACGCCTGGTCCTGG
320       E  G  L  Q  L  H  L  P  Q  V  L  A  D  A  V  S  R  L  V  L

               1030      1040      1050      1060      1070      1080
1021      GTAAGTTTGGTGACCTGACCGACAACTTCTCCTCCCCTCACGCTCGCAGAAAAGTGCTGG
340       G  K  F  G  D  L  T  D  N  F  S  S  P  H  A  R  R  K  V  L

               1090      1100      1110      1120      1130      1140
1081      CTGGAGTCGTCATGACAACAGGCACAGATGTTAAAGATGCCAAGGTGATAAGTGTTTCTA
360       A  G  V  V  M  T  T  G  T  D  V  K  D  A  K  V  I  S  V  S

               1150      1160      1170      1180      1190      1200
1141      CAGGAACAAAATGTATTAATGGTGAATACATGAGTGATCGTGGCCTTGCATTAAATGACT
380       T  G  T  K  C  I  N  G  E  Y  M  S  D  R  G  L  A  L  N  D

               1210      1220      1230      1240      1250      1260
1201      GCCATGCAGAAATAATATCTCGGAGATCCTTGCTCAGATTTCTTTATACACAACTTGAGC
400       C  H  A  E  I  I  S  R  R  S  L  L  R  F  L  Y  T  Q  L  E

               1270      1280      1290      1300      1310      1320
1261      TTTACTTAAATAACAAAGATGATCAAAAAAGATCCATCTTTCAGAAATCAGAGCGAGGGG
420       L  Y  L  N  N  K  D  D  Q  K  R  S  I  F  Q  K  S  E  R  G

               1330      1340      1350      1360      1370      1380
1321      GGTTTAGGCTGAAGGAGAATGTCCAGTTTCATCTGTACATCAGCACCTCTCCCTGTGGAG
440       G  F  R  L  K  E  N  V  Q  F  H  L  Y  I  S  T  S  P  C  G

               1390      1400      1410      1420      1430      1440
```

```
1381      ATGCCAGAATCTTCTCACCACATGAGCCAATCCTGGAAGAACCAGCAGATAGACACCCAA
460       D  A  R  I  F  S  P  H  E  P  I  L  E  E  P  A  D  R  H  P

                1450      1460      1470      1480      1490      1500
1441      ATCGTAAAGCAAGAGGACAGCTACGGACCAAAATAGAGTCTGGTGAGGGGACGATTCCAG
480       N  R  K  A  R  G  Q  L  R  T  K  I  E  S  G  E  G  T  I  P

                1510      1520      1530      1540      1550      1560
1501      TGCGCTCCAATGCGAGCATCCAAACGTGGGACGGGGTGCTGCAAGGGGAGCGGCTGCTCA
500       V  R  S  N  A  S  I  Q  T  W  D  G  V  L  Q  G  E  R  L  L

                1570      1580      1590      1600      1610      1620
1561      CCATGTCCTGCAGTGACAAGATTGCACGCTGGAACGTGGTGGGCATCCAGGGATCCCTGC
520       T  M  S  C  S  D  K  I  A  R  W  N  V  V  G  I  Q  G  S  L

                1630      1640      1650      1660      1670      1680
1621      TCAGCATTTTCGTGGAGCCCATTTACTTCTCGAGCATCATCCTGGGCAGCCTTTACCACG
540       L  S  I  F  V  E  P  I  Y  F  S  S  I  I  L  G  S  L  Y  H

                1690      1700      1710      1720      1730      1740
1681      GGGACCACCTTTCCAGGGCCATGTACCAGCGGATCTCCAACATAGAGGACCTGCCACCTC
560       G  D  H  L  S  R  A  M  Y  Q  R  I  S  N  I  E  D  L  P  P

                1750      1760      1770      1780      1790      1800
1741      TCTACACCCTCAACAAGCCTTTGCTCAGTGGCATCAGCAATGCAGAAGCACGGCAGCCAG
580       L  Y  T  L  N  K  P  L  L  S  G  I  S  N  A  E  A  R  Q  P

                1810      1820      1830      1840      1850      1860
1801      GGAAGGCCCCCAACTTCAGTGTCAACTGGACGGTAGGCGACTCCGCTATTGAGGTCATCA
600       G  K  A  P  N  F  S  V  N  W  T  V  G  D  S  A  I  E  V  I

                1870      1880      1890      1900      1910      1920
1861      ACGCCACGACTGGGAAGGATGAGCTGGGCCGCGCGTCCCGCCTGTGTAAGCACGCGTTGT
620       N  A  T  T  G  K  D  E  L  G  R  A  S  R  L  C  K  H  A  L

                1930      1940      1950      1960      1970      1980
1921      ACTGTCGCTGGATGCGTGTGCACGGCAAGGTTCCCTCCCACTTACTACGCTCCAAGATTA
640       Y  C  R  W  M  R  V  H  G  K  V  P  S  H  L  L  R  S  K  I

                1990      2000      2010      2020      2030      2040
1981      CCAAACCCAACGTGTACCATGAGTCCAAGCTGGCGGCAAAGGAGTACCAGGCCGCCAAGG
660       T  K  P  N  V  Y  H  E  S  K  L  A  A  K  E  Y  Q  A  A  K

                2050      2060      2070      2080      2090      2100
2041      CGCGTCTGTTCACAGCCTTCATCAAGGCGGGGCTGGGGGCCTGGGTGGAGAAGCCCACCG
680       A  R  L  F  T  A  F  I  K  A  G  L  G  A  W  V  E  K  P  T

                2110      2120      2130      2140      2150      2160
2101      AGCAGGACCAGTTCTCACTCACGCCCCACCATCACCATCACCAT TAATAG TCGACCTCGA
700       E  Q  D  Q  F  S  L  T  P  H  H  H  H  H  H  *  *  Sal1

                2170      2180
2161      GTCATGTAATTAGTTATGTCACGC
720
```

R/G-guideRNA synthesis

Templates for *in vitro* transcription were obtained by Phusion PCR templated with a pMG211 vector that contains the guideRNA downstream of the hammer head cassette. The forward primer (5´-GGT CAGGCCCAGGTTCTCCG) was chosen in a way that additional 280 bp were included before the T7 promotor to improve subsequent agarose gel work-up. The backward primer (5´-ACTCTGTGCTGGGGTGGTGGG) was chosen such that the guideRNA ended cleanly with no additional overhanging nucleotides. Shown is the PCR template from the T7 promotor until its 3´-end:

```
1         GCGAAATTAA TACGACTCAC TATAGGGGAA TTGTGAGCGG ATAACAATTC CCCTCTAGAA
                    T7 promotor

61        ATAATTTTGT TTAACTTTAA GAAGGAGATA TACATATGGC TAGCTATTCC ACCTGATGAG
                                                        HH-casette

121       TTTTTACGAA ACGTTCCCGT GAGGGAACGT C*GTGGAATAG TATAACAATA TGCTAAATGT
                                            *=cut          R/G-guideRNA

181       TGTTATAGTA TCCCACCACC CCAGCACAGA GT
             R/G-guideRNA
```

After urea PAGE-purification the following 61 nt guideRNA results from iv-T7 transcription of the above construct:

```
1         GUGGAAUAGU AUAACAAUAU GCUAAAUGUU GUUAUAGUAU CCCACCACCC
*C* AGCACAGAG U
*C* = counter base
```

Urea/TBE 8%-PAGE-separation of a guideRNA synthesis

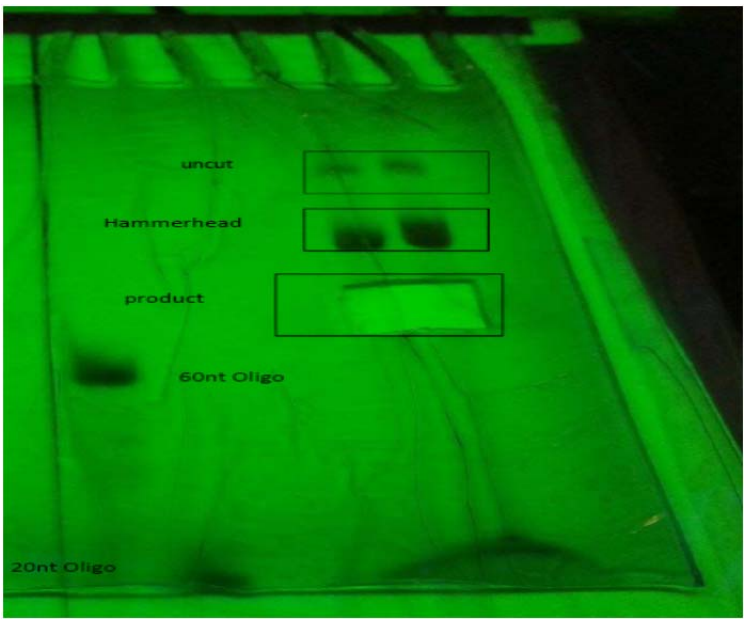


Figure S1. Preparative urea/TBE 8%-PAGE gel for the purification of the *in-vitro* transcribed R/G-guideRNA from the hammer head ribozyme. Indicated are the position of the uncut transcript (156 nt), the hammer head ribozyme (95 nt), the product (61 nt, cut out), and a 60 nt ssDNA and a 20 nt ssDNA as markers. The signal comes from UV shadowing on a TLC plate.

***In-vitro* editing**

RNA sequencing traces of the full eCFP ORF revealing different levels of off-target editing

Figure S2. Full sequencing trace corresponding to the trace shown in Figure 1B, a)

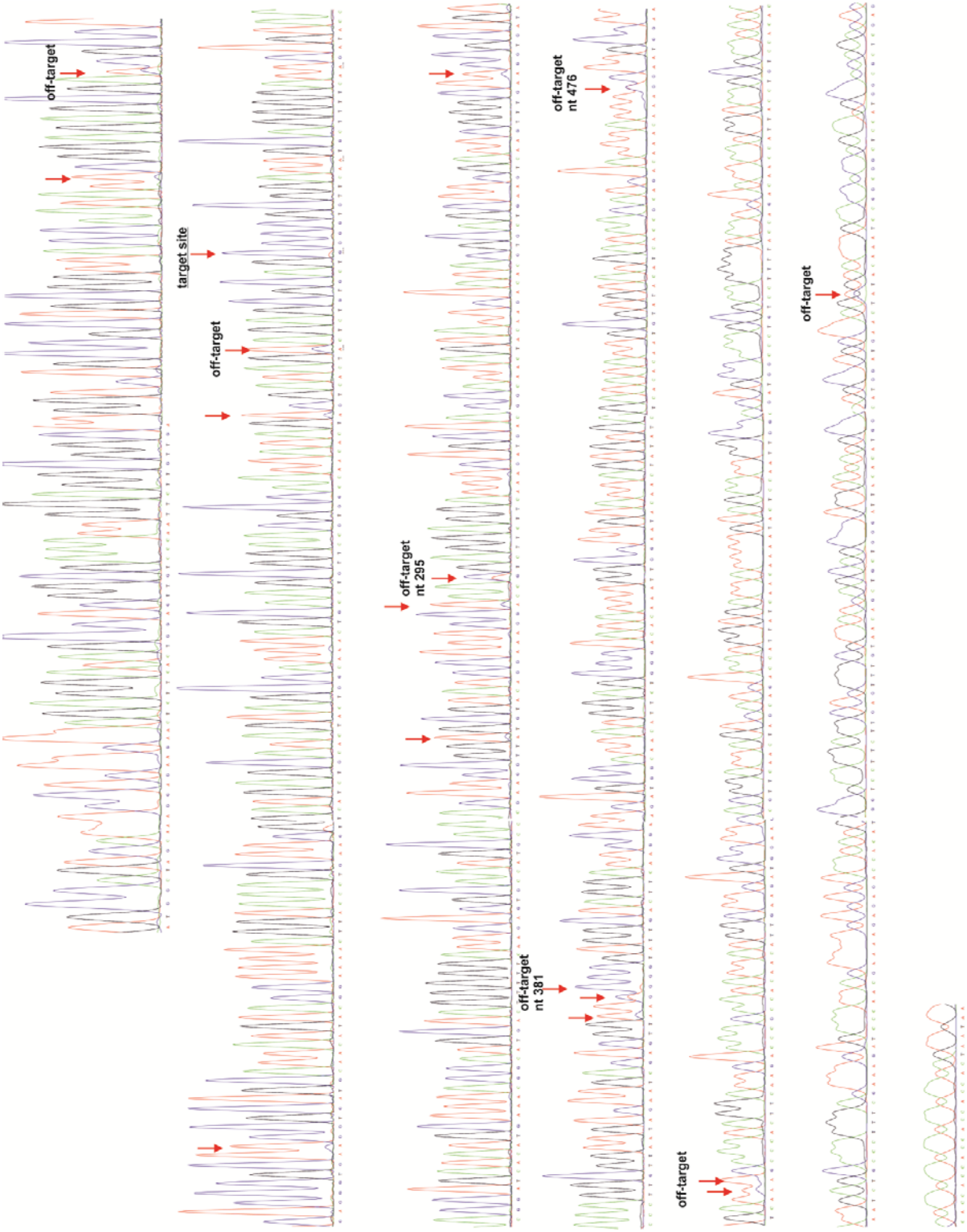


Figure S3. Full sequencing trace corresponding to the trace shown in Figure 1B, c)

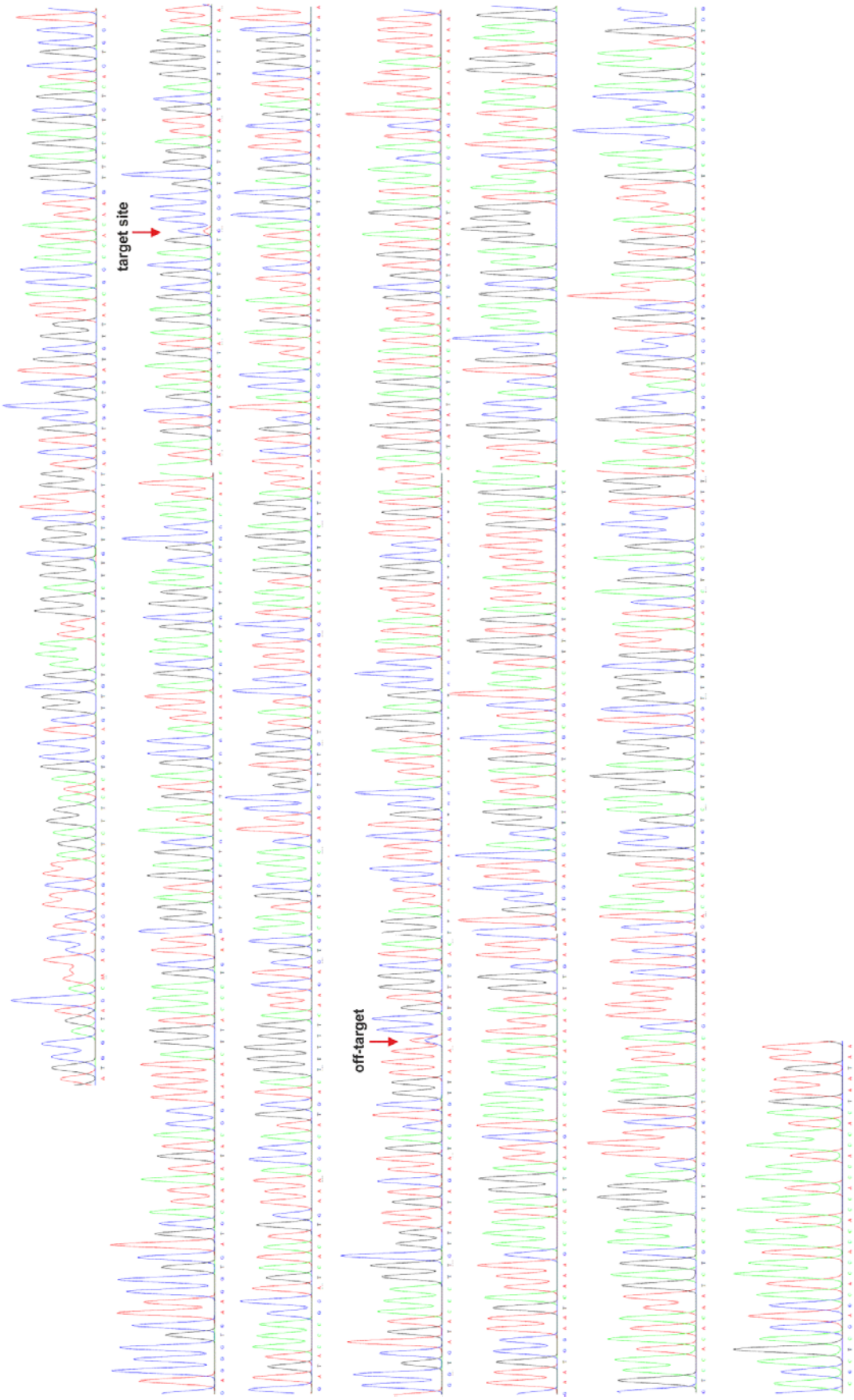


Figure S4. Full sequencing trace corresponding to the trace shown in Figure 1C, c)

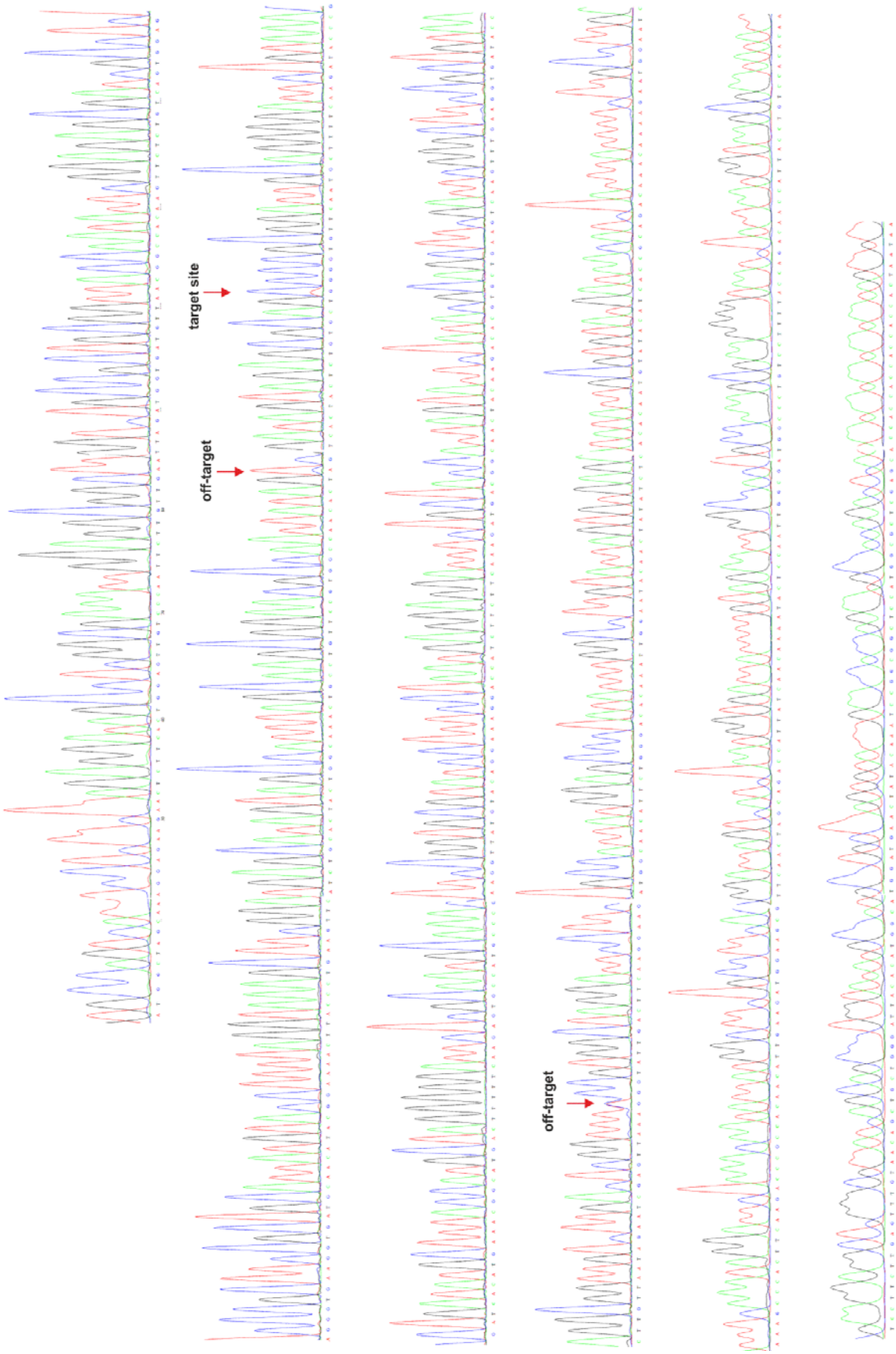
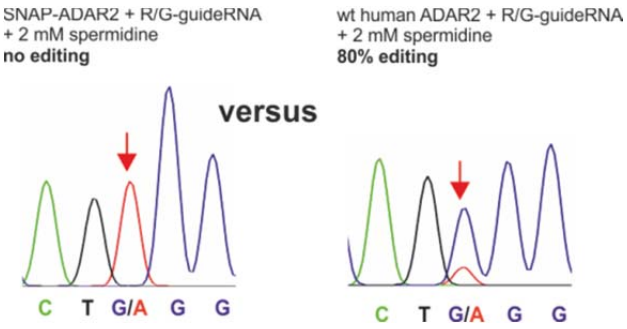


Figure S5. *Control experiment to experiments shown in Figure 1B. The R/G-guideRNA recruits only wt human ADAR2 but not SNAP-ADAR2 for editing. Conditions are identical to Figure 1B.*



Cellular editing

Gene & protein sequence of wt hADAR2 in the context of the pcDNA 3.1 vector, under control of the CMV promotor and BGH terminator:

```
                10        20        30        40        50        60
1          CTCGGATCCACCATGGATATAGAAGATGAAGAAAACATGAGTTCCAGCAGCACTGATGTG
1             BamHI     M  D  I  E  D  E  E  N  M  S  S  S  S  T  D  V

                70        80        90       100       110       120
61         AAGGAAAACCGCAATCTGGACAACGTGTCCCCCAAGGATGGCAGCACACCTGGGCCTGGC
21          K  E  N  R  N  L  D  N  V  S  P  K  D  G  S  T  P  G  P  G

               130       140       150       160       170       180
121        GAGGGCTCTCAGCTCTCCAATGGGGGTGGTGGTGGCCCCGGCAGAAAGCGGCCCCTGGAG
41          E  G  S  Q  L  S  N  G  G  G  G  G  P  G  R  K  R  P  L  E

               190       200       210       220       230       240
181        GAGGGCAGCAATGGCCACTCCAAGTACCGCCTGAAGAAAAGGAGGAAAACACCAGGGCCC
61          E  G  S  N  G  H  S  K  Y  R  L  K  K  R  R  K  T  P  G  P

               250       260       270       280       290       300
241        GTCCTCCCCAAGAACGCCCTGATGCAGCTGAATGAGATCAAGCCTGGTTTGCAGTACACA
81          V  L  P  K  N  A  L  M  Q  L  N  E  I  K  P  G  L  Q  Y  T

               310       320       330       340       350       360
301        CTCCTGTCCCAGACTGGGCCCGTGCACGCGCCTTTGTTTGTCATGTCTGTGGAGGTGAAT
101         L  L  S  Q  T  G  P  V  H  A  P  L  F  V  M  S  V  E  V  N

               370       380       390       400       410       420
361        GGCCAGGTTTTTGAGGGCTCTGGTCCCACAAAGAAAAAGGCAAAACTCCATGCTGCTGAG
121         G  Q  V  F  E  G  S  G  P  T  K  K  K  A  K  L  H  A  A  E

               430       440       450       460       470       480
421        AAGGCCTTGAGGTCTTTCGTTCAGTTTCCTAATGCCTCTGAGGCCCACCTGGCCATGGGG
141         K  A  L  R  S  F  V  Q  F  P  N  A  S  E  A  H  L  A  M  G

               490       500       510       520       530       540
481        AGGACCCTGTCTGTCAACACGGACTTCACATCTGACCAGGCCGACTTCCCTGACACGCTC
161         R  T  L  S  V  N  T  D  F  T  S  D  Q  A  D  F  P  D  T  L

               550       560       570       580       590       600
541        TTCAATGGTTTTGAAACTCCTGACAAGGCGGAGCCTCCCTTTTACGTGGGCTCCAATGGG
181         F  N  G  F  E  T  P  D  K  A  E  P  P  F  Y  V  G  S  N  G

               610       620       630       640       650       660
601        GATGACTCCTTCAGTTCCAGCGGGGACCTCAGCTTGTCTGCTTCCCCGGTGCCTGCCAGC
201         D  D  S  F  S  S  S  G  D  L  S  L  S  A  S  P  V  P  A  S

               670       680       690       700       710       720
661        CTAGCCCAGCCTCCTCTCCCTGCCTTACCACCATTCCCACCCCCGAGTGGGAAGAATCCC
221         L  A  Q  P  P  L  P  A  L  P  P  F  P  P  P  S  G  K  N  P

               730       740       750       760       770       780
721        GTGATGATCTTGAACGAACTGCGCCCAGGACTCAAGTATGACTTCCTCTCCGAGAGCGGG
241         V  M  I  L  N  E  L  R  P  G  L  K  Y  D  F  L  S  E  S  G

               790       800       810       820       830       840
781        GAGAGCCATGCCAAGAGCTTCGTCATGTCTGTGGTCGTGGATGGTCAGTTCTTTGAAGGC
261         E  S  H  A  K  S  F  V  M  S  V  V  V  D  G  Q  F  F  E  G
```

```
                850       860       870       880       890       900
841       TCGGGGAGAAACAAGAAGCTTGCCAAGGCCCGGGCTGCGCAGTCTGCCCTGGCCGCCATT
281        S  G  R  N  K  K  L  A  K  A  R  A  A  Q  S  A  L  A  A  I

                910       920       930       940       950       960
901       TTTAACTTGCACTTGGATCAGACGCCATCTCGCCAGCCTATTCCCAGTGAGGGTCTTCAG
301        F  N  L  H  L  D  Q  T  P  S  R  Q  P  I  P  S  E  G  L  Q

                970       980       990      1000      1010      1020
961       CTGCATTTACCGCAGGTTTTAGCTGACGCTGTCTCACGCCTGGTCCTGGGTAAGTTTGGT
321        L  H  L  P  Q  V  L  A  D  A  V  S  R  L  V  L  G  K  F  G

               1030      1040      1050      1060      1070      1080
1021      GACCTGACCGACAACTTCTCCTCCCCTCACGCTCGCAGAAAAGTGCTGGCTGGAGTCGTC
341        D  L  T  D  N  F  S  S  P  H  A  R  R  K  V  L  A  G  V  V

               1090      1100      1110      1120      1130      1140
1081      ATGACAACAGGCACAGATGTTAAAGATGCCAAGGTGATAAGTGTTTCTACAGGAACAAAA
361        M  T  T  G  T  D  V  K  D  A  K  V  I  S  V  S  T  G  T  K

               1150      1160      1170      1180      1190      1200
1141      TGTATTAATGGTGAATACATGAGTGATCGTGGCCTTGCATTAAATGACTGCCATGCAGAA
381        C  I  N  G  E  Y  M  S  D  R  G  L  A  L  N  D  C  H  A  E

               1210      1220      1230      1240      1250      1260
1201      ATAATATCTCGGAGATCCTTGCTCAGATTTCTTTATACACAACTTGAGCTTTACTTAAAT
401        I  I  S  R  R  S  L  L  R  F  L  Y  T  Q  L  E  L  Y  L  N

               1270      1280      1290      1300      1310      1320
1261      AACAAAGATGATCAAAAAAGATCCATCTTTCAGAAATCAGAGCGAGGGGGGTTTAGGCTG
421        N  K  D  D  Q  K  R  S  I  F  Q  K  S  E  R  G  G  F  R  L

               1330      1340      1350      1360      1370      1380
1321      AAGGAGAATGTCCAGTTTCATCTGTACATCAGCACCTCTCCCTGTGGAGATGCCAGAATC
441        K  E  N  V  Q  F  H  L  Y  I  S  T  S  P  C  G  D  A  R  I

               1390      1400      1410      1420      1430      1440
1381      TTCTCACCACATGAGCCAATCCTGGAAGAACCAGCAGATAGACACCCAAATCGTAAAGCA
461        F  S  P  H  E  P  I  L  E  E  P  A  D  R  H  P  N  R  K  A

               1450      1460      1470      1480      1490      1500
1441      AGAGGACAGCTACGGACCAAAATAGAGTCTGGTGAGGGGACGATTCCAGTGCGCTCCAAT
481        R  G  Q  L  R  T  K  I  E  S  G  E  G  T  I  P  V  R  S  N

               1510      1520      1530      1540      1550      1560
1501      GCGAGCATCCAAACGTGGGACGGGGTGCTGCAAGGGGAGCGGCTGCTCACCATGTCCTGC
501        A  S  I  Q  T  W  D  G  V  L  Q  G  E  R  L  L  T  M  S  C

               1570      1580      1590      1600      1610      1620
1561      AGTGACAAGATTGCACGCTGGAACGTGGTGGGCATCCAGGGTTCCCTGCTCAGCATTTTC
521        S  D  K  I  A  R  W  N  V  V  G  I  Q  G  S  L  L  S  I  F

               1630      1640      1650      1660      1670      1680
1621      GTGGAGCCCATTTACTTCTCGAGCATCATCCTGGGCAGCCTTTACCACGGGGACCACCTT
541        V  E  P  I  Y  F  S  S  I  I  L  G  S  L  Y  H  G  D  H  L

               1690      1700      1710      1720      1730      1740
1681      TCCAGGGCCATGTACCAGCGGATCTCCAACATAGAGGACCTGCCACCTCTCTACACCCTC
561        S  R  A  M  Y  Q  R  I  S  N  I  E  D  L  P  P  L  Y  T  L
```

```
              1750      1760      1770      1780      1790      1800
1741      AACAAGCCTTTGCTCAGTGGCATCAGCAATGCAGAAGCACGGCAGCCAGGGAAGGCCCCC
581        N  K  P  L  L  S  G  I  S  N  A  E  A  R  Q  P  G  K  A  P

              1810      1820      1830      1840      1850      1860
1801      AACTTCAGTGTCAACTGGACGGTAGGCGACTCCGCTATTGAGGTCATCAACGCCACGACT
601        N  F  S  V  N  W  T  V  G  D  S  A  I  E  V  I  N  A  T  T

              1870      1880      1890      1900      1910      1920
1861      GGGAAGGATGAGCTGGGCCGCGCGTCCCGCCTGTGTAAGCACGCGTTGTACTGTCGCTGG
621        G  K  D  E  L  G  R  A  S  R  L  C  K  H  A  L  Y  C  R  W

              1930      1940      1950      1960      1970      1980
1921      ATGCGTGTGCACGGCAAGGTTCCCTCCCACTTACTACGCTCCAAGATTACCAAACCCAAC
641        M  R  V  H  G  K  V  P  S  H  L  L  R  S  K  I  T  K  P  N

              1990      2000      2010      2020      2030      2040
1981      GTGTACCATGAGTCCAAGCTGGCGGCAAAGGAGTACCAGGCCGCCAAGGCGCGTCTGTTC
661        V  Y  H  E  S  K  L  A  A  K  E  Y  Q  A  A  K  A  R  L  F

              2050      2060      2070      2080      2090      2100
2041      ACAGCCTTCATCAAGGCGGGGCTGGGGGCCTGGGTGGAGAAGCCCACCGAGCAGGACCAG
681        T  A  F  I  K  A  G  L  G  A  W  V  E  K  P  T  E  Q  D  Q

              2110      2120      2130      2140      2150
2101      TTCTCACTCACGCCCTCTAGAGGGCCCTATTCTATAGTGTCACCTAAATGCTAG
701        F  S  L  T  P  S  R  G  P  Y  S  I  V  S  P  K  C  *
                        Xba-I
```

Gene and protein sequence of W58X eGFP in the context of the pcDNA 3.1 vector

```
                10        20        30        40        50        60
1         CTCGGATCCACCATGGCTAGCAAAGGAGAAGAACTCTTCACTGGAGTTGTCCCAATTCTT
1            BamH1     M  A  S  K  G  E  E  L  F  T  G  V  V  P  I  L

               70        80        90       100       110       120
61        GTTGAATTAGATGGTGATGTTAACGGCCACAAGTTCTCTGTCAGTGGAGAGGGTGAAGGT
21         V  E  L  D  G  D  V  N  G  H  K  F  S  V  S  G  E  G  E  G

              130       140       150       160       170       180
121       GATGCAACATACGGAAAACTTACCCTGAAGTTCATCTGCACTACTGGCAAACTGCCTGTT
41         D  A  T  Y  G  K  L  T  L  K  F  I  C  T  T  G  K  L  P  V

              190       200       210       220       230       240
181       CCGTAGCCGACACTAGTGACGACGCTCTGCTATGGCGTCCAGTGCTTTTCAAGATACCCG
61         P   *  P  T  L  V  T  T  L  C  Y  G  V  Q  C  F  S  R  Y  P
            W58x

              250       260       270       280       290       300
241       GATCACATGAAACGGCATGACTTTTTCAAGAGTGCCATGCCCGAAGGTTATGTACAGGAA
81         D  H  M  K  R  H  D  F  F  K  S  A  M  P  E  G  Y  V  Q  E

              310       320       330       340       350       360
301       AGGACCATCTTCTTCAAAGATGACGGCAACTACAAGACACGTGCTGAAGTCAAGTTTGAA
101        R  T  I  F  F  K  D  D  G  N  Y  K  T  R  A  E  V  K  F  E

              370       380       390       400       410       420
361       GGTGATACCCTTGTTAATAGAATCGAGTTAAAAGGTATTGACTTCAAGGAAGATGGCAAC
121        G  D  T  L  V  N  R  I  E  L  K  G  I  D  F  K  E  D  G  N
```

```
                430       440       450       460       470       480
421        ATTCTGGGACACAAATTGGAATACAACTATAACTCACACAATGTATACATCATGGCAGAC
141         I  L  G  H  K  L  E  Y  N  Y  N  S  H  N  V  Y  I  M  A  D

                490       500       510       520       530       540
481        AAACAAAAGAATGGAATCAAAGTGAACTTCAAGACCCGCCACAACATTGAAGATGGAAGC
161         K  Q  K  N  G  I  K  V  N  F  K  T  R  H  N  I  E  D  G  S

                550       560       570       580       590       600
541        GTTCAACTAGCAGACCATTATCAACAAAATACTCCAATTGGCGATGGCCCTGTCCTTTTA
181         V  Q  L  A  D  H  Y  Q  Q  N  T  P  I  G  D  G  P  V  L  L

                610       620       630       640       650       660
601        CCAGACAACCATTACCTGTCCACACAATCTGCCCTTTCGAAAGATCCCAACGAAAAGAGA
201         P  D  N  H  Y  L  S  T  Q  S  A  L  S  K  D  P  N  E  K  R

                670       680       690       700       710       720
661        GACCACATGGTCCTTCTTGAGTTTGTAACAGCTGCTGGGATTACACATGGCATGGATGAA
221         D  H  M  V  L  L  E  F  V  T  A  A  G  I  T  H  G  M  D  E

                730       740       750       760       770       780
721        CTATACAAATCCGGCTCTAGAGGGCCCTTCGAACAAAAACTCATCTCAGAAGAGGATCTG
241         L  Y  K  S  G  S  R  G  P  F  E  Q  K  L  I  S  E  E  D  L

                790       800       810       820       830       840
781        AATATGCATACCGGTCATCATCACCATCACCATTGAGTTTAAACCCGCTGATCAGCCTCG
261         N  M  H  T  G  H  H  H  H  H  H  *  V  *
```

Gene sequence of a 16 nt R/G-guideRNA against eGFP W58X in the context of the pSilencer vector

```
1          CGGAAGAGCG CCCAATACGC AAACCGCCTC TCCCCGCGCG TTGGCCGATT CATTAATGCA
              SapI

61         GCTGGCACGA CAGGTTTCCC GACTGGAAAG CGGGCAGTGA GCGCAACGCA ATTAATGTGA
121        GTTAGCTCAC TCATTAGGCA CCCCAGGCTT TACACTTTAT GCTTCCGGCT CGTATGTTGT
181        GTGGAATTGT GAGCGGATAA CAATTTCACA CAGGAAACAG CTATGACATG ATTACGAATT
241        GCAACGATTT AGGTGACACT ATAGAAGAGA AGGAATTAAT ACGACTCACT ATAGGGAGAG
301        AGAGAGAATT ACCCTCACTA AAGGGAGGAG AAGCATGAAT TCCCCAGTGG AAAGACGCGC
361        AGGCAAAACG CACCACGTGA CGGAGCGTGA CCGCGCGCCG AGCGCGCGCC AAGGTCGGGC
421        AGGAAGAGGG CCTATTTCCC ATGATTCCTT CATATTTGCA TATACGATAC AAGGCTGTTA
481        GAGAGATAAT TAGAATTAAT TTGACTGTAA ACACAAAGAT ATTAGTACAA AATACGTGAC
541        GTAGAAAGTA ATAATTTCTT GGGTAGTTTG CAGTTTTAAA ATTATGTTTT AAAATGGACT
601        ATCATATGCT TACCGTAACT TGAAAGTATT TCGATTTCTT GGGTTTATAT ATCTTGTGGA
661        AAGGACGCGG GATCC  *GTGGA ATAGTATAAC AATATGCTAA ATGTTGTTAT
AGTATCCCAC
                            * = transcription start   R/G-motif

721        TCGGCCACGG AACAGGTTTT TTGGAAAGCT TGG
              mRNA template U6-term.  HindIII
```

Table S1. *List of all R/G-gRNAs used in cell culture. Given is the mRNA binding site and the 3´-terminal hairpin (highlighted in gray) if applicable. The invariant R/G-motif is omitted for clarity.*

Name of R/G-gRNA	Sequence of the mRNA binding site 5'→3'	Experimental number
W58X GFP P6 16nt	UCGG**CCA**CGGAACAGG	Fig. 2B
W58X GFP P6 18nt + boxB	UCGG**CCA**CGGAACAGGCA*UCUAGAGGGCCCUGAAGAGGGCCC*	Fig. 2C / Fig. S10
W58X GFP P6 20nt + boxB	UCGG**CCA**CGGAACAGGCAGU*UCUAGAGGGCCCUGAAGAGGGCCC*	Fig. 2C / Fig. S10
W58X GFP P6 25nt + boxB	UCGG**CCA**CGGAACAGGCAGUUUGCC*UCUAGAGGGCCCUGAAGAGGGCCC*	Fig. 2C / Fig. S10
W58X GFP P6 29nt + boxB	UCGG**CCA**CGGAACAGGCAGUUUGCCAGUA*UCUAGAGGGCCCUGAAGAGGGCCC*	Fig. 2C / Fig. S10
W58X GFP P3 16 nt + boxB	GG**CCA**CGGAACAGGCA*UCUAGAGGGCCCUGAAGAGGGCCC*	Fig. 2D / Fig. S11
W58X GFP P4 16 nt + boxB	CGG**CCA**CGGAACAGGC*UCUAGAGGGCCCUGAAGAGGGCCC*	Fig. 2D / Fig. S11
W58X GFP P5 16 nt + boxB	UCGG**CCA**CGGAACAGG*UCUAGAGGGCCCUGAAGAGGGCCC*	Fig. 2D / Fig. S11
W58X GFP P7 16 nt + boxB	GUCGG**CCA**CGGAACAG*UCUAGAGGGCCCUGAAGAGGGCCC*	Fig. 2D / Fig. S11
W58X GFP P8 16 nt + boxB	UGUCGG**CCA**CGGAACA*UCUAGAGGGCCCUGAAGAGGGCCC*	Fig. 2D / Fig. S11
W58X GFP P9 16 nt + boxB	GUGUCGG**CCA**CGGAAC*UCUAGAGGGCCCUGAAGAGGGCCC*	Fig. 2D / Fig. S11
W58X GFP P10 16 nt + boxB	AGUGUCGG**CCA**CGGAA*UCUAGAGGGCCCUGAAGAGGGCCC*	Fig. 2D / Fig. S11
R407Q PINK-1 P4 16 nt + boxB	CC**CCG**AUCCACGUACC*UCUAGAGGGCCCUGAAGAGGGCCC*	Fig. S15
R407Q PINK-1 P5 16 nt + boxB	GCC**CCG**AUCCACGUAC*UCUAGAGGGCCCUGAAGAGGGCCC*	Fig. S15
R407Q PINK-1 P6 16 nt + boxB	CGCC**CCG**AUCCACGUA*UCUAGAGGGCCCUGAAGAGGGCCC*	Fig. S15
R407Q PINK-1 P7 16 nt + boxB	CCGCC**CCG**AUCCACGU*UCUAGAGGGCCCUGAAGAGGGCCC*	Fig. S15
R407Q PINK-1 P8 16 nt + boxB	UCCGCC**CCG**AUCCACG*UCUAGAGGGCCCUGAAGAGGGCCC*	Fig. S15
R407Q PINK-1 P9 16 nt + boxB	UUCCGCC**CCG**AUCCAC*UCUAGAGGGCCCUGAAGAGGGCCC*	Fig. S15
R407Q PINK-1 P10 16 nt + boxB	UUUCCGCC**CCG**AUCCA*UCUAGAGGGCCCUGAAGAGGGCCC*	Fig. S15
W437X Amber PINK-1 P8 16 nt + boxB	CACUGC**CCA**GGCAUCA*GGGCCCUCUUCAGGGCCC*	Fig. 3C / Fig. 3D / Fig. 4
3'UTR TAG#1 Actin P8 16 nt + boxB	ACGCAA**CCA**AGUCAUA*UCUAGAGGGCCCUGAAGAGGGCCC*	Fig. 3B
3'UTR TAG#2 Actin P8 16 nt + boxB	GAAUGA**CCA**UUAAAAA*UCUAGAGGGCCCUGAAGAGGGCCC*	Fig. 3B
3'UTR TAG#3 Actin P8 16 nt + boxB	GCAAUG**CCA**UCACCUC*UCUAGAGGGCCCUGAAGAGGGCCC*	Fig. 3B
3'UTR TAG#1 GAPDH P8 16 nt + boxB	AGGGGU**CCA**CAUGGCA*UCUAGAGGGCCCUGAAGAGGGCCC*	Fig. 3B
3'UTR TAG#2 GAPDH P8 16 nt	GGCUCC**CCA**GGCCCCU*UCUAGAGGGCCCUGAAGAGGGCCC*	Fig. 3B

+ boxB		
V627V TAG#1 GPI P8 16 nt + boxB	UGCCGU**CCA**CCAGGAU*UCUAGAGGGCCCUGAAGAGGGCCC*	Fig. 3B
L456L TAG#1 GusB P8 16 nt + boxB	CAGAUU**CCA**GGUGGGA*UCUAGAGGGCCCUGAAGAGGGCCC*	Fig. 3B
3'UTR TAG#2 GusB P8 16 nt + boxB	UCCCUG**CCA**GAAUAGA*UCUAGAGGGCCCUGAAGAGGGCCC*	Fig. 3B
3'UTR TAG#1 VCP P8 16 nt + boxB	CUCCGC**CCA**CCAAAUG*UCUAGAGGGCCCUGAAGAGGGCCC*	Fig. 3B
3'UTR TAG#2 VCP P8 16 nt + boxB	CCCAAA**CCA**CAACAGA*UCUAGAGGGCCCUGAAGAGGGCCC*	Fig. 3B
3'UTR TAG#3 VCP P8 16 nt + boxB	ACCCAC**CCA**CCCAGGU*UCUAGAGGGCCCUGAAGAGGGCCC*	Fig. 3B
3'UTR TAG#1 RAB7A P8 16 nt + boxB	CUGCCG**CCA**GCUGGAU*UCUAGAGGGCCCUGAAGAGGGCCC*	Fig. 3B
3'UTR TAG#2 RAB7A P8 16 nt + boxB	AGGGAA**CCA**GACAGUU*UCUAGAGGGCCCUGAAGAGGGCCC*	Fig. 3B

RNA sequencing traces of the full eGFP ORF

Figure S6. Full sequencing trace corresponding to the trace shown in Figure 2B, e), first experiment

Figure S7 *Full sequencing trace corresponding to the trace shown in Figure 2E, 50 ng experiment*

Figure S8: The prolongation of the editing time from 24 hours to 48 hours increases the editing yield in cell culture. *All editing samples included the amount of 300 ng W58X eGFP plasmid, 300 ng of ADAR2 plasmid and 1300 ng / 1600 ng of R/G-gRNA plasmid (c)-f)). The positive controls contained 300 ng eGFP plasmid, 300 ng of ADAR2 plasmid and 1300 ng of R/G-gRNA (a), b)). An increment of the fluorescent signal is obtained for the positive controls by prolonging the incubation time up to 48 hours (a-b), as well as for the two editing samples (c->e, d->f). The amount of fluorescing cells and the editing yields were increased for both editing samples. Total magnification: 100x, GFP exposure time: 50 ms.*

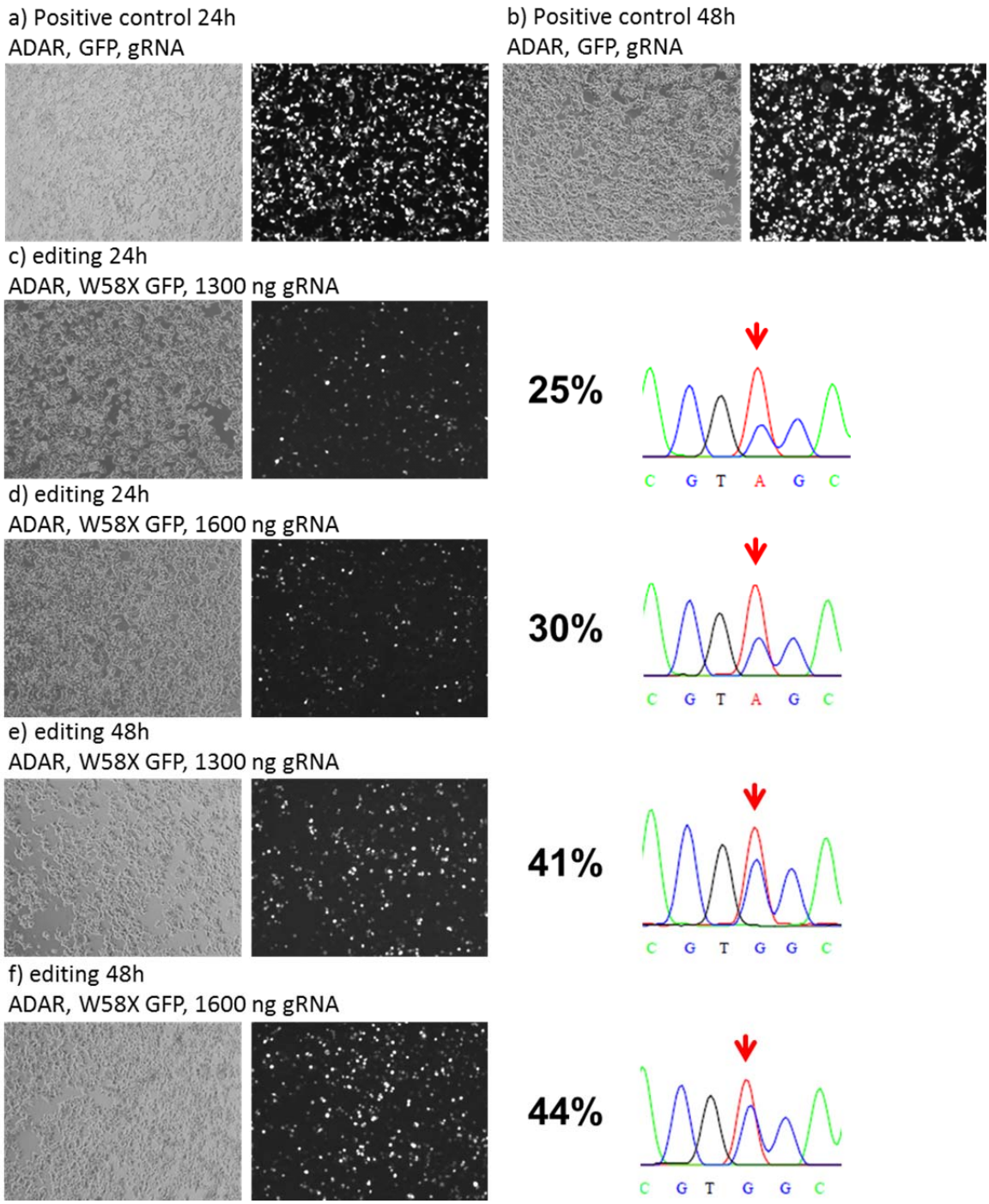
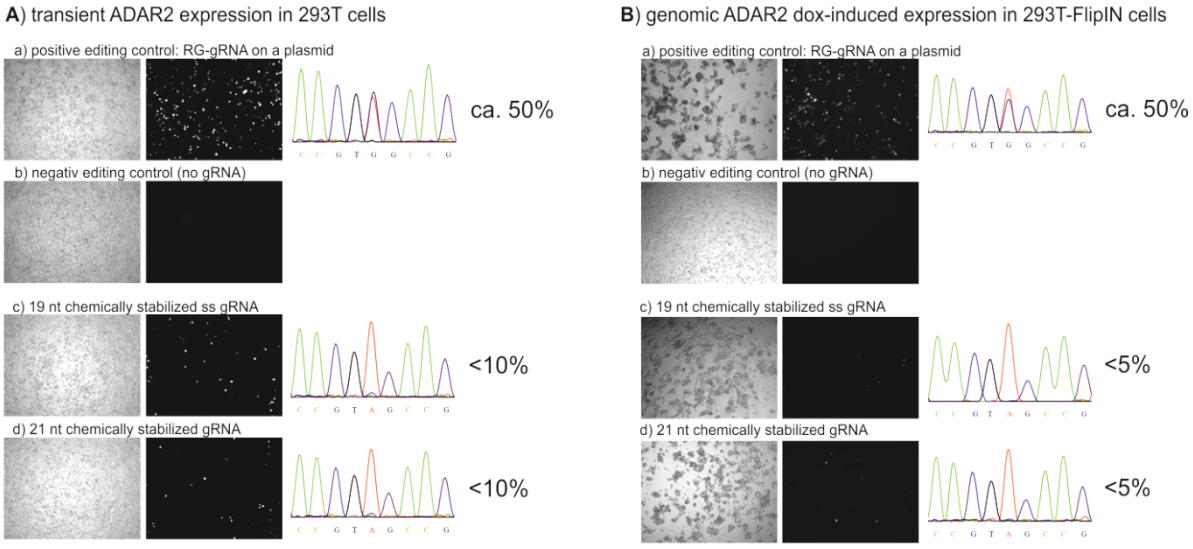


Figure S9: Comparing the editing efficiency using chemically stabilized single-stranded (ss) gRNA and U6-driven unstructured ss guideRNAs, and replacement of ADAR2 by SNAP-ADAR2. A) & **B**) Replacement of the R/G-gRNA plasmid by a ss-gRNA transfection: A) co-transfection of 500 ng W58X eGFP and 100 ng ADAR2 plasmid; B) transfection of 500 ng W58X eGFP plasmid plus induction of ADAR2 with 10 ng/mL doxycycline. 24 hours after plasmid transfection in 24-wells, the cells were detached and reseeded in a 96-well format and directly reverse transfected with 10 pmol (A) or 20 pmol (B) ss-gRNA. After 48h, fluorescence images were taken and RNA was isolated. In the controls a), 260 ng R/G-guideRNA have been transfected. The chemically stabilized guideRNAs contain 2´-O-methyl groups globally apart from a 3 nt gap around the adenosine to be edited, and terminal phosphorothioates. The full guideRNA sequences are given in Hanswillemenke et al., *JACS* **2015**. **C**) & **D**): U6-driven expression of unstructured guideRNAs. Cotransfection in 24-well format: 300 ng W58x eGFP and 1300 ng respective U6-driven guideRNA-vector, in case of C) 100 ng ADAR2 plasmid, in case of D) 10 ng/ml doxycycline. Fluorescence imaging was taken 48 hrs post transfection. The placement of the unstructured guideRNAs (C) and D)) relative to the mRNA is given, 1* stands for the edited adenosine. Negative control z) was like all four positive controls a), but with SNAP-ADAR2 instead of human ADAR2. Total magnification: 100x, GFP exposure time: 50 ms.

Editing with chemically stabilized ss guideRNAs



Editing with U6-driven, unstructured ss guideRNAs

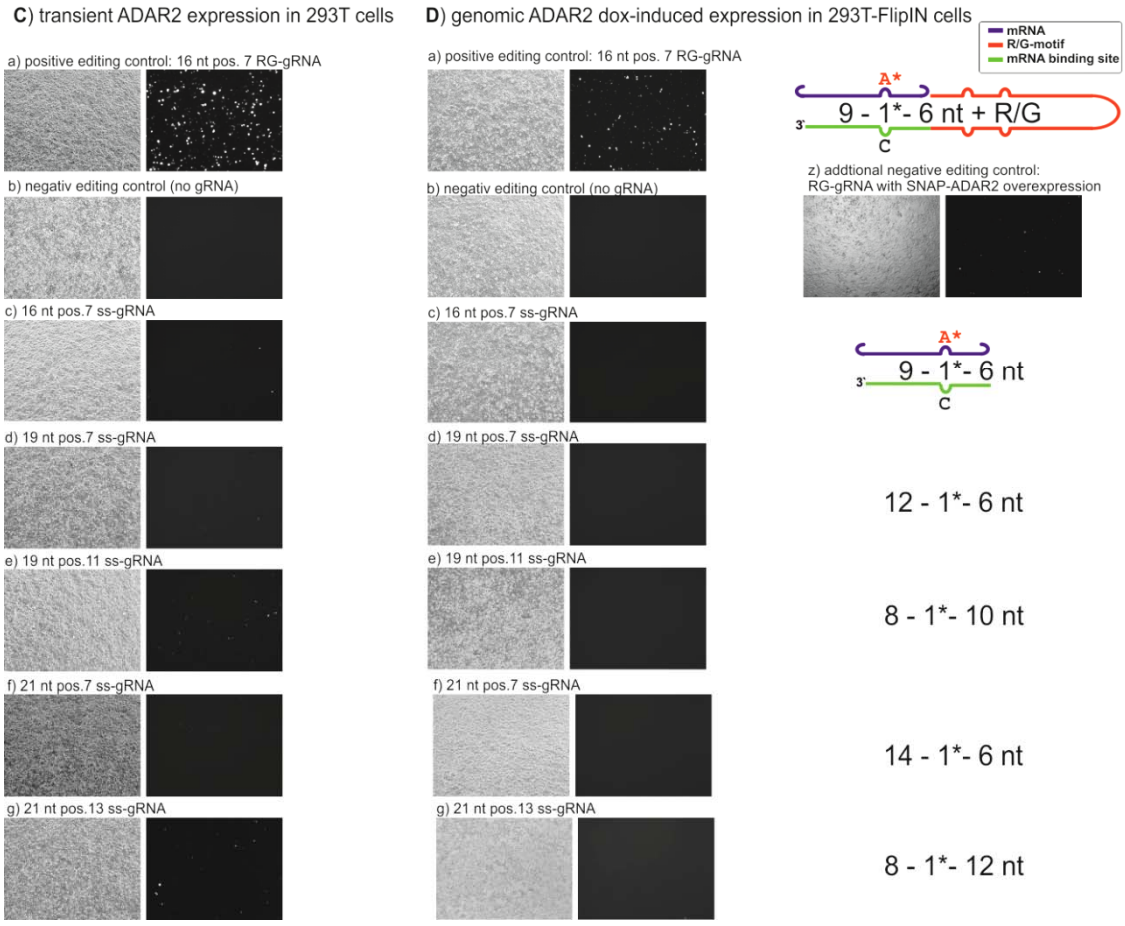


Figure S10: Microscopy analysis and editing yields of R/G-gRNAs with varying length of the flexible part and off-target editing at position 53*. The strongest fluorescent signal and highest editing yield was obtained for a R/G-gRNA with 16 or 18 nt length of the mRNA binding site. With the prolongation of the mRNA binding site of the R/G-gRNA less fluorescent signal and lower editing yields were observed. The off-target adenosine at position 53 is edited up to 10% if a 25 nt and 29 nt long R/G-gRNA is used for the editing reaction. Total magnification: 100x, GFP exposure time: 50 ms.*

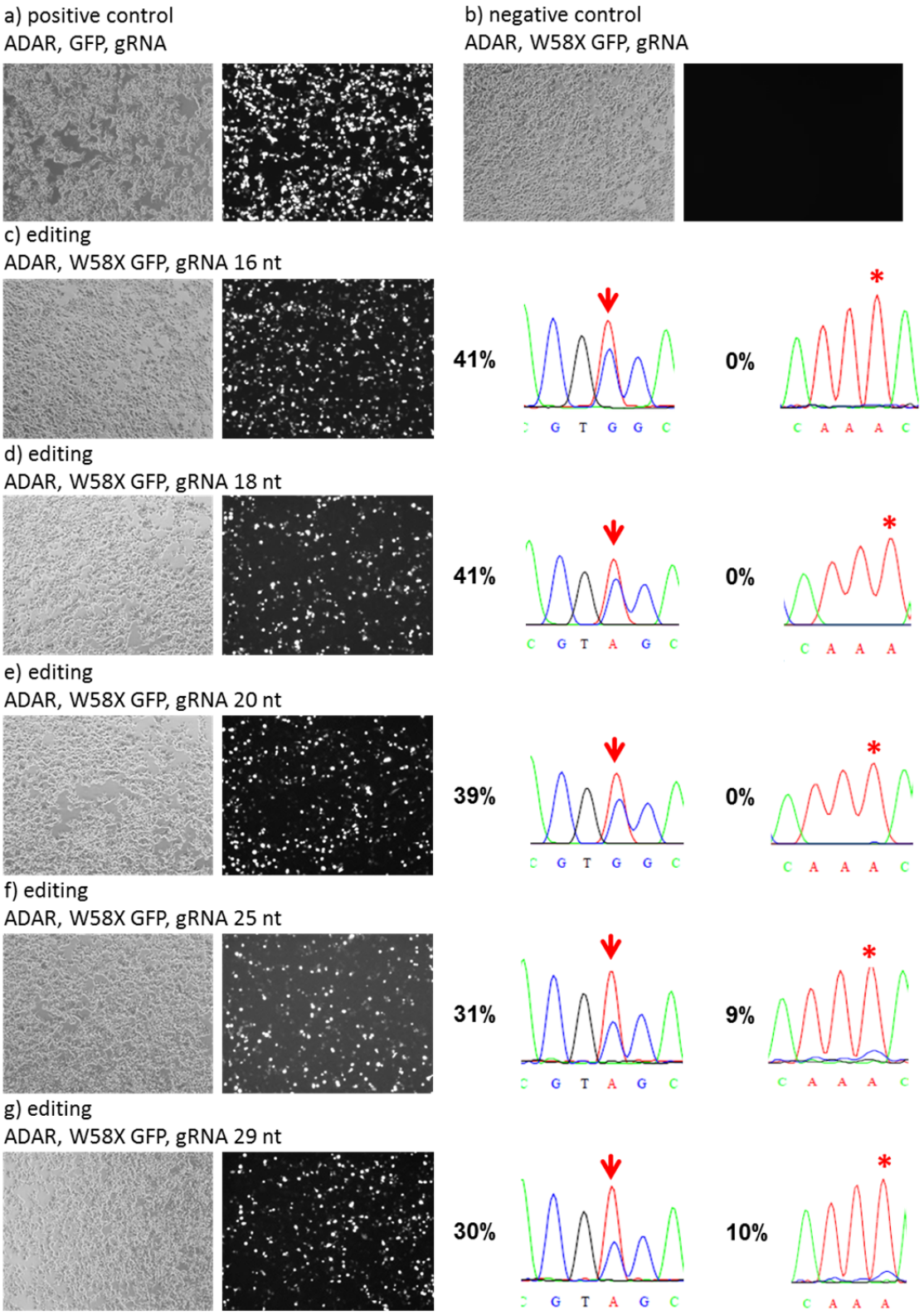
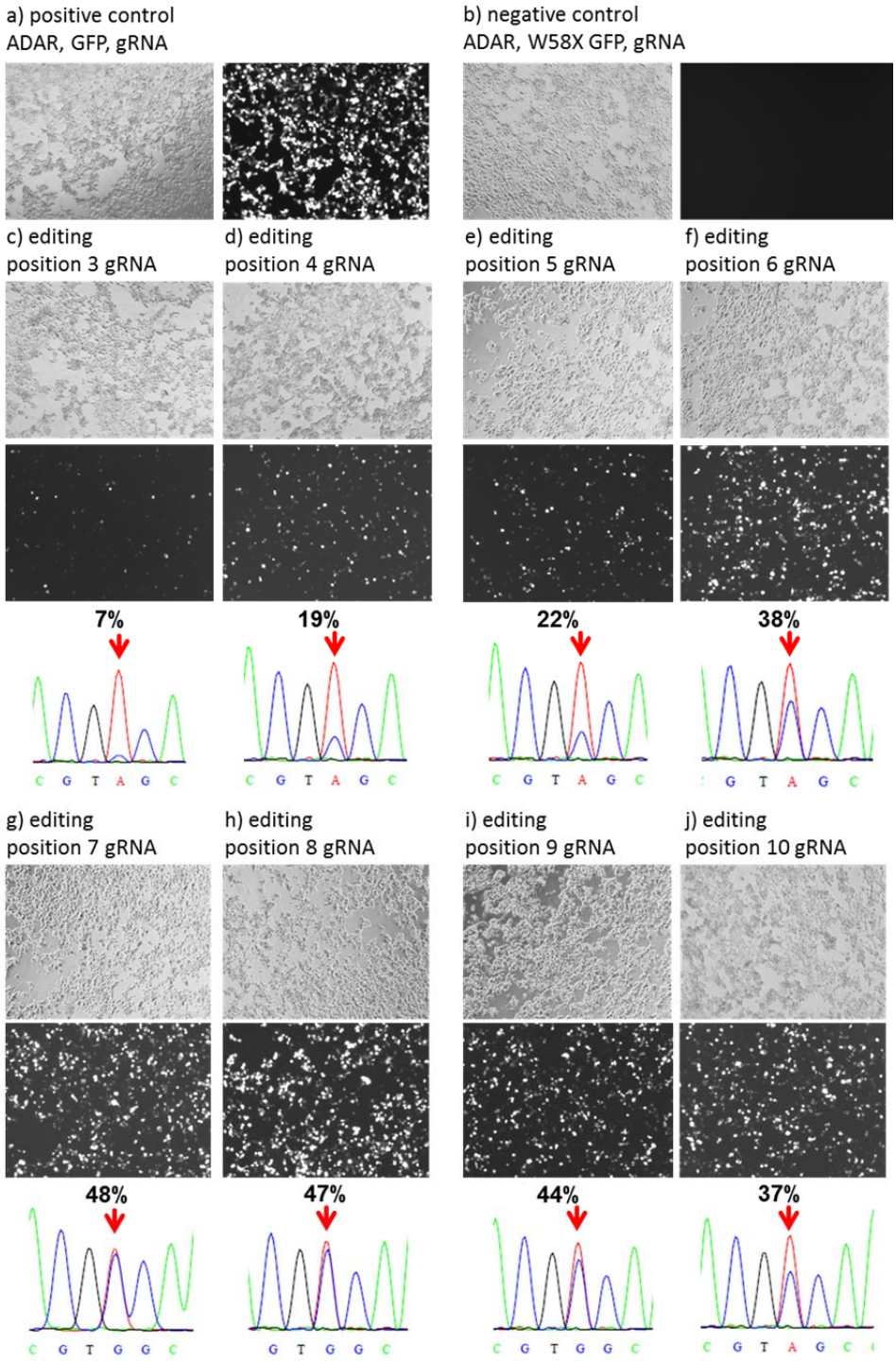


Figure S11: Effect of variable positions of the W58X eGFP R/G-gRNA towards the target adenosine in cell culture. *Co-transfection of 300 ng W58X eGFP and ADAR2 plasmid together with 1300 ng of the tested position of the R/G-gRNA was performed in a 24-well plate format. The microscopic analysis and RNA isolation for sequence analysis was performed 48 hours post transfection. The R/G-gRNAs positions 3 until 10 are abbreviated by P3 – P10 (this equals 2-9 intervening nucleotides). Starting from the R/G-gRNA position 3 an increasing fluorescent signal and amount is visible until R/G-gRNA position 8. The R/G-gRNA position 9 and 10 showed a dropping fluorescent signal. The microscopic results are confirmed by the sequence analysis and demonstrate that R/G-gRNA position 8 was the most successful guideRNA to achieve maximum editing yields. Total magnification: 100x, GFP exposure time: 50 ms.*

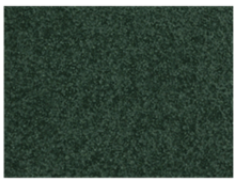


continuation Figure S11, replication of the positional effect.

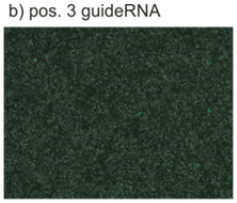
Replication 1: 24-well format, 300 ng ADAR2 plasmid, 300 ng W58X plasmid, 1600 ng R/G-guideRNA plasmid, imaging 48 hrs post transfection, total magnification 100x, 50 ms exposure

a) negative control: no guideRNA

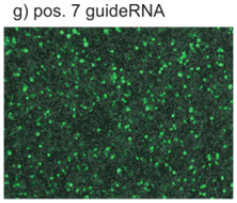
f) negative control: no ADAR



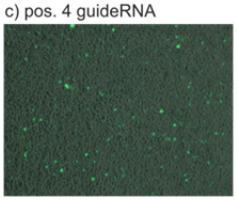
b) pos. 3 guideRNA



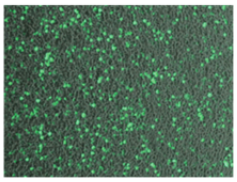
g) pos. 7 guideRNA



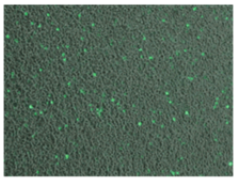
c) pos. 4 guideRNA



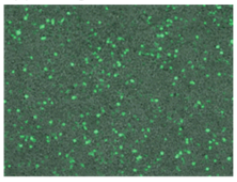
h) pos. 8 guideRNA



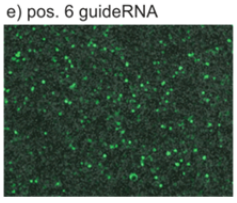
d) pos. 5 guideRNA



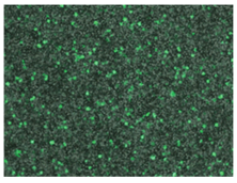
i) pos. 9 guideRNA



e) pos. 6 guideRNA



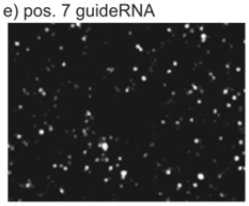
j) pos. 10 guideRNA



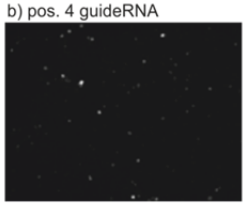
Replication 2: exactly as replication 1, but 1300 ng R/G-guideRNA

a) pos. 3 guideRNA

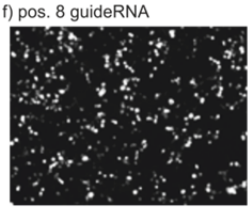
e) pos. 7 guideRNA



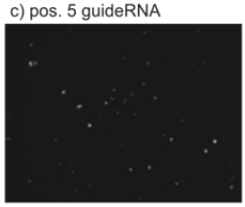
b) pos. 4 guideRNA



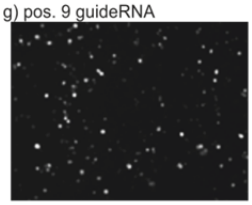
f) pos. 8 guideRNA



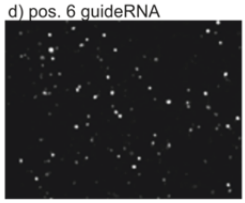
c) pos. 5 guideRNA



g) pos. 9 guideRNA



d) pos. 6 guideRNA



h) pos. 10 guideRNA

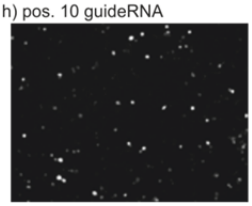


Figure S12: Effect of variable amounts of ADAR2 and R/G-gRNA plasmids in cell culture. Different amounts of R/G-gRNA P8 plasmid were transfected in 24-well plate format together with a constant amount of 300 ng W58X eGFP and ADAR2 plasmid (P8 equals 7 intervening nucleotides). Higher amounts of R/G-gRNA plasmid resulted in more and stronger fluorescence, as well as in higher editing levels. Total magnification: 100x, GFP exposure time: 50 ms.

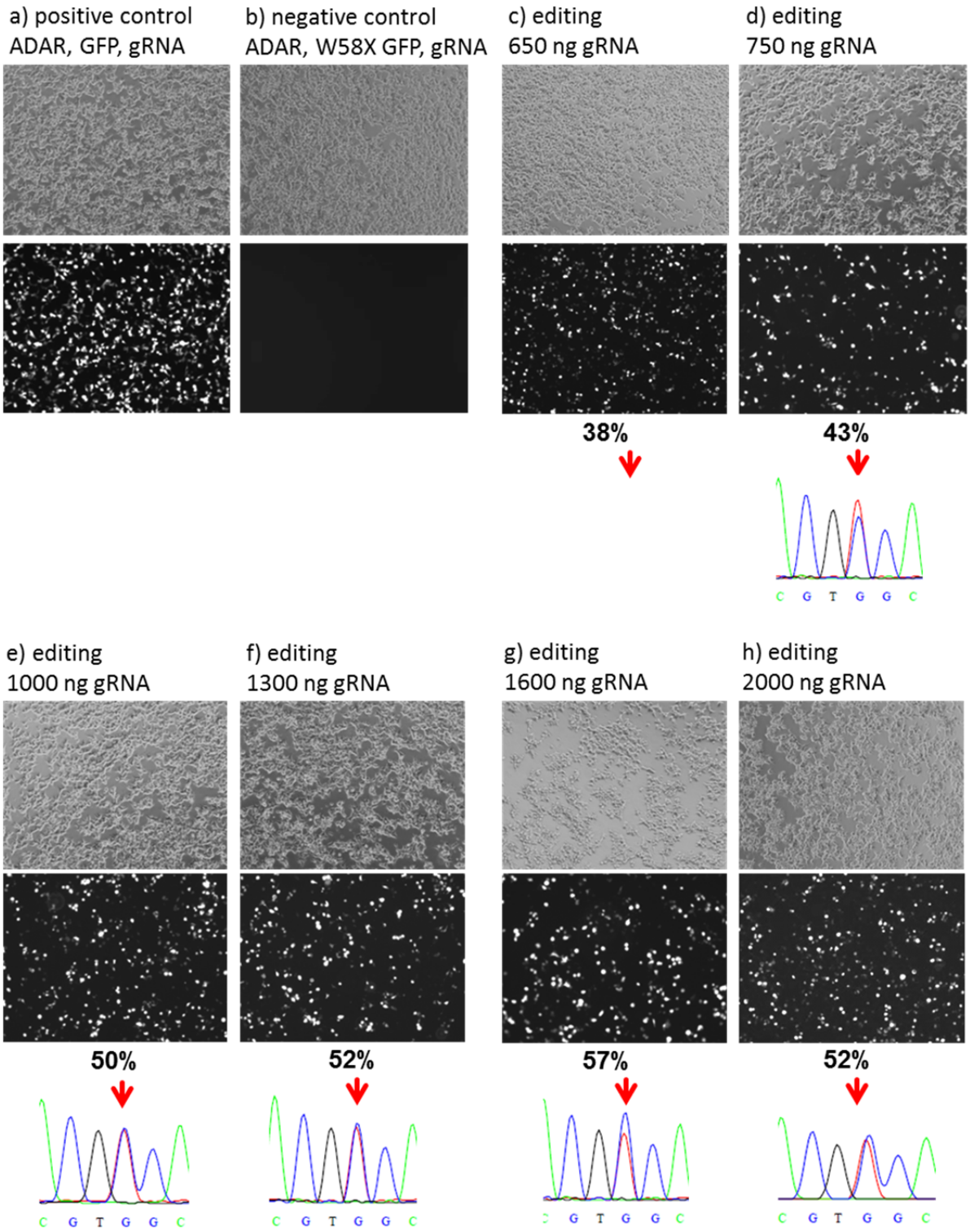
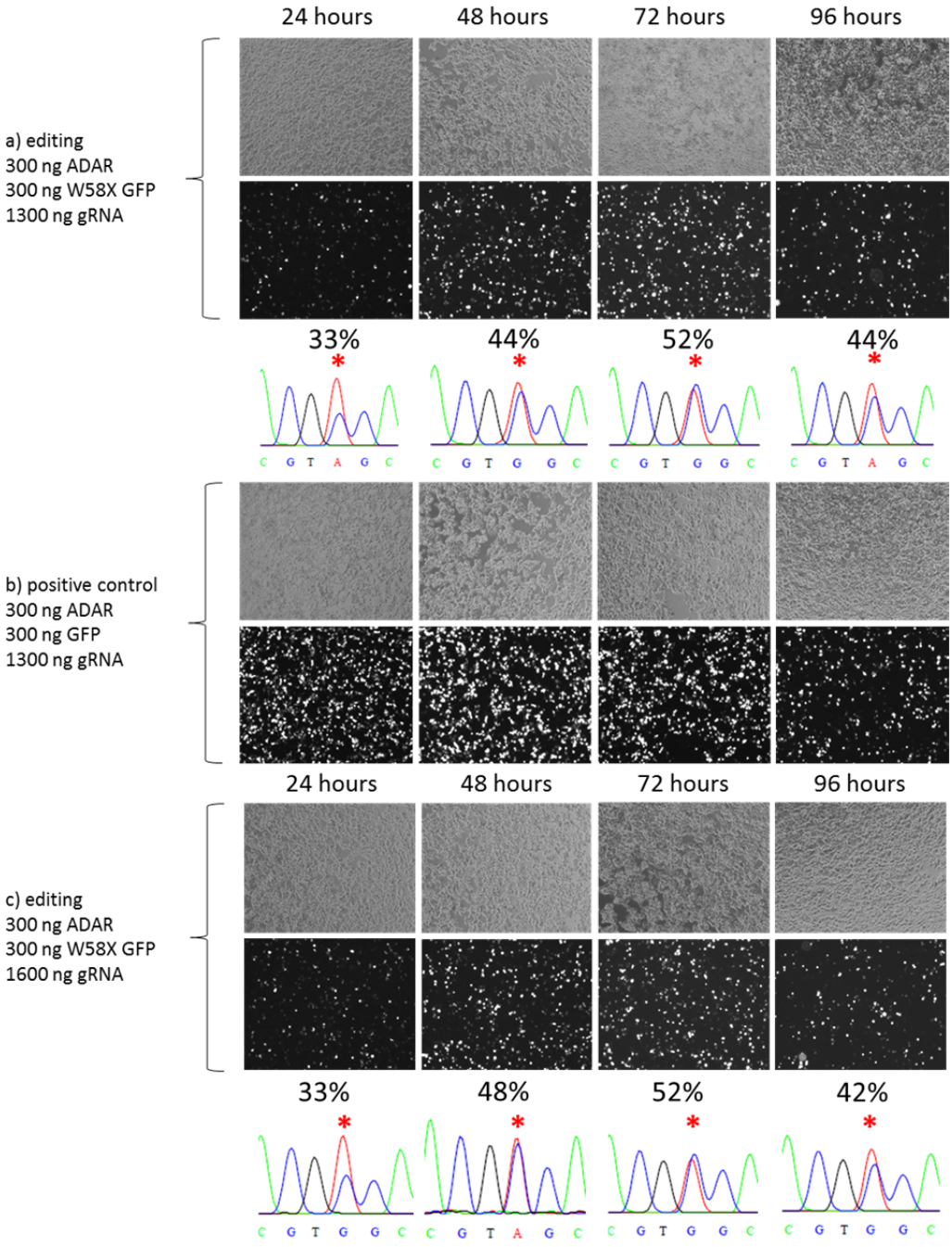

a) positive control
ADAR, GFP, gRNA
b) negative control
ADAR, W58X GFP, gRNA
c) editing
650 ng gRNA
d) editing
750 ng gRNA
38%
43%
C G T G G C
e) editing
1000 ng gRNA
f) editing
1300 ng gRNA
g) editing
1600 ng gRNA
h) editing
2000 ng gRNA
50%
52%
57%
52%
C G T G G C

Figure S13: Prolongation of the editing time for the R/G-gRNA position 8 *(position 8 equals 7 intervening nucleotides). The co-transfection experiment of 300 ng W58X eGFP plasmid, 300 ng of ADAR2 plasmid and 1300 ng or 1600 ng R/G-gRNA P8 plasmid was performed in a 24-well format. The editing efficiency was analyzed 24h, 48h, 72h and 96h post transfection. The positive control showed the strongest fluorescent signal for 48 hours of incubation. Shorter and longer incubation led to a reduced eGFP signal. For both chosen R/G-gRNA P8 plasmid amounts an increasing fluorescent signal and amount until 72h of incubation was visible. The eGFP intensity and amount of cells was declining after 96h of incubation. The sequence analysis confirmed the fluorescent microscopy: the editing yields increased until 72 hours of editing time and decreased after 96h again. Total magnification: 100x, GFP exposure time: 50 ms.*



.

Figure S14: Effect of decreasing amounts of transfected ADAR2 plasmid on the editing yield and the off-target editing at position A381. *In a 24-well format cells were co-transfected with 300 ng W58X eGFP and 1300 ng R/G-gRNA-P8 (P8 equals 7 intervening nucleotides) together with varying amounts of ADAR2 plasmid. The microscopic analysis and RNA isolation was carried out 48 h post transfection. The decrease of the ADAR2 plasmid down to 50 ng reduces the editing yield by 5 % compared to the starting concentration of 300 ng. The usage of 25 ng of ADAR2 plasmid markedly lowers the editing level down to 36% compared to 52% editing yield for 300 ng of ADAR2 plasmid. The reduction of the transfected ADAR2 plasmid amount led to a decrease at the off-target site A381 eGFP. Transfection of 100 ng or lower ADAR2 plasmid amount completely prevents the off-target editing. Total magnification: 100x, GFP exposure time: 50 ms.*

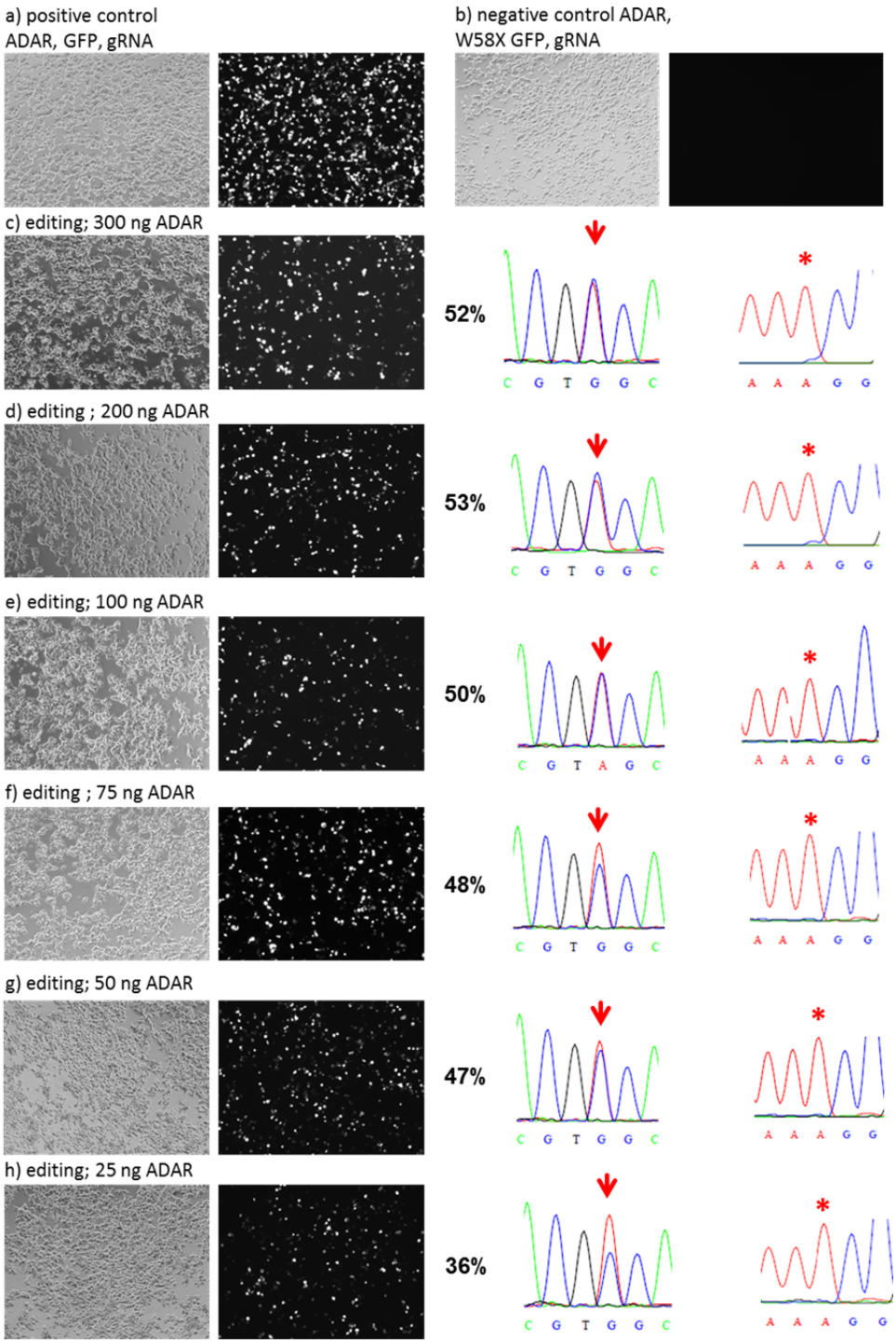


Figure S15 *Editing depends on the position of the targeted adenosine to the R/G-motif. This was also found for editing of the R407Q site in PINK1 in 293 cells analog to the experiments shown in Figure 2D. (n = 3)*

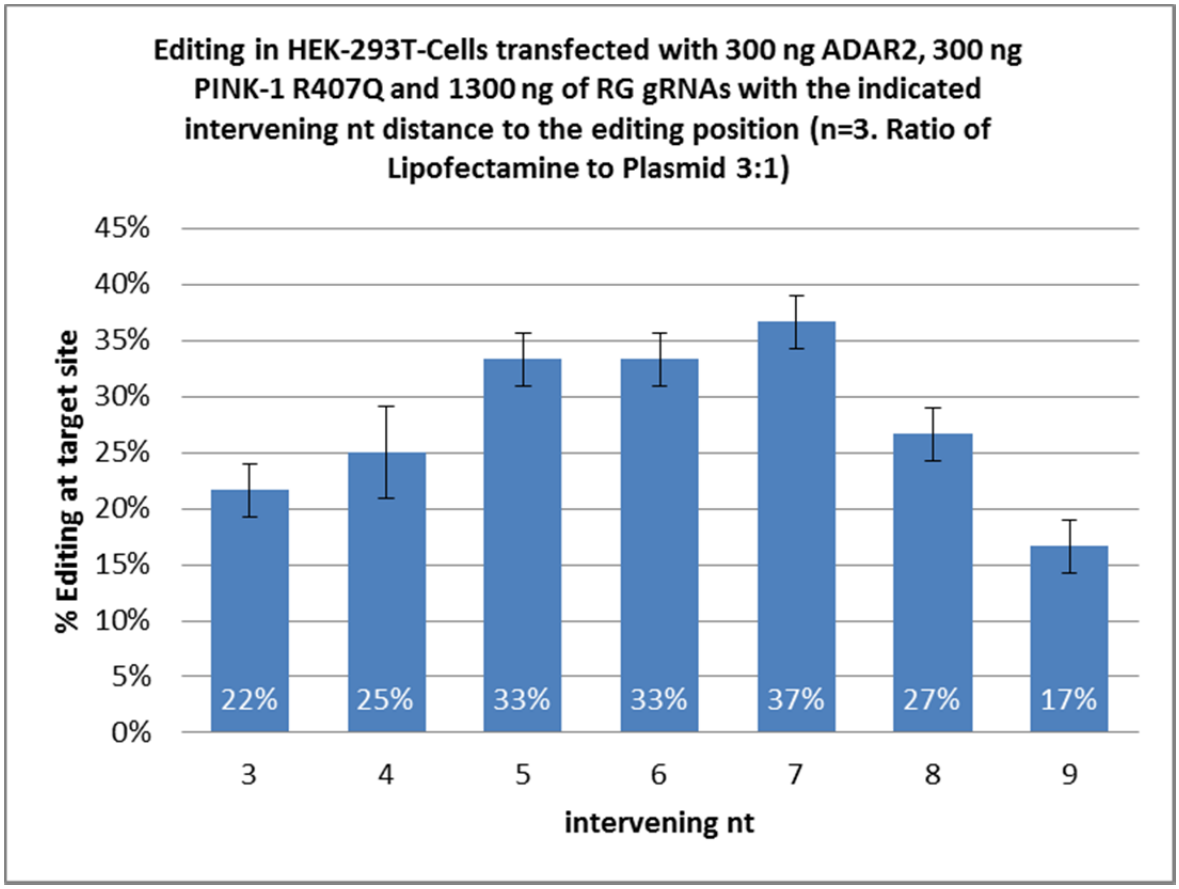
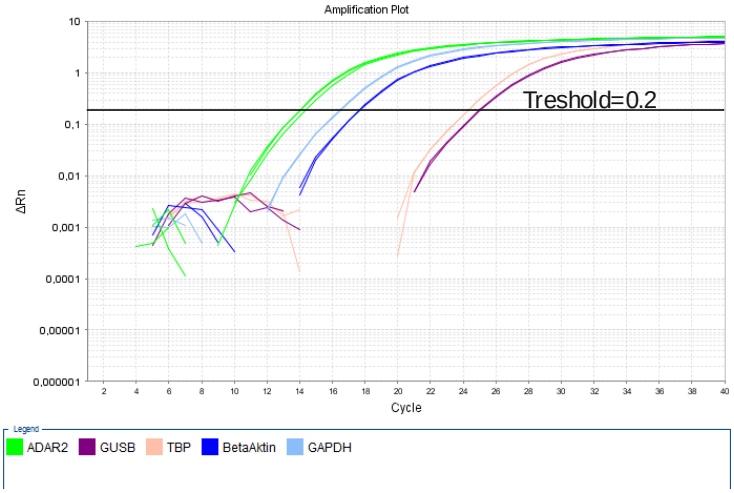


Figure S16. *qPCR analysis of ADAR2 expression. The relative ADAR2 mRNA amount in 293T cells transiently transfected (trans.) with ADAR2 and 293T cells with a genomically integrated copy of ADAR2 controlled by a CMV tet-on promoter (integr.) was determined by quantitative real-time PCR (qPCR) after 24h (doxycycline induced expression of integr. ADAR2) and 48h (expression of trans. and integr. ADAR2). For this, RNA was extracted from cell lysates (RNeasy MinElute Kit, Qiagen). After DNaseI digestion (NEB) and reverse transcription (high capacity cDNA reverse transcription kit, Applied Biosystems), 20 ng cDNA was mixed with Fast SYBR Green Master Mix (Applied Biosystems) and analyzed by the 7500 Fast Real-Time PCR System (Applied Biosystems).* ***(A)*** *For determining gene expression, primers were designed for targeting ADAR2 and the housekeeping genes ß-actin, glyceraldehyde-3-phophate dehydrogenase (GAPDH), ß-glucuronidase (GUSB) and TATA-box binding protein (TBP), (**B**) shows an example of the sybr green traces during qPCR.* ***(C)*** *qPCR of ADAR2 and the housekeeping gene was performed in triplicates and duplicates, respectively. The table displays the mean values of the cycles where the fluorescence crosses the threshold of 0.2 (ct values).* ***(D)*** *Based on these ct values, the expression of ADAR2 compared to housekeeping gene expression was determined in 293T cells after transient ADAR2 transfection or doxycycline induction by the delta ct equation.* ***(E)*** *To compare ADAR2 expression of ADAR2 transiently transfected 293T cells and ADAR2 genomically integrated 293T cells after 48h of expression, two methods were applied. In **method 1**, a calibration curve was generated from 1:5 dilutions of 20 ng cDNA of transiently transfected 293T cells (mean values from triplicates with standard derivation). For normalization, the corresponding ct (ADAR2) values were divided by the ct-value of ß-actin for 20 ng cDNA. In **method 2**, the delta-delta ct method was used to calculate difference in ADAR2 expression.*

(A) *Primers for qPCR*

Gene	***Sequence (5' to 3')***	***Product size***
ADAR2	*fw.: CGGAGATCCTTGCTCAGATT*	*99 bp*
	rev.: CCCTCGCTCTGATTTCTGAA	
ß-actin	*fw.: CGGGACCTGACTGACTAC*	*91 bp*
	rev.: TAATGTCACGCACGATTTCC	
GAPDH	*fw.: CAACAGCCTCAAGATCATCAG*	*96 bp*
	rev.: CCTTCCACGATACCAAAGTTG	
GUSB	*fw.: ACCTGTTCAAGTTGGAAGTG*	*93 bp*
	rev.: CACCTGGCACCTTAAGTTG	
TBP	*fw.: CGGAGAGTTCTGGGATTGTA*	*90 bp*
	rev.: GAAGTGCAATGGTCTTTAGGT	

(B) *An example of the raw data for transient ADAR2 expression (300 ng) in 293T cells.*



(C) *Measured ct-values of all experiments. Values are averaged from three technical replicates for ADAR2 and two technical replicates for the housekeeping genes*

	Sample	ct (ADAR2)	ct (ß-actin)	ct (GAPDH)	ct (GUSB)	ct (TBP)
Tra	293T	25.061	18.089	16.88	24.829	24.174
	293T + 300ng ADAR2, 48h	14.195	17.710	16.532	25.08	24.356
Integr.	293-pcDNA5 + Dox, 24h	23.807	17.143	16.617	23.935	23.086
	293-ADAR2 without Dox	20.916	16.733	16.199	23.696	22.72
	293-ADAR2 + Dox, 24h	18.710	17.483	16.941	24.093	23.546
	293-ADAR2 + Dox, 48h	18.633	17.654	16.766	24.651	23.784

(D) *Calculation of relative expression levels from the Δct values for ADAR2 versus four housekeeping genes*

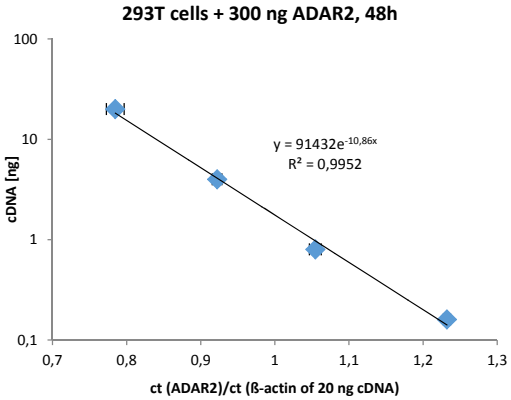
	Sample	ß-actin	GAPDH	GUSB	TPB
Tra	293T	0.008	0.003	0.851	0.541
	293T + 300ng ADAR2, 48h	11.432	5.053	1891.087	1144.895
Integr.	293-pcDNA5	0.009	0.007	1.093	0.607
	293-ADAR2 without Dox	0.055	0.038	6.869	3.492
	293-ADAR2 + 10 ng/ml Dox, 24h	0.427	0.293	41.730	28.562
	293-ADAR2 + 10ng/ml Dox, 48h	0.507	0.274	64.804	35.531

$\text{relative expression} = 2^{-\Delta\text{ct}}$, with $\Delta\text{ct} = \text{ct(ADAR2)} - \text{ct(housekeeping gene)}$

(E) *Comparison of transient versus genomic ADAR2 expression*

Method 1

A calibration curve was taken for four different ADAR2 (transient expression) dilutions (1x, 5x, 25x, 125x). Plotted is the ADAR2 ct values normalized by the ct value for beta-actin of the undiluted sample versus the amount of cDNA.



From the regression curve of the calibration plot, the relative expression of genomically expressed ADAR2 was calculated. For this genomically expressed ADAR2 was normalized to beta-actin.

ct (ADAR2)/ct (ß-actin) = 1.055 [for the experiment with 293-ADAR2 + 10 ng/ml Dox, 48h] (20 ng cDNA)

from x = 1.055 one can calculate y = 0.967, and the genomic expression to ***be 20fold*** *below that of the transient expression*

Method 2

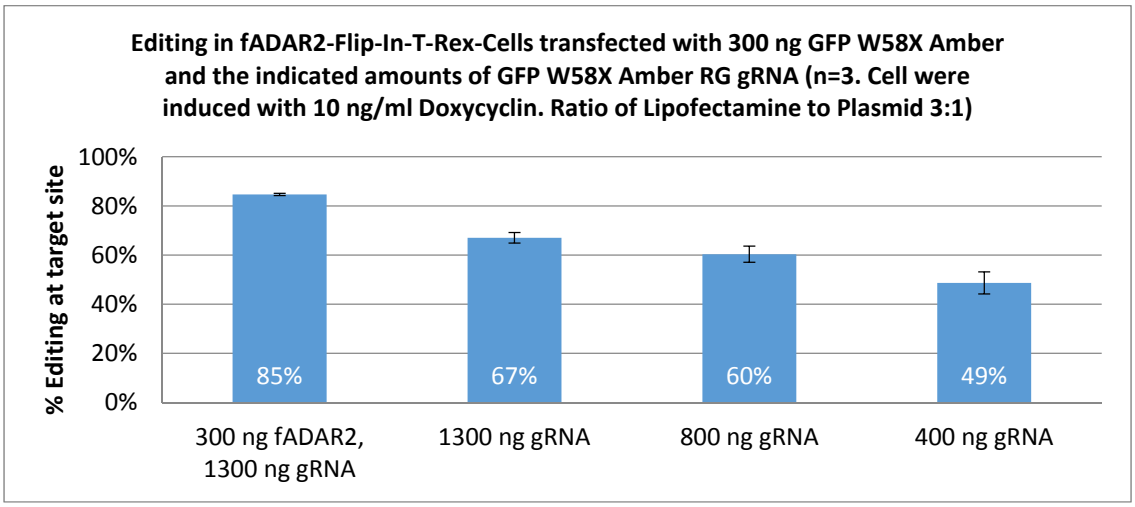
Here we estimated the relative expression level of ADAR2 transient versus genomic by the delta-delta ct method applying the following equation.

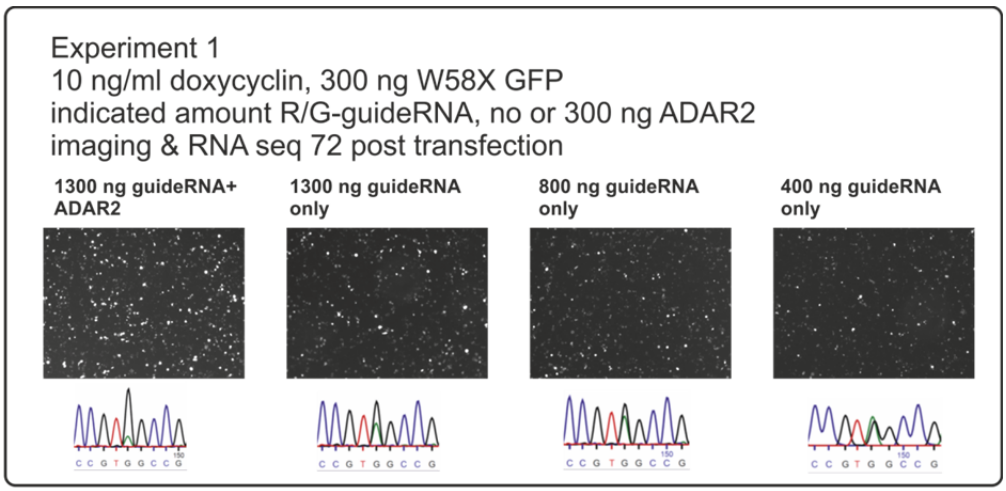
$$\Delta\Delta\mathrm{ct} = (\mathrm{ct}(\mathrm{ADAR2}) - \mathrm{ct}(\text{ßactin}))_{\mathrm{transient}} - ((\mathrm{ct}(\mathrm{ADAR2}) - \mathrm{ct}(\text{ßactin}))_{\mathrm{genomic}}$$

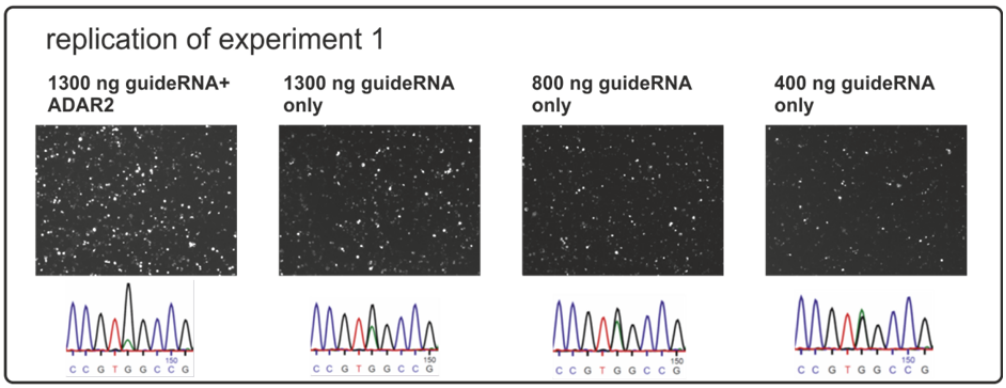
$$\text{relative expression} = 2^{-\Delta\Delta\mathrm{ct}}$$

relative expression $= 25.795$, meaning genomic expression is approx. **26fold** below transient expression

Figure S17 *Dependency of the editing yield with genomically encoded ADAR2 with varying amounts of guideRNA. Shown are the averaged yields and standard deviation for n=3.*







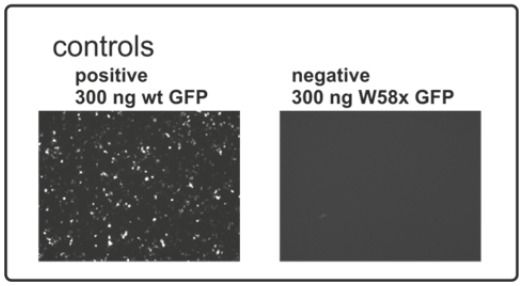


Figure S18. Full sequencing trace corresponding to the trace shown in Figure 3A, b), 1300 ng guideRNA

Editing of endogenous transcripts

Editing experiments have been carried out in duplicates exactly as described before (without further optimization), but without co-transfection of a target or reporter gene. 293 cells have been transfected with 300 ng ADAR2 and of the respective 1300 ng R/G-guideRNA in 24 well format with lipfectamine 2000, and were harvested 48h post transfection. 293-ADAR2 flip-in cells have been induced with doxycycline (10 ng/ml), then 1300 ng of the respective R/G-guideRNA was transfected with lipofectamine 2000. 72 h after transfection of the guideRNA total RNA was isolated, RT-PCR with transcript-specific primers have been done to obtain the RNA sequencing traces. The respective guideRNA sequences are listed in Table S1. Primers for RT-PCR are given in Table S2.

Target sites for site-directed RNA editing on six genes:

cDNA sequence of β-actin

```
                  10        20        30        40        50        60
1          ACCGCCGAGACCGCGTCCGCCCCGCGAGCACAGAGCCTCGCCTTTGCCGATCCGCCGCCC
1

                  70        80        90       100       110       120
61         GTCCACACCCGCCGCCAGCTCACCATGGATGATGATATCGCCGCGCTCGTCGTCGACAAC
21                                M  D  D  D  I  A  A  L  V  V  D  N

                130       140       150       160       170       180
121        GGCTCCGGCATGTGCAAGGCCGGCTTCGCGGGCGACGATGCCCCCCGGGCCGTCTTCCCC
41          G  S  G  M  C  K  A  G  F  A  G  D  D  A  P  R  A  V  F  P

                190       200       210       220       230       240
181        TCCATCGTGGGGCGCCCCAGGCACCAGGGCGTGATGGTGGGCATGGGTCAGAAGGATTCC
61          S  I  V  G  R  P  R  H  Q  G  V  M  V  G  M  G  Q  K  D  S

                250       260       270       280       290       300
241        TATGTGGGCGACGAGGCCCAGAGCAAGAGAGGCATCCTCACCCTGAAGTACCCCATCGAG
81          Y  V  G  D  E  A  Q  S  K  R  G  I  L  T  L  K  Y  P  I  E

                310       320       330       340       350       360
301        CACGGCATCGTCACCAACTGGGACGACATGGAGAAAATCTGGCACCACACCTTCTACAAT
101         H  G  I  V  T  N  W  D  D  M  E  K  I  W  H  H  T  F  Y  N

                370       380       390       400       410       420
361        GAGCTGCGTGTGGCTCCCGAGGAGCACCCCGTGCTGCTGACCGAGGCCCCCCTGAACCCC
121         E  L  R  V  A  P  E  E  H  P  V  L  L  T  E  A  P  L  N  P

                430       440       450       460       470       480
421        AAGGCCAACCGCGAGAAGATGACCCAGATCATGTTTGAGACCTTCAACACCCCAGCCATG
141         K  A  N  R  E  K  M  T  Q  I  M  F  E  T  F  N  T  P  A  M

                490       500       510       520       530       540
481        TACGTTGCTATCCAGGCTGTGCTATCCCTGTACGCCTCTGGCCGTACCACTGGCATCGTG
161         Y  V  A  I  Q  A  V  L  S  L  Y  A  S  G  R  T  T  G  I  V

                550       560       570       580       590       600
541        ATGGACTCCGGTGACGGGGTCACCCACACTGTGCCCATCTACGAGGGGTATGCCCTCCCC
181         M  D  S  G  D  G  V  T  H  T  V  P  I  Y  E  G  Y  A  L  P

                610       620       630       640       650       660
601        CATGCCATCCTGCGTCTGGACCTGGCTGGCCGGGACCTGACTGACTACCTCATGAAGATC
201         H  A  I  L  R  L  D  L  A  G  R  D  L  T  D  Y  L  M  K  I
```

```
                670       680       690       700       710       720
661       CTCACCGAGCGCGGCTACAGCTTCACCACCACGGCCGAGCGGGAAATCGTGCGTGACATT
221        L  T  E  R  G  Y  S  F  T  T  T  A  E  R  E  I  V  R  D  I

                730       740       750       760       770       780
721       AAGGAGAAGCTGTGCTACGTCGCCCTGGACTTCGAGCAAGAGATGGCCACGGCTGCTTCC
241        K  E  K  L  C  Y  V  A  L  D  F  E  Q  E  M  A  T  A  A  S

                790       800       810       820       830       840
781       AGCTCCTCCCTGGAGAAGAGCTACGAGCTGCCTGACGGCCAGGTCATCACCATTGGCAAT
261        S  S  S  L  E  K  S  Y  E  L  P  D  G  Q  V  I  T  I  G  N

                850       860       870       880       890       900
841       GAGCGGTTCCGCTGCCCTGAGGCACTCTTCCAGCCTTCCTTCCTGGGCATGGAGTCCTGT
281        E  R  F  R  C  P  E  A  L  F  Q  P  S  F  L  G  M  E  S  C

                910       920       930       940       950       960
901       GGCATCCACGAAACTACCTTCAACTCCATCATGAAGTGTGACGTGGACATCCGCAAAGAC
301        G  I  H  E  T  T  F  N  S  I  M  K  C  D  V  D  I  R  K  D

                970       980       990      1000      1010      1020
961       CTGTACGCCAACACAGTGCTGTCTGGCGGCACCACCATGTACCCTGGCATTGCCGACAGG
321        L  Y  A  N  T  V  L  S  G  G  T  T  M  Y  P  G  I  A  D  R

               1030      1040      1050      1060      1070      1080
1021      ATGCAGAAGGAGATCACTGCCCTGGCACCCAGCACAATGAAGATCAAGATCATTGCTCCT
341        M  Q  K  E  I  T  A  L  A  P  S  T  M  K  I  K  I  I  A  P

               1090      1100      1110      1120      1130      1140
1081      CCTGAGCGCAAGTACTCCGTGTGGATCGGCGGCTCCATCCTGGCCTCGCTGTCCACCTTC
361        P  E  R  K  Y  S  V  W  I  G  G  S  I  L  A  S  L  S  T  F

               1150      1160      1170      1180      1190      1200
1141      CAGCAGATGTGGATCAGCAAGCAGGAGTATGACGAGTCCGGCCCCTCCATCGTCCACCGC
381        Q  Q  M  W  I  S  K  Q  E  Y  D  E  S  G  P  S  I  V  H  R

               1210      1220      1230      1240      1250      1260
1201      AAATGCTTCTAGGCGGACTATGACTTAGTTGCGTTACACCCTTTCTTGACAAAACCTAAC
401        K  C  F  *            target 1

               1270      1280      1290      1300      1310      1320
1261      TTGCGCAGAAAACAAGATGAGATTGGCATGGCTTTATTTGTTTTTTTTGTTTTGTTTTGG
421

               1330      1340      1350      1360      1370      1380
1321      TTTTTTTTTTTTTTTTGGCTTGACTCAGGATTTAAAAACTGGAACGGTGAAGGTGACAGC
441

               1390      1400      1410      1420      1430      1440
1381      AGTCGGTTGGAGCGAGCATCCCCCAAAGTTCACAATGTGGCCGAGGACTTTGATTGCACA
461

               1450      1460      1470      1480      1490      1500
1441      TTGTTGTTTTTTTAATAGTCATTCCAAATATGAGATGCGTTGTTACAGGAAGTCCCTTGC
481                target 2

               1510      1520      1530      1540      1550      1560
1501      CATCCTAAAAGCCACCCCACTTCTCTCTAAGGAGAATGGCCCAGTCCTCTCCCAAGTCCA
501
```

```
                1570      1580      1590      1600      1610      1620
1561       CACAGGGGAGGTGATAGCATTGCTTTCGTGTAAATTATGTAATGCAAAATTTTTTTAATC
521                     target 3

                1630      1640      1650      1660      1670      1680
1621       TTCGCCTTAATACTTTTTTATTTTGTTTTATTTTGAATGATGAGCCTTCGTGCCCCCCCT
541

                1690      1700      1710      1720      1730      1740
1681       TCCCCCTTTTTTGTCCCCCAACTTGAGATGTATGAAGGCTTTTGGTCTCCCTGGGAGTGG
561

                1750      1760      1770      1780      1790      1800
1741       GTGGAGGCAGCCAGGGCTTACCTGTACACTGACTTGAGACCAGTTGAATAAAAGTGCACA
581

                1810      1820      1830      1840      1850
1801       CCTTAAAAATGAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA
601
```

cDNA sequence of GAPDH

```
                  10        20        30        40        50        60
1          GCCTCAAGACCTTGGGCTGGGACTGGCTGAGCCTGGCGGGAGGCGGGGTCCGAGTCACCG
1

                  70        80        90       100       110       120
61         CCTGCCGCCGCGCCCCCGGTTTCTATAAATTGAGCCCGCAGCCTCCCGCTTCGCTCTCTG
20

                 130       140       150       160       170       180
121        CTCCTCCTGTTCGACAGTCAGCCGCATCTTCTTTTGCGTCGCCAGCCGAGCCACATCGCT
40

                 190       200       210       220       230       240
181        CAGACACCATGGGGAAGGTGAAGGTCGGAGTCAACGGATTTGGTCGTATTGGGCGCCTGG
60                   M   G   K   V   K   V   G   V   N   G   F   G   R   I   G   R   L

                 250       260       270       280       290       300
241        TCACCAGGGCTGCTTTTAACTCTGGTAAAGTGGATATTGTTGCCATCAATGACCCCTTCA
80         V   T   R   A   A   F   N   S   G   K   V   D   I   V   A   I   N   D   P   F

                 310       320       330       340       350       360
301        TTGACCTCAACTACATGGTTTACATGTTCCAATATGATTCCACCCATGGCAAATTCCATG
100        I   D   L   N   Y   M   V   Y   M   F   Q   Y   D   S   T   H   G   K   F   H

                 370       380       390       400       410       420
361        GCACCGTCAAGGCTGAGAACGGGAAGCTTGTCATCAATGGAAATCCCATCACCATCTTCC
120        G   T   V   K   A   E   N   G   K   L   V   I   N   G   N   P   I   T   I   F

                 430       440       450       460       470       480
421        AGGAGCGAGATCCCTCCAAAATCAAGTGGGGCGATGCTGGCGCTGAGTACGTCGTGGAGT
140        Q   E   R   D   P   S   K   I   K   W   G   D   A   G   A   E   Y   V   V   E

                 490       500       510       520       530       540
481        CCACTGGCGTCTTCACCACCATGGAGAAGGCTGGGGCTCATTTGCAGGGGGGAGCCAAAA
160        S   T   G   V   F   T   T   M   E   K   A   G   A   H   L   Q   G   G   A   K

                 550       560       570       580       590       600
541        GGGTCATCATCTCTGCCCCCTCTGCTGATGCCCCCATGTTCGTCATGGGTGTGAACCATG
180        R   V   I   I   S   A   P   S   A   D   A   P   M   F   V   M   G   V   N   H
```

```
                610       620       630       640       650       660
601       AGAAGTATGACAACAGCCTCAAGATCATCAGCAATGCCTCCTGCACCACCAACTGCTTAG
200       E  K  Y  D  N  S  L  K  I  I  S  N  A  S  C  T  T  N  C  L

                670       680       690       700       710       720
661       CACCCCTGGCCAAGGTCATCCATGACAACTTTGGTATCGTGGAAGGACTCATGACCACAG
220       A  P  L  A  K  V  I  H  D  N  F  G  I  V  E  G  L  M  T  T

                730       740       750       760       770       780
721       TCCATGCCATCACTGCCACCCAGAAGACTGTGGATGGCCCCTCCGGGAAACTGTGGCGTG
240       V  H  A  I  T  A  T  Q  K  T  V  D  G  P  S  G  K  L  W  R

                790       800       810       820       830       840
781       ATGGCCGCGGGGCTCTCCAGAACATCATCCCTGCCTCTACTGGCGCTGCCAAGGCTGTGG
260       D  G  R  G  A  L  Q  N  I  I  P  A  S  T  G  A  A  K  A  V

                850       860       870       880       890       900
841       GCAAGGTCATCCCTGAGCTGAACGGGAAGCTCACTGGCATGGCCTTCCGTGTCCCCACTG
280       G  K  V  I  P  E  L  N  G  K  L  T  G  M  A  F  R  V  P  T

                910       920       930       940       950       960
901       CCAACGTGTCAGTGGTGGACCTGACCTGCCGTCTAGAAAAACCTGCCAAATATGATGACA
300       A  N  V  S  V  V  D  L  T  C  R  L  E  K  P  A  K  Y  D  D

                970       980       990      1000      1010      1020
961       TCAAGAAGGTGGTGAAGCAGGCGTCGGAGGGCCCCCTCAAGGGCATCCTGGGCTACACTG
320       I  K  K  V  V  K  Q  A  S  E  G  P  L  K  G  I  L  G  Y  T

               1030      1040      1050      1060      1070      1080
1021      AGCACCAGGTGGTCTCCTCTGACTTCAACAGCGACACCCACTCCTCCACCTTTGACGCTG
340       E  H  Q  V  V  S  S  D  F  N  S  D  T  H  S  S  T  F  D  A

               1090      1100      1110      1120      1130      1140
1081      GGGCTGGCATTGCCCTCAACGACCACTTTGTCAAGCTCATTTCCTGGTATGACAACGAAT
360       G  A  G  I  A  L  N  D  H  F  V  K  L  I  S  W  Y  D  N  E

               1150      1160      1170      1180      1190      1200
1141      TTGGCTACAGCAACAGGGTGGTGGACCTCATGGCCCACATGGCCTCCAAGGAGTAAGACC
380       F  G  Y  S  N  R  V  V  D  L  M  A  H  M  A  S  K  E  *

               1210      1220      1230      1240      1250      1260
1201      CCTGGACCACCAGCCCCAGCAAGAGCACAAGAGGAAGAGAGAGACCCTCACTGCTGGGGA
400

               1270      1280      1290      1300      1310      1320
1261      GTCCCTGCCACACTCAGTCCCCCACCACACTGAATCTCCCCTCCTCACAGTTGCCATGTA
420                                                               target 1

               1330      1340      1350      1360      1370      1380
1321      GACCCCTTGAAGAGGGGAGGGGCCTAGGGAGCCGCACCTTGTCATGTACCATCAATAAAG
440                               target 2

               1390      1400      1410      1420
1381      TACCCTGTGCTCAACCAGTTAAAAAAAAAAAAAAAAAAAA
460
```

cDNA sequence of GPI

```
                10        20        30        40        50        60
```

```
1          AATAGCCCTTACCACCAGCAGACACACATCATCTGTTGTACTTGCTTATTTGGCACATAT
1

                   70         80         90        100        110        120
61         GTATCCACAGCGCCTAGAACACTGCCTGTAACGTGGAAGGTGTTCGATCTATAGAGTTTT
20

                  130        140        150        160        170        180
121        GTCGAATGAATGAATGAAGCCGACTAGTGCACAGGGAGTGCAGCGGCGCGATGGTAGCTC
40                                                                M  V  A

                  190        200        210        220        230        240
181        TCTGCAGCCTCCAACACCTGGGCTCCAGTGATCCCCGGGCTCTGCCCACCCTCCCCACTG
60         L  C  S  L  Q  H  L  G  S  S  D  P  R  A  L  P  T  L  P  T

                  250        260        270        280        290        300
241        CCACTTCCGGGCAGAGGCCAGCAAAGCGGCGGCGCAAGAGTCCCGCCATGGCCGCTCTCA
80         A  T  S  G  Q  R  P  A  K  R  R  R  K  S  P  A  M  A  A  L

                  310        320        330        340        350        360
301        CCCGGGACCCCCAGTTCCAGAAGCTGCAGCAATGGTACCGCGAGCACCGCTCCGAGCTGA
100        T  R  D  P  Q  F  Q  K  L  Q  Q  W  Y  R  E  H  R  S  E  L

                  370        380        390        400        410        420
361        ACCTGCGCCGCCTCTTCGATGCCAACAAGGACCGCTTCAACCACTTCAGCTTGACCCTCA
120        N  L  R  R  L  F  D  A  N  K  D  R  F  N  H  F  S  L  T  L

                  430        440        450        460        470        480
421        ACACCAACCATGGGCATATCCTGGTGGATTACTCCAAGAACCTGGTGACGGAGGACGTGA
140        N  T  N  H  G  H  I  L  V  D  Y  S  K  N  L  V  T  E  D  V

                  490        500        510        520        530        540
481        TGCGGATGCTGGTGGACTTGGCCAAGTCCAGGGGCGTGGAGGCCGCCCGGGAGCGGATGT
160        M  R  M  L  V  D  L  A  K  S  R  G  V  E  A  A  R  E  R  M

                  550        560        570        580        590        600
541        TCAATGGTGAGAAGATCAACTACACCGAGGGTCGAGCCGTGCTGCACGTGGCTCTGCGGA
180        F  N  G  E  K  I  N  Y  T  E  G  R  A  V  L  H  V  A  L  R

                  610        620        630        640        650        660
601        ACCGGTCAAACACACCCATCCTGGTAGACGGCAAGGATGTGATGCCAGAGGTCAACAAGG
200        N  R  S  N  T  P  I  L  V  D  G  K  D  V  M  P  E  V  N  K
                                    target
                  670        680        690        700        710        720
661        TTCTGGACAAGATGAAGTCTTTCTGCCAGGGACCCCTCATGGTGACTGAAGCCCTTAAGC
220        V  L  D  K  M  K  S  F  C  Q  G  P  L  M  V  T  E  A  L  K

                  730        740        750        760        770        780
721        CATACTCTTCAGGAGGTCCCCGCGTCTGGTATGTCTCCAACATTGATGGAACTCACATTG
240        P  Y  S  S  G  G  P  R  V  W  Y  V  S  N  I  D  G  T  H  I

                  790        800        810        820        830        840
781        CCAAAACCCTGGCCCAGCTGAACCCCGAGTCCTCCCTGTTCATCATTGCCTCCAAGACCT
260        A  K  T  L  A  Q  L  N  P  E  S  S  L  F  I  I  A  S  K  T

                  850        860        870        880        890        900
841        TTACTACCCAGGAGACCATCACGAATGCAGAGACGGCGAAGGAGTGGTTTCTCCAGGCGG
280        F  T  T  Q  E  T  I  T  N  A  E  T  A  K  E  W  F  L  Q  A

                  910        920        930        940        950        960
901        CCAAGGATCCTTCTGCAGTGGCGAAGCACTTTGTTGCCCTGTCTACTAACACAACCAAAG
```

```
300       A  K  D  P  S  A  V  A  K  H  F  V  A  L  S  T  N  T  T  K

                970       980       990      1000      1010      1020
961       TGAAGGAGTTTGGAATTGACCCTCAAAACATGTTCGAGTTCTGGGATTGGGTGGGAGGAC
320       V  K  E  F  G  I  D  P  Q  N  M  F  E  F  W  D  W  V  G  G

               1030      1040      1050      1060      1070      1080
1021      GCTACTCGCTGTGGTCGGCCATCGGACTCTCCATTGCCCTGCACGTGGGTTTTGACAACT
340       R  Y  S  L  W  S  A  I  G  L  S  I  A  L  H  V  G  F  D  N

               1090      1100      1110      1120      1130      1140
1081      TCGAGCAGCTGCTCTCGGGGGCTCACTGGATGGACCAGCACTTCCGCACGACGCCCCTGG
360       F  E  Q  L  L  S  G  A  H  W  M  D  Q  H  F  R  T  T  P  L

               1150      1160      1170      1180      1190      1200
1141      AGAAGAACGCCCCCGTCTTGCTGGCCCTGCTGGGTATCTGGTACATCAACTGCTTTGGGT
380       E  K  N  A  P  V  L  L  A  L  L  G  I  W  Y  I  N  C  F  G

               1210      1220      1230      1240      1250      1260
1201      GTGAGACACACGCCATGCTGCCCTATGACCAGTACCTGCACCGCTTTGCTGCGTACTTCC
400       C  E  T  H  A  M  L  P  Y  D  Q  Y  L  H  R  F  A  A  Y  F

               1270      1280      1290      1300      1310      1320
1261      AGCAGGGCGACATGGAGTCCAATGGGAAATACATCACCAAATCTGGAACCCGTGTGGACC
420       Q  Q  G  D  M  E  S  N  G  K  Y  I  T  K  S  G  T  R  V  D

               1330      1340      1350      1360      1370      1380
1321      ACCAGACAGGCCCCATTGTGTGGGGGGAGCCAGGGACCAATGGCCAGCATGCTTTTTACC
440       H  Q  T  G  P  I  V  W  G  E  P  G  T  N  G  Q  H  A  F  Y

               1390      1400      1410      1420      1430      1440
1381      AGCTCATCCACCAAGGCACCAAGATGATACCCTGTGACTTCCTCATCCCGGTCCAGACCC
460       Q  L  I  H  Q  G  T  K  M  I  P  C  D  F  L  I  P  V  Q  T

               1450      1460      1470      1480      1490      1500
1441      AGCACCCCATACGGAAGGGTCTGCATCACAAGATCCTCCTGGCCAACTTCTTGGCCCAGA
480       Q  H  P  I  R  K  G  L  H  H  K  I  L  L  A  N  F  L  A  Q

               1510      1520      1530      1540      1550      1560
1501      CAGAGGCCCTGATGAGGGGAAAATCGACGGAGGAGGCCCGAAAGGAGCTCCAGGCTGCGG
500       T  E  A  L  M  R  G  K  S  T  E  E  A  R  K  E  L  Q  A  A

               1570      1580      1590      1600      1610      1620
1561      GCAAGAGTCCAGAGGACCTTGAGAGGCTGCTGCCACATAAGGTCTTTGAAGGAAATCGCC
520       G  K  S  P  E  D  L  E  R  L  L  P  H  K  V  F  E  G  N  R

               1630      1640      1650      1660      1670      1680
1621      CAACCAACTCTATTGTGTTCACCAAGCTCACACCATTCATGCTTGGAGCCTTGGTCGCCA
540       P  T  N  S  I  V  F  T  K  L  T  P  F  M  L  G  A  L  V  A

               1690      1700      1710      1720      1730      1740
1681      TGTATGAGCACAAGATCTTCGTTCAGGGCATCATCTGGGACATCAACAGCTTTGACCAGT
560       M  Y  E  H  K  I  F  V  Q  G  I  I  W  D  I  N  S  F  D  Q

               1750      1760      1770      1780      1790      1800
1741      GGGGAGTGGAGCTGGGAAAGCAGCTGGCTAAGAAAATAGAGCCTGAGCTTGATGGCAGTG
580       W  G  V  E  L  G  K  Q  L  A  K  K  I  E  P  E  L  D  G  S

               1810      1820      1830      1840      1850      1860
1801      CTCAAGTGACCTCTCACGACGCTTCTACCAATGGGCTCATCAACTTCATCAAGCAGCAGC
600       A  Q  V  T  S  H  D  A  S  T  N  G  L  I  N  F  I  K  Q  Q
```

```
              1870      1880
1861      GCGAGGCCAGAGTCCAATAA
620       R  E  A  R  V  Q  *
```

cDNA sequence of GusB

```
                10        20        30        40        50        60
1         GTCCTCAACCAAGATGGCGCGGATGGCTTCAGGCGCATCACGACACCGGCGCGTCACGCG
1

                70        80        90       100       110       120
61        ACCCGCCCTACGGGCACCTCCCGCGCTTTTCTTAGCGCCGCAGACGGTGGCCGAGCGGGG
20

               130       140       150       160       170       180
121       GACCGGGAAGCATGGCCCGGGGGTCGGCGGTTGCCTGGGCGGCGCTCGGGCCGTTGTTGT
40                M  A  R  G  S  A  V  A  W  A  A  L  G  P  L  L

               190       200       210       220       230       240
181       GGGGCTGCGCGCTGGGGCTGCAGGGCGGGATGCTGTACCCCCAGGAGAGCCCGTCGCGGG
60        W  G  C  A  L  G  L  Q  G  G  M  L  Y  P  Q  E  S  P  S  R

               250       260       270       280       290       300
241       AGTGCAAGGAGCTGGACGGCCTCTGGAGCTTCCGCGCCGACTTCTCTGACAACCGACGCC
80        E  C  K  E  L  D  G  L  W  S  F  R  A  D  F  S  D  N  R  R

               310       320       330       340       350       360
301       GGGGCTTCGAGGAGCAGTGGTACCGGCGGCCGCTGTGGGAGTCAGGCCCCACCGTGGACA
100       R  G  F  E  E  Q  W  Y  R  R  P  L  W  E  S  G  P  T  V  D

               370       380       390       400       410       420
361       TGCCAGTTCCCTCCAGCTTCAATGACATCAGCCAGGACTGGCGTCTGCGGCATTTTGTCG
120       M  P  V  P  S  S  F  N  D  I  S  Q  D  W  R  L  R  H  F  V

               430       440       450       460       470       480
421       GCTGGGTGTGGTACGAACGGGAGGTGATCCTGCCGGAGCGATGGACCCAGGACCTGCGCA
140       G  W  V  W  Y  E  R  E  V  I  L  P  E  R  W  T  Q  D  L  R

               490       500       510       520       530       540
481       CAAGAGTGGTGCTGAGGATTGGCAGTGCCCATTCCTATGCCATCGTGTGGGTGAATGGGG
160       T  R  V  V  L  R  I  G  S  A  H  S  Y  A  I  V  W  V  N  G

               550       560       570       580       590       600
541       TCGACACGCTAGAGCATGAGGGGGGCTACCTCCCCTTCGAGGCCGACATCAGCAACCTGG
180       V  D  T  L  E  H  E  G  G  Y  L  P  F  E  A  D  I  S  N  L

               610       620       630       640       650       660
601       TCCAGGTGGGGCCCCTGCCCTCCCGGCTCCGAATCACTATCGCCATCAACAACACACTCA
200       V  Q  V  G  P  L  P  S  R  L  R  I  T  I  A  I  N  N  T  L

               670       680       690       700       710       720
661       CCCCCACCACCCTGCCACCAGGGACCATCCAATACCTGACTGACACCTCCAAGTATCCCA
220       T  P  T  T  L  P  P  G  T  I  Q  Y  L  T  D  T  S  K  Y  P

               730       740       750       760       770       780
721       AGGGTTACTTTGTCCAGAACACATATTTTGACTTTTTCAACTACGCTGGACTGCAGCGGT
240       K  G  Y  F  V  Q  N  T  Y  F  D  F  F  N  Y  A  G  L  Q  R

               790       800       810       820       830       840
781       CTGTACTTCTGTACACGACACCCACCACCTACATCGATGACATCACCGTCACCACCAGCG
```

```
260       S  V  L  L  Y  T  T  P  T  T  Y  I  D  D  I  T  V  T  T  S

                850       860       870       880       890       900
841       TGGAGCAAGACAGTGGGCTGGTGAATTACCAGATCTCTGTCAAGGGCAGTAACCTGTTCA
280       V  E  Q  D  S  G  L  V  N  Y  Q  I  S  V  K  G  S  N  L  F

                910       920       930       940       950       960
901       AGTTGGAAGTGCGTCTTTTGGATGCAGAAAACAAAGTCGTGGCGAATGGGACTGGGACCC
300       K  L  E  V  R  L  L  D  A  E  N  K  V  V  A  N  G  T  G  T

                970       980       990      1000      1010      1020
961       AGGGCCAACTTAAGGTGCCAGGTGTCAGCCTCTGGTGGCCGTACCTGATGCACGAACGCC
320       Q  G  Q  L  K  V  P  G  V  S  L  W  W  P  Y  L  M  H  E  R

              1030      1040      1050      1060      1070      1080
1021      CTGCCTATCTGTATTCATTGGAGGTGCAGCTGACTGCACAGACGTCACTGGGGCCTGTGT
340       P  A  Y  L  Y  S  L  E  V  Q  L  T  A  Q  T  S  L  G  P  V

              1090      1100      1110      1120      1130      1140
1081      CTGACTTCTACACACTCCCTGTGGGGATCCGCACTGTGGCTGTCACCAAGAGCCAGTTCC
360       S  D  F  Y  T  L  P  V  G  I  R  T  V  A  V  T  K  S  Q  F

              1150      1160      1170      1180      1190      1200
1141      TCATCAATGGGAAACCTTTCTATTTCCACGGTGTCAACAAGCATGAGGATGCGGACATCC
380       L  I  N  G  K  P  F  Y  F  H  G  V  N  K  H  E  D  A  D  I

              1210      1220      1230      1240      1250      1260
1201      GAGGGAAGGGCTTCGACTGGCCGCTGCTGGTGAAGGACTTCAACCTGCTTCGCTGGCTTG
400       R  G  K  G  F  D  W  P  L  L  V  K  D  F  N  L  L  R  W  L

              1270      1280      1290      1300      1310      1320
1261      GTGCCAACGCTTTCCGTACCAGCCACTACCCCTATGCAGAGGAAGTGATGCAGATGTGTG
420       G  A  N  A  F  R  T  S  H  Y  P  Y  A  E  E  V  M  Q  M  C

              1330      1340      1350      1360      1370      1380
1321      ACCGCTATGGGATTGTGGTCATCGATGAGTGTCCCGGCGTGGGCCTGGCGCTGCCGCAGT
440       D  R  Y  G  I  V  V  I  D  E  C  P  G  V  G  L  A  L  P  Q

              1390      1400      1410      1420      1430      1440
1381      TCTTCAACAACGTTTCTCTGCATCACCACATGCAGGTGATGGAAGAAGTGGTGCGTAGGG
460       F  F  N  N  V  S  L  H  H  H  M  Q  V  M  E  E  V  V  R  R

              1450      1460      1470      1480      1490      1500
1441      ACAAGAACCACCCCGCGGTCGTGATGTGGTCTGTGGCCAACGAGCCTGCGTCCCACCTAG
480       D  K  N  H  P  A  V  V  M  W  S  V  A  N  E  P  A  S  H  L
                                                                target 1
              1510      1520      1530      1540      1550      1560
1501      AATCTGCTGGCTACTACTTGAAGATGGTGATCGCTCACACCAAATCCTTGGACCCCTCCC
500       E  S  A  G  Y  Y  L  K  M  V  I  A  H  T  K  S  L  D  P  S

              1570      1580      1590      1600      1610      1620
1561      GGCCTGTGACCTTTGTGAGCAACTCTAACTATGCAGCAGACAAGGGGGCTCCGTATGTGG
520       R  P  V  T  F  V  S  N  S  N  Y  A  A  D  K  G  A  P  Y  V

              1630      1640      1650      1660      1670      1680
1621      ATGTGATCTGTTTGAACAGCTACTACTCTTGGTATCACGACTACGGGCACCTGGAGTTGA
540       D  V  I  C  L  N  S  Y  Y  S  W  Y  H  D  Y  G  H  L  E  L

              1690      1700      1710      1720      1730      1740
1681      TTCAGCTGCAGCTGGCCACCCAGTTTGAGAACTGGTATAAGAAGTATCAGAAGCCCATTA
560       I  Q  L  Q  L  A  T  Q  F  E  N  W  Y  K  K  Y  Q  K  P  I
```

```
                1750      1760      1770      1780      1790      1800
1741        TTCAGAGCGAGTATGGAGCAGAAACGATTGCAGGGTTTCACCAGGATCCACCTCTGATGT
580         I  Q  S  E  Y  G  A  E  T  I  A  G  F  H  Q  D  P  P  L  M

                1810      1820      1830      1840      1850      1860
1801        TCACTGAAGAGTACCAGAAAAGTCTGCTAGAGCAGTACCATCTGGGTCTGGATCAAAAAC
600         F  T  E  E  Y  Q  K  S  L  L  E  Q  Y  H  L  G  L  D  Q  K

                1870      1880      1890      1900      1910      1920
1861        GCAGAAAATACGTGGTTGGAGAGCTCATTTGGAATTTTGCCGATTTCATGACTGAACAGT
620         R  R  K  Y  V  V  G  E  L  I  W  N  F  A  D  F  M  T  E  Q

                1930      1940      1950      1960      1970      1980
1921        CACCGACGAGAGTGCTGGGGAATAAAAAGGGGATCTTCACTCGGCAGAGACAACCAAAAA
640         S  P  T  R  V  L  G  N  K  K  G  I  F  T  R  Q  R  Q  P  K

                1990      2000      2010      2020      2030      2040
1981        GTGCAGCGTTCCTTTTGCGAGAGAGATACTGGAAGATTGCCAATGAAACCAGGTATCCCC
660         S  A  A  F  L  L  R  E  R  Y  W  K  I  A  N  E  T  R  Y  P

                2050      2060      2070      2080      2090      2100
2041        ACTCAGTAGCCAAGTCACAATGTTTGGAAAACAGCCTGTTTACTTGAGCAAGACTGATAC
680         H  S  V  A  K  S  Q  C  L  E  N  S  L  F  T  *

                2110      2120      2130      2140      2150      2160
2101        CACCTGCGTGTCCCTTCCTCCCCGAGTCAGGGCGACTTCCACAGCAGCAGAACAAGTGCC
700

                2170      2180      2190      2200      2210      2220
2161        TCCTGGACTGTTCACGGCAGACCAGAACGTTTCTGGCCTGGGTTTTGTGGTCATCTATTC
720

                2230      2240      2250      2260      2270      2280
2221        TAGCAGGGAACACTAAAGGTGGAAATAAAAGATTTTCTATTATGGAAATAAAGAGTTGGC
740
          target 2
                2290      2300      2310      2320
2281        ATGAAAGTGGCTACTGAAAAAAAAAAAAAAAAAAAAAAAAA
760
```

cDNA sequence of RAB7A

```
                  10        20        30        40        50        60
1           GTCTCGTGACAGGTACTTCCGCTCGGGGCGGCGGCGGTGGCGGAAGTGGGAGCGGGCCTG
1

                  70        80        90       100       110       120
61          GAGTCTTGGCCATAAAGCCTGAGGCGGCGGCAGCGGCGGAGTTGGCGGCTTGGAGAGCTC
21

                 130       140       150       160       170       180
121         GGGAGAGTTCCCTGGAACCAGAACTTGGACCTTCTCGCTTCTGTCCTCCGTTTAGTCTCC
41

                 190       200       210       220       230       240
181         TCCTCGGCGGGAGCCCTCGCGACGCGCCCGGCCCGGAGCCCCCAGCGCAGCGGCCGCGTT
61

                 250       260       270       280       290       300
241         TGAAGGATGACCTCTAGGAAGAAAGTGTTGCTGAAGGTTATCATCCTGGGAGATTCTGGA
```

```
81                M  T  S  R  K  K  V  L  L  K  V  I  I  L  G  D  S  G

                 310        320        330        340        350        360
301        GTCGGGAAGACATCACTCATGAACCAGTATGTGAATAAGAAATTCAGCAATCAGTACAAA
101         V  G  K  T  S  L  M  N  Q  Y  V  N  K  K  F  S  N  Q  Y  K

                 370        380        390        400        410        420
361        GCCACAATAGGAGCTGACTTTCTGACCAAGGAGGTGATGGTGGATGACAGGCTAGTCACA
121         A  T  I  G  A  D  F  L  T  K  E  V  M  V  D  D  R  L  V  T

                 430        440        450        460        470        480
421        ATGCAGATATGGGACACAGCAGGACAGGAACGGTTCCAGTCTCTCGGTGTGGCCTTCTAC
141         M  Q  I  W  D  T  A  G  Q  E  R  F  Q  S  L  G  V  A  F  Y

                 490        500        510        520        530        540
481        AGAGGTGCAGACTGCTGCGTTCTGGTATTTGATGTGACTGCCCCCAACACATTCAAAACC
161         R  G  A  D  C  C  V  L  V  F  D  V  T  A  P  N  T  F  K  T

                 550        560        570        580        590        600
541        CTAGATAGCTGGAGAGATGAGTTTCTCATCCAGGCCAGTCCCCGAGATCCTGAAAACTTC
181         L  D  S  W  R  D  E  F  L  I  Q  A  S  P  R  D  P  E  N  F

                 610        620        630        640        650        660
601        CCATTTGTTGTGTTGGGAAACAAGATTGACCTCGAAAACAGACAAGTGGCCACAAAGCGG
201         P  F  V  V  L  G  N  K  I  D  L  E  N  R  Q  V  A  T  K  R

                 670        680        690        700        710        720
661        GCACAGGCCTGGTGCTACAGCAAAAACAACATTCCCTACTTTGAGACCAGTGCCAAGGAG
221         A  Q  A  W  C  Y  S  K  N  N  I  P  Y  F  E  T  S  A  K  E

                 730        740        750        760        770        780
721        GCCATCAACGTGGAGCAGGCGTTCCAGACGATTGCACGGAATGCACTTAAGCAGGAAACG
241         A  I  N  V  E  Q  A  F  Q  T  I  A  R  N  A  L  K  Q  E  T

                 790        800        810        820        830        840
781        GAGGTGGAGCTGTACAACGAATTTCCTGAACCTATCAAACTGGACAAGAATGACCGGGCC
261         E  V  E  L  Y  N  E  F  P  E  P  I  K  L  D  K  N  D  R  A

                 850        860        870        880        890        900
841        AAGGCCTCGGCAGAAAGCTGCAGTTGCTGAGGGGGCAGTGAGAGTTGAGCACAGAGTCCT
281         K  A  S  A  E  S  C  S  C  *

                 910        920        930        940        950        960
901        TCACAAACCAAGAACACACGTAGGCCTTCAACACAATTCCCCTCTCCTCTTCCAAACAAA
301

                 970        980        990       1000       1010       1020
961        ACATACATTGATCTCTCACATCCAGCTGCCAAAAGAAAACCCCATCAAACACAGTTACAC
321

               1030       1040       1050       1060       1070       1080
1021       CCCACATATCTCTCACACACACACACACACGCACACACACACACACAGATCTGACGTAAT
341

               1090       1100       1110       1120       1130       1140
1081       CAAACTCCAGCCCTTGCCCGTGATGGCTCCTTGGGGTCTGCCTGCCCACCCACATGAGCC
361

               1150       1160       1170       1180       1190       1200
1141       CGCGAGTATGGCAGCAGGACAAGCCAGCGGTGGAAGTCATTCTGATATGGAGTTGGCATT
381
```

```
                1210      1220      1230      1240      1250      1260
1201      GGAAGCTTATTCTTTTTGTTCACTGGAGAGAGAGAGAACTGTTTACAGTTAATCTGTGTC
401

                1270      1280      1290      1300      1310      1320
1261      TAATTATCTGATTTTTTTTATTGGTCTTGTGGTCTTTTTACCCCCCCTTTCCCCTCCCTC
421

                1330      1340      1350      1360      1370      1380
1321      CTTGAAGGCTACCCCTTGGGAAGGCTGGTGCCCCATGCCCCATTACAGGCTCACACCCAG
441

                1390      1400      1410      1420      1430      1440
1381      TCTGATCAGGCTGAGTTTTGTATGTATCTATCTGTTAATGCTTGTTACTTTTAACTAATC
461

                1450      1460      1470      1480      1490      1500
1441      AGATCTTTTTACAGTATCCATTTATTATGTAATGCTTCTTAGAAAAGAATCTTATAGTAC
481

                1510      1520      1530      1540      1550      1560
1501      ATGTTAATATATGCAACCAATTAAAATGTATAAATTAGTGTAAGAAATTCTTGGATTATG
501

                1570      1580      1590      1600      1610      1620
1561      TGTTTAAGTCCTGTAATGCAGGCCTGTAAGGTGGAGGGTTGAACCCTGTTTGGATTGCAG
521

                1630      1640      1650      1660      1670      1680
1621      AGTGTTACTCAGAATTGGGAAATCCAGCTAGCGGCAGTATTCTGTACAGTAGACACAAGA
541                                 target 1

                1690      1700      1710      1720      1730      1740
1681      ATTATGTACGCCTTTTATCAAAGACTTAAGAGCCAAAAAGCTTTTCATCTCTCCAGGGGG
561

                1750      1760      1770      1780      1790      1800
1741      AAAACTGTCTAGTTCCCTTCTGTGTCTAAATTTTCCAAAACGTTGATTTGCATAATACAG
581         target 2

                1810      1820      1830      1840      1850      1860
1801      TGGTATGTGCAATGGATAAATTGCCGTTATTTCAAAAATTAAAATTCTCATTTTCTTTCT
601

                1870      1880      1890      1900      1910      1920
1861      TTTTTTTCCCCCCTGCTCCACACTTCAAAACTCCCGTTAGATCAGCATTCTACTACAAGA
621

                1930      1940      1950      1960      1970      1980
1921      GTGAAAGGAAAACCCTAACAGATCTGTCCTAGTGATTTTACCTTTGTTCTAGAAGGCGCT
641                                   target 3

                1990      2000      2010      2020      2030      2040
1981      CCTTTCAGGGTTGTGGTATTCTTAGGTTAGCGGAGCTTTTTCCTCTTTTCCCCACCCATC
661

                2050      2060      2070      2080      2090      2100
2041      TCCCCAATATTGCCCATTATTAATTAACCTCTTTCTTTGGTTGGAACCCTGGCAGTTCTG
681
```

```
                2110       2120       2130       2140       2150       2160
2101        CTCCCTTCCTAGGATCTGCCCCTGCATTGTAGCTTGCTTAACGGAGCACTTCTCCTTTTT
701

                2170       2180       2190       2200       2210       2220
2161        CCAAAGGTCTACATTCTAGGGTGTGGGCTGAGTTCTTCTGTAAAGAGATGAACGCAATGC
721

                2230       2240
2221        CAATAAAATTGAACAAGAACAATG
741
```

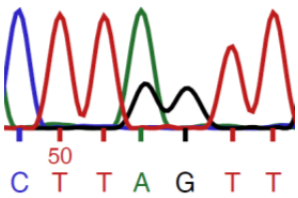
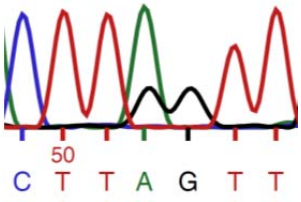
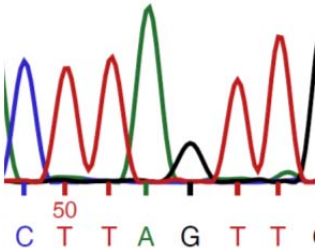
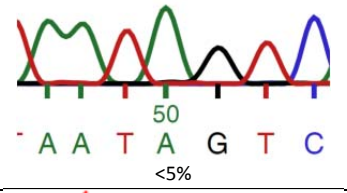
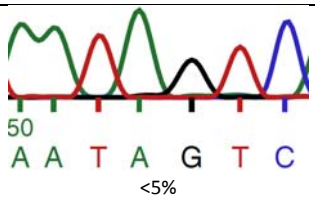
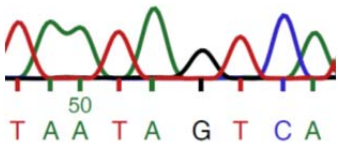
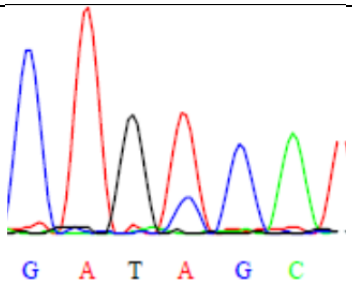
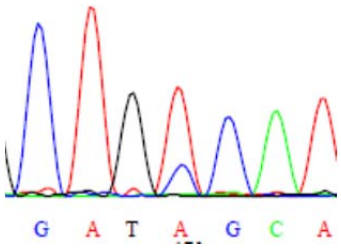
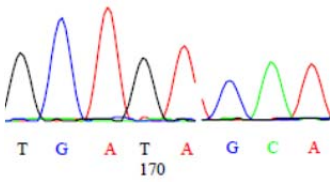
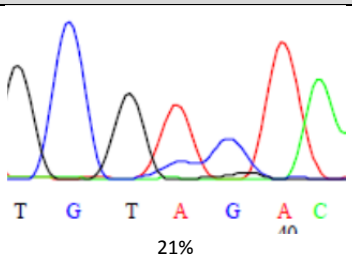
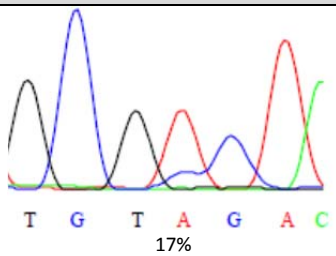
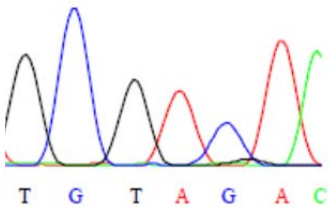
Only 3’ UTR sequence of VCP (very long transcript, thus only this part is shown)

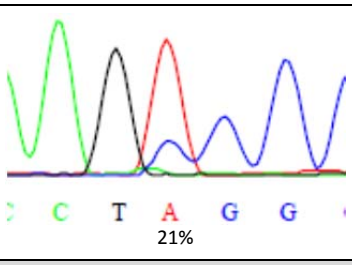
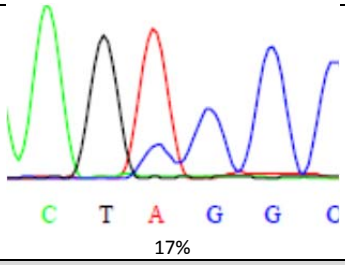
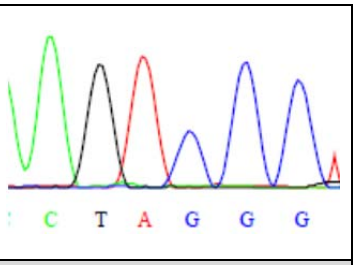
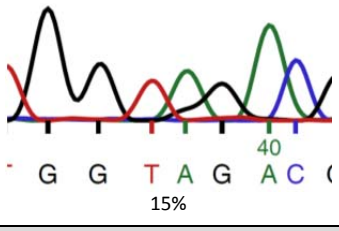
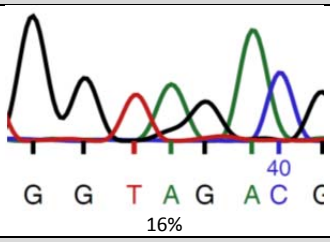
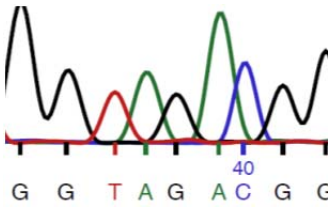
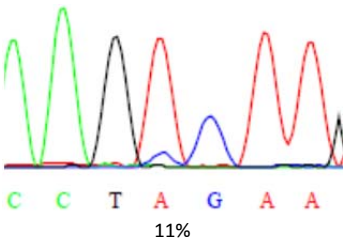
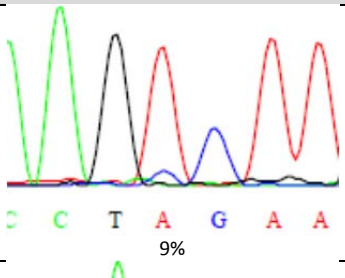
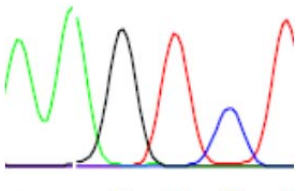
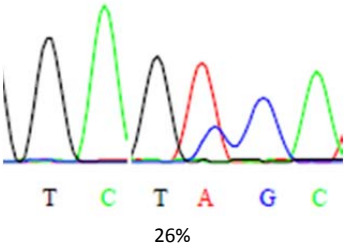
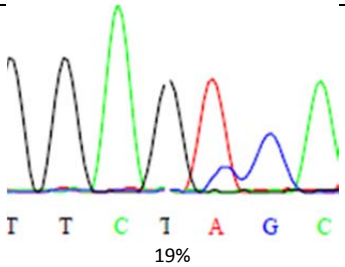
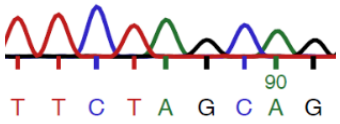
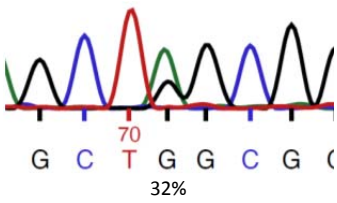
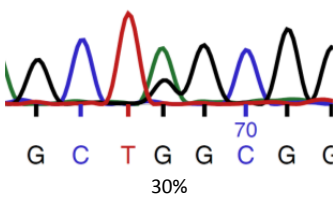
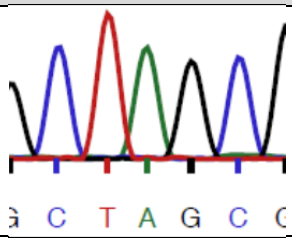
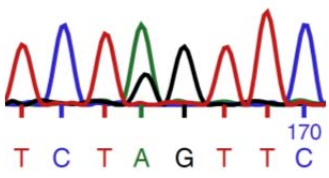
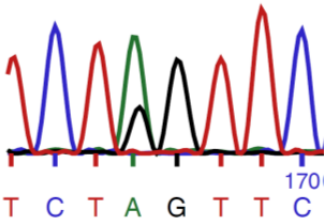
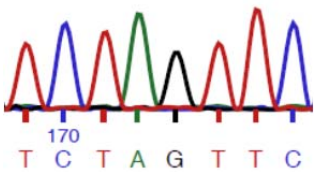
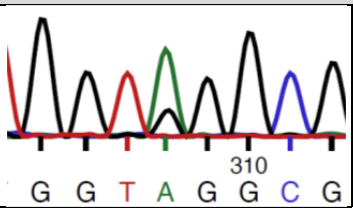
```
                  10         20         30         40         50         60
1           …ATTCCCTTCAGGGAACCAGGGTGGAGCTGGCCCCAGTCAGGGCAGTGGAGGCGGCACAGG
1
                  70         80         90        100        110        120
61          TGGCAGTGTATACACAGAAGACAATGATGATGACCTGTATGGCTAAGTGGTGGTGGCCAG
21
                 130        140        150        160        170        180
121         CGTGCAGTGAGCTGGCCTGCCTGGACCTTGTTCCCTGGGGGTGGGGGCGCTTGCCCAGGA
41
                 190        200        210        220        230        240
181         GAGGGACCAGGGGTGCGCCCACAGCCTGCTCCATTCTCCAGTCTGAACAGTTCAGCTACA
61
                 250        260        270        280        290        300
241         GTCTGACTCTGGACAGGGGGTTTCTGTTGCAAAAATACAAAACAAAAGCGATAAAATAAA
81
                 310        320        330        340        350        360
301         AGCGATTTTCATTTGGTAGGCGGAGAGTGAATTACCAACAGGGAATTGGGCCTTGGGCCT
101                    target 1
                 370        380        390        400        410        420
361         ATGCCATTTCTGTTGTAGTTTGGGGCAGTGCAGGGGACCTGTGTGGGGTGTGAACCAAGG
121                    target 2
                 430        440        450        460        470        480
421         CACTACTGCCACCTGCCACAGTAAAGCATCTGCACTTGACTCAATGCTGCCCGAGCCCTC
141
                 490        500        510        520        530        540
481         CCTTCCCCCTATCCAACCTGGGTAGGTGGGTAGGGGCCACAGTTGCTGGATGTTTATATA
161                          target 3
                 550        560        570        580        590        600
541         GAGAGTAGGTTGATTTATTTTACATGCTTTTGAGTTAATGTTGGAAAACTAATCACAAGC
181
                 610        620        630        640        650        660
601         AGTTTCTAAACCAAAAAATGACATGTTGTAAAAGGACAATAAACGTTGGGTCAAAATGGA
201
                 670        680        690        700        710        720
661         GCCTGAGTCCTGGGCCCTGTGCCTGCTTCTTTTCCTGGGAACAGCCTTGGGCTACCCACC
221
                 730        740        750        760        770        780
721         ACTCCCAAGGCATTCTTCCAAATGTGAAATCCTGGAAGTAAGATTGCACCTTCTTCCTCT
241
                 790        800        810        820        830        840
781         CCTGATCAACATCGGTATGATGTCTCCTGTTGCCTCACCCTTTGTCTGCAGTATCACTGG
261
                 850        860        870        880        890        900
841         ATAGGACTGGTGGAAAGGGAGCAGCCTGACAGAGCTCCAAATGTGGAGAATATGGCATCC
281
                 910        920        930        940        950        960
901         CTCCACCTATATTTGATGTGGACGGTAAGGCTAGGCCTGCAGGATCCCTTATCCTGACCA
```

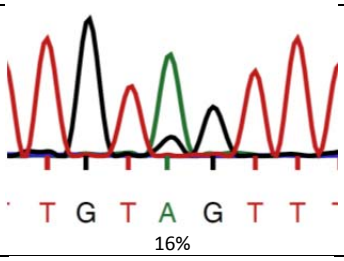
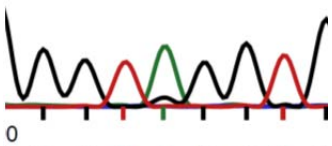
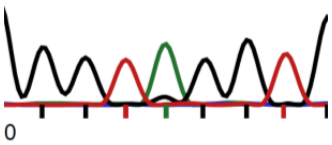
```
301
                970        980        990       1000       1010       1020
961         AAGACTGTGTTGGGGTGCCATTTGAAAATCGCAGGGTTGCAAAAGAATACAATCTTACTT
321
               1030       1040       1050       1060       1070       1080
1021        GCAGGTGGATATTCTCTATACTCTCTTTTAATGCATCTAAAAATCCCAAACATCCCCTGG
341
               1090       1100       1110       1120       1130       1140
1081        TTGGTGATCACTTACAGTTGTGTCCACCTTTATTTTATGTACTTTGATTAAAAAAAAAAA
361
               1150
1141        ACTTTTTGTTAATATAAAA
381
```

Figure S19. Sequencing results. *Not all RT-PCR reactions were successful or gave sufficient sequencing quality. Those are indicated here with n.d. Primers used in RT-PCR are given on Table S2.*

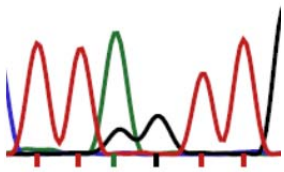
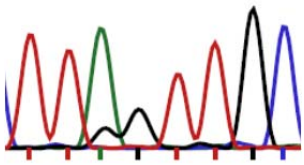
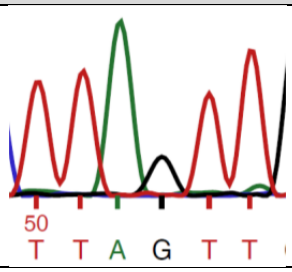
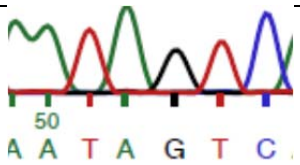
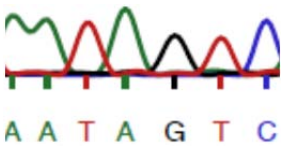
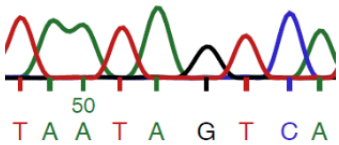
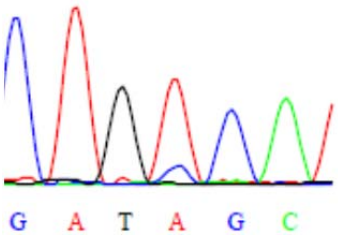
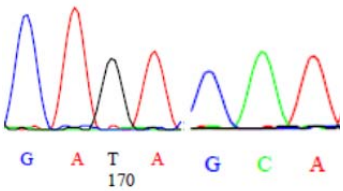
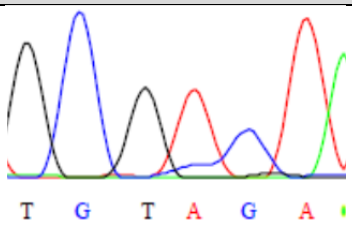
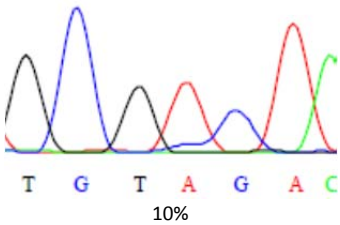
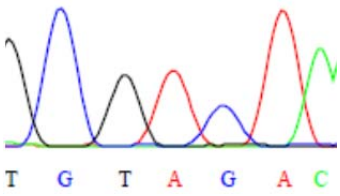
a) Editing in 293T-Cells with ADAR2 overexpression

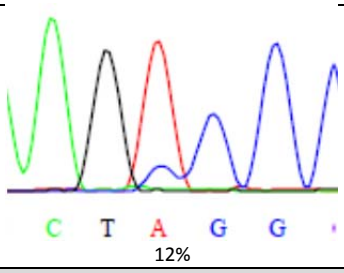
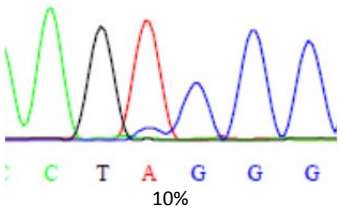
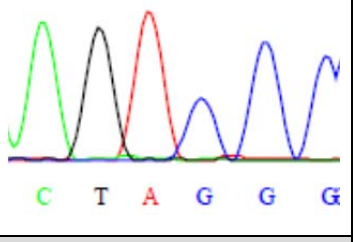
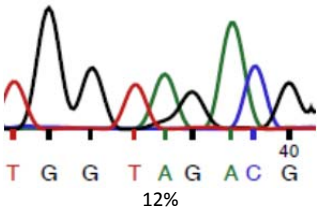
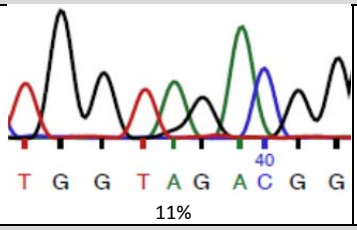
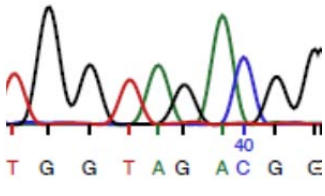
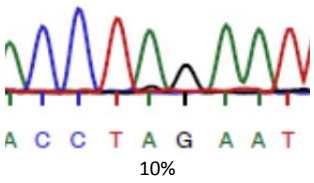
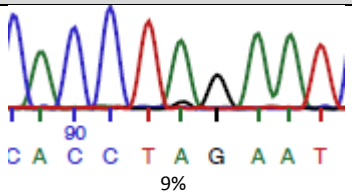
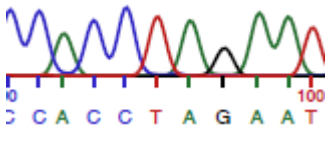
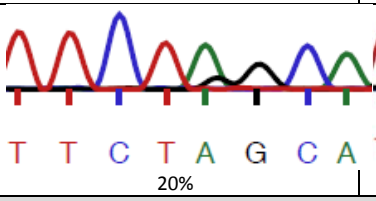
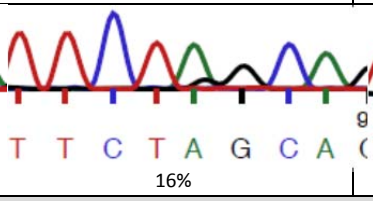
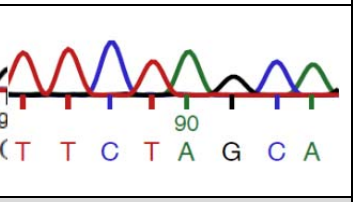
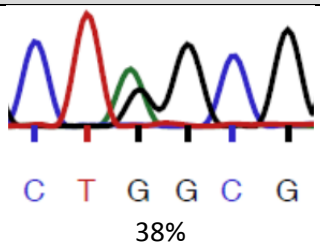
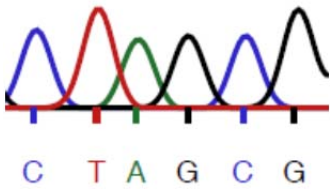
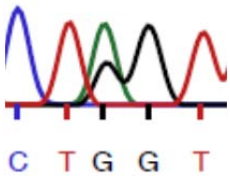
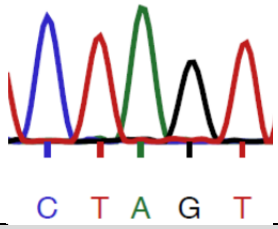
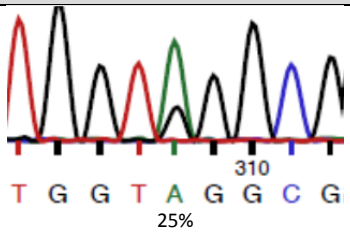
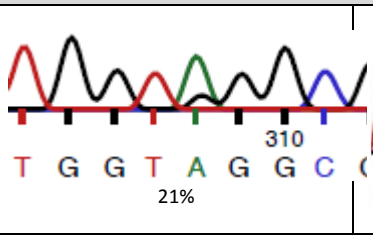
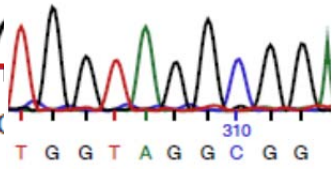
settings	experiment A:	experiment B:	negative control:
General Settings: 300 ng pTS57 (ADAR2), 1300 ng gRNA, DNA to Lipofectamine Ratio 1:3, negative control just without guideRNA the rest is treated equally			
B-actin			
TAG#1			
TAG#2			
TAG#3			
GAPDH			
TAG#1			

TAG#2			
GPI			
TAG#1			
GUSB			
TAG#1			
TAG#2			
RAB7A			
TAG#1			
TAG#2			
VCP			
TAG#1	n.d.		n.d.

		25%	
TAG#2	 16%	n.d.	n.d.
TAG#3	 12%	 11%	n.d.

cont. Figure S19 b) Editing in 293-ADAR2-Flip-In-Cells

settings	experiment A:	experiment B:	negative control:
General Settings: 1300 ng gRNA, DNA to Lipofectamine Ratio 1:3; negative control just empty transfection			
B-actin			
TAG#1	 17%	 14%	
TAG#2	 <5%	 <5%	
TAG#3	 14%	n.d.	
GAPDH			
TAG#1	 10%	 10%	

TAG#2			
GPI			
TAG#1			
GUSB			
TAG#1			
TAG#2			
RAB7A			
TAG#1		n.d.	
TAG#2		n.d.	
VCP			
TAG#1			

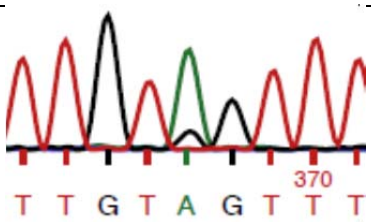
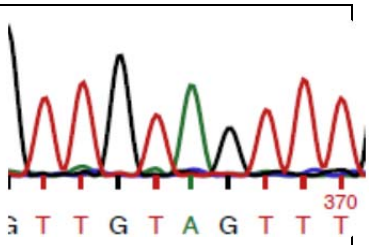
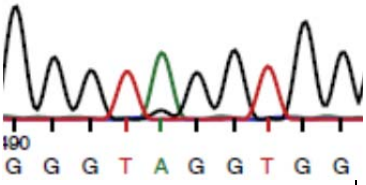
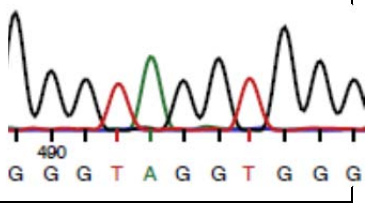
TAG#2	 15%	n.d.	
TAG#3	 13%	n.d	

Table S2.

Primers used for reverse-transcription and PCR:	Sequence 5'->3':
Beta-Aktin_fw	CAGCAGATGTGGATCAGCAAGCAGGAG
Beta-Aktin_bw	GGAAGGGGGGGCACGAAGGCTCATC
Beta-Aktin_Seq_fw	GGTGACAGCAGTCGGTTGGAGCGAGC
GAPDH_fw	CTCAAGATCATCAGCAATGCCTCCTGC
GAPDH_bw	GAGCACAGGGTACTTTATTGATGGTACATGACAAGG
GAPDH_Seq_fw	CACTGCTGGGGAGTCCCTGCCAC
GPI_fw_1	CTACACCGAGGGTCGAGCCGTGCTG
GPI_fw_2	CCTGGACACCACCCAGAGCACCCTC
GPI_fw_3	GGGAGTACAGGCACCTGCCACCATG
GPI_bw_1	CAGGGCAACAAAGTGCTTCGCCACTGC
GPI_bw_2	CAGAGGTGAGGAGTGGAAAACAGTCTTGGG
GPI_bw_3	CAGCTATGATTGTATCACTGCAGTCCAGCCTG
GUSB_fw_1	CAACCAAGATGGCGCGGATGGCTTCAGG
GUSB_fw_2	CGCTGCCGCAGTTCTTCAACAACGTTTCTCTG
GUSB_bw_1	GTTGATGGCGATAGTGATTCGGAGCCGGG
GUSB_bw_2	CAGTAGCCACTTTCATGCCAACTCTTTATTTCCATAATAG
GUSB_seq_fw1	CCTGCGTGTCCCTTCCTCCCCG
GUSB_seq_fw2	CCGGCCTGTGACCTTTGTGAGCAACTC
RAB7A_fw	CTTGGATTATGTGTTTAAGTCCTGTAATGCAGGCC
RAB7A_bw	GGAGCAGAACTGCCAGGGTTCCAACC
RAB7A_Seq_fw	CAGTGGTATGTGCAATGGATAAATTGCCG
VCP_3'UTR_Exon-Junction_fw	GTCGGGGCTTTGGCAGCTTCAGATTCC
VCP_3'UTR_bw	CCTACTCTCTATATAAACATCCAGCAACTGTGGCC

PINK1 Editing

Gene sequence of **wt PINK1** with a C-terminal V5 and His_6-tag in the context of the pcDNA3.1 vector. PINK1 is controlled by the CMV promoter and the BGH terminator:

```
                10        20        30        40        50        60
1          CTG**GCTAGCATG**GCGGTGCGACAGGCGCTGGGCCGCGGCCTGCAGCTGGGTCGAGCGCTG
1              Nhe-I M  A  V  R  Q  A  L  G  R  G  L  Q  L  G  R  A  L

                70        80        90       100       110       120
61         CTGCTGCGCTTCACGGGCAAGCCCGGCCGGGCCTACGGCTTGGGGCGGCCGGGCCCGGCG
21          L  L  R  F  T  G  K  P  G  R  A  Y  G  L  G  R  P  G  P  A

               130       140       150       160       170       180
121        GCGGGCTGTGTCCGCGGGGAGCGTCCAGGCTGGGCCGCAGGACCGGGCGCGGAGCCTCGC
41          A  G  C  V  R  G  E  R  P  G  W  A  A  G  P  G  A  E  P  R

               190       200       210       220       230       240
181        AGGGTCGGGCTCGGGCTCCCTAACCGTCTCCGCTTCTTCCGCCAGTCGGTGGCCGGGCTG
61          R  V  G  L  G  L  P  N  R  L  R  F  F  R  Q  S  V  A  G  L

               250       260       270       280       290       300
241        GCGGCGCGGTTGCAGCGGCAGTTCGTGGTGCGGGCCTGGGGCTGCGCGGGCCCTTGCGGC
81          A  A  R  L  Q  R  Q  F  V  V  R  A  W  G  C  A  G  P  C  G

               310       320       330       340       350       360
301        CGGGCAGTCTTTCTGGCCTTCGGGCTAGGGCTGGGCCTCATCGAGGAAAAACAGGCGGAG
101         R  A  V  F  L  A  F  G  L  G  L  G  L  I  E  E  K  Q  A  E

               370       380       390       400       410       420
361        AGCCGGCGGGCGGTCTCGGCCTGTCAGGAGATCCAGGCAATTTTTACCCAGAAAAGCAAG
121         S  R  R  A  V  S  A  C  Q  E  I  Q  A  I  F  T  Q  K  S  K

               430       440       450       460       470       480
421        CCGGGGCCTGACCCGTTGGACACGAGACGCTTGCAGGGCTTTCGGCTGGAGGAGTATCTG
141         P  G  P  D  P  L  D  T  R  R  L  Q  G  F  R  L  E  E  Y  L

               490       500       510       520       530       540
481        ATAGGGCAGTCCATTGGTAAGGGCTGCAGTGCTGCTGTGTATGAAGCCACCATGCCTACA
161         I  G  Q  S  I  G  K  G  C  S  A  A  V  Y  E  A  T  M  P  T

               550       560       570       580       590       600
541        TTGCCCCAGAACCTGGAGGTGACAAAGAGCACCGGGTTGCTTCCAGGGAGAGGCCCAGGT
181         L  P  Q  N  L  E  V  T  K  S  T  G  L  L  P  G  R  G  P  G

               610       620       630       640       650       660
601        ACCAGTGCACCAGGAGAAGGGCAGGAGCGAGCTCCGGGGGCCCCTGCCTTCCCCTTGGCC
201         T  S  A  P  G  E  G  Q  E  R  A  P  G  A  P  A  F  P  L  A

               670       680       690       700       710       720
661        ATCAAGATGATGTGGAACATCTCGGCAGGTTCCTCCAGCGAAGCCATCTTGAACACAATG
221         I  K  M  M  W  N  I  S  A  G  S  S  S  E  A  I  L  N  T  M

               730       740       750       760       770       780
721        AGCCAGGAGCTGGTCCCAGCGAGCCGAGTGGCCTTGGCTGGGGAGTATGGAGCAGTCACT
241         S  Q  E  L  V  P  A  S  R  V  A  L  A  G  E  Y  G  A  V  T

               790       800       810       820       830       840
781        TACAGAAAATCCAAGAGAGGTCCCAAGCAACTAGCCCCTCACCCCAACATCATCCGGGTT
```

```
261        Y  R  K  S  K  R  G  P  K  Q  L  A  P  H  P  N  I  I  R  V

                 850        860        870        880        890        900
841        CTCCGCGCCTTCACCTCTTCCGTGCCGCTGCTGCCAGGGGCCCTGGTCGACTACCCTGAT
281         L  R  A  F  T  S  S  V  P  L  L  P  G  A  L  V  D  Y  P  D

                 910        920        930        940        950        960
901        GTGCTGCCCTCACGCCTCCACCCTGAAGGCCTGGGCCATGGCCGGACGCTCTTTCTAGTC
301         V  L  P  S  R  L  H  P  E  G  L  G  H  G  R  T  L  F  L  V

                 970        980        990       1000       1010       1020
961        ATGAAGAACTATCCCTGTACCCTGCGCCAGTACCTTTGTGTGAACACACCCAGCCCCCGC
321         M  K  N  Y  P  C  T  L  R  Q  Y  L  C  V  N  T  P  S  P  R

               1030       1040       1050       1060       1070       1080
1021       CTCGCCGCCATGATGCTGCTGCAGCTGCTGGAAGGCGTGGACCATCTGGTTCAACAGGGC
341         L  A  A  M  M  L  L  Q  L  L  E  G  V  D  H  L  V  Q  Q  G

               1090       1100       1110       1120       1130       1140
1081       ATCGCGCACAGAGACCTGAAATCCGACAACATCCTTGTGGAGCTGGACCCAGACGGCTGC
361         I  A  H  R  D  L  K  S  D  N  I  L  V  E  L  D  P  D  G  C

               1150       1160       1170       1180       1190       1200
1141       CCCTGGCTGGTGATCGCAGATTTTGGCTGCTGCCTGGCTGATGAGAGCATCGGCCTGCAG
381         P  W  L  V  I  A  D  F  G  C  C  L  A  D  E  S  I  G  L  Q

               1210       1220       1230       1240       1250       1260
1201       TTGCCCTTCAGCAGCTGGTACGTGGATCGGGGCGGAAACGGCTGTCTGATGGCCCCAGAG
401         L  P  F  S  S  W  Y  V  D  R  G  G  N  G  C  L  M  A  P  E

               1270       1280       1290       1300       1310       1320
1261       GTGTCCACGGCCCGTCCTGGCCCCAGGGCAGTGATTGACTACAGCAAGGCTGATGCCTGG
421         V  S  T  A  R  P  G  P  R  A  V  I  D  Y  S  K  A  D  A  W

               1330       1340       1350       1360       1370       1380
1321       GCAGTGGGAGCCATCGCCTATGAAATCTTCGGGCTTGTCAATCCCTTCTACGGCCAGGGC
441         A  V  G  A  I  A  Y  E  I  F  G  L  V  N  P  F  Y  G  Q  G

               1390       1400       1410       1420       1430       1440
1381       AAGGCCCACCTTGAAAGCCGCAGCTACCAAGAGGCTCAGCTACCTGCACTGCCCGAGTCA
461         K  A  H  L  E  S  R  S  Y  Q  E  A  Q  L  P  A  L  P  E  S

               1450       1460       1470       1480       1490       1500
1441       GTGCCTCCAGACGTGAGACAGTTGGTGAGGGCACTGCTCCAGCGAGAGGCCAGCAAGAGA
481         V  P  P  D  V  R  Q  L  V  R  A  L  L  Q  R  E  A  S  K  R

               1510       1520       1530       1540       1550       1560
1501       CCATCTGCCCGAGTAGCCGCAAATGTGCTTCATCTAAGCCTCTGGGGTGAACATATTCTA
501         P  S  A  R  V  A  A  N  V  L  H  L  S  L  W  G  E  H  I  L

               1570       1580       1590       1600       1610       1620
1561       GCCCTGAAGAATCTGAAGTTAGACAAGATGGTTGGCTGGCTCCTCCAACAATCGGCCGCC
521         A  L  K  N  L  K  L  D  K  M  V  G  W  L  L  Q  Q  S  A  A

               1630       1640       1650       1660       1670       1680
1621       ACTTTGTTGGCCAACAGGCTCACAGAGAAGTGTTGTGTGGAAACAAAAATGAAGATGCTC
541         T  L  L  A  N  R  L  T  E  K  C  C  V  E  T  K  M  K  M  L

               1690       1700       1710       1720       1730       1740
1681       TTTCTGGCTAACCTGGAGTGTGAAACGCTCTGCCAGGCAGCCCTCCTCCTCTGCTCATGG
561         F  L  A  N  L  E  C  E  T  L  C  Q  A  A  L  L  L  C  S  W
```

```
                1750      1760      1770      1780      1790      1800
1741       AGGGCAGCCCTGCTCGAGTCTAGAGGGCCCTTCGAAGGTAAGCCTATCCCTAACCCTCTC
581         R  A  A  L  L  E  S  R  G  P  F  E  G  K  P  I  P  N  P  L

                1810      1820      1830      1840      1850
1801       CTCGGTCTCGATTCTACGCGTACCGGTCATCATCACCATCACCATTGAGTTTAAACCCG
601         L  G  L  D  S  T  R  T  G  H  H  H  H  H  H  *    Pme-I
```

Gene & protein sequence of PINK1 W437X amber with a C-terminal V5- and His_6-tag in the context of the pcDNA 3.1 vector, under control of the the CMV promoter and the BGH terminator. The amber Stop codon is highlighted in yellow.:

```
                  10        20        30        40        50        60
1          CTGGCTAGCATGGCGGTGCGACAGGCGCTGGGCCGCGGCCTGCAGCTGGGTCGAGCGCTG
1             Nhe-I  M  A  V  R  Q  A  L  G  R  G  L  Q  L  G  R  A  L

                  70        80        90       100       110       120
61         CTGCTGCGCTTCACGGGCAAGCCCGGCCGGGCCTACGGCTTGGGGCGGCCGGGCCCGGCG
21          L  L  R  F  T  G  K  P  G  R  A  Y  G  L  G  R  P  G  P  A

                 130       140       150       160       170       180
121        GCGGGCTGTGTCCGCGGGGAGCGTCCAGGCTGGGCCGCAGGACCGGGCGCGGAGCCTCGC
41          A  G  C  V  R  G  E  R  P  G  W  A  A  G  P  G  A  E  P  R

                 190       200       210       220       230       240
181        AGGGTCGGGCTCGGGCTCCCTAACCGTCTCCGCTTCTTCCGCCAGTCGGTGGCCGGGCTG
61          R  V  G  L  G  L  P  N  R  L  R  F  F  R  Q  S  V  A  G  L

                 250       260       270       280       290       300
241        GCGGCGCGGTTGCAGCGGCAGTTCGTGGTGCGGGCCTGGGGCTGCGCGGGCCCTTGCGGC
81          A  A  R  L  Q  R  Q  F  V  V  R  A  W  G  C  A  G  P  C  G

                 310       320       330       340       350       360
301        CGGGCAGTCTTTCTGGCCTTCGGGCTAGGGCTGGGCCTCATCGAGGAAAAACAGGCGGAG
101         R  A  V  F  L  A  F  G  L  G  L  G  L  I  E  E  K  Q  A  E

                 370       380       390       400       410       420
361        AGCCGGCGGGCGGTCTCGGCCTGTCAGGAGATCCAGGCAATTTTTACCCAGAAAAGCAAG
121         S  R  R  A  V  S  A  C  Q  E  I  Q  A  I  F  T  Q  K  S  K

                 430       440       450       460       470       480
421        CCGGGGCCTGACCCGTTGGACACGAGACGCTTGCAGGGCTTTCGGCTGGAGGAGTATCTG
141         P  G  P  D  P  L  D  T  R  R  L  Q  G  F  R  L  E  E  Y  L

                 490       500       510       520       530       540
481        ATAGGGCAGTCCATTGGTAAGGGCTGCAGTGCTGCTGTGTATGAAGCCACCATGCCTACA
161         I  G  Q  S  I  G  K  G  C  S  A  A  V  Y  E  A  T  M  P  T

                 550       560       570       580       590       600
541        TTGCCCCAGAACCTGGAGGTGACAAAGAGCACCGGGTTGCTTCCAGGGAGAGGCCCAGGT
181         L  P  Q  N  L  E  V  T  K  S  T  G  L  L  P  G  R  G  P  G

                 610       620       630       640       650       660
601        ACCAGTGCACCAGGAGAAGGGCAGGAGCGAGCTCCGGGGGCCCCTGCCTTCCCCTTGGCC
201         T  S  A  P  G  E  G  Q  E  R  A  P  G  A  P  A  F  P  L  A

                 670       680       690       700       710       720
```

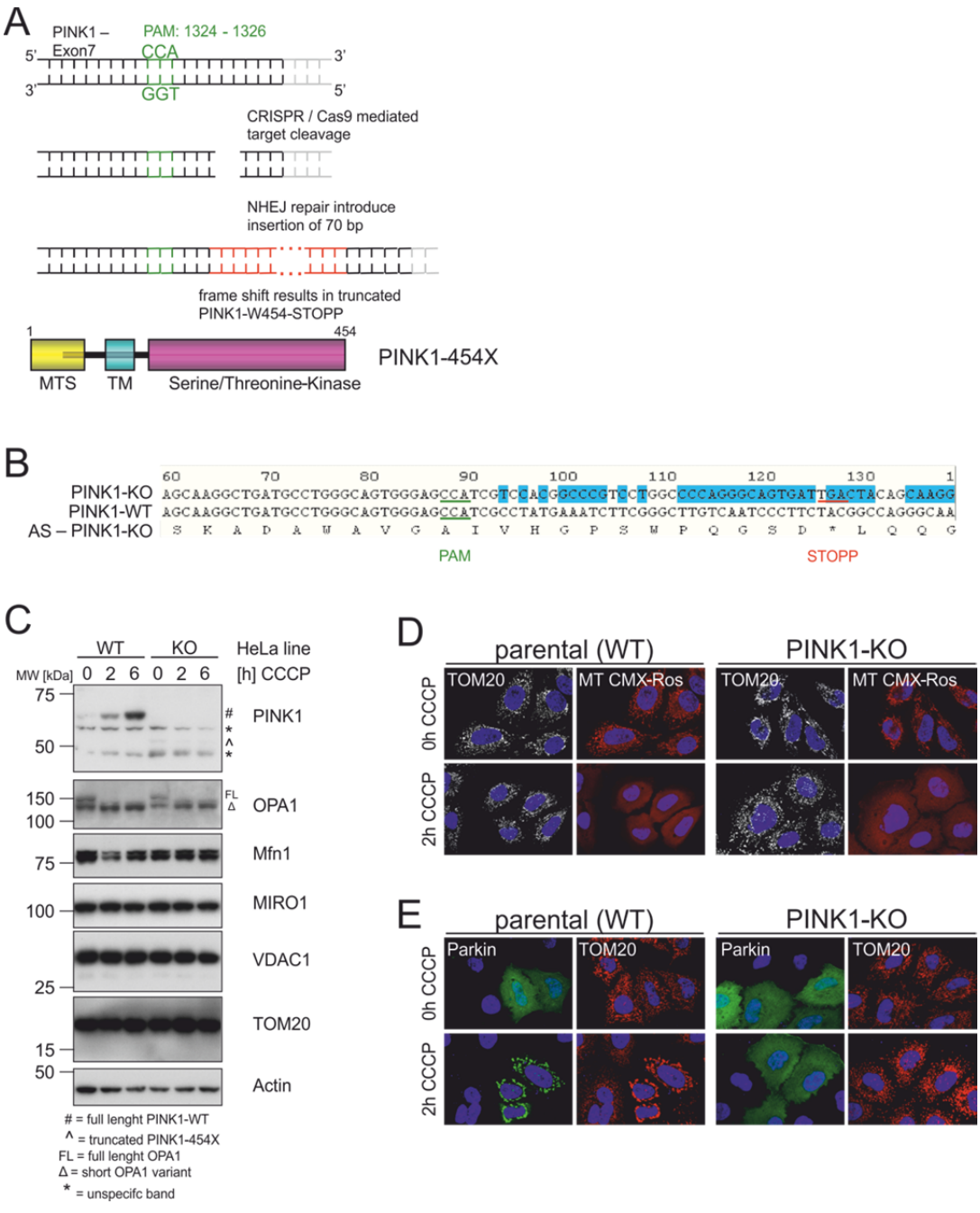
```
661       ATCAAGATGATGTGGAACATCTCGGCAGGTTCCTCCAGCGAAGCCATCTTGAACACAATG
221        I  K  M  M  W  N  I  S  A  G  S  S  S  E  A  I  L  N  T  M

                 730       740       750       760       770       780
721       AGCCAGGAGCTGGTCCCAGCGAGCCGAGTGGCCTTGGCTGGGGAGTATGGAGCAGTCACT
241        S  Q  E  L  V  P  A  S  R  V  A  L  A  G  E  Y  G  A  V  T

                 790       800       810       820       830       840
781       TACAGAAAATCCAAGAGAGGTCCCAAGCAACTAGCCCCTCACCCCAACATCATCCGGGTT
261        Y  R  K  S  K  R  G  P  K  Q  L  A  P  H  P  N  I  I  R  V

                 850       860       870       880       890       900
841       CTCCGCGCCTTCACCTCTTCCGTGCCGCTGCTGCCAGGGGCCCTGGTCGACTACCCTGAT
281        L  R  A  F  T  S  S  V  P  L  L  P  G  A  L  V  D  Y  P  D

                 910       920       930       940       950       960
901       GTGCTGCCCTCACGCCTCCACCCTGAAGGCCTGGGCCATGGCCGGACGCTCTTTCTAGTC
301        V  L  P  S  R  L  H  P  E  G  L  G  H  G  R  T  L  F  L  V

                 970       980       990      1000      1010      1020
961       ATGAAGAACTATCCCTGTACCCTGCGCCAGTACCTTTGTGTGAACACACCCAGCCCCCGC
321        M  K  N  Y  P  C  T  L  R  Q  Y  L  C  V  N  T  P  S  P  R

               1030      1040      1050      1060      1070      1080
1021      CTCGCCGCCATGATGCTGCTGCAGCTGCTGGAAGGCGTGGACCATCTGGTTCAACAGGGC
341        L  A  A  M  M  L  L  Q  L  L  E  G  V  D  H  L  V  Q  Q  G

               1090      1100      1110      1120      1130      1140
1081      ATCGCGCACAGAGACCTGAAATCCGACAACATCCTTGTGGAGCTGGACCCAGACGGCTGC
361        I  A  H  R  D  L  K  S  D  N  I  L  V  E  L  D  P  D  G  C

               1150      1160      1170      1180      1190      1200
1141      CCCTGGCTGGTGATCGCAGATTTTGGCTGCTGCCTGGCTGATGAGAGCATCGGCCTGCAG
381        P  W  L  V  I  A  D  F  G  C  C  L  A  D  E  S  I  G  L  Q

               1210      1220      1230      1240      1250      1260
1201      TTGCCCTTCAGCAGCTGGTACGTGGATCGGGGCGGAAACGGCTGTCTGATGGCCCCAGAG
401        L  P  F  S  S  W  Y  V  D  R  G  G  N  G  C  L  M  A  P  E

               1270      1280      1290      1300      1310      1320
1261      GTGTCCACGGCCCGTCCTGGCCCCAGGGCAGTGATTGACTACAGCAAGGCTGATGCCTAG
421        V  S  T  A  R  P  G  P  R  A  V  I  D  Y  S  K  A  D  A  *

               1330      1340      1350      1360      1370      1380
1321      GCAGTGGGAGCCATCGCCTATGAAATCTTCGGGCTTGTCAATCCCTTCTACGGCCAGGGC
441        A  V  G  A  I  A  Y  E  I  F  G  L  V  N  P  F  Y  G  Q  G

               1390      1400      1410      1420      1430      1440
1381      AAGGCCCACCTTGAAAGCCGCAGCTACCAAGAGGCTCAGCTACCTGCACTGCCCGAGTCA
461        K  A  H  L  E  S  R  S  Y  Q  E  A  Q  L  P  A  L  P  E  S

               1450      1460      1470      1480      1490      1500
1441      GTGCCTCCAGACGTGAGACAGTTGGTGAGGGCACTGCTCCAGCGAGAGGCCAGCAAGAGA
481        V  P  P  D  V  R  Q  L  V  R  A  L  L  Q  R  E  A  S  K  R

               1510      1520      1530      1540      1550      1560
1501      CCATCTGCCCGAGTAGCCGCAAATGTGCTTCATCTAAGCCTCTGGGGTGAACATATTCTA
501        P  S  A  R  V  A  A  N  V  L  H  L  S  L  W  G  E  H  I  L

               1570      1580      1590      1600      1610      1620
1561      GCCCTGAAGAATCTGAAGTTAGACAAGATGGTTGGCTGGCTCCTCCAACAATCGGCCGCC
```

```
521        A  L  K  N  L  K  L  D  K  M  V  G  W  L  L  Q  Q  S  A  A

                1630       1640       1650       1660       1670       1680
1621       ACTTTGTTGGCCAACAGGCTCACAGAGAAGTGTTGTGTGGAAACAAAAATGAAGATGCTC
541         T  L  L  A  N  R  L  T  E  K  C  C  V  E  T  K  M  K  M  L

                1690       1700       1710       1720       1730       1740
1681       TTTCTGGCTAACCTGGAGTGTGAAACGCTCTGCCAGGCAGCCCTCCTCCTCTGCTCATGG
561         F  L  A  N  L  E  C  E  T  L  C  Q  A  A  L  L  L  C  S  W

                1750       1760       1770       1780       1790       1800
1741       AGGGCAGCCCTGCTCGAGTCTAGAGGGCCCTTCGAAGGTAAGCCTATCCCTAACCCTCTC
581         R  A  A  L  L  E  S  R  G  P  F  E  G  K  P  I  P  N  P  L

                1810       1820       1830       1840       1850
1801       CTCGGTCTCGATTCTACGCGTACCGGTCATCATCACCATCACCATTGA**GTTTAAACCCG**
601         L  G  L  D  S  T  R  T  G  H  H  H  H  H  H  *    Pme-I
```

Generation and Characterization of the PINK1-Knock out HeLa cell line

To genomically engineer a PINK1-knock out HeLa cell line, the CRISPR/Cas9 system was used. Therefore, the guideRNA sequence 5'-**CCA**TCGCCTATGAAATC-3' (with the PAM (protospacer adjacent motif) in bold) within exon 7 of PINK1, was inserted into the Cas9 expression plasmid pSpCas9(BB)-2A-Puro (F. A. Ran et al. Genome engineering using the CRISPR-Cas9 system. Nat. Protocols, 2013, doi:10.1038/nprot.2013.143) to yield pSpCas9(PINK1-KO)-2A-PURO. The parental HeLa cells were transfected with 2 µg of the pSpCas9(PINK1-KO)-2A-PURO plasmid using the Human Stem Cell Nucleofector Kit (Lonza) and a nucleofector device (Amaxa). 48h after transfection, the cells were selected using 2 µg/ml puromycin in DMEM+10% FCS for 48h. The genomic DNA from clones was isolated using the the ReliaPrep gDNA Tissue Miniprep System (Promega). The PINK1 exon 7 was amplified by flanking intronic primers (forward: 5' TGGATCAGGTGATGTGCAGGA 3' and reverse: 5' AGGATCTGTCACTGTG GCTCT 3'). The CRISPR/Cas9 genomic editing resulted in an introduction of a 70 bp DNA fragment into PINK1 exon 7, 4 bp upstream of the PAM sequence. This disrupts the functional PINK1 allele by introducing a premature Stop codon at position 454 inside the kinase domain (A, B). Due to the premature Stop codon, the resulting PINK1 protein is non-functional. While in parental wt Hela cells a stabilization of full length PINK1 is seen after CCCP treatment (#, Fig. S21 C), no PINK1 protein expression in mutant HeLa cells was detected (Fig. S21 C). Indeed, only faint amounts of the unstable truncated PINK1-454X were detected (^, Fig. S21 C). We thus call this cell line "PINK1 knock-out (KO)" even though a truncated PINK1 protein of similar length like the pathogenic W437X mutant could potentially be expressed. The characterization of the PINK1-KO cells highlights no differences in mitochondrial protein levels compared to wt HeLa cells (Fig. S21 C). In addition, PINK1 functional null did not interfere with the disruption of the mitochondrial membrane potential by CCCP, since the loss of the long OPA1 isoform (Fig. S21 C) as well as the loss of mitochondrial staining, with the membrane potential sensitive MitoTracker Red CMX-ROS was observed (Fig. S21 D). The mutated PINK1 is in addition also unable to induce perinuclear co-localization of Parkin and mitochondria (see microscopy pictures, Fig. S21 D and E). Thus, we considered that the PINK1-KO Hela cells are functional null regarding PINK1. Important, due to the insertion mutation this non-functional transcript is not repaired by site-directed RNA editing.

Figure S20. Generation and chracterization of the PINK1-KO HeLa cell line
Besides the antibodies described on page 61-62, the following additional antibodies were used:
rabbit anti-OPA1 (BD Transduction Laboratories, #612606)
mouse anti-Mfn1 (Abnova, H00055669-M04)
rabbit anti-VDAC1 (Millipore, Ab10527)
mouse anti-MIRO1 (Sigma, WH0055288M1).



PINK1 Functional Assay

HeLa cells (PINK1 wt or KO) were cultured under standard conditions (DMEM + 10% FBS, 37°C, 5% CO_2). The mitophagy assay was performed in 24-wells. Each well contained a cover-slip coated with poly-D-lysine (Sigma Aldrich). The cells were seeded at $2.5x10^4$/well. After 24h the cells were transfected with the indicated plasmids using FuGene6 (Promega). If the editing vector was transfected a 1-to-6 ratio was used. If ADAR2 or guideRNA alone were transfected, a 1-to-3 ratio was applied. If not indicated, the plasmid amounts/well has been 300 ng for EGFP-Parkin, 300 ng for PINK1 wt or PINK1 W437amber, 300 ng editing vector. In control experiment d) 1300 ng of a guideRNA plasmid based on pSilencer lacking ADAR2 was co-transfected instead of the editing vector. In control e) 200 ng of an editing vector lacking any guideRNAs but containing ADAR2 was co-transfected instead of the original editing vector. Treatment with 10 µM CCCP (in DMEM + 10% FBS) was either performed 46h after transfection for 2h or 24h after transfection for 24 h. The depolarization of the mitochondrial membrane potential with 10 µM CCCP (in DMEM + 10% FBS) was either performed 46h after transfection for 2h or 24h after transfection for 24 h. To visualize the mitochondria with a membrane potential sensitive dye, like MitoTracker Red CMXROS, a CCCP wash out was performed. For this, the depolarizing agent CCCP was washed out by changing the media twice every 15 min. Then the cells were incubated with 100 nM MitoTracker Red CMXROS (Invitrogen, M7512) in DMEM for 30 min at 37°C directly prior fixation or harvesting. Then 48 h after transfection, the cells were washed once with PBS and then either fixated for immunocytochemical staining (A) or harvested for RNA isolation (B).

(A) After fixation with 4% PFA/PBS for 20 minutes at RT and three wash steps with PBS, the cells were permeabilized with 1% Triton X-100/PBS for 5 minutes at RT followed by three wash steps with PBS. Then, the cells were blocked with 10% FCS/PBS for 1h at RT and stained with following antibodies diluted in 5% FCS/PBS for 2h at RT: mouse anti-ADAR2 (Santa Cruz, SC-73409, 1:1000), and rabbit anti-PINK1 (Novus Biologicals, BC 100-949, 1:750). The cells were then washed three times with PBS and incubated with the following secondary antibodies diluted in 10% FCS/PBS: goat anti-mouse or rabbit Alexa Fluor-488, 568 or 647 (Invitrogen, 1:1000). After two washing steps with PBS the nuclei of the cells where stained with Hoechst33342 (Thermo Fisher, 1:5000) for 5 minutes at RT. The cover-slips where mounted onto glass-slides using the Dako fluorescent mounting medium. Confocal images with a slice thickness of 0,7 µm were obtained with an AxioImager microscope equipped with an ApoTome imaging system (Carl Zeiss) using a 63x objective. The images were processed using the AxioVision software 4.8.1. For the quantification of Parkin clustering, double- (EGFP-Parkin and PINK1) and triple- (EGFP-Parkin, PINK1 and hADAR2) positive cells were evaluated. More than 100 cells per cover slip and condition were analyzed for quantification.

(B) The cells were harvested by trypsination using 60 µl Trypsin-EDTA. After inactivation with 440 µl DMEM+10%FBS the cells were peletted at 300 g for 5 minutes at 4°C. After removing the supernatant the cells where washed once in 500 µl ice-cold PBS and centrifuged again at 300 g for 5 minutes at 4°C. The cell pellet was snap frozen in liquid nitrogen prior RNA isolation using the RNeasy Mini Kit (Qiagen). Remaining Plasmid DNA that could interfere with following PCRs was removed by DNase-I (NEB) digestion for 10 min at 37°C. Then, the RNA was reverse transcribed using 1 µl M-MuLV reverse transcriptase (NEB), 0,25 µl murine RNase inhibitor (NEB) and a BHG backward primer (5´-CTAGAAGGCACAGTCGAGGC). The cDNA was isolated using the nucleospin gel and PCR-cleanup kit (Machery-Nagel). The following Phusion-PCR was performed with 4% DMSO, 0.8 M betaine and a PINK1 specific primer pair 5´-CTAGA AGGCACAGTCGAGGC and 5´-GAGCCAGGAGCTGGTCCCAGCGAGCCGAG". The 1167 bp

PINK1 DNA fragment was isolated from the 1,4% agarose gel using the nucleospin gel and PCR-cleanup kit (Machery-Nagel). The sequencing was performed by the company Eurofins using the sequencing primer 5´-GGTACGTGGATCAGGGCG GAAACGGC.

Figure S21. Western Blot analysis *of PINK1-W437X-V5 editing. HeLa WT or PINK1-KO cells were transfected with plasmids for wt PINK1-V5 or PINK1-W437X-V5, and with or without the editing vector for 48h as indicated. The cells were treated with 10 µM CCCP 6h prior lysis with 8M Urea buffer (10 mM Tris (pH 8.0), 100 mM NaH_2PO_4, 8M urea). 10 µg of total lysates were separated by SDS-PAGE and transferred onto PVDF membrane using the wet blotting method. The following primary antibodies were used: mouse anti-ADAR2 (Santa Cruz, sc-73409), rabbit anti-PINK1 (Novus Biologicals, BC 100-494), mouse anti-V5 (Invitrogen, R960-25, rabbit anti-TOM20 (Santa Cruz, sc-11415), and mouse anti beta-actin (Sigma, clone AC-15). In panel* ***A****, the potential PINK1 variants are shown.In panel* ***B****, the PINK1 antibody stains all processed and truncated versions of the wt and the overexpressed PINK1 variants. Upon editing a small fraction of fulllength PINK1 (FL & ΔFL) appears beside an excess of truncated PINK1 (X & ΔX) which is in accordance with an editing yield of 10%. The V5 antibody detects fulllength V5-tagged PINK1 (FL & ΔFL) only in presence of the editing vector, which again indicates the successful editing of the transfected PINK1-W437X-V5. After editing and CCCP treatment TOM20 ubiquitinylation (Ub-1, Ub-2, Ub-3) is visible similar to wt PINK1.*

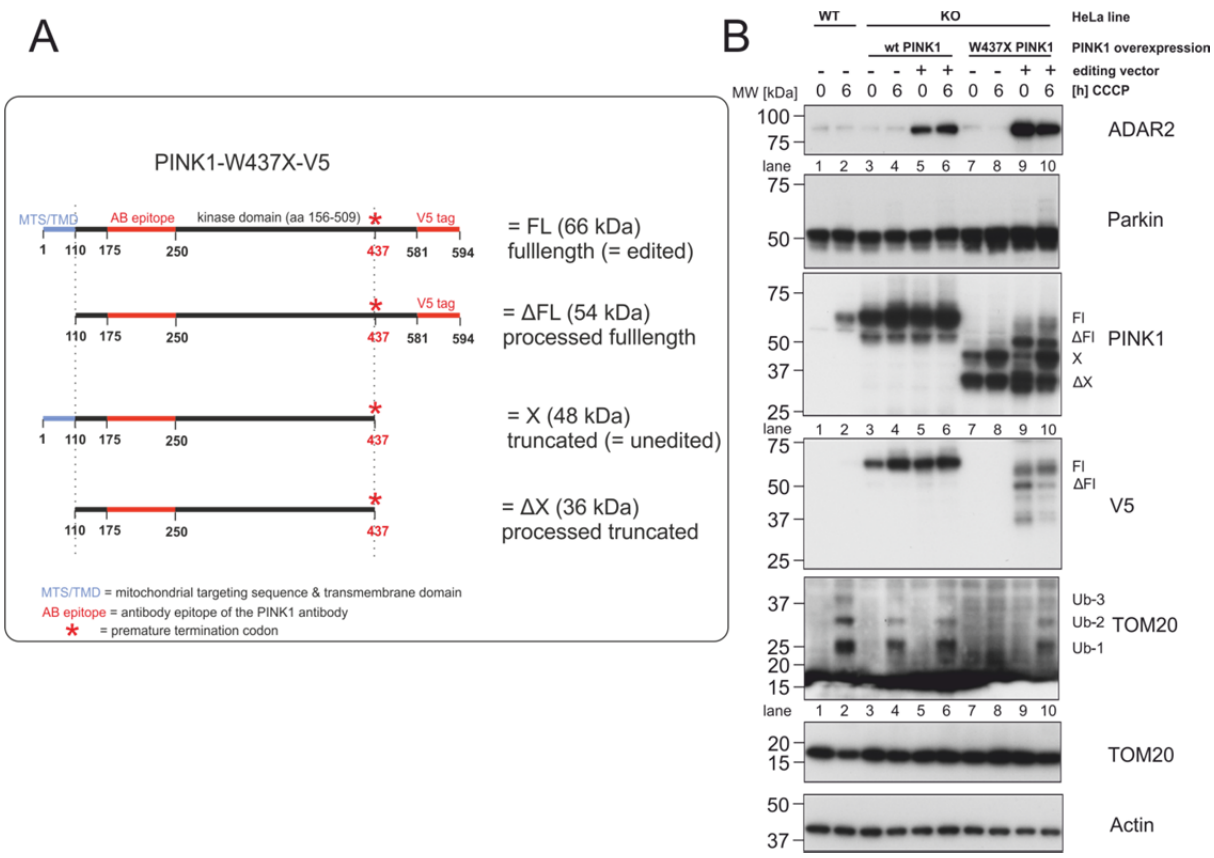


Figure S22. *Full sequencing trace of the PINK1 W437amber editing experiment in HeLa cells, corresponding to the experiment shown in Figure 4A, 4B c).*

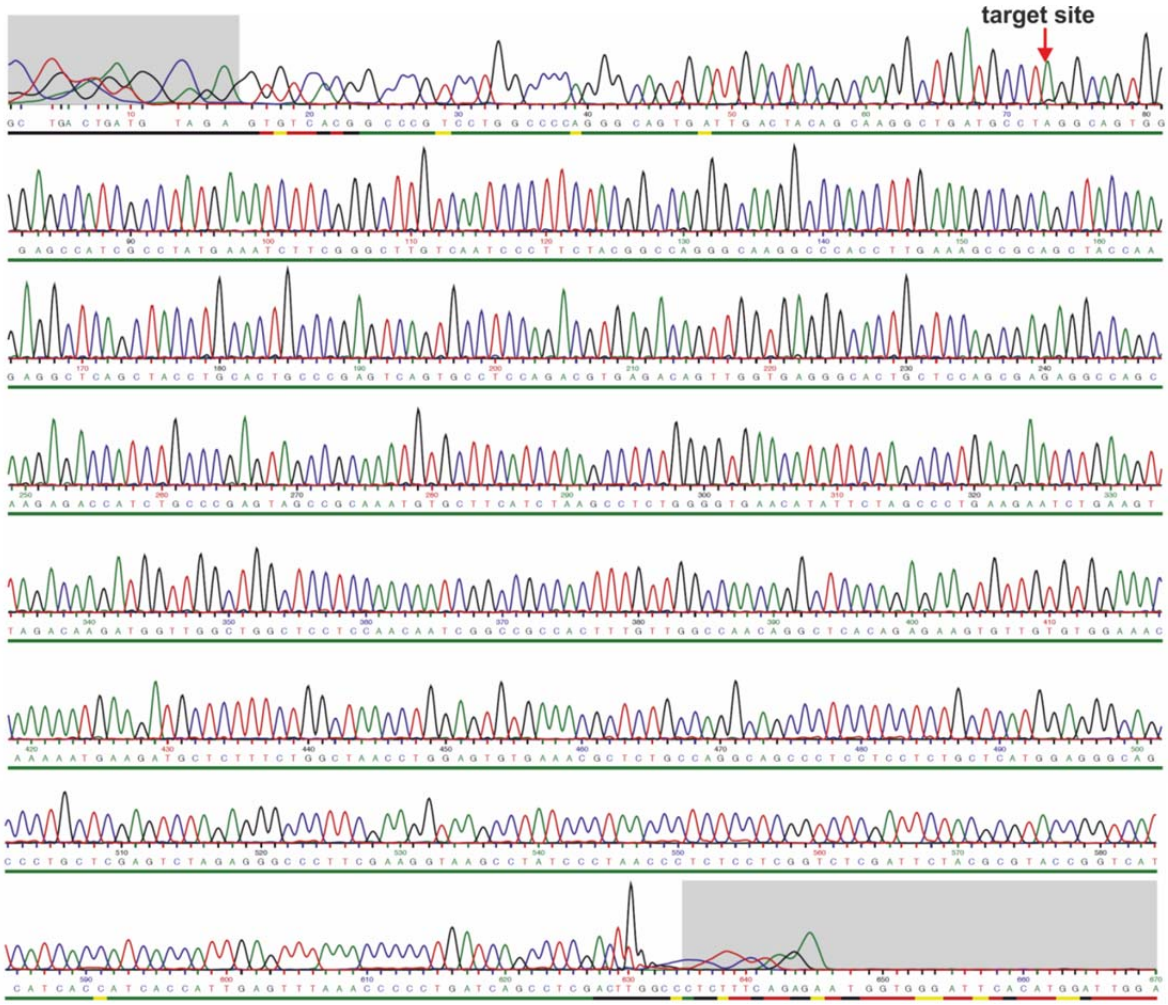
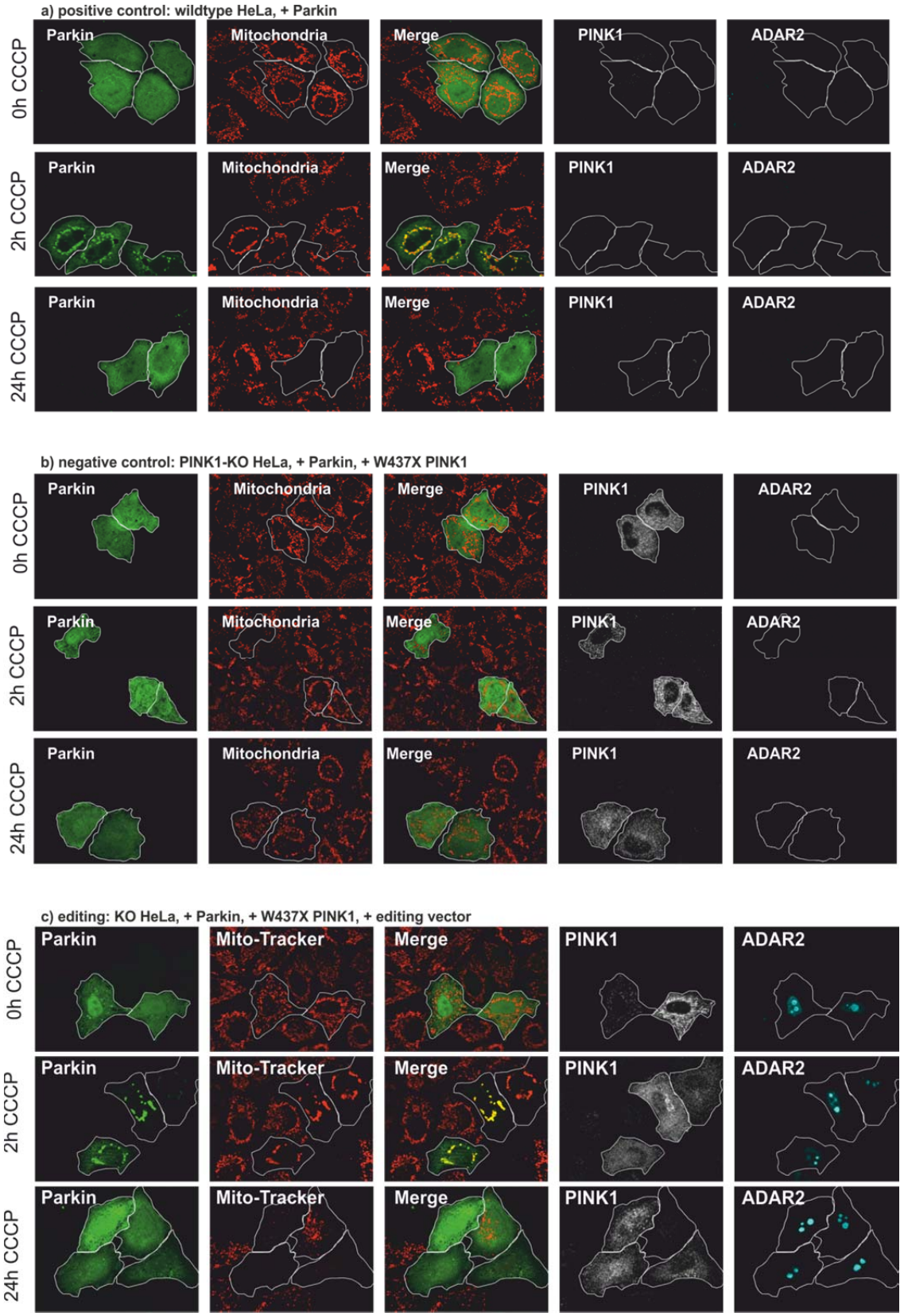
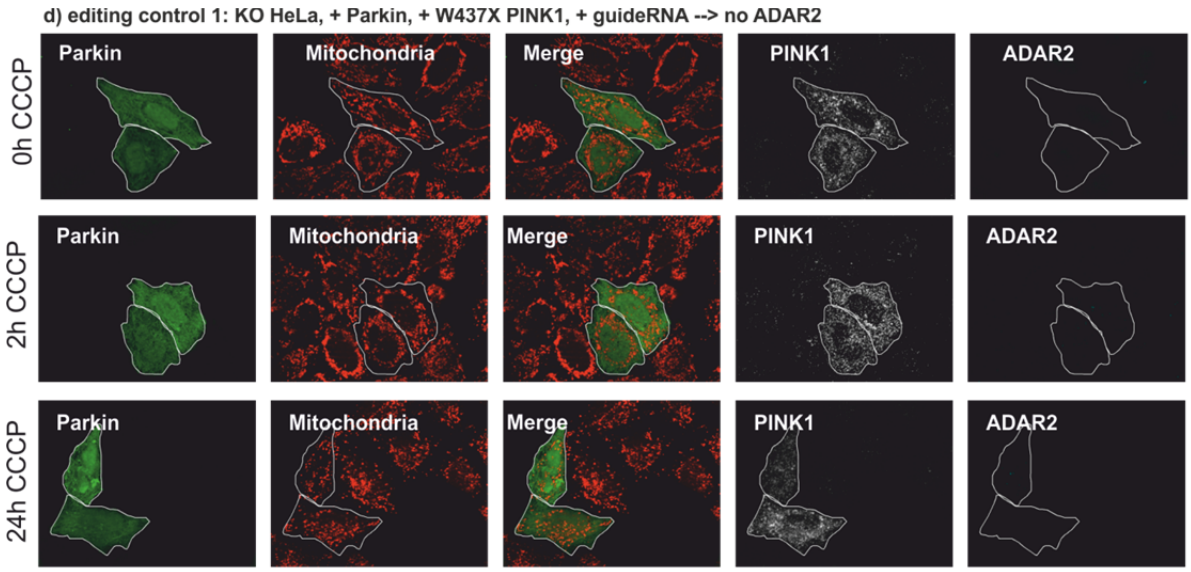
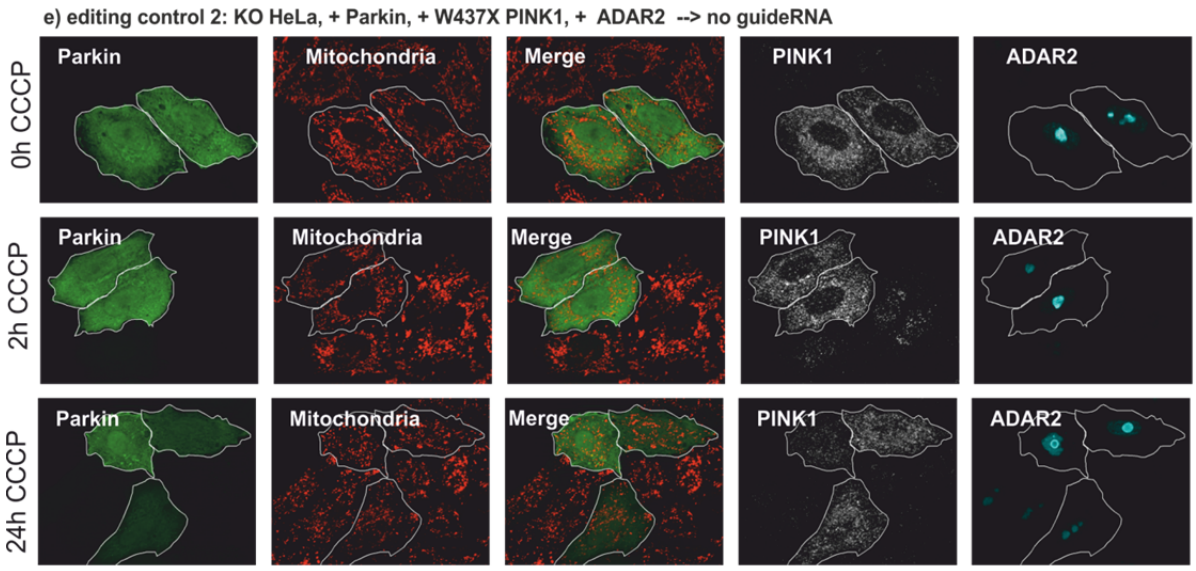
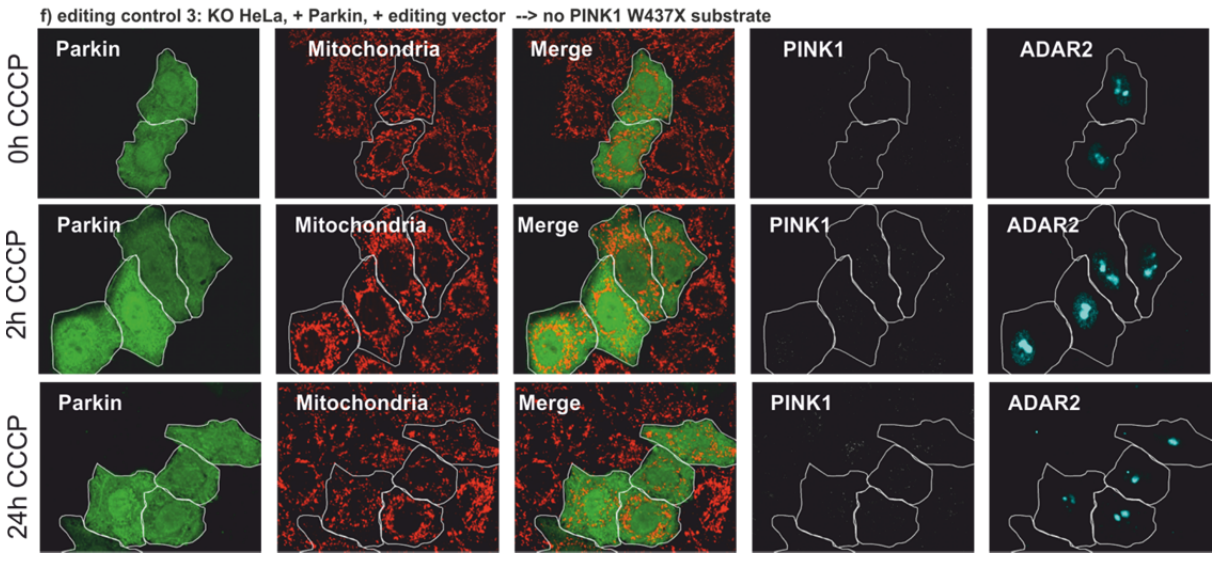


Figure S23. Mitophagy assay. Shown are the complete mitophagy assays that relate to the Parkin-clustering experiments a) – f) shown on Figure 4A in the main text. The mitophagy assay shown in c) is the same as shown in the main text Figure 4C. For better visualization of mitophagy, the Parkin-positive cells were encircled.




d) editing control 1: KO HeLa, + Parkin, + W437X PINK1, + guideRNA --> no ADAR2
0h CCCP
2h CCCP
24h CCCP
Parkin
Mitochondria
Merge
PINK1
ADAR2


e) editing control 2: KO HeLa, + Parkin, + W437X PINK1, + ADAR2 --> no guideRNA
0h CCCP
2h CCCP
24h CCCP
Parkin
Mitochondria
Merge
PINK1
ADAR2


f) editing control 3: KO HeLa, + Parkin, + editing vector --> no PINK1 W437X substrate
0h CCCP
2h CCCP
24h CCCP
Parkin
Mitochondria
Merge
PINK1
ADAR2

6.6.3 PUBLICATION 3: HEEP M., MACH P., REAUTSCHNIG P., WETTENGEL J., STAFFORST T. APPLYING HUMAN ADAR1P110 AND ADAR1P150 FOR SITE-DIRECTED RNA EDITING-G/C SUBSTITUTION STABILIZES GUIDERNAS AGAINST EDITING. GENES (BASEL), 2017. 8(1).

genes

MDPI

Article

Applying Human ADAR1p110 and ADAR1p150 for Site-Directed RNA Editing—G/C Substitution Stabilizes GuideRNAs against Editing

Madeleine Heep, Pia Mach, Philipp Reautschnig, Jacqueline Wettengel and Thorsten Stafforst *

Interfaculty Institute of Biochemistry, University of Tübingen, Auf der Morgenstelle 15, 72076 Tübingen, Germany; mad.heep@web.de (M.H.); pia.mach@student.uni-tuebingen.de (P.M.); philipp.reautschnig@uni-tuebingen.de (P.R.); jacqueline.wettengel@uni-tuebingen.de (J.W.)
* Correspondence: thorsten.stafforst@uni-tuebingen.de; Tel.: +49-7071-2975-376

Academic Editor: H. Ulrich Göringer
Received: 25 November 2016; Accepted: 6 January 2017; Published: 14 January 2017

Abstract: Site-directed RNA editing is an approach to reprogram genetic information at the RNA level. We recently introduced a novel guideRNA that allows for the recruitment of human ADAR2 to manipulate genetic information. Here, we show that the current guideRNA design is already able to recruit another human deaminase, ADAR1, in both isoforms, p110 and p150. However, further optimization seems necessary as the current design is less efficient for ADAR1 isoforms. Furthermore, we describe hotspots at which the guideRNA itself is edited and show a way to circumvent this auto-editing without losing editing efficiency at the target. Both findings are important for the advancement of site-directed RNA editing as a tool in basic biology or as a platform for therapeutic editing.

Keywords: site-directed RNA editing; ADAR; guideRNA; genetic disease; RNA repair

1. Introduction

Site-directed RNA editing is a method to recode genetic information at the RNA level [1]. The approach is based on the enzymatic conversion of adenosine (A) to inosine (I). Inosine is interpreted as guanosine in many biochemical processes including translation, thus A-to-I RNA editing allows the recoding of amino acids, splice elements, miRNAs, and miRNA binding sites among others [2]. We and others have recently developed methods, called site-directed RNA editing, that employ engineered deaminases in combination with short guideRNAs to recode single adenosine bases at specific sites in any user-defined transcript [3,4]. Due to the usage of guideRNAs, the target selection and specificity is easily and rationally programmed based on simple Watson–Crick base pairing rules. We and others have shown the functioning of site-directed editing in the PCR tube, in human cell lines and even in simple organisms for the repair of reporter genes, but also disease-relevant genes like CFTR [4–6]. However, with respect to the ease of the procedure and in particular with respect to application in medicine, the requirement for the expression of an engineered deaminase is a limiting factor. Hence, we have recently developed a guideRNA architecture derived from the natural R/G-site of the GluR2 transcript that enables the recruitment of human ADAR2 for site-directed RNA editing (Figure 1B) [7]. Applying such guideRNAs allows for the editing of endogenous housekeeping genes by expression of a guideRNA only. Furthermore, we demonstrated the recoding of a nonsense mutation in PINK1 to a level sufficient to rescue the mitophagy phenotype that links dysfunctional PINK1 to the etiology of Parkinson's disease.

Human cells express three different forms of ADAR (adenosine deaminase acting on RNA), called ADAR1, 2, and 3, which are potentially re-addressable for site-directed RNA editing (Figure 1A) [2,8].

Besides ADAR2, the recruitment of ADAR1 is particularly interesting. On one hand ADAR1 is highly active for the deamination reaction and has shown to exhibit a slightly different substrate preference [9]. On the other hand, ADAR1 is more ubiquitously and to a higher level expressed as compared to ADAR2 [10], which is mainly expressed in neurons and which is believed to be exclusively localized to the nucleus [11]. ADAR1 is expressed in two isoforms, an interferon-inducible long form, called p150, and a constitutive, short form, called p110 (Figure 1A). The long form (p150) comprises an additional N-terminal stretch containing the Z-DNA/Z-RNA binding domain Zα and a nuclear export signal (NES) [12]. Due to the latter, ADAR1p150 is mainly found in the cytosol. ADAR1 and ADAR2 have distinct but also overlapping substrates. Both are essential as we know from knockout studies in mice [13–15]. Alteration of RNA editing is linked to various neurological diseases including behavioral disorders, epilepsy, and the Prader-Willi syndrome [16–19]. Mutations in ADAR1 are linked to the Aicardi-Goutieres syndrome [20], an autoimmune disorder, and others [21]. Both, hyper- [22] and hypoediting [23] have been associated with cancer [24–26].

Here, we demonstrate that a trans-acting guideRNA that has recently been developed for the recruitment of ADAR2 is also capable of recruiting both ADAR1 isoforms. Furthermore, we studied the auto-editing of the guideRNA itself, and present novel guideRNA sequences that are less prone to auto-editing but still allow for the recruitment of ADAR2.

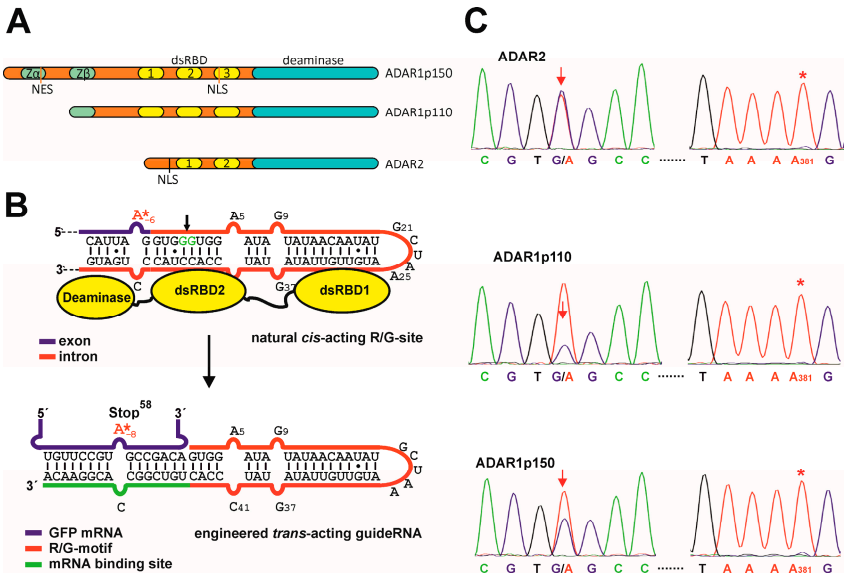


Figure 1. Site-directed RNA editing with R/G-guideRNAs. (**A**) Scheme of the three human ADARs used in this study; (**B**) Design principle: the R/G-guideRNAs have been developed as trans-acting guideRNA from the natural cistronic R/G-motif of the GluR2 transcript. The binding sites of the dsRBDs (dsRNA-binding domains) of ADAR2 are indicated; (**C**) Sanger sequencing of the editing experiments when applying the R/G-guideRNA with ADAR1p110 or p150 compared to ADAR2 in the repair the W58x codon in eGFP. Shown is the sequence around the editing site (arrow) and around a typical off-target site, A381 (*). For full Sanger sequences see Figures S2 and S3.

2. Materials and Methods

2.1. ADAR1-Expressing Cells

The 293 Flp-In T-REx system (Life Technologies) was used for stable integration of a single copy of the cDNA of ADAR1p110 or ADAR1p150 at a genomic FRT-site in engineered 293 cells. Briefly, 4×10^6 cells were seeded on a 10 cm dish. After one day, 1 µg of the respective ADAR1-form in a pcDNA5 vector under control of the tet-on CMV promoter, and 9 µg of pOG44 expressing the Flp recombinase were transfected with lipofectamine 2000 (30 µL). One day later, the medium was changed

for at least two weeks to selection medium (DMEM, 10% FBS, 100 µg/mL hygromycin B, 15 µg/mL blasticidin S). Cells were kept in selection medium prior to the editing experiment, which was then performed in the absence of antibiotics.

2.2. Editing of W58amber eGFP

Protocol for experiments shown in Figures 1 and 3B: The R/G-guideRNAs were subcloned into the pSilencer2.1-U6hygro vector under control of U6 promoter and terminator, the W58x eGFP gene was delivered on a pcDNA3.1 vector as described before [7]. ADAR-expressing 293T cells were cultivated with DMEM, 10% FBS, 1% H/B, 37 °C, 5% CO_2. Cells (3×10^5/well) were seeded into poly-D-lysin-coated 24-well plates and induced with doxycycline (10 ng/mL). Transfection was carried out 24 h later with lipofectamine 2000, applying 300 ng of W58x GFP and 1300 ng of guideRNA per well. Editing was evaluated 72 h after transfection by isolation and sequencing analysis of the target mRNA. For the latter, total RNA from the cells (NucleoSpin RNA Plus kit, Macherey Nagel, Düren, Germany) was DNaseI-digested, followed by reverse transcription, Taq-PCR amplification, and Sanger sequencing.

2.3. Defining Hotspots of Auto-Editing

Protocol for the experiment shown in Figure 2: The respective eCFP W66x mRNA containing the guideRNA in *cis* was in-vitro-transcribed with T7 RNA pol as described earlier [7]. ADAR2 was produced and purified as described earlier [7]. Editing was carried out as described earlier [7] with (mRNA) = 25 nM, (ADAR2) = 350 nM, (Mg^{2+}) = 3 mM, no heparin, no spermidine, and incubation in editing buffer (12.5 mM Tris, 12.5 mM Tris-HCl, 75 mM KCl, 2 mM DTT) for 180 min while cycling between 30 °C and 37 °C.

2.4. Testing Auto-Editing Inside the R/G-Motif for Variants 1–4

Protocol for experiments shown in Figure 3A: The R/G-guideRNAs were subcloned into the 3′-UTR of wild-type eGFP in a pcDNA3.1 vector. GFP expression served as a transfection control. Editing was carried out in 293T cells under transient expression of one of the respective three ADAR forms from plasmid vectors (CMV promotor). For this, 293T cells (2×10^5/well) were seeded into 24-well plates. Transfection was carried out 24 h later with lipofectamine 2000, applying 300 ng of wt eGFP-guideRNA and 600 ng of the respective ADAR per well. RNA was isolated for sequencing 48 h after co-transfection as described above.

3. Results

3.1. Site-Directed RNA Editing with ADAR1

In this study, we started from a recently developed guideRNA architecture [7] that is based on the cistronic R/G-motif of the GluR2 transcript (Figure 1B). Compared to the natural R/G-site, the targeted adenosine in the mRNA is put to position −8, thus 2 nt further upstream of the R/G-motif (Figure 1B). We had shown the functioning of the guideRNA by co-transfection of the plasmid-borne guideRNA together with ADAR2 on a plasmid or by transfection of the guideRNA-plasmid into engineered ADAR2-expressing 293 Flp-In cells [7]. To test if the original guideRNA design is able to recruit ADAR1 isoforms, we have now created 293 Flp-In cells that express ADAR1p110 or ADAR1p150 from a single genomic copy under control of a CMV-tet on promoter. In such cell lines, the maximal induction-level of all three ADAR forms was similar and comparable to the level of β-actin mRNA (Figure S1). After induction (24 h) of the respective ADAR with doxycycline, the respective guideRNA construct was co-transfected together with a GFP reporter construct containing a single W58amber missense codon for site-directed repair. 72 h after co-transfection, the editing yield was determined by Sanger sequencing of the RNA (Figure 1), as described earlier [7]. In ADAR2-expressing cells, editing yields around 50% are typically obtained, with no detectable off-target editing in the ORF of the reporter

gene [7]. This was again confirmed here (Figure 1C). The editing reaction in ADAR1-expressing cells (Figure 1C) gave slightly reduced yields for the p150 form (35%–40%), and markedly reduced yields (20%) for the p110 form. Again, no off-target editing was detectable for either ADAR1 form in the ORF of the reporter gene. One site, adenosine 381, which was prone to off-target editing under transient overexpression of ADAR2 before [7], is shown in Figure 1C to exemplify the lack of off-target editing in the ORF of eGFP under genomic expression of all three ADARs (see Figures S2 and S3 for full Sanger sequencing traces).

3.2. Auto-Editing inside the GuideRNA

The original guideRNA design is relatively AU-rich. Editing could suffer from the auto-editing of the guideRNA itself thus destabilizing the hairpin. We tested this in an initial editing experiment on an in-vitro transcribed RNA substrate that contained the eCFP W66amber sequence in frame with a cis-acting guideRNA inside the PCR tube (Figure 2). Under these conditions, we found five sites in the R/G-motif prone to auto-editing: A8, A14, A34, A36, and A39. All five editing events can be assumed to destabilize the helix. Notably, the editing yields at the sites in the 3′-half of the R/G hairpin were higher compared to those in the 5′-half of the hairpin. We further tested the auto-editing by co-transfection of the same cis-acting CFP-guideRNA construct with ADAR2 in 293T cells (Figure S4). All auto-editing sites were confirmed, the editing yields stayed mostly unchanged, only at A-16 and A34 was the editing yield increased, and, at one new site, A11, auto-editing occurred additionally.

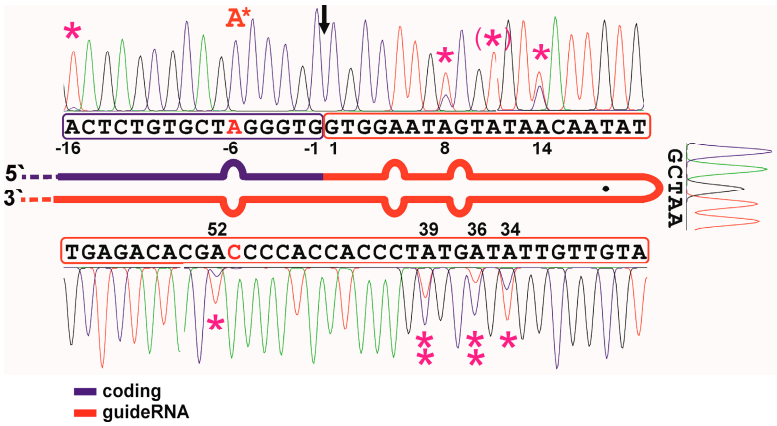


Figure 2. Identification of auto-editing hotspots. An in-vitro transcribed RNA substrate containing a part of the eCFP ORF around the W66x site (blue) in *cis* with an R/G-guideRNA (red) was edited with purified ADAR2 enzyme in a PCR tube. Hotspots for auto-editing have been marked by magenta asterisks; double asterisks mark strongly edited sites; asterisk in brackets mark an editing site only found in a similar experiment inside the cell (see Figure S4). The red A* marks the targeted editing site, full conversion was achieved. The black arrow shows the site where the cistronic motif is cut when the guideRNA is applied in *trans*.

3.3. G/C-Substitution inside the R/G-Motif Reduces Auto-Editing

To reduce auto-editing of the guideRNA, 3, 6, or 13 A/U base pairs in the original R/G-motif have been substituted by G/C base pairs, creating variants 2, 3, and 4 (Figure 3A). To assess auto-editing inside these motif variants, we constructed for every guideRNA an eGFP reporter transcript that contains the complete guideRNA in the 3′-UTR of the transcript, thus ca. 550 nt downstream the targeted editing site. This allows for testing the guideRNAs in a situation that resembles the trans-acting guideRNA under concurrent possibility to access auto-editing by Sanger sequencing. Editing was tested by co-transfection of these constructs (300 ng) with each of the three forms of ADAR (600 ng) into 293T cells. Editing in the R/G-motif was analyzed 48 h after co-transfection by RNA isolation

and Sanger sequencing. The editing yields inside the R/G-motif are given in color-coded circles in Figure 3A.

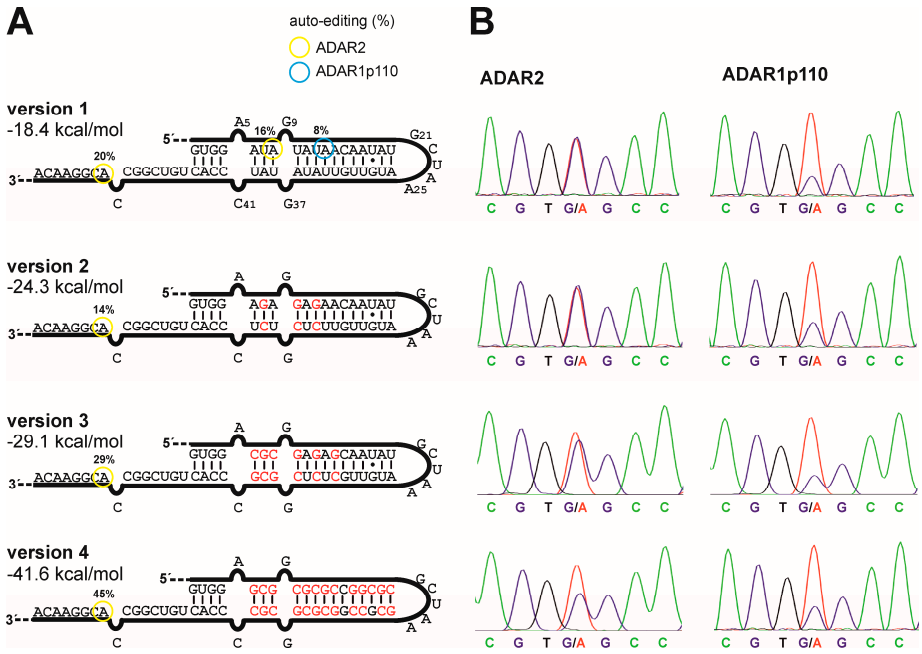


Figure 3. GuideRNA variants that avoid auto-editing. (**A**) Three guideRNA versions with differing degree of G/C substitution were tested for auto-editing. The shown sequence was introduced into the 3′-UTR of the eGFP transcript for easier Sanger sequencing. The folding energies have been estimated using mfold. For ADAR1p150, no auto-editing site was detectable; (**B**) Performance of the three new guideRNA versions for the editing of W58x with ADAR2 or ADAR1p110, respectively, when applied in *trans* in the respective ADAR-expressing cell line. For fluorescence imaging, see Figure S5.

Under these assay conditions, the starting guideRNA architecture, version 1, gives only subtle auto-editing, and only at positions A8 by ADAR2 and A13 by ADAR1p110. With ADAR2, we found an additional off-target editing at A52, which is part of the mRNA binding sequence. For ADAR1p150, no auto-editing was detectable. It is reasonable to infer that the 36 bp RNA substrate , given in Figure 2, is much more heavily edited in the R/G-motif compared to the lone-standing, 20 bp R/G-motif, given in Figure 3, as the ADAR enzymes are known to accept dsRNA with $\geq$30 bp much better than those $\leq$30 bp [2,11]. When introducing G/C-substitutions into the R/G-motif, no auto-editing was detectable with any of the three enzymes. This was already achieved with only 3 G/C substitutions, which have been chosen in a way to target the main auto-editing sites (A8, A14, A34, A36, and A39) simultaneously with a minimal number of substitutions.

The G/C substitutions do not only remove editable adenosines and put adenosines into contexts non-preferred for editing (e.g., 5′-GA), but they rather increase the free energy gain of folding from -18 kcal/mol up to -42 kcal/mol as estimated by the software mfold [27] (Figure 3A). We tested how the new guideRNA variants 2 to 4 behave compared to the original version 1 for the editing of the reporter W58x eGFP in ADAR2-expressing cells. For versions 2 and 3, virtually no change in the editing yield was observed. Only the highly substituted version 4 showed a reduction in editing yield (Figure 3B, Figure S5). For ADAR1p110, we found a virtually unchanged editing yield with all versions but at comparably low editing efficiency (Figure 3B).

4. Discussion

This is the first report about site-directed RNA editing with human ADAR1 isoforms, the constitutive (p110) and the inducible one (p150). As the R/G-guideRNA was developed from a classical

ADAR2 substrate for the recruitment of ADAR2, it could be expected that the recruitment of ADAR1 isoforms is less efficient. However, we achieved site-directed RNA editing with the R/G-guideRNAs with both ADAR1 isoforms, in particular with the p150 form; thus, the general strategy is feasible with ADAR1, and our R/G-guideRNA provides a starting point for the development of more appropriate guideRNAs for ADAR1.

Furthermore, we show the first results on auto-editing of the guideRNA itself. We have defined the hotspots for auto-editing and show a simple strategy to circumvent it by minimal substitution in the hairpin secondary structure. The respective optimized guideRNAs are much less auto-edited; however, our results show that a high substitution degree may interfere with editing efficiency at the target. Even though the confinement of auto-editing was not improving the overall editing yield under the given continuous expression of the guideRNA, it could improve editing when drug-like, chemically stabilized R/G-guideRNAs are explored in the future. Notably, when defining the hotspots of auto-editing by applying an artificial eCFP mRNA containing the guideRNA in *cis*, we found full conversion at the targeted adenosine even in cell culture (Figure S4). In contrast, when applying the same guideRNA in *trans*, editing levels above 60% are difficult to obtain. It is likely that there is still space for further optimization of the guideRNA′s design to obtain better editing results with all three enzymes. In summary, both findings are important for the advancement of site-directed RNA editing as a tool in basic biology or as a platform for therapeutic editing [7,28].

Supplementary Materials: The following are available online at www.mdpi.com/2073-4425/8/1/34/s1, Figure S1: qPCR analysis of ADAR expression in engineered cells, Figure S2: Sanger sequencing trace of editing in ADAR1p110-expressing cells, Figure S3: Sanger sequencing trace of editing in ADAR1p150-expressing cells, Figure S4: Defining auto-editing hotspots in cellular editing; Figure S5: Fluorescence imaging data belonging to experiment shown in Figure 3B.

Acknowledgments: We gratefully acknowledge support from the University of Tübingen and the Deutsche Forschungsgemeinschaft (STA 1053/3-2). This work has received funding from the European Research Council (ERC) under the European Union′s Horizon 2020 research and innovation program (grant agreement No 647328). Open access fees have been paid from University and DFG funds.

Author Contributions: P.R., J.W., and T.S. conceived and designed the experiments; P.R., M.H., P.M., and J.W. performed cloning; M.H. and J.W. performed editing experiments; P.M. and M.H. performed RT-qPCR; all authors analyzed the data; P.M. engineered the ADAR1-expressing cells; T.S. wrote the paper.

Conflicts of Interest: The authors have filed a patent on the technology described in this paper.

References

1. Vogel, P.; Stafforst, T. Site-directed RNA editing with antagomir deaminases—A tool to study protein and RNA function. *ChemMedChem* **2014**, *9*, 2021–2025. [CrossRef] [PubMed]
2. Nishikura, K. Functions and regulation of RNA editing by ADAR deaminases. *Annu. Rev. Biochem.* **2010**, *79*, 321–349. [CrossRef] [PubMed]
3. Stafforst, T.; Schneider, M.F. An RNA–Deaminase conjugate selectively repairs point mutations. *Angew. Chem. Int. Ed.* **2012**, *51*, 11166–11169. [CrossRef] [PubMed]
4. Montiel-Gonzalez, M.F.; Guillermo, I.; Yudowski, A.; Rosenthal, J.J.C. Correction of mutations within the cystic fibrosis transmembrane conductance regulator by site-directed RNA editing. *Proc. Natl. Acad. Sci. USA* **2013**, *110*, 18285–18290. [CrossRef] [PubMed]
5. Vogel, P.; Schneider, M.F.; Wettengel, J.; Stafforst, T. Improving site-directed RNA editing in vitro and in cell culture by chemical modification of the guideRNA. *Angew. Chem. Int. Ed.* **2014**, *53*, 6267–6271. [CrossRef] [PubMed]
6. Hanswillemenke, A.; Kuzdere, T.; Vogel, P.; Jékely, G.; Stafforst, T. Site-directed RNA editing in vivo can be triggered by the light-driven assembly of an artificial riboprotein. *J. Am. Chem. Soc.* **2015**, *137*, 15875–15881. [CrossRef] [PubMed]
7. Wettengel, J.; Reautschnig, J.; Geisler, S.; Kahle, P. J.; Stafforst, T. Harnessing human ADAR2 for RNA repair—Recoding a PINK1 mutation rescues mitophagy. *Nucl. Acids Res.* **2016**. [CrossRef] [PubMed]
8. Bass, B.L. RNA editing by adenosine deaminases that act on RNA. *Annu. Rev. Biochem.* **2002**, *71*, 817–846. [CrossRef] [PubMed]

9. Schneider, M.F.; Wettengel, J.; Hoffmann, P.C.; Stafforst, T. Optimal guideRNAs for re-directing deaminase activity of hADAR1 and hADAR2 in trans. *Nucl. Acids Res.* **2004**. [CrossRef] [PubMed]
10. Picardi, E.; Manzari, C.; Mastropasqua, F.; Aiello, I.; D'Erchia, A.M.; Pesole, G. Profiling RNA editing in human tissues: towards the inosinome Atlas. *Sci. Rep.* **2015**. [CrossRef] [PubMed]
11. Nishikura, K. A-to-I editing of coding and non-coding RNAs by ADARs. *Nat. Rev. Mol. Cell Biol.* **2016**, *17*, 83–96. [CrossRef] [PubMed]
12. Barraud, P.; Allain, F.H.-T. ADAR Proteins: Double-stranded RNA and Z-DNA binding domains. *Curr. Topics Microbiol. Immun.* **2012**, *353*, 35–60.
13. Higuchi, M.; Maas, S.; Single, F.N.; Hartner, J.; Rozov, A.; Burnashev, N.; Feldmeyer, D.; Sprengel, R.; Seeburg, P.H. Point mutation in an AMPA receptor gene rescues lethality in mice deficient in the RNA-editing enzyme ADAR2. *Nature* **2000**, *406*, 78–81. [PubMed]
14. Hartner, J.C.; Schmittwolf, C.; Kispert, A.; Müller, A.M.; Higuchi, M.; Seeburg, P.H. Liver disintegration in the mouse embryo caused by deficiency in the RNA-editing enzyme ADAR1. *J. Biol. Chem.* **2004**, *279*, 4894–4902. [CrossRef] [PubMed]
15. Wang, Q.; Miyakoda, M.; Yang, W.; Khillan, J.; Stachura, D.L.; Weiss, M.J.; Nishikura, K. Stress-induced apoptosis associated with null mutation of ADAR1 RNA editing deaminase gene. *J. Biol. Chem.* **2004**, *279*, 4952–4961. [CrossRef] [PubMed]
16. Morabito, M.V.; Abbas, A.I.; Hood, J.L.; Kesterson, R.A.; Jacobs, M.M.; Kump, D.S.; Hachey, D.L.; Roth, B.L.; Emeson, R.B. Mice with altered serotonin 2C receptor RNA editing display characteristics of Prader-Willi syndrome. *Neurobiol. Dis.* **2010**, *39*, 169–180. [CrossRef] [PubMed]
17. Maas, S.; Kawahara, Y.; Tamburro, K.M.; Nishikura, K. A-to-I RNA editing and human disease. *RNA Biol.* **2006**, *3*, 1–9. [CrossRef] [PubMed]
18. Slotkin, W.; Nishikura, K. Adenosine-to-inosine RNA editing and human disease. *Genome Med.* **2013**. [CrossRef] [PubMed]
19. Silberberg, G.; Lundin, D.; Navon, R.; Öhman, M. Deregulation of the A-to-I RNA editing mechanism in psychiatric disorders. *Hum. Mol. Genet.* **2012**, *21*, 311–321. [CrossRef] [PubMed]
20. Rice, G.I.; Kasher, P.R.; Forte, G.M.; Mannion, N.M.; Greenwood, S.M.; Szynkiewicz, M.; Dickerson, J.E.; Bhaskar, S.S.; Zampini, M.; Briggs, T.A.; et al. Mutations in ADAR1 cause Aicardi-Goutières syndrome associated with a type I interferon signature. *Nat. Genet.* **2012**, *44*, 1243–1248. [CrossRef] [PubMed]
21. Zhang, X.J.; He, P.P.; Li, M.; He, C.D.; Yan, K.L.; Cui, Y.; Yang, S.; Zhang, K.Y.; Gao, M.; Chen, J.J.; et al. Seven novel mutations of the ADAR gene in Chinese families and sporadic patients with dyschromatosis symmetrica hereditaria (DSH). *Hum. Mutat.* **2004**, *23*, 629–630. [CrossRef] [PubMed]
22. Chen, L.; Li, Y.; Lin, C.H.; Chan, T.H.M.; Chow, R.K.K.; Song, Y.; Liu, M.; Yuan, Y.F.; Fu, L.; Kong, K.L.; et al. Recoding RNA editing of antizyme inhibitor 1 predisposes to hepatocellular carcinoma. *Nat. Med.* **2013**, *19*, 209–216. [CrossRef] [PubMed]
23. Shimokawa, T.; Rahman, M.F.; Tostar, U.; Sonkoly, E.; Ståhle, M.; Pivarcsi, A.; Palaniswamy, R.; Zaphiropoulos, P.G. RNA editing of the GLI1 transcription factor modulates the output of Hedgehog signaling. *RNA Biol.* **2013**, *10*, 321–333. [CrossRef] [PubMed]
24. Gallo, A. RNA editing enters the limelight in cancer. *Nat. Med.* **2013**, *19*, 130–131. [CrossRef] [PubMed]
25. Paz-Yaacov, N.; Bazak, L.; Buchumenski, I.; Porath, H.T.; Danan-Gotthold, M.; Knisbacher, B.A.; Eisenberg, E.; Levanon, E.Y. Elevated RNA editing activity is a major contributor to transcriptomic diversity in tumors. *Cell Rep.* **2015**, *13*, 267–276. [CrossRef] [PubMed]
26. Han, L.; Diao, L.; Yu, S.; Xu, X.; Li, J.; Zhang, R.; Yang, Y.; Werner, H.M.; Eterovic, A.K.; Yuan, Y.; et al. The genomic landscape and clinical relevance of A-to-I RNA editing in human cancers. *Cancer Cell* **2015**, *28*, 515–528. [CrossRef] [PubMed]
27. Zuker, M. Mfold web server for nucleic acid folding and hybridization prediction. *Nucl. Acids Res.* **2003**, *31*, 3406–3415.
28. Reautschnig, P.; Vogel, P.; Stafforst, T. The notorious RNA in the spotlight—Drug or target for the treatment of disease. *RNA Biol.* **2016**. [CrossRef]

Supplementary Materials: Applying Human ADAR1p110 and ADAR1p150 for Site-Directed RNA Editing—G/C Substitution Stabilizes GuideRNAs Against Editing

Madeleine Heep, Pia Mach, Philipp Reautschnig, Jacqueline Wettengel and Thorsten Stafforst

Gene	*Sequence (5' to 3')*	*Product size*
ADAR1	*fw.: GCATTTGAGGATGGACTACG*	*101 bp*
	rev.: TCCTTAGTCTTCCCGGATTG	
ADAR2	*fw.: CGGAGATCCTTGCTCAGATT*	*99 bp*
	rev.: CCCTCGCTCTGATTTCTGAA	
ß-actin	*fw.: CGGGACCTGACTGACTAC*	*91 bp*
	rev.: TAATGTCACGCACGATTTCC	

A. Primers for Qpcr.

Cell line	Mean C_T (β-actin)	Mean C_T (ADAR)	$2^{-\Delta C_T}$
ADAR2 -dox (control)	18.596	24.574	0.02
ADAR1 (p150) - dox	18.836	22.634	0.07
ADAR1 (p110) - dox	19.153	22.057	0.13
ADAR2 + dox (control)	18.597	19.026	0.74
ADAR1 (p150) + dox	18.821	19.587	0.59
ADAR1 (p110) + dox	18.584	19.290	0.61

B. Measured ct-values of all experiments. Values are averaged from three technical replicates. Calculation of relative expression levels from the Δct values for ADAR2 versus the housekeeping gene ß-actin. (relative expression $= 2^{-\Delta ct}$, with $\Delta ct = ct(ADAR1/2) - ct(ß - actin)$).

Figure S1. qPCR analysis of ADAR expression in engineered cells. The relative amount of ADAR mRNA in 293T cells with a genomically integrated copy of ADAR controlled by a CMV tet-on promoter (integr.) was determined by quantitative real-time PCR (qPCR) after 72h (doxycycline induced expression of integr. ADAR). For this, RNA was extracted from cell lysates (RNeasy MinElute Kit, Qiagen, Hilden, Germany). After DNaseI digestion (NEB) and reverse transcription (high capacity cDNA reverse transcription kit, Applied Biosystems, Foster City, CA, USA), 20 ng cDNA was mixed with Fast SYBR Green Master Mix (Applied Biosystems) and analyzed by the 7500 Fast Real-Time PCR System (Applied Biosystems). (**A**) For determining gene expression, primers were designed for targeting ADAR1, ADAR2 and the housekeeping gene ß-actin; (**B**) qPCR of ADAR1, ADAR2 and the housekeeping gene was performed in triplicates. The table displays the mean values of the cycles where the fluorescence crosses the threshold of 0.35 for ADAR and 0.15 for ß-actin (ct values).

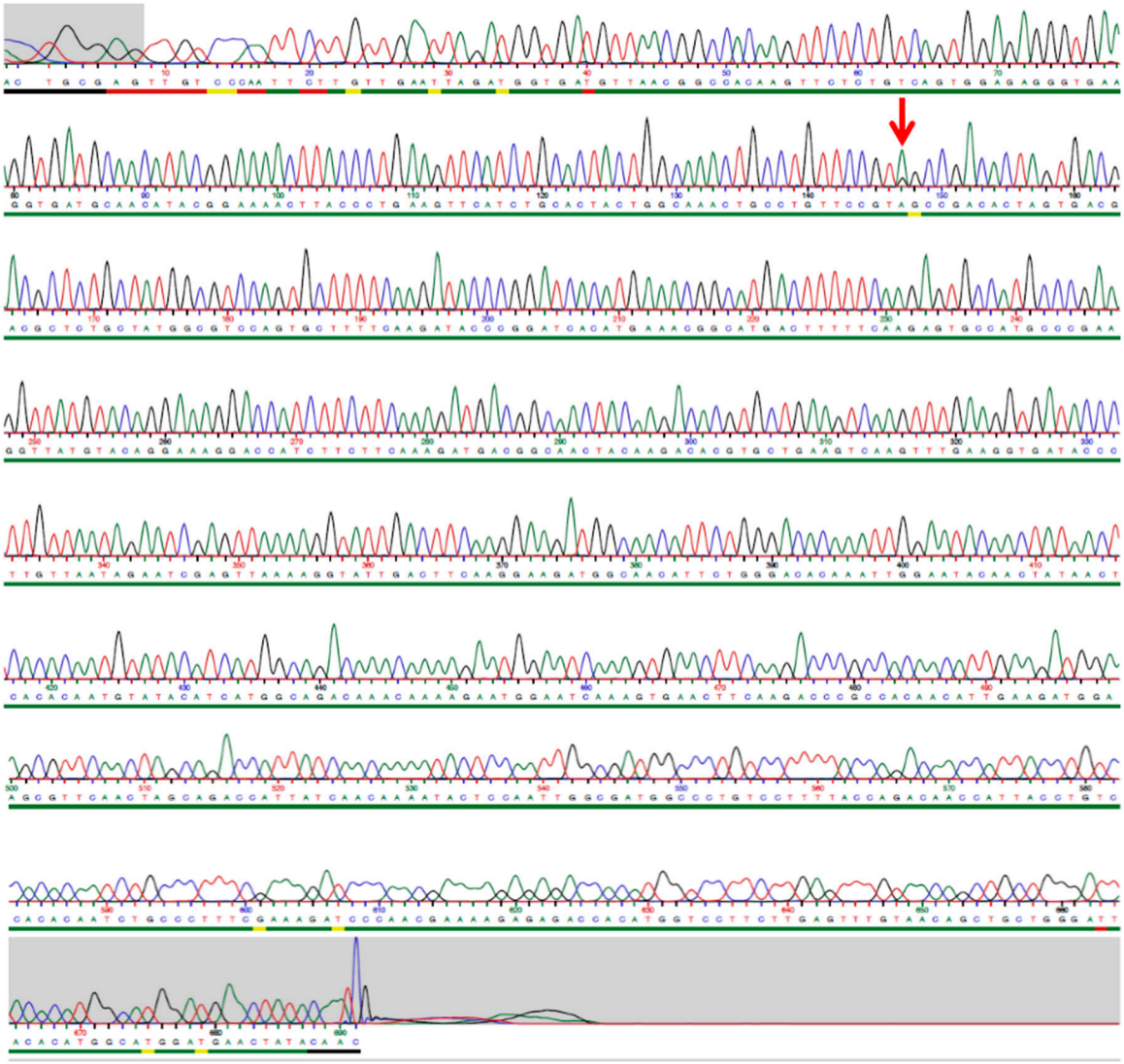


Figure S2. Sanger sequencing trace of editing in ADAR1p110-expressing cells. The target site is marked by a red arrow.

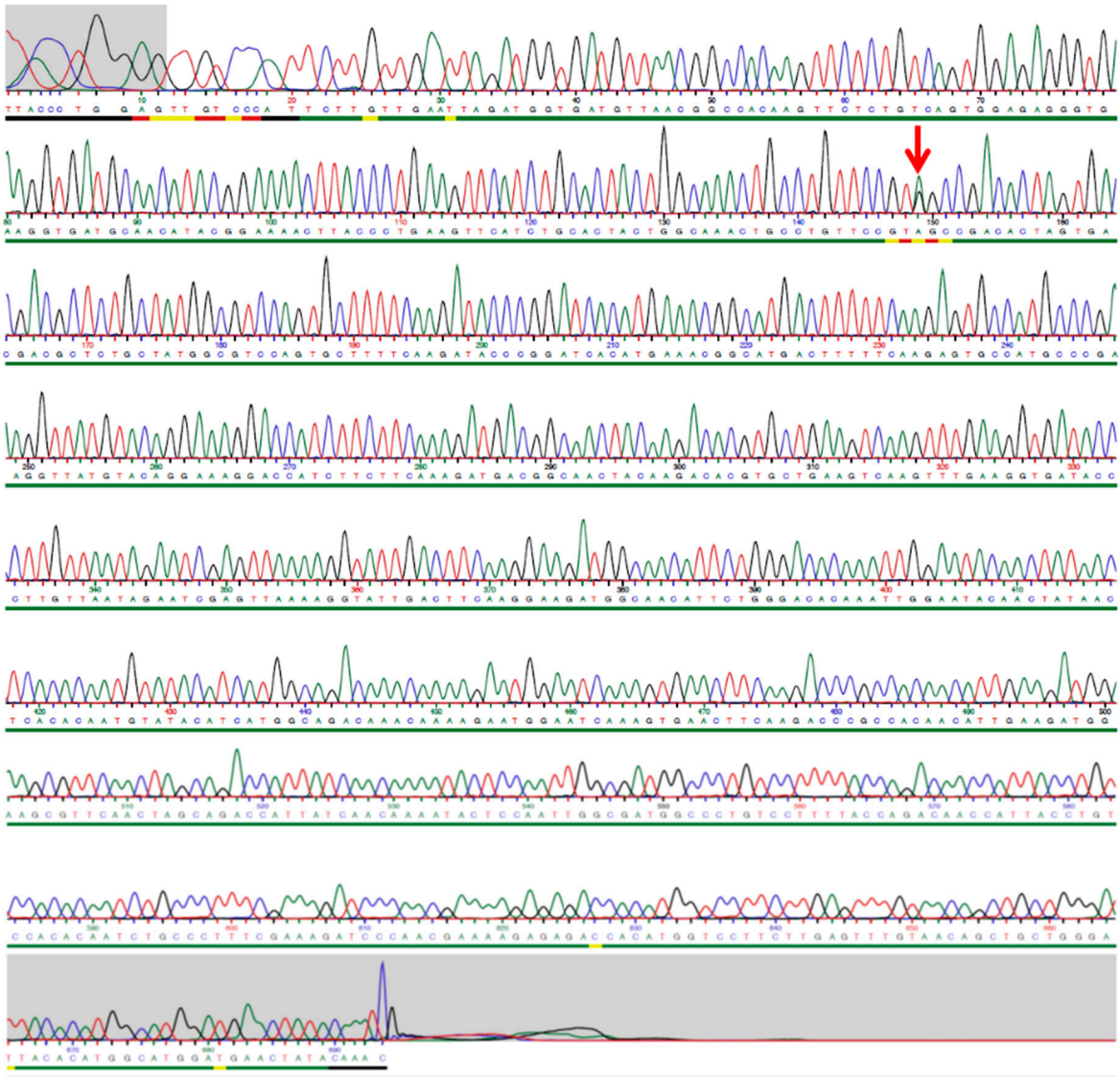


Figure S3. Sanger sequencing trace of editing in ADAR1p150-expressing cells. The target site is marked by a red arrow.

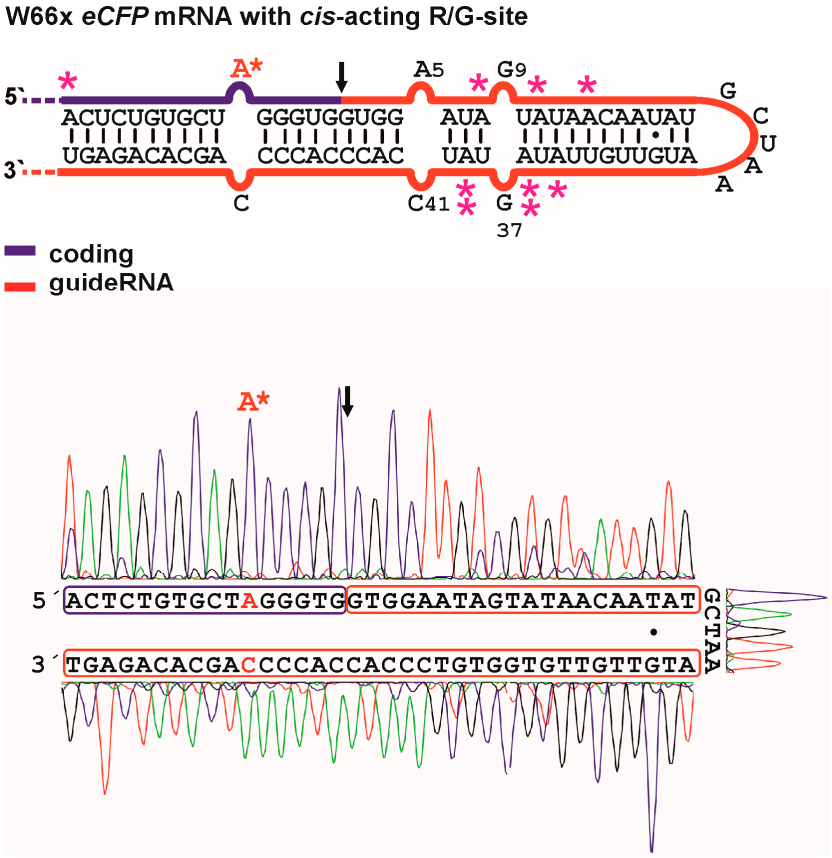


Figure S4. Defining auto-editing hotspots in cellular editing. The same construct as described in Figure 2 main text was used as an auto-editing probe inside the 293 cell. Cells were co-transfected in a 24-well format with a pcDNA3.1 vectors containing the transcript for editing (300 ng) and a pcDNA3.1 vector containing ADAR2 (300 ng). After 24 h, RNA was isolated and sequenced.

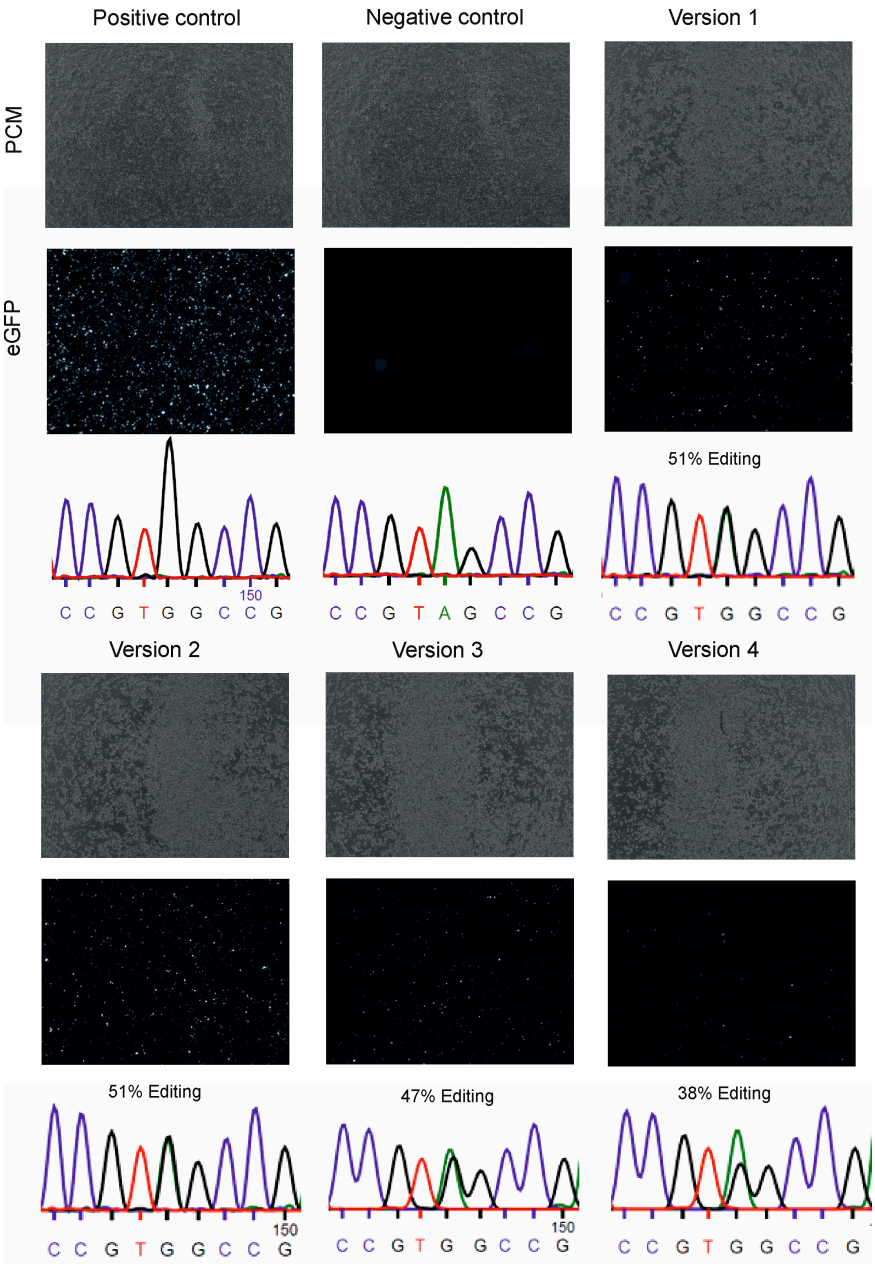


Figure S5. Fluorescence imaging data belonging to the experiment shown in Figure 3B, main text. Fluorescence images (50x magnification, 70 ms exposure for GFP) were taken 72 h after co-transfection with the respective gRNA (version 1 to 4, 1300 ng) and GFP W58X amber (300 ng). Transfection with 300 ng wild-type GFP and 300 ng GFP W58X amber served as positive and negative controls respectively. Sanger sequencing was performed after RNA isolation (NucleoSpin RNA Plus Kit, Macherey-Nagel, Dueren, Germany), reverse transcription and Taq-PCR. (eGFP = fluorescence channel, PCM = phase contrast microscopy).

6.6.4 **PUBLICATION 4:** MERKLE T., MERZ S., REAUTSCHNIG P., BLAHA A., LI Q., VOGEL P., WETTENGEL J., LI J.B., STAFFORST T. PRECISE RNA EDITING BY RECRUITING ENDOGENOUS ADARS WITH ANTISENSE OLIGONUCLEOTIDES. NAT BIOTECHNOL, 2019. 37(2): P. 133-138.

nature biotechnology

LETTERS

https://doi.org/10.1038/s41587-019-0013-6

Precise RNA editing by recruiting endogenous ADARs with antisense oligonucleotides

Tobias Merkle[1], Sarah Merz[1], Philipp Reautschnig[1], Andreas Blaha[1], Qin Li[2], Paul Vogel[1], Jacqueline Wettengel[1], Jin Billy Li[2] and Thorsten Stafforst[1]*

Site-directed RNA editing might provide a safer or more effective alternative to genome editing in certain clinical scenarios. Until now, RNA editing has relied on overexpression of exogenous RNA editing enzymes or of endogenous human ADAR (adenosine deaminase acting on RNA) enzymes. Here we describe the engineering of chemically optimized antisense oligonucleotides that recruit endogenous human ADARs to edit endogenous transcripts in a simple and programmable way, an approach we call RESTORE (recruiting endogenous ADAR to specific transcripts for oligonucleotide-mediated RNA editing). We observed almost no off-target editing, and natural editing homeostasis was not perturbed. We successfully applied RESTORE to a panel of standard human cell lines and human primary cells and demonstrated repair of the clinically relevant PiZZ mutation, which causes α1-antitrypsin deficiency, and editing of phosphotyrosine 701 in STAT1, the activity switch of the signaling factor. RESTORE requires only the administration of an oligonucleotide, circumvents ectopic expression of proteins, and represents an attractive approach for drug development.

Adenosine-to-inosine editing in RNA diversifies the transcriptome by recoding of amino acid codons, Start codons and Stop codons, and by alteration of splicing, among other mechanisms[1]. Steering such enzymes to specific sites at selected transcripts, a strategy called site-directed RNA editing[2,3], holds great promise for the treatment of disease and as a tool to study protein and RNA function. Unlike DNA editing, RNA editing manipulates genetic information in a reversible and tunable manner. These properties may enable manipulations that are either lethal or quickly compensated when done at the genome level[4]. Furthermore, RNA editing could be safer because potential adverse effects should be reversible and dose-dependent.

We and others have recently published several RNA editing strategies based on expression of exogenous engineered deaminases[5–7]. However, in a therapeutic setting, harnessing of the widely expressed endogenous ADARs, including ADAR1 and ADAR2, would be preferable[8] as it would replace ectopic expression of an engineered protein with administration of an oligonucleotide drug. Recently, we engineered a plasmid-borne guide RNA (gRNA) that recruits human ADAR2 to elicit programmable, site-specific RNA editing[9]. Such gRNAs comprise two parts: an invariant ADAR-recruiting domain and a programmable specificity domain (Fig. 1a). The ADAR-recruiting domain forms an imperfect 20-bp hairpin (Fig. 1a) and was adapted from a well-known ADAR2 target site in the GluR2 mRNA, and thus was called the R/G motif. The specificity domain is a programmable, short (~18 nt), single-stranded sequence reverse complementary to the target mRNA (Fig. 1a). We optimized the gRNA for ADAR2 recruitment and demonstrated its expression from a U6 promotor[9]. However, sufficient editing of endogenous transcripts such as *GAPDH* (glyceraldehyde-3-phosphate dehydrogenase) or *ACTB* (β-actin) always required co-overexpression of ADAR2, whereas expression of the gRNA alone failed to achieve editing[9]. In the present study, we have engineered antisense oligonucleotides (ASOs) that recruit endogenous ADAR to edit endogenous transcripts in cancer cell lines and in primary human cells.

We applied a plasmid-borne approach[9] to screen for better ADAR-recruiting domains (Supplementary Fig. 1). While testing 15 different designs, we found sequence variant 9.4 (with an additional 5 bp at the 5′ site of the R/G motif). Although less effective with ADAR2, variant 9.4 almost doubled editing with the ADAR1 isoform p110. Using ADAR1 for RNA editing could be beneficial as its expression is particularly widespread[10].

To further enhance editing efficiency, we tested chemically stabilized ASOs[11] (RESTORE) instead of plasmid-borne[9] gRNA expression. In the first round, we tested three ASO designs (v1, v4, v9.4). The ASOs comprised an ADAR-recruiting domain composed entirely of natural ribonucleotides and a specificity domain that was chemically modified (2′-O-methylations, phosphorothioate, Fig. 1b), containing a modification gap opposite the editing site, much like what was described before[12].

Using ASOs targeting a 5′ UAG site in the 3′ untranslated region (3′ UTR) of either *ACTB* or *GAPDH*, we assessed the ADAR preferences of the ASOs. We lipofected them into engineered Flp-In T-REx 293 cells expressing a specific ADAR isoform (ADAR2, ADAR1 p110 or ADAR1 p150)[9,13]. We found the highest editing efficiency (75%–85%) in ADAR1 p150-expressing cells (Fig. 1c). Editing yields were lower for ADAR1 isoform p110 (12%–50%), but showed a strong (two- to threefold) benefit of ASO v9.4 compared to ASO v1. Editing with ADAR2 was similar to editing with ADAR1 p110, however, ASO v9.4 was inferior to ASO v1. Chemical modification of the ASO was required to obtain high editing yields (Supplementary Figs. 2 and 3). Also, the presence of the ADAR-recruiting domain was essential (Supplementary Fig. 2). Finally, we tested the concurrent editing of both transcripts by cotransfection of two ASOs (Fig. 1c, right). Notably, the editing yields stayed virtually unchanged, demonstrating that site-directed RNA editing can be carried out at several transcripts simultaneously.

In HeLa cells, targeting 5′ UAG codons in the 3′ UTRs of *GAPDH* and *ACTB*, ASO v1 and v4 gave some editing (Fig. 2a). However, the ASO v9.4 gave clearly higher editing of both transcripts (~40%). A control ASO lacking the ADAR-recruiting domain did not elicit

[1]Interfaculty Institute of Biochemistry, University of Tübingen, Tübingen, Germany. [2]Department of Genetics, Stanford University, Stanford, CA, USA.
*e-mail: thorsten.stafforst@uni-tuebingen.de

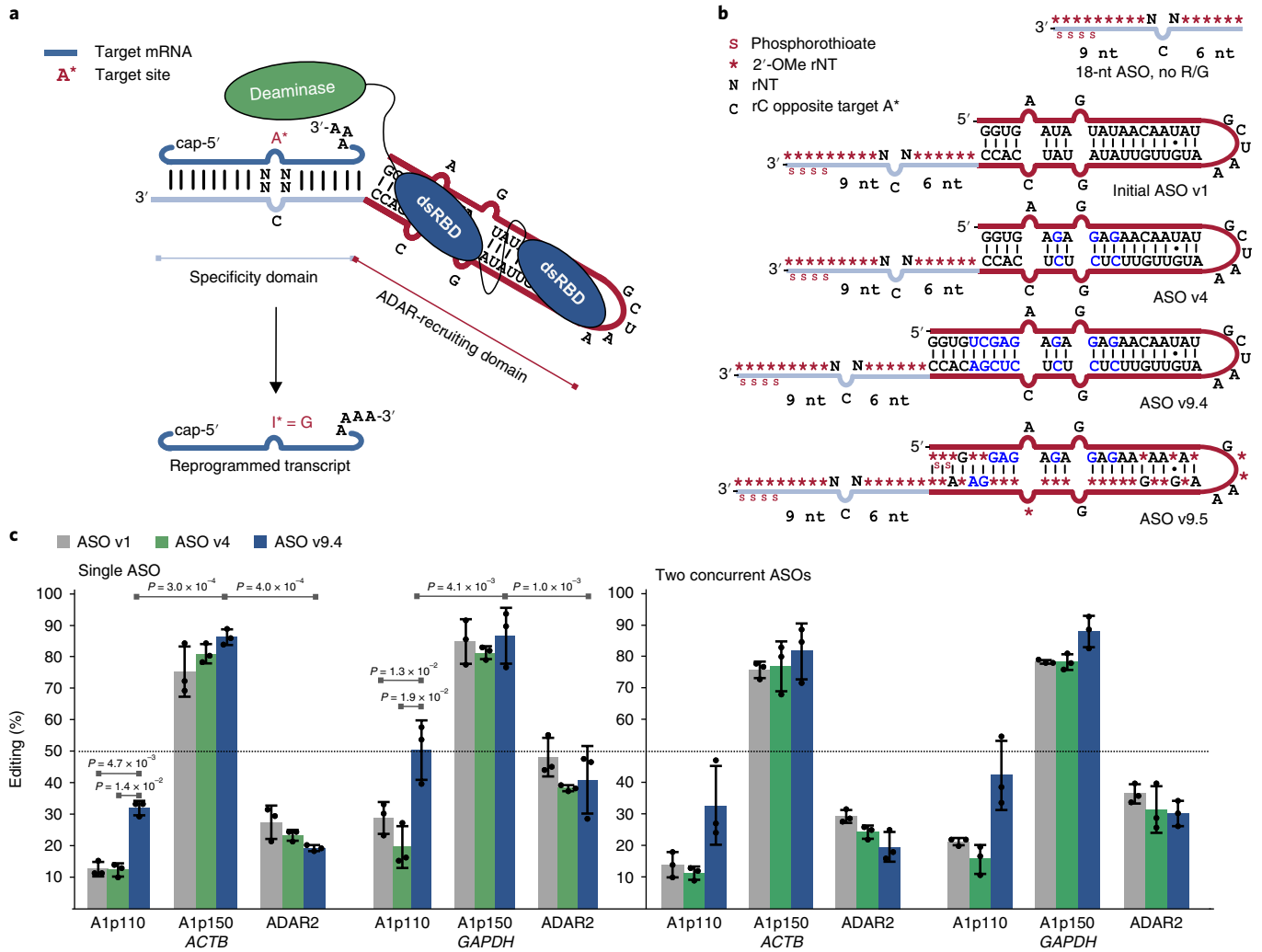


Fig. 1 | Design of ADAR-directing ASOs and characterization in engineered ADAR-expressing cell lines (293 Flp-In T-REx). a, Principle of RESTORE: ASOs comprise a programmable specificity domain that determines target mRNA binding and an invariant ADAR-recruiting domain to steer endogenous ADAR to the ASO:mRNA hybrid. Site-directed RNA editing at the mRNA is controlled by the chemically modified ASO and results in a specific adenosine-to-inosine change (functionally equivalent to an adenosine-to-guanosine change). dsRBD, double-stranded RNA-binding domain; A* or I*, adenosine or inosine base at target site, respectively. **b**, Sequences and chemical modifications of ASOs (see also Supplementary Table 1). rNT, natural ribonucleotide; rC, cytidine. **c**, Comparative editing of two endogenous transcripts (*ACTB*, *GAPDH*) by transfection of the respective chemically modified ASOs into the indicated ADAR-expressing cell line. Either a single ASO (against *GAPDH* or *ACTB*) or two ASOs (against *GAPDH* and *ACTB*) were transfected. Data in **c** are shown as the mean ± s.d., $N = 3$ independent experiments; significance (P) was calculated with a two-tailed paired t-test. A1p110 represents the ADAR1 p110 isoform, A1p150 the ADAR1 p150 isoform; n.d., no editing was detectable. All targets are given in Supplementary Note 1.

editing (Fig. 2a). As we knew ASO v9.4 to prefer ADAR1 p150, we repeated the experiment in the presence of interferon (IFN)-α, which is known to induce ADAR1 p150 expression[14]. Indeed, IFN-α treatment almost doubled the editing yields for all ASO designs (v1, v4, v9.4) and both transcripts (up to 70% with v.9.4). Again, when targeting both transcripts simultaneously by cotransfecting two ASOs, editing yields stayed unchanged (Fig. 2a, right). To assess the impact of chemical modifications, we compared the ASOs with RNAs of the same sequence transcribed in vitro, and found the latter substantially inferior (for example, relative reduction of editing yield by 37%–87% for v9.4), which might explain why the RESTORE approach works better than plasmid-borne gRNAs of the same sequence (Supplementary Figs. 2 and 3).

We thus extended the chemical modification to the ADAR-recruiting domain. Specifically, we stabilized the 5′ terminus (2′-O-methylation and phosphorothioate) and substituted all pyrimidines with their 2′-O-methylated analogs. Even though heavily modified, this ASO design, v9.5, was equal or even better in recruiting endogenous ADAR in HeLa cells (Fig. 2b), demonstrating that ADARs' double-stranded RNA–binding domains accept extensive chemical modification.

First we tested which ADAR isoform was recruited by ASO v9.5 in HeLa cells. On western blot, ADAR1 p110 was well expressed, whereas ADAR1 p150 was faintly visible but clearly inducible by IFN-α (Fig. 2d and Supplementary Fig. 4). ADAR2 was not detectable (data not shown). We applied RNA interference to knock down specific ADAR isoforms. When transfecting an siRNA against ADAR2 or a mock siRNA, editing was unaffected (Fig. 2c). However, the knockdown of ADAR1 p150 resulted in a decrease of editing (down to 10%–20%). The concurrent knockdown of both ADAR1 isoforms completely abolished editing. Both ADAR1 isoforms contributed to editing; however, the much more weakly

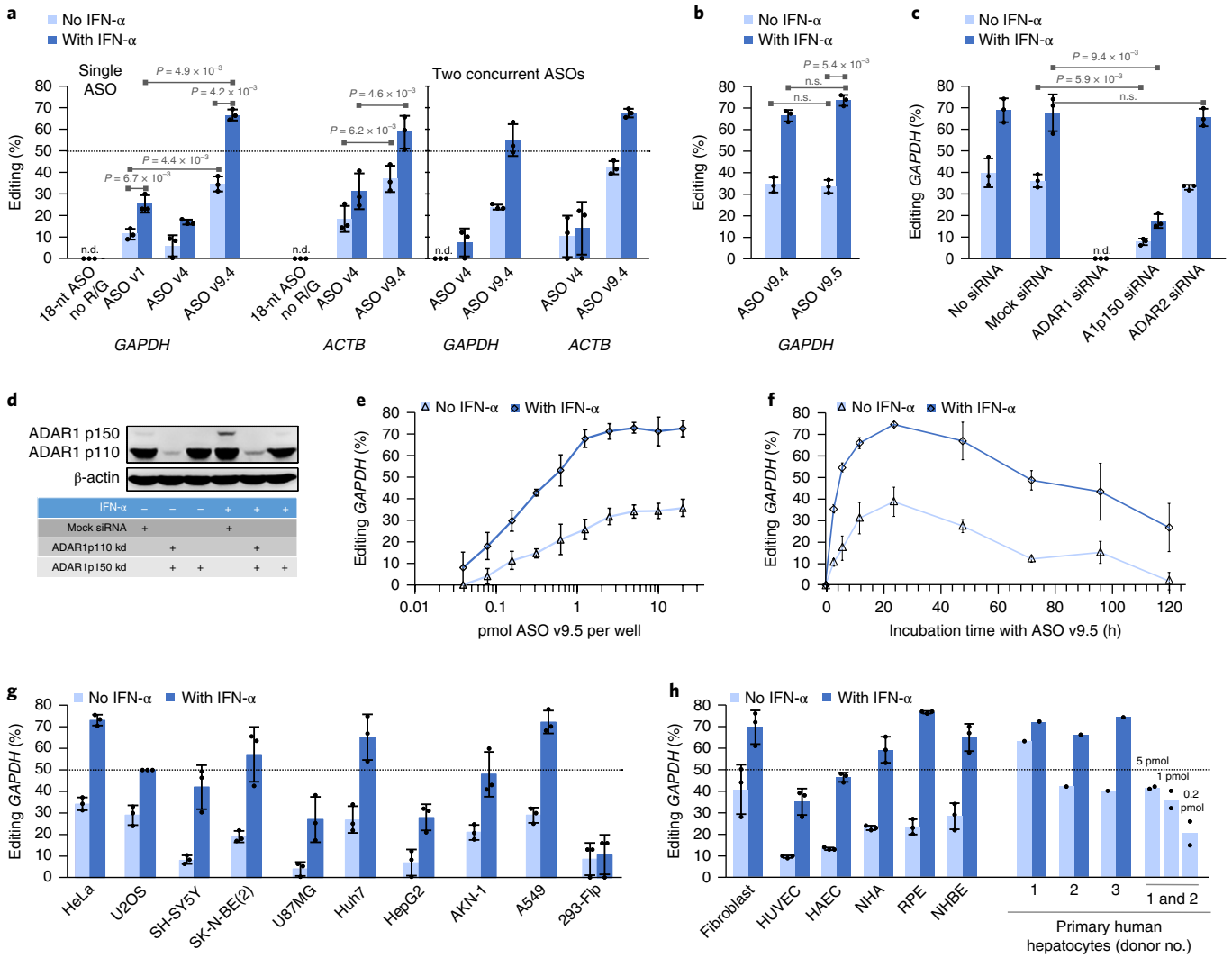


Fig. 2 | Applying RESTORE to edit endogenous transcripts (*GAPDH* and *ACTB*, each with a targeted 5′ UAG triplet in the 3′ UTR) in various cell lines by transfection with ASOs, performed in presence or absence of IFN-α, as indicated. a, Comparing ASO designs for the recruitment of endogenous ADAR in HeLa cells. Either a single ASO (against *GAPDH* or *ACTB*) or both ASOs (against *GAPDH* and *ACTB*) were transfected. "no R/G" indicates an ASO lacking the ADAR-recruiting domain. **b**, Comparative editing of ASO v9.4 and v9.5 on *GAPDH*. **c**, Effect of isoform-specific ADAR knockdown on the *GADPH* editing yield in HeLa cells. **d**, The knockdown efficiency was verified by western blot in technical duplicate. The western blot is composed of two images with different exposure times. The full blots are given in Supplementary Fig. 4. **e**, Determination of the half-maximal effective dose (ED_{50}) of ASO v9.5 for editing *GAPDH* in HeLa cells. **f**, Time course of *GAPDH* editing yields in HeLa cells. **f**, *GAPDH* editing yields with ASO v9.5 in various standard (cancer) cells lines. **h**, *GAPDH* editing yields with ASO v9.5 in various primary human cells. HUVEC, human umbilical vein endothelial cells; HAEC, human aortic endothelial cells; NHA, normal human astrocytes; RPE, human retinal pigment epithelium; NHBE, normal human bronchial epithelium. Data in **a**–**h** are shown as the mean ± s.d., $N = 3$ independent experiments; experiments in hepatocytes are single determinations for each donor (donors 1–3) as indicated. Significance (P) was calculated with a two-tailed paired t-test; n.s., $P > 0.05$; A1p150, ADAR1 p150; n.d., no editing was detectable.

expressed p150 isoform contributed more. This is in good agreement with the observed positive effect of IFN-α (Fig. 2) and the better performance of the ASO in ADAR1 p150-expressing 293 Flp-In T-REx cells (Fig. 1c). It remains unclear why the weakly expressed p150 isoform is more effective than the more strongly expressed p110 isoform. Reasons could be the different intracellular localization, different regulation, or the additional N-terminal part of the p150 isoform—for example, the Z-DNA binding α domain[1].

We found a sigmoidal dose dependency for ASO v9.5-mediated RNA editing, reaching half-maximum editing yield at 0.2 pmol ASO per well of a 96-well plate with IFN-α and 0.4 pmol/well without IFN-α (Fig. 2e and Supplementary Fig. 5). The maximum editing yield was obtained at ≥2 pmol/well, a dose similar to that used for siRNA duplexes in RNA interference[15]. As additional controls, we tested the effect of a nontargeting ASO v9.5 and of an ADAR-recruiting domain v9.5 lacking any specificity domain on the on-target editing of *GAPDH* with ASO v9.5 (Supplementary Figs. 6 and 7). The on-target yield was not affected by the cotransfected components, indicating that the endogenous editing capacity is not limiting. We further assayed the time profile of the editing yield over 5 d in rapidly dividing HeLa cells (10% FBS, 5 pmol/well ASO). The maximum editing yield was observed 12–48 h after transfection and dropped slowly (Fig. 2f).

To assess the application scope of RESTORE, we applied ASO v9.5 in a panel of ten immortalized human cell lines (Fig. 2g). Editing yield was cell line dependent, with yields ranging from 4% to 34% (average 18.5%). Yields were two- to threefold higher after IFN-α treatment, ranging from 11% to 73% (average 46.8%). As ADAR

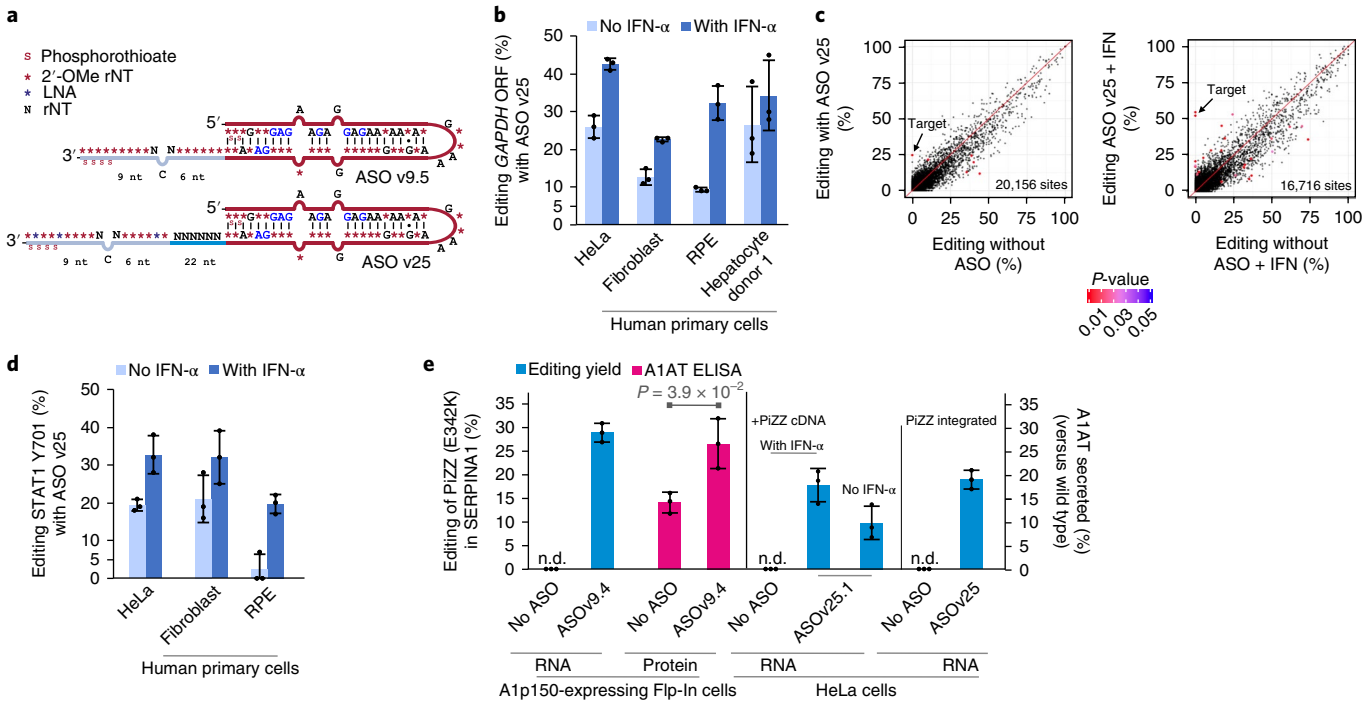


Fig. 3 | Applying RESTORE for ORF editing with ASO v25, off-target analysis, and editing of disease-relevant sites. a, ASO design v25. **b**, Editing of 5′ UAG site no. 1 in the ORF of *GAPDH* with ASO v25 in HeLa and human primary cells. **c**, Analysis of off-target editing in the poly(A)$^+$ transcriptome when recruiting endogenous ADAR from HeLa cells to 5′ UAG site no. 1 in the ORF of *GAPDH* with ASO v25, in absence (left) or presence (right) of IFN-α. Scatter plots show differential editing at ~18,000 sites per experiment comparing editing levels in cells treated with ASO v25 compared to empty transfected cells. Experiments were done in two independent replicates. The on-target editing is indicated by an arrow. Significantly differently edited sites ($P < 0.01$, Fisher's exact test, two-sided, $N > 50$) are highlighted in red. **d**, Editing of the Tyr701 site (5′ UAU codon) of STAT1 in HeLa and primary cells. **e**, Editing of the PiZZ mutation causing α1-antitrypsin deficiency (E342K in SERPINA1, 5′ CAA codon) either in ADAR1 p150–expressing 293 Flp-In T-REx cells with v9.4 ASO or in HeLa cells with v25 ASO (3-nt gap) or v25.1 (2-nt gap). The *SERPINA1* E342K cDNA was either cotransfected or genetically integrated into HeLa cells. α1-Antitrypsin (A1AT) secretion was normalized to the secretion when transfecting wild-type *SERPINA1*. Data in **b**,**d**,**e** are shown as the mean ± s.d., $N = 3$ independent experiments; significance (P) in **e** was calculated with a two-tailed paired t-test. n.d., no editing was detectable.

expression can differ between cancer and normal cells[16], we further tested a panel of seven primary cells from different tissues, including patient fibroblasts[17] and commercially acquired astrocytes, hepatocytes, retinal pigment epithelium cells (RPE), bronchial epithelial cells, and endothelial cells from arterial and venous vessels (Fig. 2h). We found higher editing levels in primary cells than in immortalized cells, obtaining editing levels of 10%–63% (average 31.5%). Notably, in all hepatocyte samples and in the fibroblasts, the editing levels were higher than in HeLa cells. Again, editing yields increased after IFN-α treatment (35%–77%, average 62.6%). We transfected a series of ASO dilutions (0.2–25 pmol ASO v9.5 per well of a 24-well plate, no IFN-α treatment) into hepatocytes of donors 1 and 2 and found a clear dose dependency (Fig. 2h).

We then tested the editing of a 5′ UAG triplet in the open reading frame (ORF) of *GAPDH* in ADAR-expressing 293 Flp-In T-REx cells with an ASO v9.4. The editing in the ORF followed the same trend as in the 3′ UTR (ADAR1 p150 > ADAR1 p110 ≈ ADAR2), but with generally lower yields (11%–55%, Supplementary Fig. 8). Editing required the presence of the ADAR-recruiting domain (Supplementary Fig. 9). ASO v9.4 did not achieve editing in the ORF of *GAPDH* with endogenous ADAR in HeLa or A549 cells.

Thus, we further optimized the ASO design. We assumed that editing in the ORF might be kinetically limited by translation, as we had observed before[18,19]. To improve on-target binding kinetics, we increased the length of specificity domain and included locked nucleic acid (LNA)[20,21] modifications. We tested stepwise elongation of the specificity domain and found elongation at the 5′ site to improve performance. Finally, we identified ASO v25, which contains a 40-nt specificity domain partly modified with 2′-O-methylation, phosphorothioate, and three LNA bases (Fig. 3a). After transfection into HeLa cells, ASO v25 achieved editing yields of 26 ± 3% (without IFN-α) and 42.7 ± 1.5% (with IFN-α; Fig. 3b). The chemical modification of the otherwise unchanged ADAR-recruiting domain was important. Without chemical modifications, v25 gave no editing in the absence and only moderate editing (14 ± 4%) in the presence of IFN-α (Supplementary Fig. 10). We then tested ASO v25 in several primary cells for the editing of the 5′ UAG site in the ORF of *GAPDH*. Editing levels of 12.7 ± 2.1% (fibroblast), 9.3 ± 0.6% (RPE), and 27 ± 10% (hepatocyte) were obtained. As before, IFN-α increased the editing levels, to 22.7 ± 0.6% (fibroblast), 32.3 ± 4.5% (RPE) and 34 ± 9% (hepatocyte).

Off-target editing is a major problem of recent editing strategies. Also ADAR-directing ASOs could potentially elicit off-target editing or perturb the natural editing homeostasis. We conducted deep RNA sequencing (50 Mio 2 × 100 nt paired end reads per experiment) for the editing of the *GAPDH* ORF with ASO v25 in HeLa cells with and without IFN-α. The editing was precise, producing little off-target editing and keeping the natural editing homeostasis intact. In absence of IFN-α, only 3 out of 20,156 sites were significantly differently edited ($P < 0.01$, Fisher's exact test) compared to the control lacking ASO transfection (Fig. 3c, left, and Supplementary Datasets 1–3). All off-target sites were known sites, in noncoding regions (introns, 3′ UTR). With IFN-α, the on-target yield increased from 25% to 52% (Fig. 3c, right), and 14 significantly differently edited off-target sites were detected, all in noncoding regions (yields 17–55%). Most sites (11 of 14) were known.

Both the 3 novel sites and the 11 known sites represented ASO-dependent off-targets effects as supported by sequence alignment with the ASO (Supplementary Fig. 11). Notably, 5 of 14 off-target sites showed attenuated editing in presence of the ASO. Sequence analysis suggests that this was due to a steric blockade of those specific natural editing sites by the ASO (Supplementary Fig. 12) and was not due to a general sequestering of ADAR by the ADAR-recruiting domain of the ASO. For comparison, the effect of IFN-α on ADAR1 expression and on the editing homeostasis was clearly visible (Supplementary Fig. 13), whereas no effects on ADAR1 expression (Supplementary Fig. 13) and global editing homeostasis (Fig. 3c) were detectable for the transfection of the ASO under both conditions (with or without IFN-α).

To illustrate the therapeutic potential of RESTORE, we give two examples. First, we targeted the functionally important phosphotyrosine 701 in endogenous signal transducer and activator of transcription 1 (STAT1)[22]. With an ASO v25 we achieved editing yields of 21.0 ± 6.2% in primary fibroblasts and up to 7% in RPE cells without IFN-α (Fig. 3d). With IFN-α, the yields increased to 32 ± 7% (fibroblasts) and 19.7 ± 2.5% (RPE). Overall, editing of the endogenous *STAT1* transcript was possible in moderate yields in primary cell lines and HeLa cells. Second, we edited the PiZZ mutation (E342K) in SERPINA1 (serpin family A member 1), the most common cause of α1-antitrypsin (A1AT) deficiency[23]. Loss of functional antitrypsin due to the PiZZ allele causes severe damage to the lungs and the liver. Initially, we edited the E342K mutation (5′ CAA triplet) by overexpression of the mutated *SERPINA1* cDNA in ADAR1 p150–expressing 293 Flp-In T-REx cells. When applying an ASO v9.4, we achieved an editing yield of 29 ± 2% (Fig. 3e). The secretion of A1AT was measured by ELISA and was normalized to the secretion by cells transfected with wild-type *SERPINA1* cDNA. The secretion level was elevated from 14 ± 1.8% before to 27 ± 4.3% after repair. The 5′ CAA triplet contains an additional editable adenosine in closest proximity to the targeted A. We indeed found off-target editing at this proximal site (Supplementary Fig. 14); however, this was strongly reduced by further chemical modification of the ASO (Supplementary Fig. 15), as described before in the SNAP-ADAR system[19]. To test the repair of the PiZZ mutation with endogenous ADAR, we created a HeLa cell line stably expressing mutated *SERPINA1* cDNA using the piggyBac[24] system or by plasmid-borne overexpression of the cDNA. With an ASO v25, we obtain editing yields of 19 ± 2% (piggyBac, with IFN-α), 18 ± 4% (plasmid-borne, with IFN-α) and 10 ± 4% (plasmid-borne, without IFN-α).

Several strategies for site-directed adenosine-to-inosine RNA editing have been described so far, including SNAP-ADAR[5], λ N-ADAR[6] and Cas13b-ADAR[7]. However, they all have severe limitations with respect to therapy. First, all systems require the codelivery of an artificial deaminase together with a gRNA in appropriate stoichiometry. Second, they all suffer from massive off-target editing (tens of thousands of sites) due to the overexpression of ADAR fusions[7,19,25], an unsolved problem[26]. By contrast, our RESTORE approach simplifies the delivery and only a few off-target editing events were observed in our experiments. Our ASOs recruit endogenous ADARs to edit endogenous transcripts in good to moderate yields in many primary human cells. The editing yields are in the range of or even better than those achieved with the recently published Cas13b-ADAR strategy[7] in HEK293 cells. The codon scope of RESTORE is probably limited by the codon preferences of natural ADARs[27], but we have already demonstrated here the editing of three different codons. The codon scope can be extended when using engineered hyperactive deaminases[26]; however, this is hampered by massive off-target editing[26]. In contrast, our data suggest that RESTORE allows editing with minimal off-target effects and without perturbing the natural editing homeostasis, unlike the other strategies.

ASOs have been developed as drugs to interact with RNase H, RNA interference, and splicing[11]. RESTORE now adds the reprogramming of genetic information at specific sites by interaction with ADARs. We demonstrated the editing of two disease-relevant transcripts, *SERPINA1* and *STAT1*, with v25 ASOs. Notably, the delivery of therapeutically effective, chemically stabilized siRNAs and ASOs into human liver has been achieved recently[28,29]. We found primary hepatocytes comparably suitable for the RESTORE approach, and good editing has already been achieved in absence of IFN-α. Hepatocytes would also be the target for many inherited genetic diseases, including α1-antitrypsin deficiency.

In the past, optimization of ASO sequence and chemistry was crucial to creating drugs that are effective in the clinic[11,27,29]. We found here that our ASOs accept dense chemical modification and outcompete plasmid-borne gRNAs to recruit endogenous ADARs. There is still a large sequence and chemistry space to further improve the pharmacological properties of ADAR-directing ASOs—for example, to make the ASO shorter, to recruit ADARs more efficiently, and to expand the approach to other ADAR isoforms. This last might allow good editing without IFN-α-driven induction of ADAR1 p150 in the future. However, we expect IFN-α treatment to be more suitable in a therapeutic setting than ectopic expression of ADARs[9,30], as the latter could be difficult to deliver and control, whereas IFN-α is an approved drug[31]. Together, this work sets the stage for the development of a new drug system to reprogram the transcriptome using only antisense oligonucleotides.

Online content

Any methods, additional references, Nature Research reporting summaries, source data, statements of data availability, and associated accession codes are available at https://doi.org/10.1038/s41587-019-0013-6.

Received: 1 June 2018; Accepted: 11 December 2018;
Published online: 28 January 2019

References

1. Nishikura, K. A-to-I editing of coding and non-coding RNAs by ADARs. *Nat. Rev. Mol. Cell Biol.* **17**, 83–96 (2016).
2. Vogel, P. & Stafforst, T. Site-directed RNA editing with antagomir deaminases-a tool to study protein and RNA function. *ChemMedChem.* **9**, 2021–2025 (2014).
3. Gagnidze, K., Rayon-Estrada, V., Harroch, S., Bulloch, K. & Papavasiliou, F. N. A new chapter in genetic medicine: RNA editing and its role in disease pathogenesis. *Trends Mol. Med.* **24**, 294–303 (2018).
4. Rossi, A. et al. Genetic compensation induced by deleterious mutations but not gene knockdowns. *Nature* **524**, 230–233 (2015).
5. Stafforst, T. & Schneider, M. F. An RNA-deaminase conjugate selectively repairs point mutations. *Angew. Chem. Int. Ed.* **51**, 11166–11169 (2012).
6. Montiel-Gonzalez, M. F., Vallecillo-Viejo, I., Yudowski, G. A. & Rosenthal, J. J. C. Correction of mutations within the cystic fibrosis transmembrane conductance regulator by site-directed RNA editing. *Proc. Natl Acad. Sci. USA* **110**, 18285–18290 (2013).
7. Cox, D. B. T. et al. RNA editing with CRISPR-Cas13. *Science* **358**, 1019–1027 (2017).
8. Woolf, T. M., Chase, J. M. & Stinchcomb, D. T. Toward the therapeutic editing of mutated RNA sequences. *Proc. Natl Acad. Sci. USA* **92**, 8298–8302 (1995).
9. Wettengel, J., Reautschnig, P., Geisler, S., Kahle, P. J. & Stafforst, T. Harnessing human ADAR2 for RNA repair—recoding a PINK1 mutation rescues mitophagy. *Nucleic Acids Res.* **45**, 2797–2808 (2017).
10. Picardi, E. et al. Profiling RNA editing in human tissues: towards the inosinome atlas. *Sci. Rep.* **5**, 14941 (2015).
11. Bennett, C. F., Baker, B. F., Pham, N., Swayze, E. & Geary, R. S. Pharmacology of Antisense Drugs. *Annu. Rev. Pharmacol. Toxicol.* **57**, 81–105 (2017).
12. Vogel, P., Schneider, M. F., Wettengel, J. & Stafforst, T. Improving site-directed RNA editing in vitro and in cell culture by chemical modification of the gRNA. *Angew. Chem. Int. Ed. Engl.* **53**, 6267–6271 (2014).
13. Heep, M., Mach, P., Reautschnig, P., Wettengel, J. & Stafforst, T. Applying human ADAR1p110 and ADAR1p150 for site-directed RNA editing-G/C substitution stabilizes gRNAs against editing. *Genes (Basel)* **8**, 34 (2017).

14. Patterson, J. B., Thomis, D. C., Hans, S. L. & Samuel, C. E. Mechanism of interferon action: double-stranded RNA-specific adenosine deaminase from human cells is inducible by alpha and gamma interferons. *Virology* **210**, 508–511 (1995).
15. Kim, D.-H. et al. Synthetic dsRNA Dicer substrates enhance RNAi potency and efficacy. *Nat. Biotechnol.* **23**, 222–226 (2005).
16. Xu, X., Wang, Y. & Liang, H. The role of A-to-I RNA editing in cancer development. *Curr. Opin. Genet. Dev.* **48**, 51–56 (2018).
17. Valente, E. M. et al. Hereditary early-onset Parkinson's disease caused by mutations in PINK1. *Science* **304**, 1158–1160 (2004).
18. Vogel, P., Hanswillemenke, A. & Stafforst, T. Switching protein localization by site-directed RNA editing under control of light. *ACS Synth. Biol.* **6**, 1642–1649 (2017).
19. Vogel, P. et al. Efficient and precise editing of endogenous transcripts with SNAP-tagged ADARs. *Nat. Methods* **15**, 535–538 (2018).
20. Singh, S. K., Koshkin, A. A., Wengel, J. & Nielsen, P. LNA (locked nucleic acids): synthesis and high-affinity nucleic acid recognition. *Chem. Commun. (Camb).* 455–456 (1998).
21. Obika, S. et al. Synthesis of 2′-O,4′-C-methyleneuridine and -cytidine. Novel bicyclic nucleosides having a fixed C3, -endo sugar puckering. *Tetrahedr. Lett.* **38**, 8735–8738 (1997).
22. O'Shea, J. J. et al. The JAK-STAT pathway: impact on human disease and therapeutic intervention. *Annu. Rev. Med.* **66**, 311–328 (2015).
23. Lomas, D. A. & Mahadeva, R. α1-antitrypsin polymerization and the serpinopathies: pathobiology and prospects for therapy. *J. Clin. Invest.* **110**, 1585–1590 (2002).
24. Woodard, L. E. & Wilson, M. H. piggyBac-ing models and new therapeutic strategies. *Trends Biotechnol.* **33**, 525–533 (2015).
25. Vallecillo-Viejo, I. C. et al. Abundant off-target edits from site-directed RNA editing can be reduced by nuclear localization of the editing enzyme. *RNA Biol.* **15**, 104–114 (2018).
26. Vogel, P. & Stafforst, T. Critical review on engineering deaminases for site-directed RNA editing. *Curr. Opin. Biotechnol.* **55**, 74–80 (2018).
27. Eggington, J. M., Greene, T. & Bass, B. L. Predicting sites of ADAR editing in double-stranded RNA. *Nat. Commun.* **2**, 319 (2011).
28. Prakash, T. P. et al. Targeted delivery of antisense oligonucleotides to hepatocytes using triantennary N-acetyl galactosamine improves potency 10-fold in mice. *Nucleic Acids Res.* **42**, 8796–8807 (2014).
29. Fitzgerald, K. et al. A highly durable RNAi therapeutic inhibitor of PCSK9. *N. Engl. J. Med.* **376**, 41–51 (2017).
30. Fukuda, M. et al. Construction of a guide-RNA for site-directed RNA mutagenesis utilising intracellular A-to-I RNA editing. *Sci. Rep.* **7**, 41478 (2017).
31. Antonelli, G., Scagnolari, C., Moschella, F. & Proietti, E. Twenty-five years of type I interferon-based treatment: a critical analysis of its therapeutic use. *Cytokine Growth Factor Rev.* **26**, 121–131 (2015).

Acknowledgements

We gratefully acknowledge the donation of primary fibroblasts from E. M. Valente (Università degli Studi di Salerno, Fisciano, Italy), of the hepatocytes cell line AKN-1 from A. Nüssler (BG Klinik, Tübingen, Germany), and of the U2OS Flp-In cell line from E. Schiebel (Universität Heidelberg, Germany). We gratefully acknowledge support from the Deutsche Forschungsgemeinschaft to T.S. (STA 1053/3-2; STA 1053/7-1). This work is supported by the Institutional Strategy of the University of Tübingen (Deutsche Forschungsgemeinschaft, ZUK 63) with an intramural innovation grant for J.W. This work is supported by National Institutes of Health grants R01GM102484 and R01GM124215 to J.B.L.

Author contributions

T.M., S.M., A.B., P.R., J.W., P.V. and T.S. conceived, performed and analyzed the experiments. Q.L. and J.B.L. analyzed and all authors interpreted next-generation sequencing data. All authors contributed to writing the manuscript.

Competing interests

T.S., J.W. and P.V. hold a patent on site-directed RNA editing (PCT/DE2016/000309). T.S., J.W., P.R. and T.M. are inventors of a filed patent based on the work published here.

Additional information

Supplementary information is available for this paper at https://doi.org/10.1038/s41587-019-0013-6.

Reprints and permissions information is available at www.nature.com/reprints.

Correspondence and requests for materials should be addressed to T.S.

Publisher's note: Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Methods

Antisense oligonucleotides. Unmodified RNA oligonucleotides were produced by in vitro transcription from linear synthetic DNA templates (purchased from Sigma-Aldrich, Germany) with T7 RNA polymerase (Thermo Scientific, USA) at 37 °C overnight. The resulting RNA was precipitated in ethanol and purified via urea (7 M) polyacrylamide (15%) gel electrophoresis (PAGE), extracted into water, precipitated with ethanol and resuspended and stored in nuclease-free water. All chemically modified RNA oligonucleotides purchased from Biospring (Germany), Eurogentec (Belgium) or Dharmacon (USA). Long sequences were assembled from two pieces by ligation. Sequences and modification patterns of all ASO are given in Supplementary Table 1.

Analysis of RNA editing. Total RNA was extracted from the cells with the RNeasy MinElute Kit (Qiagen, Germany). After DNase I (NEB, USA) treatment and reverse transcription with M-MuLV reverse transcriptase (NEB, USA), a subsequent PCR with Taq DNA polymerase (NEB) was performed. The resulting DNA was purified on an agarose gel and analyzed by Sanger sequencing (Eurofins Genomics, Germany). Adenosine-to-inosine editing yields were quantified by measuring the height of the guanosine and adenosine peaks at the respective site and dividing the guanosine peak height by the sum of the guanosine and adenosine peak heights. If the reverse primer was used for sequencing, cytidine and thymidine peaks were treated accordingly.

Cloning and editing with the plasmid-borne approach. Firefly luciferase was expressed under control of a CMV promotor from a pShuttle-CMV plasmid (see Supplementary Note 1). The W417X amber mutation was introduced via overlap PCR. Sequences of the cloned products were verified by Sanger sequencing. The R/G gRNAs were expressed under control of the U6 promotor from a modified pSilencer backbone as described before[9]. Sequences of the cloned products were verified by Sanger sequencing. Sequences of all applied R/G gRNAs are given in Supplementary Table 1. Flp-In 293 T-REx cells (R78007, Thermo Fisher Scientific) containing the respective genomically integrated ADAR version were generated previously[9,13]. Cells were cultured in DMEM plus 10% FBS plus 100 µg/mL hygromycin B plus 15 µg/mL blasticidin S. For editing, 2.5×10^5 cells/well (ADAR1p110, ADAR1p150) or 3×10^5 cells/well (ADAR2) were seeded into poly-D-lysine-coated 24-well plates in 500 µL DMEM plus 10% FBS plus 10 ng/mL doxycycline. Twenty-four hours later, transfection was performed with the luciferase reporter plasmid (300 ng) and the R/G gRNA (1,300 ng) using a ratio of Lipofectamine 2000 to plasmid of 3:1. The medium was changed every 24 h until harvest. RNA was isolated and sequenced 72 h after transfection, as described above. Results are reported in Supplementary Fig. 1.

Editing procedure with ASOs in ADAR-expressing 293 cells. Forty-eight hours before ASO transfection, 2×10^5 of the respective ADAR-Flp-In 293 T-REx cells per well were seeded in 24-well plates in DMEM plus 10%FBS containing 10 ng/mL doxycycline for induction of ADAR gene expression. After 48 h cells were detached and reverse-transfected in 96-well plates. For this, the respective ASO (5 pmol/well unless stated otherwise) and Lipofectamine 2000 (0.75 µL/well) were each diluted with OptiMEM to a volume of 10 µL in separate tubes. After 5 min, the two solutions were mixed and 100 µL cell suspension (5×10^4 cells) in DMEM plus 10%FBS plus 10 ng/mL doxycycline was added to the transfection mixture inside 96-well plates. Twenty-four hours later, cells were harvested for RNA isolation and sequencing as described above. Results are reported in Fig. 1c and Supplementary Figs. 2, 3, 6 and 8a.

Editing procedure with ASO in HeLa cells. HeLa cells (cat. no. ATCC CCL-2) were cultured in DMEM plus 10% FBS plus P/S (100 U/mL penicillin and 100 µg/mL streptomycin). 5×10^4 cells in 100 µL DMEM plus 10% FBS (plus 600 units IFN-α , Merck, cat. no. IF007, lot number 2937858) were added to a transfection mix of 0.5 µL Lipofectamine 2000 and 5 pmol gRNA/well in a 96-well format. For concurrent editing with two different ASOs, 2.5 pmol of each respective ASO were cotransfected. After 24 h cells were harvested for RNA isolation and sequencing. Results are reported in Fig. 2a–f and Supplementary Fig. 7.

siRNA knockdown of ADAR isoforms and western blot. HeLa cells were reverse transfected in 12-well format with 2.5 pmol siRNA against ADAR1 (both isoforms, Dharmacon, SMARTpool: ON-TARGETplus ADAR (103) siRNA, L-008630-00-0005), ADAR1p150 (Ambion (Life Technologies), sense strand: 5′-GCCUCGCGGGCGCAAUGAAtt; antisense strand: 5′-UUCAUUGCGCCCGCGAGGCat), ADAR2 (Dharmacon, SMARTpool: ON-TARGETplus ADARB1 (104) siRNA, L-009263-01-0005) or mock (Dharmacon, siGENOME Non-Targeting siRNA Pool #2, D-001206-14-05). For this, 200 µL of transfection mix, containing 2.5 µL of the respective siRNA (1 nM) and 3 µL HiPerFect (Qiagen, Germany) and OptiMEM, were distributed evenly in each well before adding 800 µL DMEM plus 10% FBS containing 1.2×10^5 HeLa cells. Medium was changed every 24 h. For RNA editing experiments, cells were detached 48 h after siRNA transfection and were reverse-transfected with the respective ASO as described above. For western blotting, cells were harvested and lysed in urea lysis buffer (8 M urea, 100 mM NaH_2PO_4, 10 mM Tris, pH 8.0) 72 h after reverse transfection of the siRNA. Shear force was applied using a 23-gauge syringe, and the cell debris was removed by centrifugation at 30,000 *g* for 15 min at 4 °C. Then a Bradford assay was used to normalize total protein amounts, and appropriate amounts of protein lysate in 1× Laemmli buffer were loaded for SDS–PAGE (4% stacking, 12% separating gel). Proteins were transferred on a PVDF membrane using a tank-blotting system at 30 V overnight. The membrane was blocked in 5% nonfat dry milk TBST plus 50 µg/mL avidin for 2 h at room temperature, and was afterwards incubated with the primary antibodies (5% nonfat dry milk TBST plus 1:1,000 anti-ADAR1, Santa Cruz, sc-73408 or anti-ADAR2, Santa Cruz, sc-73409, plus 1:40,000 anti-beta-actin, Sigma Aldrich, A5441) at 4 °C overnight. The secondary antibodies (5% nonfat dry milk TBST plus 1:10,000 anti-mouse–HRP plus 1:50,000 Precision Protein StrepTactin-HRP Conjugate, Bio-Rad, cat. no. 1610381) were incubated for 1.5 h at room temperature. After each antibody incubation, the membrane was washed three times for 5 min with TBST. Detection was performed using 1 mL of Clarity Western ECL Substrate (Bio-Rad) and a Fusion SL Vilber Lourmat (Vilber). For antibodies, see also Supplementary Table 2. Results are reported in Fig. 2c,d and Supplementary Fig. 4.

Potency determination. For potency determination, HeLa cells were transfected as described above with varying ASO amounts (39 fmol–20 pmol per well of a 96-well plate). Results are reported in Fig. 2e.

Time course. For time course experiments, HeLa cells were transfected as described above. Prior transfection cells were treated with IFN-α for 24 h (where indicated). Cells were harvested for RNA isolation at the respective time points indicated. For time points later than 24 h after transfection, cells were detached after 24 h and transferred into 24-well plates to avoid overgrowth. Medium (containing IFN-α where indicated) was changed every 24 h. Results are reported in Fig. 2f.

Screening of immortalized cell lines. ASO transfection was not systematically optimized. All cells were cultured in DMEM plus 10% FBS plus P/S. 5×10^4 cells of the respective cell line per well of a 96-well plate (HeLa cells (cat. no. ATCC CCL-2), U2OS-Flp-In T-REx[32] (kind donation from Elmar Schiebel), SK-N-BE(2) (cat. no. ATCC CRL-2271), U87MG (cat. no. ATCC HTB-14), Huh7 (CLS GmbH, Heidelberg, cat. no. 300156), HepG2 (DSMZ, Braunschweig, Germany, cat. no. ACC180), AKN-1(kind donation from the Nüssler laboratory[33]), empty HEK-Flp-In T-REx (R78007, Thermo Fisher scientific, stably transfected with empty pcDNA5 vector) and A549 (European Collection of Authenticated Cell Cultures ECACC 86012804)) were reverse transfected with the respective ASO (5 pmol per well of a 96-well plate) as described above for HeLa cells without further optimization. Only SH-SY5Y (cat. no. ATCC CRL-2266) cells were reverse transfected differently, in a 24-well format: to 100 µL transfection mix consisting of 2.5 µL Lipofectamine 2000 and 25 pmol ASO in OptiMEM, 5×10^5 cells in 500 µL medium (plus 3,000 U IFN-α) were added. Results are reported in Fig. 2g.

Screening of human primary cell lines. ASO transfection was not systematically optimized. All primary cells were purchased from Lonza except for the primary fibroblasts, which were a kind gift from the Valente laboratory[17]. Primary fibroblasts were cultured in DMEM plus 20%FBS. The other cell lines were cultured in their respective commercial media as indicated: human umbilical vein endothelial cells (HUVEC, Lonza cat. no. CC-2517) and human aortic endothelial cells (HAEC, Lonza cat. no. CC-2535) in medium 200PRF (Thermo Fisher Scientific cat. no. M200PRF500) with Low Serum Growth Supplement (Thermo Fisher Scientific cat. no. S00310), normal human astrocytes (NHA, Lonza cat. no. CC-2565) in ABM Basal Medium (Lonza cat. no. CC-3187) with AGM SingleQuot Kit Supplementary & Growth Factors (Lonza cat. no. CC-4123), human retinal pigment epithelial cells (H-RPE, Lonza cat. no. 194987) in EpiLife Medium (Thermo Fisher Scientific cat. no. MEPI500CA) with Human Corneal Growth Supplement (Thermo Fisher Scientific cat. no. S0095), and normal human bronchial epithelial cells (NHBE, Lonza cat. no. CC-2540) in Airway Epithelial Cell Basal Medium (LGC Standard cat. no. ATCC-PCS-300-030) with the Bronchial Epithelial Cell Growth Kit (LGC Standard cat. no. ATCC-PCS-300-040). Primary human hepatocytes (PHH, Lonza cat. no. HUCPI) were thawed in Cryo HH thawing medium (Lonza cat. no. MCHT50), seeded in Hepatocyte Plating Medium with Supplement (Lonza cat. no. MP100) and, 6 h after seeding, cultured in Hepatocyte Maintenance Medium with Supplement (Lonza cat. no. MM250). 3.5×10^4 HUVEC and HAEC, 1×10^5 NHA, H-RPE and NHBE and 4.5×10^5 PHH cells were seeded 24 h before ASO transfection in 24-well format. For PHH, rat collagen I-coated 24-well plates (GreinerBioOne) were used. Shortly before forward transfection, medium was changed; 3,000 U IFN-α in 500 µL medium per well was included if indicated. For each 24-well, 1.5 µL Lipofectamine RNAiMAX (Thermo Fisher Scientific) and 25 pmol ASO were diluted separately in a total volume of 50 µL OptiMEM, respectively. After 5 min incubation the two solutions were combined, and after another 20 min incubation, the 100 µL transfection mix was evenly distributed in one well. After 24 h cells were harvested for RNA isolation and sequencing. Results are reported in Fig. 2h.

ORF editing. If not indicated, ORF editing experiments were performed the same as editing experiments in the 3′ UTR for the respective cell lines as described

above. For PHH, 7.5 µL RNAiMAX per well were used. Before reverse transcription of RNA from cells treated with design v25 ASOs, total RNA was incubated with an RNA strand reverse complementary to the respective ASO and heated to 95 °C for 3 min. Results are reported in Fig. 3b and Supplementary Figs. 8–10.

Next-generation RNA sequencing experiment. The RNA editing experiment was done by transfection of 5 pmol ASO against 5′-UAG ORF site #1 in the ORF of *GAPDH* into HeLa cells as described above. For samples with IFN-α, HeLa cells were treated with IFN-α 24 h before reverse transfection as described above. Overall, four settings were carried out, each with an independent duplicate. Those settings include (1) empty lipofection, (2) empty lipofection plus IFN-α, (3) ASO transfection, and (4) ASO transfection plus IFN-α. RNA was isolated with the RNeasy MinElute Kit, treated with DNase I, incubated with an RNA strand reverse complementary to the respective ASO and heated to 95 °C for 3 min and purified again with the RNeasy MinElute Kit. Purified RNA was delivered to CeGaT (Germany) for poly(A)$^+$ mRNA sequencing. The library was prepared from 200ng RNA with the TruSeq Stranded mRNA Library Prep Kit (Illumina, USA) and sequenced with a NovaSeq 6000 (50M reads, 2 × 100 bp paired end, Illumina, USA). Results are reported in Fig. 3c and Supplementary Figs. 11–13.

Mapping of RNA-seq and reads. We adopted a previously published pipeline to accurately align RNA-seq reads onto the genome[34,35]. We used BWA (version 0.7.10)[36] to align the reads to a combination of the reference genome sequences and exonic sequences surrounding known splicing junctions from known gene models. Each of the paired-end reads was mapped separately using the commands "bwa aln fastqfile" and "bwa samse -n4". We then chose the length of the splicing junction to be slightly shorter than the RNA-seq reads to prevent redundant alignment (i.e., 95 bp for reads of 100 bp length). The reference genomes used were hg19 and the gene models were obtained through the UCSC Genome Browser for Gencode, RefSeq, Ensembl, and UCSC Genes. We considered only uniquely mapped reads with mapping quality $q > 10$ and used Picard[v] to remove clonal reads (PCR duplicates) mapped to the same location. Of these identical reads, only the read with the highest mapping quality was kept for downstream analysis. Unique and nonduplicate reads were subjected to local realignment and base score recalibration using the IndelRealigner and TableRecalibration from the Genome Analysis Toolkit (GATK, version 3.6)[37]. The above steps were applied separately to each of the RNA-seq samples.

Identification of editing sites from RNA-seq data. We used the UnifiedGenotyper from GATK[37] to call variants from the mapped RNA-seq reads. In contrast to the usual practice of variant calling, we identified the variants with relatively loose criteria by using the UnifiedGenotyper tool with options stand_call_conf 0, stand_emit_conf 0, and output mode EMIT_VARIANTS_ONLY. Variants from nonrepetitive and repetitive non-Alu regions were required to be supported by at least three reads containing mismatches between the reference genome sequences and RNA-seq. Supporting of one mismatch read was required for variants in Alu regions. This set of variant candidates was subject to several filtering steps to increase the accuracy of editing site calling. We first removed all known human SNPs present in dbSNP build 137 (except SNPs of molecular type "cDNA"; database version 135; http://www.ncbi.nlm.nih.gov/SNP/), the 1000 Genomes Project, and the University of Washington Exome Sequencing Project (http://evs.gs.washington.edu/EVS/). To remove false-positive RNA-seq variant calls due to technical artifacts, further filters were applied as previously described[34,35]. In brief, we required a variant call quality $Q > 20$ (refs. [34,35]), discarded variants if they occurred in the first 6 bases of a read[36], removed variants in simple repeats[38], removed intronic variants that were within 4 bp of splice junctions, and discarded[37] variants in homopolymers. Moreover, we removed reads mapped to highly similar regions of the transcriptome by BLAT[39]. Finally, variants were annotated using ANNOVAR (version 11122013)[40] based on gene models from Gencode, RefSeq, Ensembl and UCSC.

Assignment of known versus novel sites. The resulting sets of sites identified from RNA-seq data were compared with all sites available in the RADAR database[41] and were subsequently referred to as 'known' sites if found in RADAR or 'novel' sites if not found.

Identification of significantly differently edited sites. We quantified editing levels of edited sites with ≥50 reads coverage (combined coverage of both replicates) and performed Fisher's exact tests followed by Benjamini–Hochberg's multiple test correction (adjusted $P < 0.01$) to identify significantly differently edited sites across the samples (absolute editing difference > 10%). Additional next-generation sequencing quality data are given in the Supplementary Information.

***SERPINA1* editing and A1AT-ELISA.** To obtain *SERPINA1* cDNA for cloning, total RNA was isolated from HepG2 cells and reverse transcribed. The E342K mutation was inserted into the cDNA by PCR and both *SERPINA1* wild-type and the E342K mutant were each cloned on a pcDNA3.1 vector under control of the CMV promotor using HindIII and ApaI restriction. For genomic integration of *SERPINA1* using the piggyBac transposon system, the wild-type and mutant cDNA was cloned on a PB-CA vector using the same restriction sites as above. 1×10^6 HeLa cells were seeded in a six-well plate 24 h before transfection. 1 µg of the piggyBac transposase vector (Transposagen Biopharmaceuticals) and 2.5 µg of the SERPINA1 PB-CA vector were cotransfected using 10.5 µL FuGENE6 (Promega) according to the manufacturer's protocol. After 24 h, cells were selected for 2 weeks in DMEM plus 10% FBS containing 10 µg/mL puromycin. For editing, stably transfected or plasmid transfected (300 ng plasmid/0.9 µL FuGENE6 for Hela and 100 ng plasmid/0.3 µL Lipofectamine 2000 for Flp-ADAR1p150 cells) cells were reverse transfected with the respective ASO as described above. After 24 h, cell culture supernatant was collected for the A1AT ELISA and cells were harvested for RNA isolation and sequencing. The A1AT ELISA was performed with a commercial kit (cat. no. ab108799, Abcam) according to the manufacturer's protocol. Samples from three biological replicates were measured in technical duplicates. The A1AT protein amount was calculated from a standard curve using linear regression. ASO v25 refers to the ASO with the common 3-nt gap around the editing site; ASO v25.1 refers to an ASO of the same sequence but with an additional chemical modification (2′ O-methyl) close to the editing site (2-nt gap; see also Supplementary Fig. 15 and Supplementary Table 1). Results are reported in Fig. 3e and Supplementary Figs. 14 and 15.

Reporting Summary. Further information on research design is available in the Nature Research Reporting Summary linked to this article.

Data availability

This manuscript provides Supplementary Information on primary data and further controls (Supplementary Figs. 1–15), and it contains a table of ASOs (Supplementary Table 1), a list of target sequences (Supplementary Note 1), and spreadsheets with significantly differently edited sites (Supplementary Datasets 1–3). The original next-generation sequencing data have been deposited in the NCBI GEO database under accession code GSE121573. Code is available at http://lilab.stanford.edu/SNPiR/.

References

32. Malecki, M. J. et al. Leukemia-associated mutations within the NOTCH1 heterodimerization domain fall into at least two distinct mechanistic classes. *Mol. Cell. Biol.* **26**, 4642–4651 (2006).
33. Nüssler, A. K. et al. Isolation and characterization of a human hepatic epithelial-like cell line (AKN-1) from a normal liver. *In Vitro Cell. Dev. Biol. Anim.* **35**, 190–197 (1999).
34. Ramaswami, G. et al. Accurate identification of human Alu and non-Alu RNA editing sites. *Nat. Methods* **9**, 579–581 (2012).
35. Ramaswami, G. et al. Identifying RNA editing sites using RNA sequencing data alone. *Nat. Methods* **10**, 128–132 (2013).
36. Li, H. & Durbin, R. Fast and accurate long-read alignment with Burrows-Wheeler transform. *Bioinformatics* **26**, 589–595 (2010).
37. McKenna, A. et al. The Genome Analysis Toolkit: a MapReduce framework for analyzing next-generation DNA sequencing data. *Genome Res.* **20**, 1297–1303 (2010).
38. Li, H. et al. The Sequence Alignment/Map format and SAMtools. *Bioinformatics* **25**, 2078–2079 (2009).
39. Kent, W. J. BLAT–the BLAST-like alignment tool. *Genome Res.* **12**, 656–664 (2002).
40. Wang, K., Li, M. & Hakonarson, H. ANNOVAR: functional annotation of genetic variants from high-throughput sequencing data. *Nucleic Acids Res.* **38**, e164 (2010).
41. Ramaswami, G. & Li, J. B. RADAR: a rigorously annotated database of A-to-I RNA editing. *Nucleic Acids Res.* **42**, D109–D113 (2014).

nature biotechnology

SUPPLEMENTARY INFORMATION

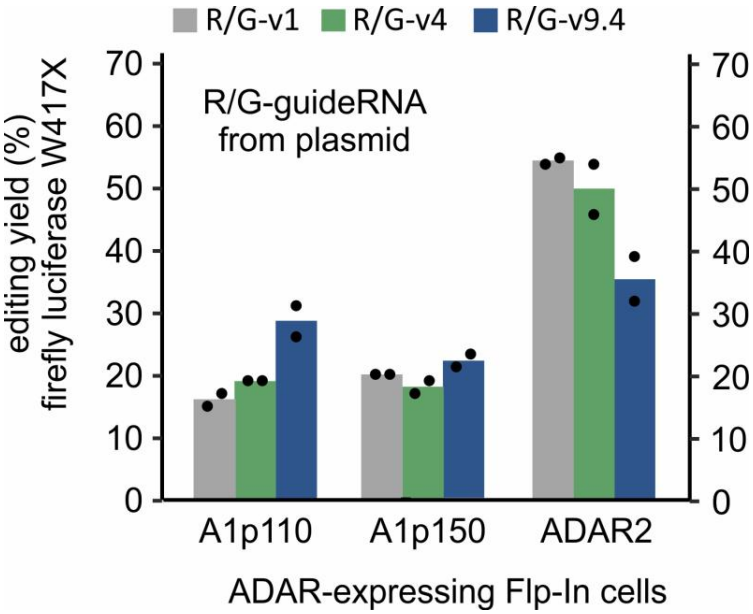
https://doi.org/10.1038/s41587-019-0013-6

In the format provided by the authors and unedited.

Precise RNA editing by recruiting endogenous ADARs with antisense oligonucleotides

Tobias Merkle[1], Sarah Merz[1], Philipp Reautschnig[1], Andreas Blaha[1], Qin Li[2], Paul Vogel[1], Jacqueline Wettengel[1], Jin Billy Li[2] and Thorsten Stafforst[1]*

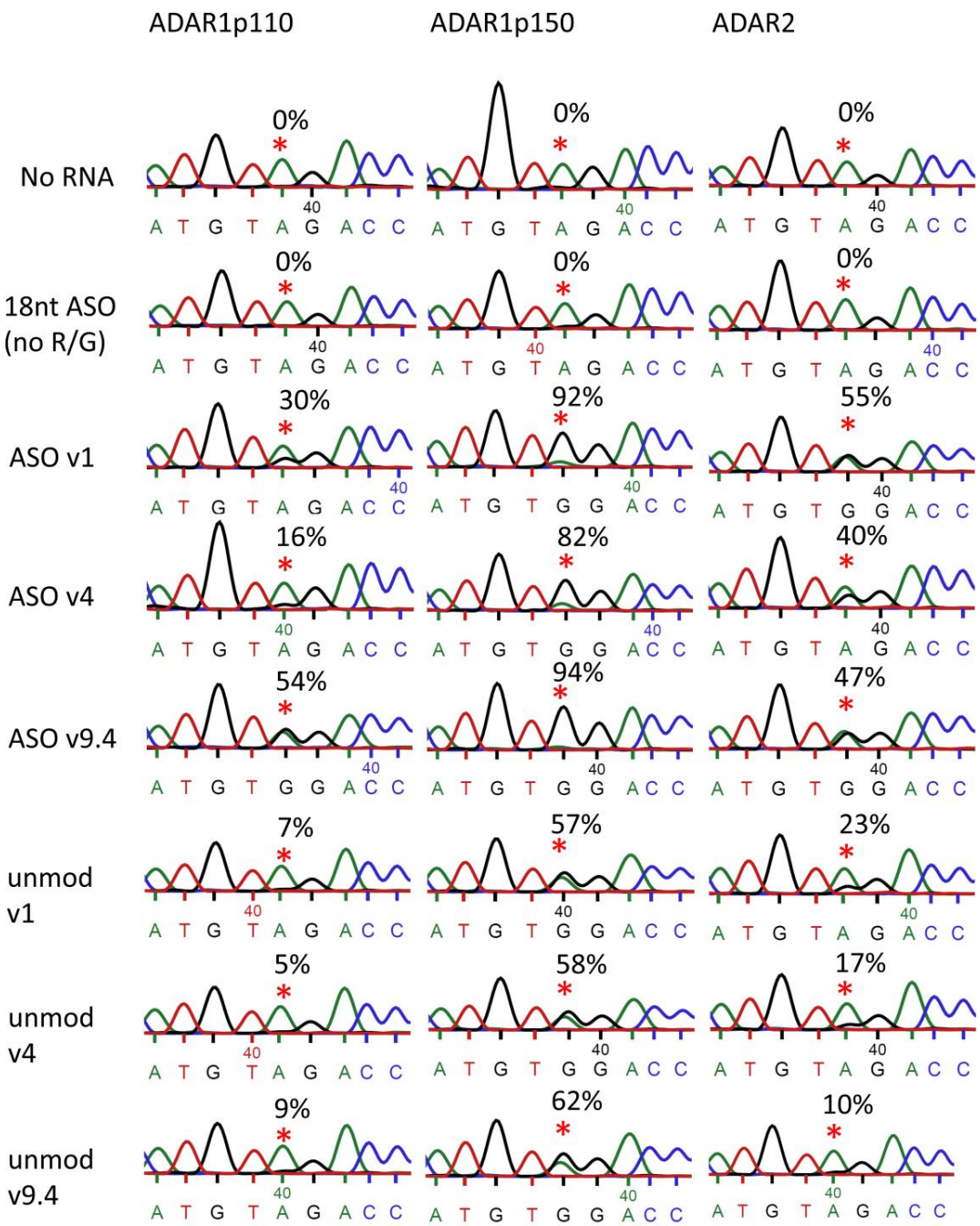
[1]Interfaculty Institute of Biochemistry, University of Tübingen, Tübingen, Germany. [2]Department of Genetics, Stanford University, Stanford, CA, USA.
*e-mail: thorsten.stafforst@uni-tuebingen.de



Supplementary Figure 1

Screening to improve the ADAR-recruiting domain

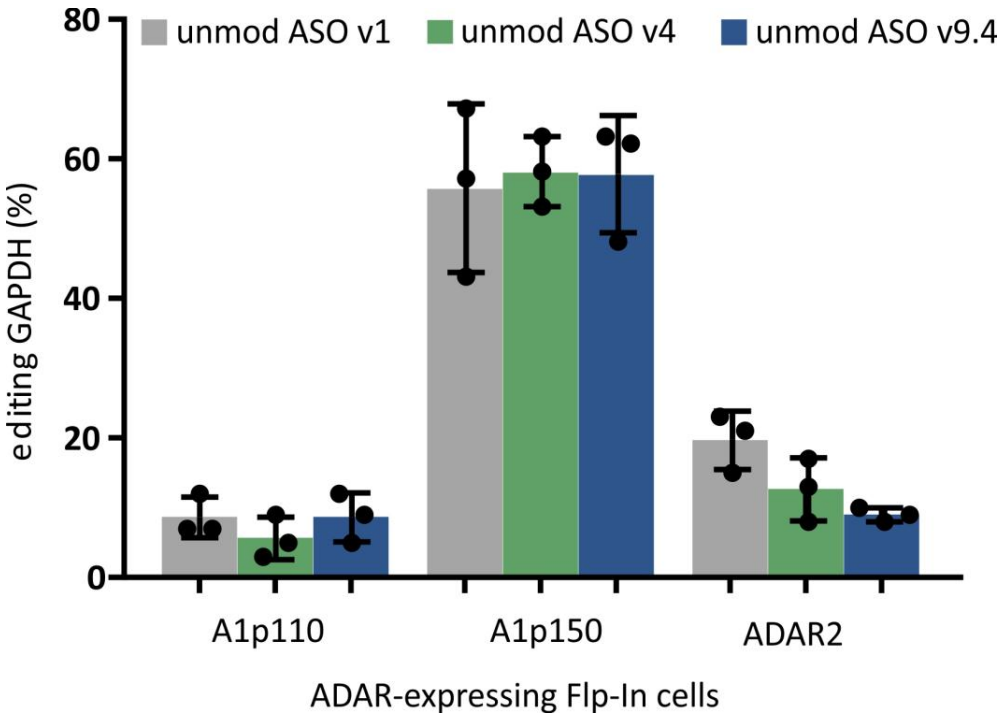
A plasmid borne screening assay was applied to screen for improved ADAR-recruiting domains. For this, plasmids expressing the respective ASO as a chemically unmodified guideRNA from a U6 promotor were prepared. The guide RNA plasmids were co-overexpressed together with a reporter contruct (firefly luciferase) in 293 Flp-In T-REx cells expressing a specific ADAR isoform (A1p110 = ADAR1p110; A1p150 = ADAR1p150). Editing yields were determined by Sanger sequencing. Data are shown as the mean±SD, N=2 independent experiments



Supplementary Figure 2

Sequencing traces for editing of a 5′ UAG site in the 3′ UTR of GAPDH in 293 Flp-In T-REx ADAR cells

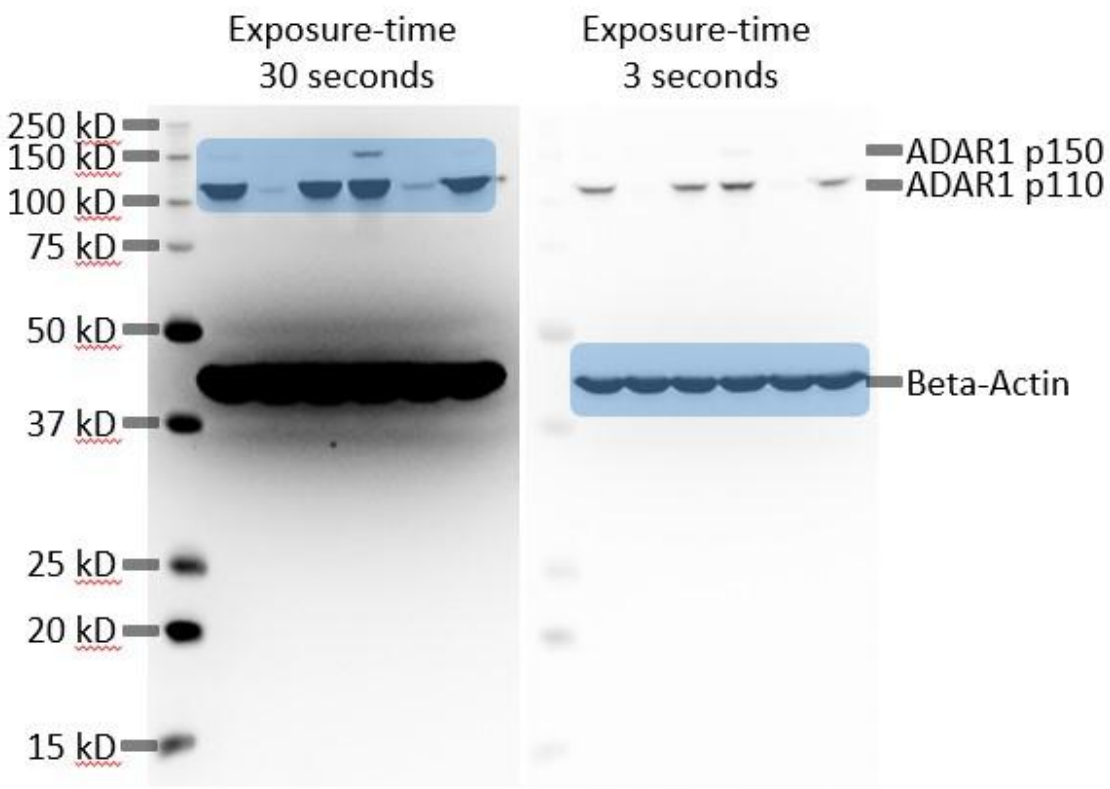
Exemplary editing traces for the editings shown in Figure 1C in the manuscript, but including additional controls (“No RNA” = empty transfection; “18nt ASO no R/G” = ASO lacking the ADAR-recruiting domain; “unmod” means chemically unmodified, *in-vitro* transcribed ASOs of the indicated design v1, v4 or v9.4. Red asterisks indicate the editing sites.



Supplementary Figure 3

Editing yields for targeting a 5′ UAG codon in the 3′ UTR of *GAPDH* with chemically unmodified, in vitro transcribed ASOs in 293 Flp-In T-REx ADAR cells

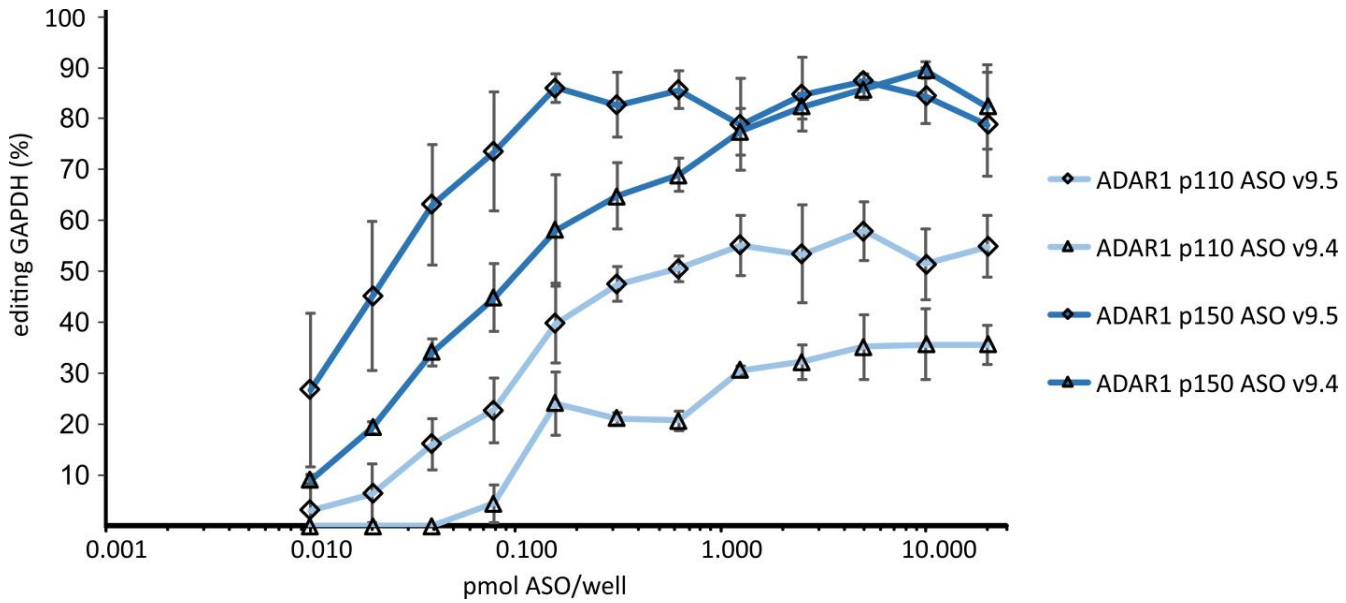
Unmodified *in-vitro* transcribed ASOs v1, v4 and v9.4 (5 pmol / 96well) were transfected into the respective ADAR-expressing Flp-In cell line. Data are shown as the mean±SD, N=3 independent experiments. A1p110 = ADAR1p110; A1p150 = ADAR1p150



Supplementary Figure 4

Western blot analysis of ADAR knockdown

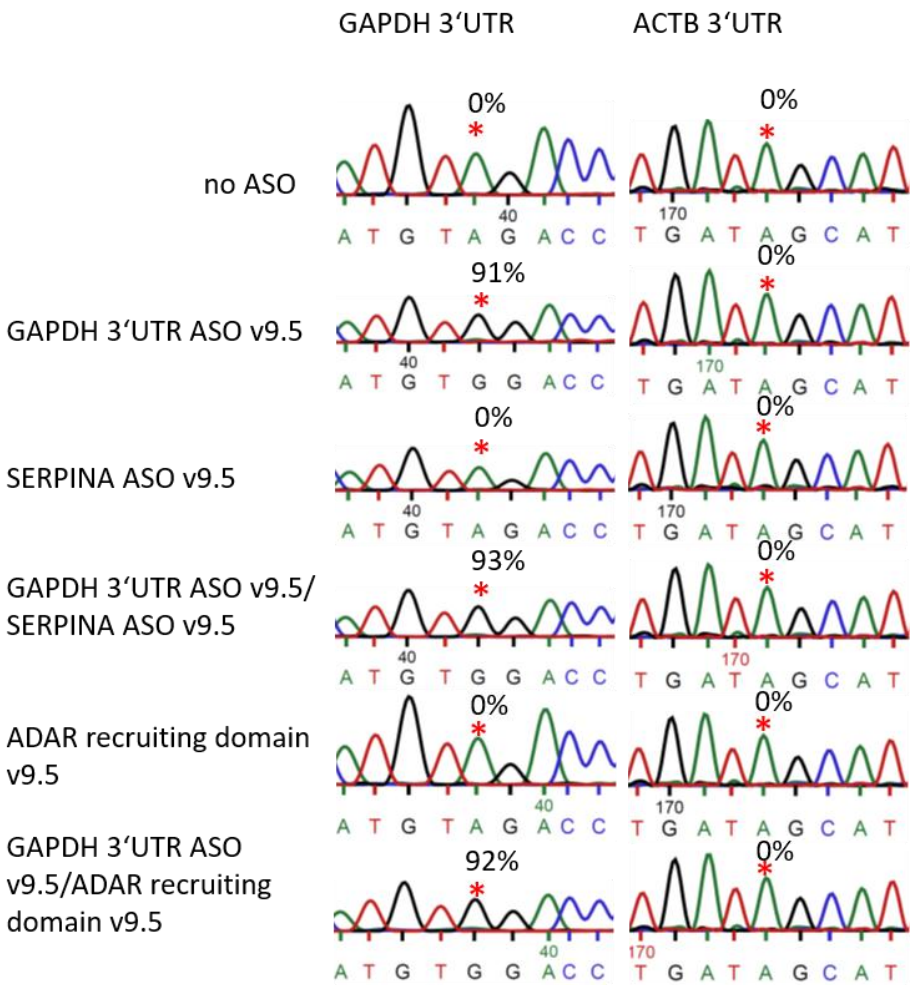
The western blot shown in Figure 2D in the manuscript was merged from images generated with two different exposure times. The part showing the ADAR bands comes from a 30 second exposure. The part showing β-actin from a 3 second exposure. The pictures were captured by the FusionCapt Advance SL4 (16.09b) software installed on the Fusion SL Vilber Lourmat (Vilber) western blot analyzer. No further image processing with respect to contrast or brightness was done. The western blot was done in technical duplicate.



Supplementary Figure 5

Determination of the effective dose (ED_{50}) of the respective ASO for editing *GAPDH* in the respective 293 Flp-In T-REx ADAR cells

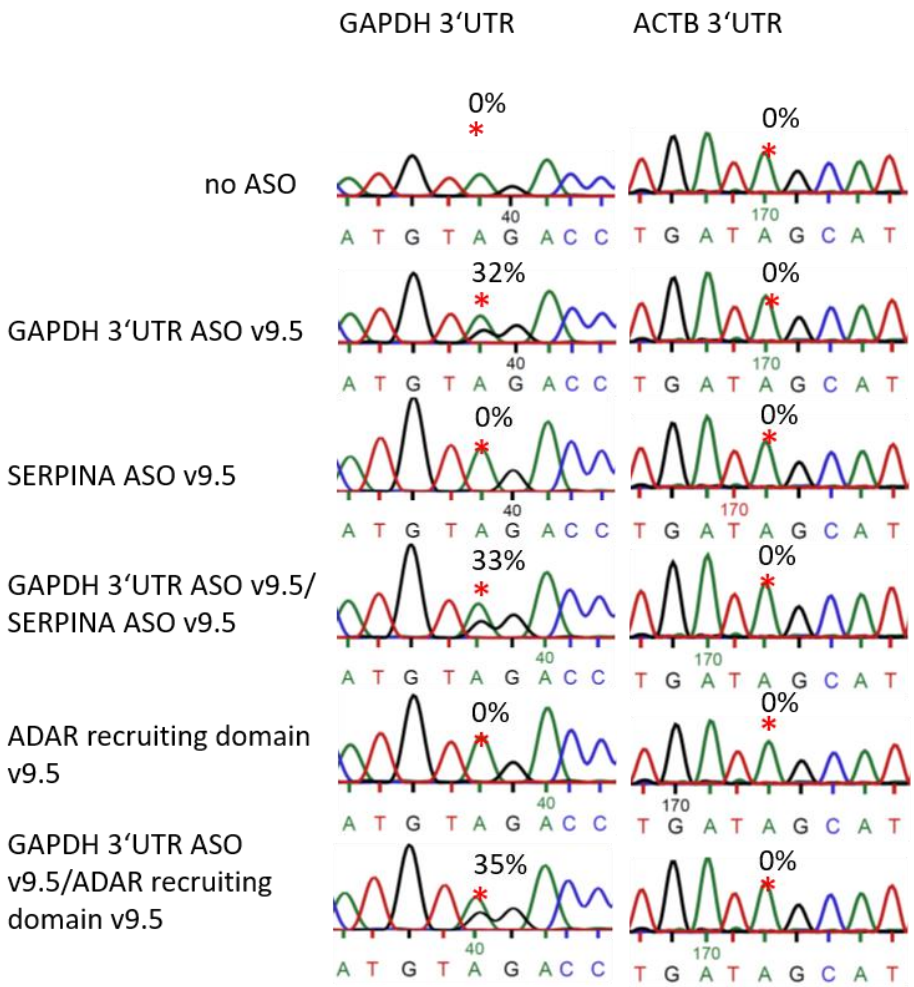
Shown is an experiment completely analog to that shown in the manuscript in Figure 2E, but in the indicated ADAR-expressing 293 Flp-In T-REx cell. Data are shown as the mean±SD, N=3 independent experiments.



Supplementary Figure 6

Effect of cotransfection of a nontargeting ASO v9.5 or the chemically stabilized ADAR-recruiting domain v9.5 alone on the *GAPDH* 3′-UTR editing with ASO v9.5 in ADAR1p150-expressing 293 Flp-In T-REx cells

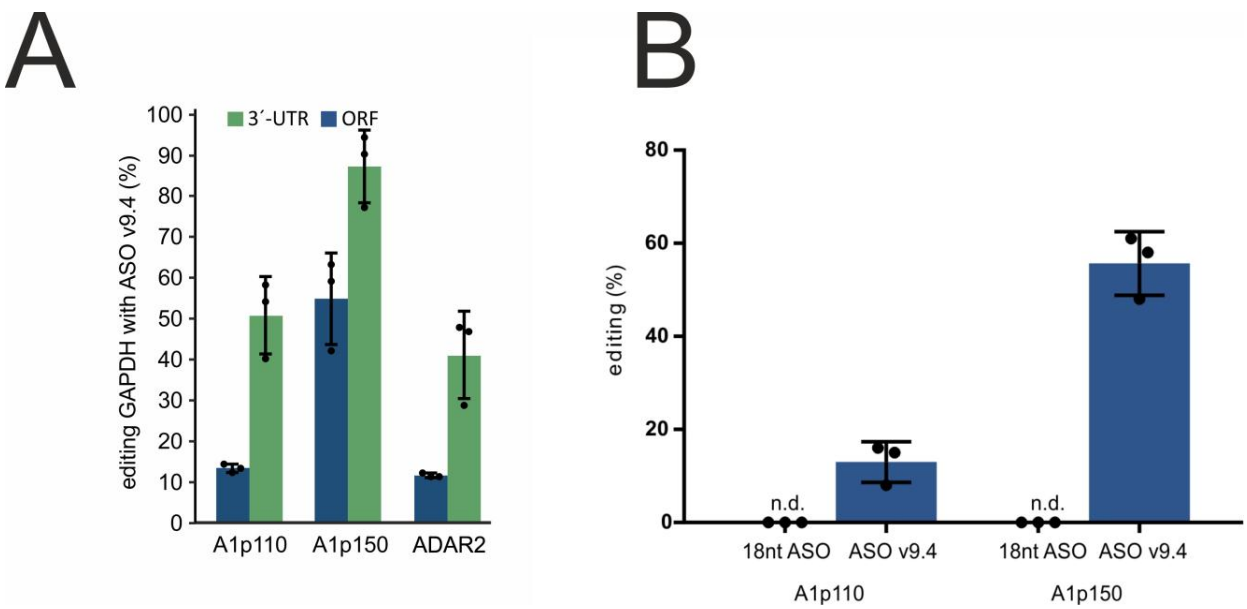
This is an additional control experiment. The on-target is the 5´-UAG codon in the 3´-UTR of GAPDH. A surveyed potential off-target is the 5´-UAG site in the 3´-UTR of ACTB. SERPINA ASO v9.5 acts as a non-targeting control, as the target (SERPINA1) is not expressed in this cell line. Another control is the ADAR-recruiting domain v9.5. This is the isolated, chemically stabilized ADAR-recruiting domain lacking any specificity domain. An ASO v9.5 against the on-target was co-transfected with either the non-targeting control or the control lacking a specificity domain. On-target editing requires the presences of the matching ASO. The surveyed potential off-target (ACTB) was not edited to detectable level under any condition. The on-target yield was not perturbed by the presence of the non-targeting ASO or the ADAR-recruiting domain alone, suggesting that only the combination of matching specificity and ADAR-recruiting domain enables site-directed RNA editing. It further suggests that the natural editing capacity is not limiting the editing reaction. (5 pmol ASO/96 well have been used)



Supplementary Figure 7

Effect of cotransfection of a nontargeting ASO v9.5 or the chemically stabilized ADAR-recruiting domain v9.5 alone on the *GAPDH* 3′-UTR editing with ASO v9.5, but for the recruitment of endogenous ADAR in HeLa cells without IFN-α

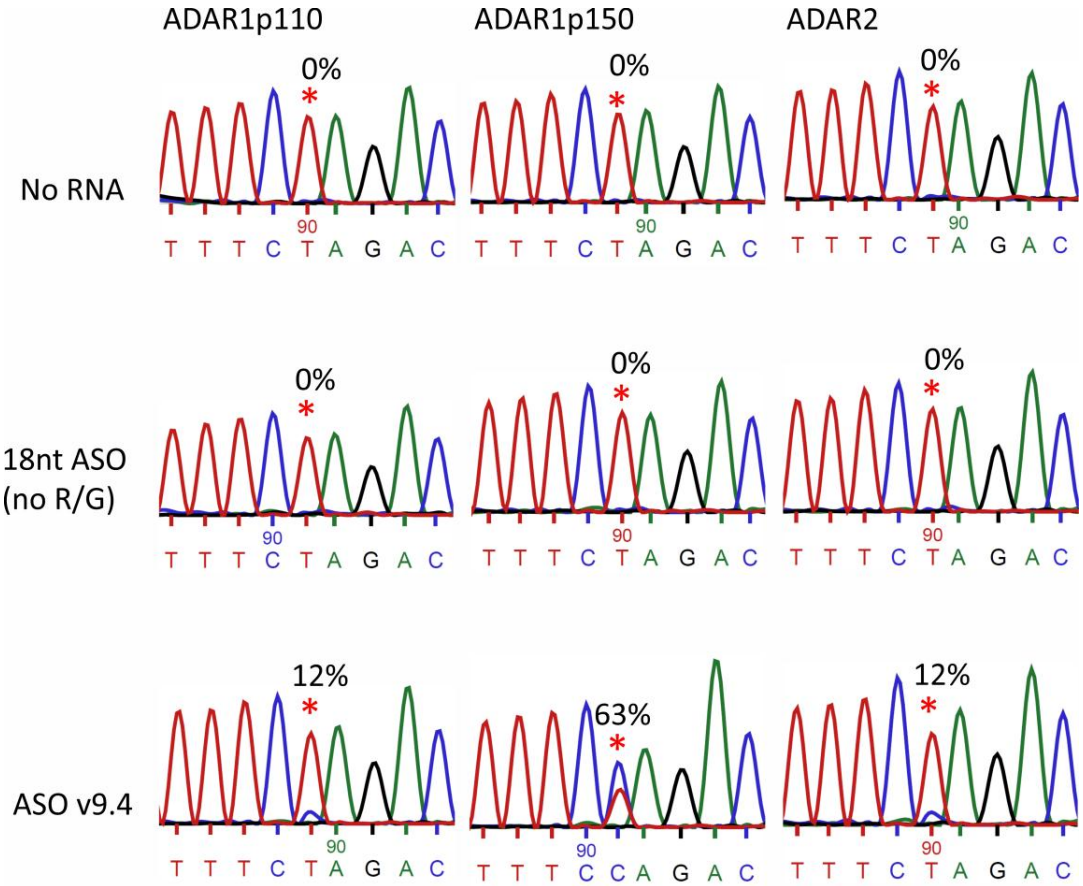
This control experiment is the exact copy of the expriment shown in the preceding Supplementary Figure but was carried out in HeLa cells, recruiting endogenous ADAR. Exactly the same results have been observed and the same conclusions can be drawn.



Supplementary Figure 8

Editing of 5′ UAG codons in the ORF of GAPDH versus 3′ UTR in ADAR-expressing 293 Flp-In T-REx cells

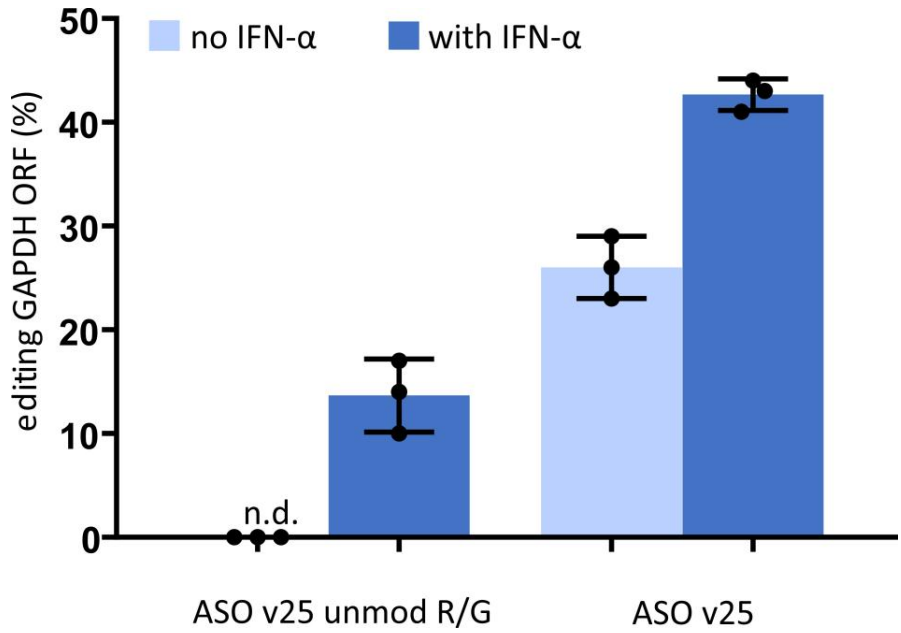
Editing of two different 5´-UAG codons in the ORF of GAPDH in 293 Flp-In T-REx ADAR cells (ORF #1 and #2). A) ORF site #2; here the comparison was made to the editing of the 5´-UAG codon in the 3´-UTR; and all three ADAR-expressing 293 Flp-In T-REx cell lines are included. B) The editing of the ORF site #1 was only tested in ADAR1-expressing Flp-In T-REx cell lines. The results are very similar. Further editing experiments, as shown in Figure 3B, target ORF site #1. Data in A) and B) are shown as the mean±SD, N=3 independent experiments. A1p110 = ADAR1p110; A1p150 = ADAR1p150.



Supplementary Figure 9

Sequencing traces for editing of a 5′ UAG site (ORF site 2) in the ORF of *GAPDH* in 293 Flp-In T-REx ADAR cells

Exemplary editing traces for the editings shown in Supplementary Figure 8A, but including additional controls (“No RNA” = empty transfection; “18nt ASO no R/G” = ASO lacking the ADAR-recruiting domain. Red asterisks indicate the editing sites. A reverse primer was used for sequencing.

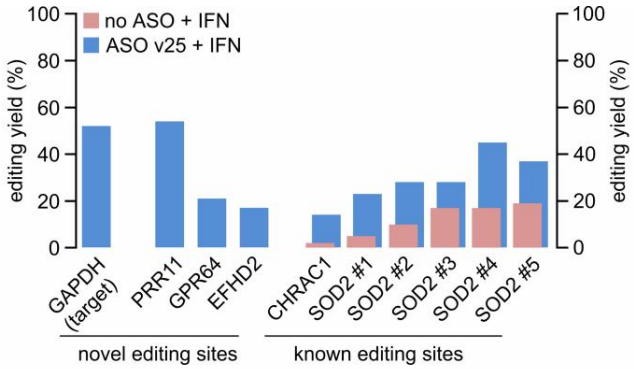


Supplementary Figure 10

Editing yields for the editing of a 5′ UAG codon in the ORF of *GAPDH* in HeLa cells with ASO v25 containing a chemically unmodified versus modified ADAR-recruiting domain

Here, an ASO v25 with a chemically unmodified ADAR-recruiting domain (unmod R/G), was compared to an ASO of the same sequence with addititional chemical modifiaction (all pyrimidine nucleotides in the ADAR-recruiting domain are backbone 2'-O-methylated). ASOs were transfected in HeLa cells. Data are shown as the mean±SD, N=3 independent experiments.

A



B

Edited site	Editing level		Sequence
	ASO v25+IFN	no gRNA+IFN	
GAPDH	53%	0%	CCAACTGCTTAGCACCCCTGGCCAAGGTCATCCATGACAA
PRR11	55%	0%	TGCCAACTCTGCCTCACTCCATGAGATGATCCATAACAAATACAGATAAAAACACCCAGCTGGGTGCAG
GPR64	21%	0%	TGCACCAAGATGTTTAGCTTTATACCTTGGCCACAGAGAGGGATGAACTG
EFHD2	17%	0%	GAGACACCGCTGCTTAGCACCCCCAGCCAGAACACCCTGAGGGTCTCGGG
CHARC1	14%	2%	ATTTTTGTATTTTTTAGAGACGGGGTTCACCATGTTGGCCAGGCTTGTCT
SOD2#1	23%	5%	ATGCTACCATGCCTGGCTACTTTTTGTATTTTTAGCAGAGACAGGGTTTCACCATGTTGGCCAGGGTG
SOD2#2	28%	10%	TGGCTCAGCCTCCCAAAGTGCTGGGATTACAGGTGTGAGCCACTGCACCT
SOD2#3	28%	17%	GGCCAAGGCAGGCACTTGAAGGAGTTCAAGACCAGCGTGGCCAACGTGGTGAAACCCTGTCT
SOD2#4	45%	17%	GGCTCAGCCTCCCAAAGGTGCTGGGATTACAGGTGTGAGCCACTGCACCT
SOD2#5	37%	19%	TGATCCGTCTGGCTCAGCCTCCCAAAGTGCTGGGATTACAGGTGTGAGCC

C

```
GAPDH      CCAACTGCTTAGCACCCCTGGCCAAGGTCATCCATGACAA
PRR11   TGCCAACTCTGCCTCACTCCATG---AGATGATCCATAACAAATACAGATAAAAACACCCAGCTGGGTGCAG

GAPDH        CCAACTGCTTAGCAC----CCCTGGCCAAGGTCATCCATG-ACAA
GPR64   TGCACCAAGATGTTTAGCTTTATACCTTGGCCACAGAGAGGGATGAACTG

GAPDH        CCAACTGCTTAGCACCCCTGGCCAAGGTCATCCATGACAA
EFHD2   GAGACACCGCTGCTTAGCACCCCCAGCC-AGAACACCC-TGAGGGTCTCGGG

GAPDH             CCAACTGCTTAG----------CACCC--CTGGCCAAGGTCATCCATGACAA
CHARC1       ATTTTTGTATTTTTTAGAGACGGGGTTCACCATGTTGGCCAGGCTTGTCT

GAPDH              CCAACTGC--TTAGCACCCCTGGCCAAG---------------GTCATCCATGACAA
SOD2#1  ATGCTACCATGCCTGGCTACTTTTTGTATTTTTAGCAGAGACAGGGTTTCACCATGTTGGCCAGGGTG

GAPDH   CCAACTGCTTAGCACCCC------------TGGCCAAGGTCATCCATGACAA
SOD2#2      TGGCTCAGCCTCCCAAAGTGCTGGGATTACAGGTGTGAGCCACTGCACCT

GAPDH      CCA--------ACT--------GCTTAGCACC--CCTGGCCAAGGTCAT----CCATGACAA
SOD2#3  GGCCAAGGCAGGCACTTGAAGGAGTTCAAGACCAGCGTGGCCAACGTGGTGAAACCCTGTCT

GAPDH   CCAACTGCTTAGCACCCC-------------TGGCCAAGGTCATCCATGACAA
SOD2#4       GGCTCAGCCTCCCAAAGGTGCTGGGATTACAGGTGTGAGCCACTGCACCT

GAPDH   CCAACTGCTTAGCACCCCTGGCCAAGGT----------CATCCATGACAA
SOD2#5  TGATCCGTCTGGCTCAGCCTCCCAAAGTGCTGGGATTACAGGTGTGAGCC
```

D

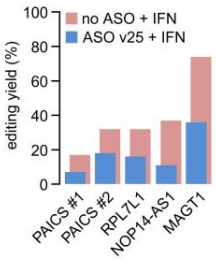
```
GAPDH                CCAA--------------------------------CTGCTTAGCACCCC-----------
SOD2 nt2100ff .  AGGAAACTCAAGATTCTCCTTTATTTCTGTGCTTGTGGGAATCCCCCTGGCACACCCCAAAGAGGGG

GAPDH            ------------------------------------TGGCCAAG-GTCATCCATGACAA
SOD2 nt2100ff .  TCCCTGCTCCGTCTCACAGGGATCTTTTTGTATATTTGGCTTAGCATCATACA------
```

Supplementary Figure 11

Analysis of off-target editing sites with increased editing yield upon ASO treatment

A) Besides the targeted site in GAPDH, 9 off-target editing sites were identified in ASO + IFN-α-treated cells compared to the control (no ASO + IFN-α). Six of them (CHARC1, SOD2 #1-#5) were known editing sites and found to be already edited in the control, N=2 independent experiments. **B)** and **C)** The regions around the off-target sites were aligned to the ASO-interacting region (40 nt) of the GAPDH transcript using MUSCLE (ebi.ac.uk/tools/msa/muscle/). The red A indicates the edited site and nucleotides matching to the target sequence of the ASO in GAPDH are highlighted in turquois. The sequence alignment suggests that the editing at the three novel editing sites (PRR11, GPR64, EFHD2) is clearly induced by misguiding through the ASO. Notably, the strongest off-target (PRR11) might be controllable by further chemical modification of the specificity domain of the ASO. Four of the nine off-target sites (SOD2 #2-5) lack any strong homology to target region, also the edited codon is different from 5´-UAG. This makes it very unlikely that the off-target editing at such sites was induced by the ASO via direct binding to the off-target site, also because those sites all reside in secondary RNA structure (Alu elements). However, we found a potential ASO binding site in the 3´-UTR of SOD2 (panel **D)** around nt 2100ff. (refering to NM_000636.3) that resides around 300 nt 5´ to the first Alu element (nt 2380-2670) and around 1300 nt 5´ to the second Alu element (nt 3400-3525). Since all SOD2 off-target sites reside in the two Alu elements one could imagine an ASO-dependent induction of the editing by either increase of the local ADAR concentration or by assisting the formation of an editable RNA secondary structure in the Alu element.

A



B

Edited site	Editing level		Sequence
	ASO v25+IFN	no ASO+IFN	
GAPDH	53%	0%	CCAACTGCTTAGCACCCCTGGCCAAGGTCATCCATGACAA
PAICS#1	7%	17%	TAGCTGGGCACAGTGGCTCACACCTACGATCACAGCACTTCGGGAGGCTGAAGCAAGCAGATC
PAICS#2	18%	32%	TGACAGGCACCTGTAATCCCAGCTACTCGGGAGGCTGAGGCATGAGAATTGAA
RPL7L1	16%	32%	AGAATTGCTTGAACCCAGGGAGGCAGAGGTTGCAGTGAGCCGAGATCACCATTGCACTCCACCTGGGCAACAAAAG
NOP14-AS1	11%	37%	ACCAGCCTGGGCAACACAGGAAGACCCCGTCTCTACCAAAAAAACATAAAAATTAGCCAAGTGTGGTGGCAC
MAGT1	36%	74%	GAGGCCAGGAGTTAGAGACCAGCCTGGCCAACAAGGCAAAACCCCGTCTCTACTAAAATATGAAAATTAGCTGG

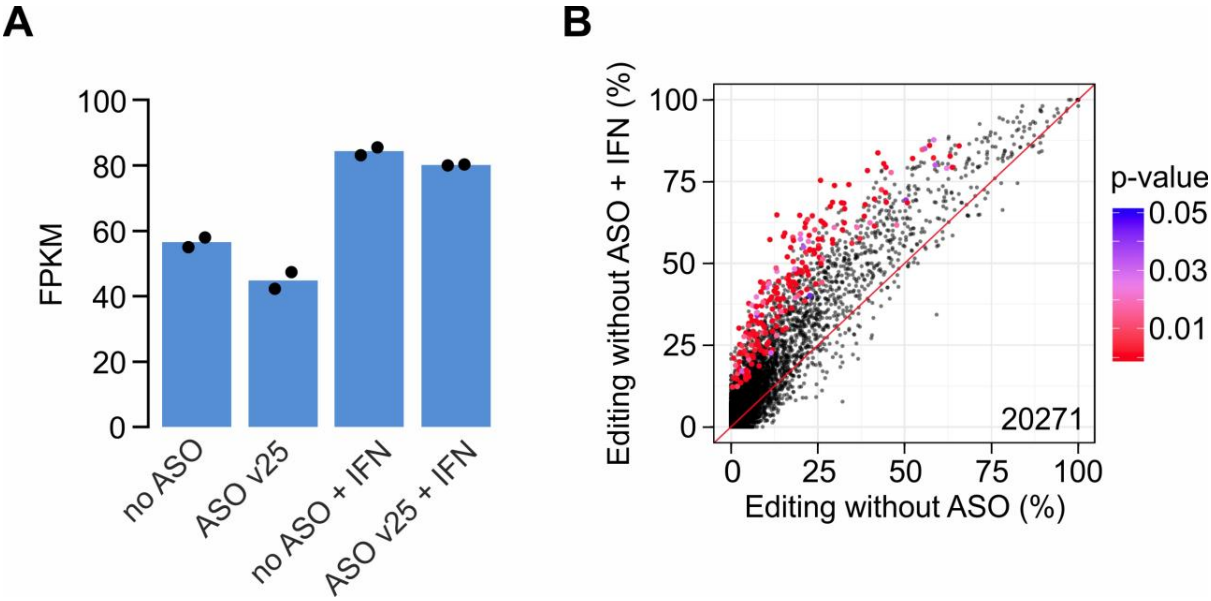
C

```
GAPDH            CCAACTGCTTAGCACCCCTGGCCAAGGTCATCCATGA--------CAA
PAICS#1   TAGCTGGGCACAGTGGCTCACACCTACGATCACAGCACTTCGGGAGGCTGAAGCAAGCAGATC

GAPDH                     CCAACTGCTTAGCACCCCTGGCCAAGGTCATCCATGACAA
PAICS#2   TGACAGGCACCTGTAATCCCAGCTACTCGGGA-----GGCTGAGG-----CATGAGAATTGAA

GAPDH     CCAACTGCTTAGCACCCCTG-------------------GCCAAGGTCATCCATG----------------ACAA
RPL7L1    AGAATTGCTT-GAACCCAGGGAGGCAGAGGTTGCAGTGAGCCGAGATCACCATTGCACTCCACCTGGGCAACAAAAG

GAPDH      CCAA-CTGCTTAGCAC-------CCCTGGCC----------------AAGGTCATCCA-------TGACAA
NOP14-AS1 ACCAGCCTGGGCAACACAGGAAGACCCCGTCTCTACCAAAAAAACATAAAAATTAGCCAAGTGTGGTGGCAC

GAPDH         CCAACTGCTTAGCACC--CCTGGCCAA----------------------GGTCATCCATGACAA
MAGT1     GAGGCCAGGAGTTAGAGACCAGCCTGGCCAACAAGGCAAAACCCCGTCTCTACTAAAATATGAAAATTAGCTGG
```

Supplementary Figure 12

Analysis of off-target editing sites with attenuated editing upon ASO treatment

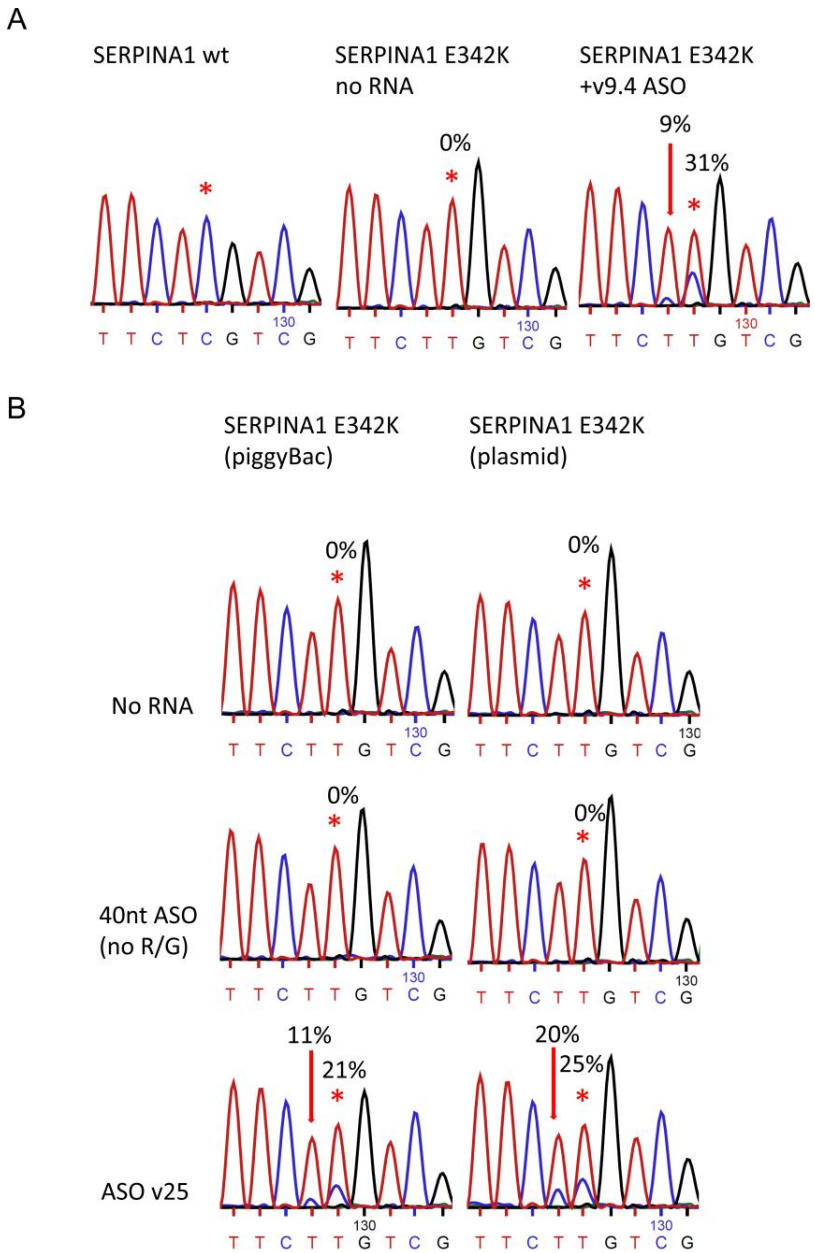
A) Five editing sites, all located in Alu sequences, were found to be significantly less edited in ASO-transfected, IFN-α-treated cells compared to the control lacking ASO transfection (but treated with IFN-α), N=2 independent experiments. **B)** and **C)** The regions around the off-target sites were aligned to the ASO-interacting region (40 nt) of the GAPDH transcript using MUSCLE (ebi.ac.uk/tools/msa/muscle/). The red A indicates the respective edited site and nucleotides matching with the ASO target sequence on GAPDH are highlighted in turquois. For the most strongly affected site (MAGT1), but also for the other four sites, the ASO seems to be able to bind tightly in proximity to the respective editing sites and therefore interrupt the dsRNA secondary structure of the Alu repeat, which is required for editing. This suggests that the attenuated editing found at those sites is caused by direct interaction of the ASO with the off-target transcript and is not due to a global sequestering of the ADAR enzyme by the ASO.



Supplementary Figure 13

Effect of IFN-α and ASO treatment on ADAR1 expression and the natural editing homeostasis

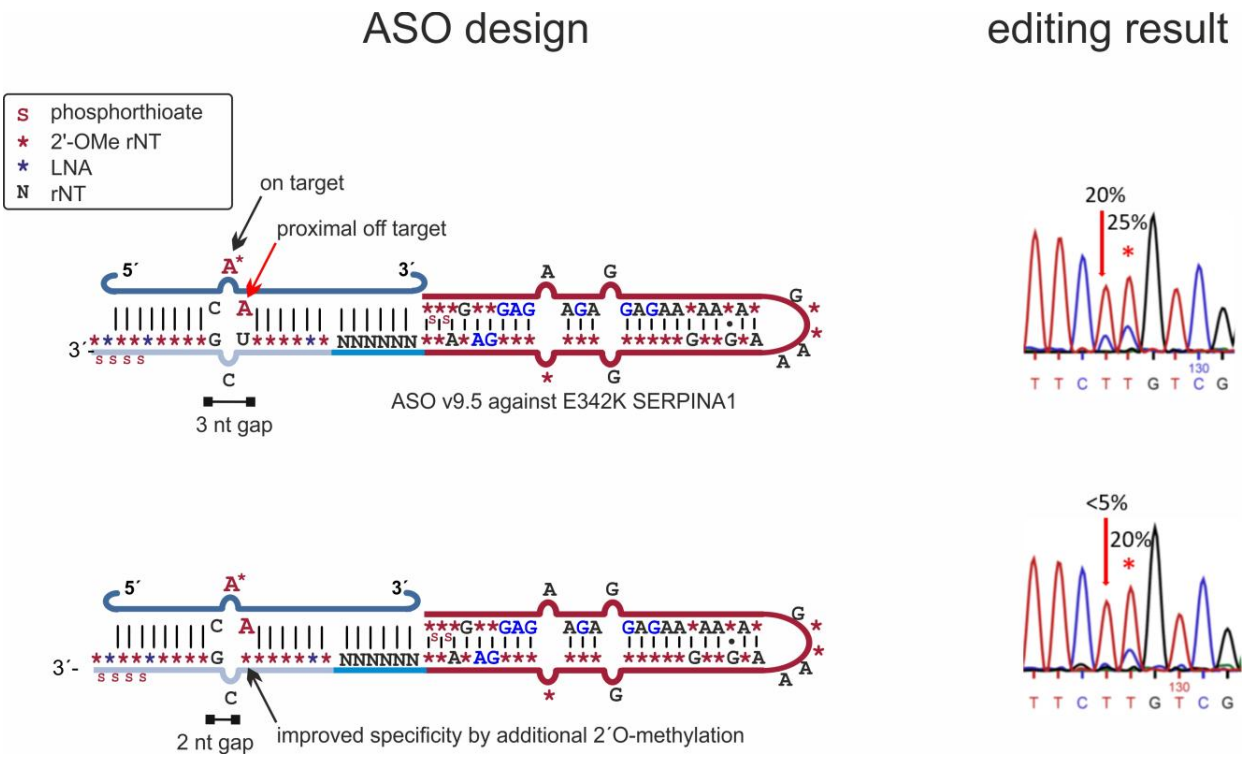
A) FPKM values describing overall ADAR1 (p110+p150) expression following IFN-α treatment and ASO administration. IFN-α treatment induced ADAR1 expression in HeLa cells in a similar manner independent of ASO transfection. N=2 independent experiments. **B)** Analysis of significantly differently edited sites after IFN-α treatment in HeLa cells (no ASO transfection). Editing appears globally increased following IFN-α treatment. Significance of 20271 edited sites was tested using Fisher's exact test (two-sided, $p<0.01$, $N\geq50$); 116 sites were detected as significantly differently edited. The NGS expriment was done in independent duplicate.



Supplementary Figure 14

Sequencing traces of editing the PiZZ mutation in *SERPINA1*, showing on- and off-target editing in the A-rich 5′ CAA codon

Exemplary sequencing traces of the 3 experimental conditions shown in Figure 3E of the manuscript. Red arrows indicate off-target, red asterisks indicate on-target editing sites, a reverse primer was used for sequencing. Shown are additional controls for empty transfection, and for transfecting an ASO lacking the ADAR-recruiting domain (no R/G). A) Editing in ADAR1p150-expressing 293 Flp-In T-REx cells; B) Editing of SERPINA1 PiZZ in HeLa cells expressing SERPINA1 PiZZ either genomically integrated (piggyBac) or transiently overexpressed (plasmid). In particular ASO v25 gave substantial off-target editing with the proximal adenosine in the targeted 5´-CAA codon.



Supplementary Figure 15

Improvement of editing specificity in the ASO:mRNA hybrid

Editing of the 5´-CAA codon to restore the E342K mutation in SERPINA1 (a reverse primer was used here!) comes along with off target editing at the nearest neighboring adenosine (see also preceding Supplementary Figure, panel B). To reduce proximal off-target editing, the 3 nt gap in the modification pattern of the ASO was reduced to a 2 nt gap by putting an additional chemical modification (2´-O-methyl uridine) opposite the off-target nucleotide. Two representative sequencing traces were selected from three very similar replicates which show that the additional chemical modification strongly reduces the proximal off-target edit while only modestly influencing the on target editing. (editing was performed in HeLa cells with SERPINA1 PiZZ cDNA overexpressed from a plasmid, 5 pmol/96 well ASO was transfected)

Supplementary Notes and Tables

Supplementary Table 1: List of guideRNAs, ASOs and further oligonucleotides used in this study.

(N)=RNA base, [N]=2'-OMe RNA base, *=Phosphorothioate linkage, {N}=LNA base.

R/G guide RNAs expressed from plasmid	**5'-3' sequence**	**Figure**
Luciferase R/G-v1	(GUGGAAUAGUAUAACAAUAUGCUAAAUGUUGUUAUAGUAUCCCACGUGCAGC CAGCCGUCCUCUAGAGGGCCCUGAAGAGGGCCC)	SI1
Luciferase R/G-v4	(GUGGAAGAGGAGAACAAUAUGCUAAAUGUUGUUCUCGUCUCCCACGUGCAGC CAGCCGUCCUCUAGAGGGCCCUGAAGAGGGCCC)	SI1
Luciferase R/G-v9.4	(GUGGUCGAGAAGAGGAGAACAAUAUGCUAAAUGUUGUUCUCGUCUCCUCGAC CACGUGCAGCCAGCCGUCCUCUAGAGGGCCCUGAAGAGGGCCC)	SI1
Chemically synthesized ASOs	**5'-3' sequence**	**Figure**
ACTB 3'UTR 18nt	[GC AAU G] (CCA) [UC AC] [C*][U*][C*][C*][C] Propandiol	2A
ACTB 3'UTR ASO v1	(GGUGA AUAGUAUAAC AAUAUGCUAA AUGUUGUUAU AGUAUCCACC) [GC AAU G] (CCA) [UC AC] [C*][U*][C*][C*][C] Propandiol	1C,2A
ACTB 3'UTR ASO v4	(GGUGAAG AGGAGAACAA UAUGCUAAAU GUUGUUCUCG UCUCCACC)[GC AAU G](CCA) [UC AC][C*][U*][C*][C*][C] Propandiol	1C,2A
ACTB 3'UTR ASO v9.4	(GGU GUC GAG AAG AGG AGA AC AAU AUG CUA AAU GUU GUU CUCGUC UCC UCG ACA CC) [GC AAU G] (CCA) [UC AC] [C*][U*][C*][C*][C] Propandiol	1C,2A
GAPDH 3'UTR 18nt	[AG GGG U] (CCA) [CA UG] [G*][C*][A*][A*][C] Propandiol	2A, SI2
GAPDH 3'UTR ASO v1	(GGUGA AUAGUAUAAC AAUAUGCUAA AUGUUGUUAU AGUAUCCACC) [AG GGG U] (CCA) [CA UG] [G*][C*][A*][A*][C] Propandiol	1C,2A, SI2
GAPDH 3'UTR ASO v4	(GGUGAAG AGGAGAACAA UAUGCUAAAU GUUGUUCUCG UCUCCACC)[AG GGG U](CCA)[CA UG][G*][C*][A*][A*][C] Propandiol	1C,2A, SI2
GAPDH 3'UTR ASO v9.4	(GGU GUC GAG AAG AGG AGA AC AAU AUG CUA AAU GUU GUU CUCGUC UCC UCG ACA CC) [AG GGG U] (CCA) [CA UG] [G*][C*][A*][A*][C] Propandiol	1C,2A,2B, SI2, SI5
GAPDH 3'UTR ASO v9.5	[G*][G*][U] (G)[U][C] (GAG AAG AGG AGA A)[C] (AA)[U] (A)[U](G) [C][U](A AA)[U] (G)[U][C](G)[UU](G)[UUCUC](G)[UCUCCUC](G A)[C](A) [CCAGGGGU] (CCA) [CAUG][G*][C*][A*] [A*][C]	2B, 2C, 2E, 2F, 2G, 2H, SI5, SI6, SI7, SI8A
GAPDH ORF1 ASO 18nt	[GGG GUG] (CCA) [AG CA] [G*][U*][U*][G*][G] Propandiol	SI8B
GAPDH ORF1 ASO v9.4	(GGU GUC GAG AAG AGG AGA AC AAU AUG CUA AAU GUU GUU CUCGUC UCC UCG ACA CC)[GGG GUG](CCA)[AG CA] [G*][U*][U*][G*][G] Propandiol	SI8B
GAPDH ORF2 ASO 18nt	[GGG GUG](CCA)[AG CA] [G*][U*][U*][G*][G] Propandiol	SI9
GAPDH ORF2 ASO v9.4	(GGU GUC GAG AAG AGG AGA AC AAU AUG CUA AAU GUU GUU CUCGUC UCC UCG ACA CC)[GU UUU U] (CCA) [GA CG] [G*][C*][A*][G*][G] Propandiol	SI8A, SI9
GAPDH ORF1 ASO v25	[G]*[G]*[U] (G)[U][C] (GAG AAG AGG AGA A)[C] (AA) [U] (A)[U](G) [C][U](A AA)[U] (G)[U][U] (G)[U][U] [C][U][C](G)[U][C] [U][C][C] [U][C](G A)[C](A) [C][C] (UUGUCAUGGAUGACCUU GGCCA) [G] {G} [GG UG] (CCA) [AGCA] {G*}[U*][U*]{G*}[G] Aminolinker	3B,3C, SI10-13
GAPDH ORF1 ASO R/G unmod v25	[G]*[G]*[U] (GUCGAG AAG AGG AGA ACAAUAUGCUA AAUGUUGUUCUCGUCUCCUCG ACACC UUGUCAUGGAUGACCUU GGCCA) [G] {G} [GG UG] (CCA) [AGCA] {G*}[U*][U*]{G*}[G] Aminolinker	SI10
SERPINA ASO v9.4	(GGU GUC GAG AAG AGG AGA AC AAU AUG CUA AAU GUU GUU CUCGUC UCC UCG ACA CC) [CCU UUC] (UCG) [UCG A] [U*][G*][G*][U*][C] Propandiol	3E, SI6, SI7, SI14A
SERPINA ASO 40nt	(CAUGGCCCCAGCAGCUUCAGUC) [C] {C}[UUUC] (UCG) [UCGA]{T*}[G*] [G*] {T*} [C] Aminolinker	SI14B
SERPINA ASO v25	[G*] [G*] [U] (G)[U][C] (GAG AAG AGG AGA A)[C] (AA) [U] (A)[U](G) [C][U](A AA)[U] (G)[U][U] (G)[U][U] [C][U][C](G)[U][C] [U][C][C] [U][C](G A C A C C CAUGGCCCCAGCAGCUUCAGUC) [C] {C}[UUUC] (UCG) [UCGA]{T*}[G*] [G*] {T*} [C] Aminolinker	3E, SI14B, SI15
STAT1 ASO v25	[G*] [G*] [U] (G)[U][C] (GAG AAG AGG AGA A)[C] (AA) [U] (A)[U](G) [C][U](A AA)[U] (G)[U][U] (G)[U][U] [C][U][C](G)[U][C][U][C][C] [U][C](GACACCCA GACACAGAAAUCAACUCAGU) [C] {T} [UGAU] (ACA) [UCCA] {G*} [U*] [U*] {C*}[C] Aminolinker	3D
GAPDH 3'UTR unmod ASO v1	(GGUGA AUAGUAUAAC AAUAUGCUAA AUGUUGUUAU AGUAUCCACC AG GGG UCCACA UG GCAAC)	SI2, SI3
GAPDH 3'UTR unmod ASO v4	(GGUGAAG AGGAGAACAA UAUGCUAAAU GUUGUUCUCG UCUCCACCAG GGG UCCACA UGGCAAC)	SI2, SI3
GAPDH 3'UTR unmod ASO v9.4	(GGU GUC GAG AAG AGG AGA AC AAU AUG CUA AAU GUU GUU CUCGUC UCC UCG ACA CCAG GGG U CCACA UGGCAAC)	SI2, SI3
SERPINA ASO v25 2nt gap also called ASO v25.1 in Fig. 3E	[G*] [G*] [U] (G)[U][C] (GAG AAG AGG AGA A)[C] (AA) [U] (A)[U](G) [C][U](A AA)[U] (G)[U][U] (G)[U][U] [C][U][C](G)[U][C] [U][C][C] [U][C](G A C A C C CAUGGCCCCAGCAGCUUCAGUC) [C] {C}[UUUCU] (CG) [UCGA]{T*}[G*] [G*] {T*} [C] Aminolinker	3E, SI15
Sense guideRNAs for RT PCR	**5'-3' sequence**	**Figure**
GAPDH sense	(GGACCAACUGCUUGGCACCCCUGGCCAAGGUCAUCCAUGACAACUUUGGUAUCGUGGAAGGACC)	3B, 3C

STAT1 sense	(GGGAACUGGAUCUAUCAAGACUGAGUUGAUUUCUGUGUCUGAAGUGUAAGUGAACACAGAA)	3D
SERPINA1 sense	(GGACCATCGACGAGAAAGGGACTGAAGCTGCTGGGGCCATGTTTTTAGAGGCCATACCCAT)	3E,SI14B, SI15

Supplementary Table 2: List of the antibodies used to generate the western blot illustrated in figure 2D.

Antibody	Target Protein	Produced in	Immunoglobulin Class	Dilution used	Supplier	Order #	Against	Validation
ADAR1 antibody	α-ADAR1	Mouse	monoclonal IgG	1:1000	Santa Cruz	sc-73408	amino acids 440-826 corresponding to the middle region of ADAR1 of human origin	Validated in our lab via siRNA KO and Western Blot PMID: # 28669490 PMID: # 28278381 PMID: # 27573237
ADAR2 antibody	α-ADAR2	Mouse	monoclonal IgG	1:1000	Santa Cruz	sc-73409	N-terminal region corresponding to amino acids 2-179 of ADAR2 of human origin	Validated in our lab via overexpression and Western Blot PMID: # 26601943 PMID: # 24345557 PMID: # 27907896
Clone AC-15	α-Beta-Actin	Mouse	monoclonal IgG	1:40.000	Sigma Aldrich	A5441	Actin N-terminal peptide, Ac-Asp-Asp-Asp-Ile-Ala-Ala-Leu-Val-Ile-Asp-Asn-Gly-Ser-Gly-Lys	PMID: # 15809369 PMID: # 15048076 PMID: # 21217779

Supplementary Note 1. List of gene sequences and target sequences

Sequence of dual Luciferase Renilla 2A Firefly W417X reporter cDNA with chosen editing site (Firefly-Luciferase W417X, yellow).

```
                10        20        30        40        50        60
1         ATGACTTCGAAAGTTTATGATCCAGAACAAAGGAAACGGATGATAACTGGTCCGCAGTGG
1          M  T  S  K  V  Y  D  P  E  Q  R  K  R  M  I  T  G  P  Q  W

                70        80        90       100       110       120
61        TGGGCCAGATGTAAACAAATGAATGTTCTTGATTCATTTATTAATTATTATGATTCAGAA
21         W  A  R  C  K  Q  M  N  V  L  D  S  F  I  N  Y  Y  D  S  E

               130       140       150       160       170       180
121       AAACATGCAGAAAATGCTGTTATTTTTTTACATGGTAACGCGGCCTCTTCTTATTTATGG
41         K  H  A  E  N  A  V  I  F  L  H  G  N  A  A  S  S  Y  L  W

               190       200       210       220       230       240
181       CGACATGTTGTGCCACATATTGAGCCAGTAGCGCGGTGTATTATACCAGACCTTATTGGT
61         R  H  V  V  P  H  I  E  P  V  A  R  C  I  I  P  D  L  I  G

               250       260       270       280       290       300
241       ATGGGCAAATCAGGCAAATCTGGTAATGGTTCTTATAGGTTACTTGATCATTACAAATAT
81         M  G  K  S  G  K  S  G  N  G  S  Y  R  L  L  D  H  Y  K  Y

               310       320       330       340       350       360
301       CTTACTGCATGGTTTGAACTTCTTAATTTACCAAAGAAGATCATTTTTGTCGGCCATGAT
101        L  T  A  W  F  E  L  L  N  L  P  K  K  I  I  F  V  G  H  D

               370       380       390       400       410       420
361       TGGGGTGCTTGTTTGGCATTTCATTATAGCTATGAGCATCAAGATAAGATCAAAGCAATA
121        W  G  A  C  L  A  F  H  Y  S  Y  E  H  Q  D  K  I  K  A  I

               430       440       450       460       470       480
421       GTTCACGCTGAAAGTGTAGTAGATGTGATTGAATCATGGGATGAATGGCCTGATATTGAA
141        V  H  A  E  S  V  V  D  V  I  E  S  W  D  E  W  P  D  I  E

               490       500       510       520       530       540
481       GAAGATATTGCGTTGATCAAATCTGAAGAAGGAGAAAAAATGGTTTTGGAGAATAACTTC
161        E  D  I  A  L  I  K  S  E  E  G  E  K  M  V  L  E  N  N  F

               550       560       570       580       590       600
541       TTCGTGGAAACCATGTTGCCATCAAAAATCATGAGAAAGTTAGAACCAGAAGAATTTGCA
181        F  V  E  T  M  L  P  S  K  I  M  R  K  L  E  P  E  E  F  A

               610       620       630       640       650       660
601       GCATATCTTGAACCATTCAAAGAGAAAGGTGAAGTTCGTCGTCCAACATTATCATGGCCT
201        A  Y  L  E  P  F  K  E  K  G  E  V  R  R  P  T  L  S  W  P

               670       680       690       700       710       720
661       CGTGAAATCCCGTTAGTAAAAGGTGGTAAACCTGACGTTGTACAAATTGTTAGGAATTAT
221        R  E  I  P  L  V  K  G  G  K  P  D  V  V  Q  I  V  R  N  Y

               730       740       750       760       770       780
721       AATGCTTATCTACGTGCAAGTGATGATTTACCAAAAATGTTTATTGAATCGGACCCAGGA
241        N  A  Y  L  R  A  S  D  D  L  P  K  M  F  I  E  S  D  P  G

               790       800       810       820       830       840
781       TTCTTTTCCAATGCTATTGTTGAAGGTGCCAAGAAGTTTCCTAATACTGAATTTGTCAAA
261        F  F  S  N  A  I  V  E  G  A  K  K  F  P  N  T  E  F  V  K

               850       860       870       880       890       900
841       GTAAAAGGTCTTCATTTTTCGCAAGAAGATGCACCTGATGAAATGGGAAAATATATCAAA
281        V  K  G  L  H  F  S  Q  E  D  A  P  D  E  M  G  K  Y  I  K

               910       920       930       940       950       960
901       TCGTTCGTTGAGCGAGTTCTCAAAAATGAACAAGGAAGCGGAGCTACTAACTTCAGCCTG
301        S  F  V  E  R  V  L  K  N  E  Q  G  S  G  A  T  N  F  S  L

               970       980       990      1000      1010      1020
961       CTGAAGCAGGCTGGAGACGTGGAGGAGAACCCTGGACCTATGGAAGATGCCAAAAACATT
```

```
321       L  K  Q  A  G  D  V  E  E  N  P  G  P  M  E  D  A  K  N  I

               1030      1040      1050      1060      1070      1080
1021      AAGAAGGGCCCAGCGCCATTCTACCCACTCGAAGACGGGACCGCCGGCGAGCAGCTGCAC
341       K  K  G  P  A  P  F  Y  P  L  E  D  G  T  A  G  E  Q  L  H

               1090      1100      1110      1120      1130      1140
1081      AAAGCCATGAAGCGCTACGCCCTGGTGCCCGGCACCATCGCCTTTACCGACGCACATATC
361       K  A  M  K  R  Y  A  L  V  P  G  T  I  A  F  T  D  A  H  I

               1150      1160      1170      1180      1190      1200
1141      GAGGTGGACATTACCTACGCCGAGTACTTCGAGATGAGCGTTCGGCTGGCAGAAGCTATG
381       E  V  D  I  T  Y  A  E  Y  F  E  M  S  V  R  L  A  E  A  M

               1210      1220      1230      1240      1250      1260
1201      AAGCGCTATGGGCTGAATACAAACCATCGGATCGTGGTGTGCAGCGAGAATAGCTTGCAG
401       K  R  Y  G  L  N  T  N  H  R  I  V  V  C  S  E  N  S  L  Q

               1270      1280      1290      1300      1310      1320
1261      TTCTTCATGCCCGTGTTGGGTGCCCTGTTCATCGGTGTGGCTGTGGCCCCAGCTAACGAC
421       F  F  M  P  V  L  G  A  L  F  I  G  V  A  V  A  P  A  N  D

               1330      1340      1350      1360      1370      1380
1321      ATCTACAACGAGCGCGAGCTGCTGAACAGCATGGGCATCAGCCAGCCCACCGTCGTATTC
441       I  Y  N  E  R  E  L  L  N  S  M  G  I  S  Q  P  T  V  V  F

               1390      1400      1410      1420      1430      1440
1381      GTGAGCAAGAAAGGGCTGCAAAAGATCCTCAACGTGCAAAAGAAGCTACCGATCATACAA
461       V  S  K  K  G  L  Q  K  I  L  N  V  Q  K  K  L  P  I  I  Q

               1450      1460      1470      1480      1490      1500
1441      AAGATCATCATCATGGATAGCAAGACCGACTACCAGGGCTTCCAAAGCATGTACACCTTC
481       K  I  I  I  M  D  S  K  T  D  Y  Q  G  F  Q  S  M  Y  T  F

               1510      1520      1530      1540      1550      1560
1501      GTGACTTCCCATTTGCCACCCGGCTTCAACGAGTACGACTTCGTGCCCGAGAGCTTCGAC
501       V  T  S  H  L  P  P  G  F  N  E  Y  D  F  V  P  E  S  F  D

               1570      1580      1590      1600      1610      1620
1561      CGGGACAAAACCATCGCCCTGATCATGAACAGTAGTGGCAGTACCGGATTGCCCAAGGGC
521       R  D  K  T  I  A  L  I  M  N  S  S  G  S  T  G  L  P  K  G

               1630      1640      1650      1660      1670      1680
1621      GTAGCCCTACCGCACCGCACCGCTTGTGTCCGATTCAGTCATGCCCGCGACCCCATCTTC
541       V  A  L  P  H  R  T  A  C  V  R  F  S  H  A  R  D  P  I  F

               1690      1700      1710      1720      1730      1740
1681      GGCAACCAGATCATCCCCGACACCGCTATCCTCAGCGTGGTGCCATTTCACCACGGCTTC
561       G  N  Q  I  I  P  D  T  A  I  L  S  V  V  P  F  H  H  G  F

               1750      1760      1770      1780      1790      1800
1741      GGCATGTTCACCACGCTGGGCTACTTGATCTGCGGCTTTCGGGTCGTGCTCATGTACCGC
581       G  M  F  T  T  L  G  Y  L  I  C  G  F  R  V  V  L  M  Y  R

               1810      1820      1830      1840      1850      1860
1801      TTCGAGGAGGAGCTATTCTTGCGCAGCTTGCAAGACTATAAGATTCAATCTGCCCTGCTG
601       F  E  E  E  L  F  L  R  S  L  Q  D  Y  K  I  Q  S  A  L  L

               1870      1880      1890      1900      1910      1920
1861      GTGCCCACACTATTTAGCTTCTTCGCTAAGAGCACTCTCATCGACAAGTACGACCTAAGC
621       V  P  T  L  F  S  F  F  A  K  S  T  L  I  D  K  Y  D  L  S

               1930      1940      1950      1960      1970      1980
1921      AACTTGCACGAGATCGCCAGCGGCGGGGCGCCGCTCAGCAAGGAGGTAGGTGAGGCCGTG
641       N  L  H  E  I  A  S  G  G  A  P  L  S  K  E  V  G  E  A  V

               1990      2000      2010      2020      2030      2040
1981      GCCAAACGCTTCCACCTACCAGGCATCCGCCAGGGCTACGGCCTGACAGAAACAACCAGC
661       A  K  R  F  H  L  P  G  I  R  Q  G  Y  G  L  T  E  T  T  S

               2050      2060      2070      2080      2090      2100
2041      GCCATTCTGATCACCCCCGAAGGGGACGACAAGCCTGGCGCAGTAGGCAAGGTGGTGCCC
681       A  I  L  I  T  P  E  G  D  D  K  P  G  A  V  G  K  V  V  P
```

```
              2110      2120      2130      2140      2150      2160
2101      TTCTTCGAGGCTAAGGTGGTGGACTTGGACACCGGTAAGACACTGGGTGTGAACCAGCGC
701        F  F  E  A  K  V  V  D  L  D  T  G  K  T  L  G  V  N  Q  R

              2170      2180      2190      2200      2210      2220
2161      GGCGAGCTGTGCGTCCGTGGCCCCATGATCATGAGCGGCTACGTTAACAACCCCGAGGCT
721        G  E  L  C  V  R  G  P  M  I  M  S  G  Y  V  N  N  P  E  A

              2230      2240      2250      2260      2270      2280
2221      ACAAACGCTCTCATCGACAAGGACGGCTAGCTGCACAGCGGCGACATCGCCTACTGGGAC
741        T  N  A  L  I  D  K  D  G  *  L  H  S  G  D  I  A  Y  W  D

              2290      2300      2310      2320      2330      2340
2281      GAGGACGAGCACTTCTTCATCGTGGACCGGCTGAAGAGCCTGATCAAATACAAGGGCTAC
761        E  D  E  H  F  F  I  V  D  R  L  K  S  L  I  K  Y  K  G  Y

              2350      2360      2370      2380      2390      2400
2341      CAGGTAGCCCCAGCCGAACTGGAGAGCATCCTGCTGCAACACCCCAACATCTTCGACGCC
781        Q  V  A  P  A  E  L  E  S  I  L  L  Q  H  P  N  I  F  D  A

              2410      2420      2430      2440      2450      2460
2401      GGGGTCGCCGGCCTGCCCGACGACGATGCCGGCGAGCTGCCCGCCGCAGTCGTCGTGCTG
801        G  V  A  G  L  P  D  D  D  A  G  E  L  P  A  A  V  V  V  L

              2470      2480      2490      2500      2510      2520
2461      GAACACGGTAAAACCATGACCGAGAAGGAGATCGTGGACTATGTGGCCAGCCAGGTCACA
821        E  H  G  K  T  M  T  E  K  E  I  V  D  Y  V  A  S  Q  V  T

              2530      2540      2550      2560      2570      2580
2521      ACCGCCAAGAAGCTGCGCGGTGGTGTTGTGTTCGTGGACGAGGTGCCTAAAGGACTGACC
841        T  A  K  K  L  R  G  G  V  V  F  V  D  E  V  P  K  G  L  T

              2590      2600      2610      2620      2630      2640
2581      GGCAAGTTGGACGCCCGCAAGATCCGCGAGATTCTCATTAAGGCCAAGAAGGGCGGCAAG
861        G  K  L  D  A  R  K  I  R  E  I  L  I  K  A  K  K  G  G  K

              2650
2641      ATCGCCGTGTAA
881        I  A  V  *
```

Sequence of GAPDH mRNA isoform 1 (NM_002046.5) with chosen editing sites (yellow).

```
              10        20        30        40        50        60
1         GCCTCAAGACCTTGGGCTGGGACTGGCTGAGCCTGGCGGGAGGCGGGGTCCGAGTCACCG
1
               70        80        90       100       110       120
61        CCTGCCGCCGCGCCCCCGGTTTCTATAAATTGAGCCCGCAGCCTCCCGCTTCGCTCTCTG
20
              130       140       150       160       170       180
121       CTCCTCCTGTTCGACAGTCAGCCGCATCTTCTTTTGCGTCGCCAGCCGAGCCACATCGCT
40
              190       200       210       220       230       240
181       CAGACACCATGGGGAAGGTGAAGGTCGGAGTCAACGGATTTGGTCGTATTGGGCGCCTGG
60                 M  G  K  V  K  V  G  V  N  G  F  G  R  I  G  R  L
              250       260       270       280       290       300
241       TCACCAGGGCTGCTTTTAACTCTGGTAAAGTGGATATTGTTGCCATCAATGACCCCTTCA
80        V  T  R  A  A  F  N  S  G  K  V  D  I  V  A  I  N  D  P  F
              310       320       330       340       350       360
301       TTGACCTCAACTACATGGTTTACATGTTCCAATATGATTCCACCCATGGCAAATTCCATG
100       I  D  L  N  Y  M  V  Y  M  F  Q  Y  D  S  T  H  G  K  F  H
              370       380       390       400       410       420
361       GCACCGTCAAGGCTGAGAACGGGAAGCTTGTCATCAATGGAAATCCCATCACCATCTTCC
120       G  T  V  K  A  E  N  G  K  L  V  I  N  G  N  P  I  T  I  F
              430       440       450       460       470       480
421       AGGAGCGAGATCCCTCCAAAATCAAGTGGGGCGATGCTGGCGCTGAGTACGTCGTGGAGT
140       Q  E  R  D  P  S  K  I  K  W  G  D  A  G  A  E  Y  V  V  E
              490       500       510       520       530       540
481       CCACTGGCGTCTTCACCACCATGGAGAAGGCTGGGGCTCATTTGCAGGGGGGAGCCAAAA
160       S  T  G  V  F  T  T  M  E  K  A  G  A  H  L  Q  G  G  A  K
```

```
                550       560       570       580       590       600
541       GGGTCATCATCTCTGCCCCCTCTGCTGATGCCCCCATGTTCGTCATGGGTGTGAACCATG
180       R  V  I  I  S  A  P  S  A  D  A  P  M  F  V  M  G  V  N  H
                610       620       630       640       650     ORF site #1
601       AGAAGTATGACAACAGCCTCAAGATCATCAGCAATGCCTCCTGCACCACCAACTGCTTAG
200       E  K  Y  D  N  S  L  K  I  I  S  N  A  S  C  T  T  N  C  L
                670       680       690       700       710       720
661       CACCCCTGGCCAAGGTCATCCATGACAACTTTGGTATCGTGGAAGGACTCATGACCACAG
220       A  P  L  A  K  V  I  H  D  N  F  G  I  V  E  G  L  M  T  T
                730       740       750       760       770       780
721       TCCATGCCATCACTGCCACCCAGAAGACTGTGGATGGCCCCTCCGGGAAACTGTGGCGTG
240       V  H  A  I  T  A  T  Q  K  T  V  D  G  P  S  G  K  L  W  R
                790       800       810       820       830       840
781       ATGGCCGCGGGGCTCTCCAGAACATCATCCCTGCCTCTACTGGCGCTGCCAAGGCTGTGG
260       D  G  R  G  A  L  Q  N  I  I  P  A  S  T  G  A  A  K  A  V
                850       860       870       880       890       900
841       GCAAGGTCATCCCTGAGCTGAACGGGAAGCTCACTGGCATGGCCTTCCGTGTCCCCACTG
280       G  K  V  I  P  E  L  N  G  K  L  T  G  M  A  F  R  V  P  T
                910       920       930 ORF site #2               960
901       CCAACGTGTCAGTGGTGGACCTGACCTGCCGTCTAGAAAAACCTGCCAAATATGATGACA
300       A  N  V  S  V  V  D  L  T  C  R  L  E  K  P  A  K  Y  D  D
                970       980       990      1000      1010      1020
961       TCAAGAAGGTGGTGAAGCAGGCGTCGGAGGGCCCCCTCAAGGGCATCCTGGGCTACACTG
320       I  K  K  V  V  K  Q  A  S  E  G  P  L  K  G  I  L  G  Y  T
               1030      1040      1050      1060      1070      1080
1021      AGCACCAGGTGGTCTCCTCTGACTTCAACAGCGACACCCACTCCTCCACCTTTGACGCTG
340       E  H  Q  V  V  S  S  D  F  N  S  D  T  H  S  S  T  F  D  A
               1090      1100      1110      1120      1130      1140
1081      GGGCTGGCATTGCCCTCAACGACCACTTTGTCAAGCTCATTTCCTGGTATGACAACGAAT
360       G  A  G  I  A  L  N  D  H  F  V  K  L  I  S  W  Y  D  N  E
               1150      1160      1170      1180      1190      1200
1141      TTGGCTACAGCAACAGGGTGGTGGACCTCATGGCCCACATGGCCTCCAAGGAGTAAGACC
380       F  G  Y  S  N  R  V  V  D  L  M  A  H  M  A  S  K  E  *
               1210      1220      1230      1240      1250      1260
1201      CCTGGACCACCAGCCCCAGCAAGAGCACAAGAGGAAGAGAGAGACCCTCACTGCTGGGGA
400
               1270      1280      1290      1300      1310     3´-UTR site
1261      GTCCCTGCCACACTCAGTCCCCCACCACACTGAATCTCCCCTCCTCACAGTTGCCATGTA
420
               1330      1340      1350      1360      1370      1380
1321      GACCCCTTGAAGAGGGGAGGGGCCTAGGGAGCCGCACCTTGTCATGTACCATCAATAAAG
440
               1390      1400      1410      1420
1381      TACCCTGTGCTCAACCAGTTAAAAAAAAAAAAAAAAAAAAA
```

Sequence of ß-actin mRNA (NM_001101.3) with chosen editing site (yellow).

```
                 10        20        30        40        50        60
1         ACCGCCGAGACCGCGTCCGCCCCGCGAGCACAGAGCCTCGCCTTTGCCGATCCGCCGCCC
1          T  A  E  T  A  S  A  P  R  A  Q  S  L  A  F  A  D  P  P  P
                 70        80        90       100       110       120
61        GTCCACACCCGCCGCCAGCTCACCATGGATGATGATATCGCCGCGCTCGTCGTCGACAAC
21         V  H  T  R  R  Q  L  T  M  D  D  D  I  A  A  L  V  V  D  N
                130       140       150       160       170       180
121       GGCTCCGGCATGTGCAAGGCCGGCTTCGCGGGCGACGATGCCCCCCGGGCCGTCTTCCCC
41         G  S  G  M  C  K  A  G  F  A  G  D  D  A  P  R  A  V  F  P
                190       200       210       220       230       240
181       TCCATCGTGGGGCGCCCCAGGCACCAGGGCGTGATGGTGGGCATGGGTCAGAAGGATTCC
61         S  I  V  G  R  P  R  H  Q  G  V  M  V  G  M  G  Q  K  D  S
                250       260       270       280       290       300
241       TATGTGGGCGACGAGGCCCAGAGCAAGAGAGGCATCCTCACCCTGAAGTACCCCATCGAG
81         Y  V  G  D  E  A  Q  S  K  R  G  I  L  T  L  K  Y  P  I  E
                310       320       330       340       350       360
301       CACGGCATCGTCACCAACTGGGACGACATGGAGAAAATCTGGCACCACACCTTCTACAAT
101        H  G  I  V  T  N  W  D  D  M  E  K  I  W  H  H  T  F  Y  N
                370       380       390       400       410       420
361       GAGCTGCGTGTGGCTCCCGAGGAGCACCCCGTGCTGCTGACCGAGGCCCCCCTGAACCCC
121        E  L  R  V  A  P  E  E  H  P  V  L  L  T  E  A  P  L  N  P
```

```
                430       440       450       460       470       480
421       AAGGCCAACCGCGAGAAGATGACCCAGATCATGTTTGAGACCTTCAACACCCCAGCCATG
141        K  A  N  R  E  K  M  T  Q  I  M  F  E  T  F  N  T  P  A  M
                490       500       510       520       530       540
481       TACGTTGCTATCCAGGCTGTGCTATCCCTGTACGCCTCTGGCCGTACCACTGGCATCGTG
161        Y  V  A  I  Q  A  V  L  S  L  Y  A  S  G  R  T  T  G  I  V
                550       560       570       580       590       600
541       ATGGACTCCGGTGACGGGGTCACCCACACTGTGCCCATCTACGAGGGGTATGCCCTCCCC
181        M  D  S  G  D  G  V  T  H  T  V  P  I  Y  E  G  Y  A  L  P
                610       620       630       640       650       660
601       CATGCCATCCTGCGTCTGGACCTGGCTGGCCGGGACCTGACTGACTACCTCATGAAGATC
201        H  A  I  L  R  L  D  L  A  G  R  D  L  T  D  Y  L  M  K  I
                670       680       690       700       710       720
661       CTCACCGAGCGCGGCTACAGCTTCACCACCACGGCCGAGCGGGAAATCGTGCGTGACATT
221        L  T  E  R  G  Y  S  F  T  T  T  A  E  R  E  I  V  R  D  I
                730       740       750       760       770       780
721       AAGGAGAAGCTGTGCTACGTCGCCCTGGACTTCGAGCAAGAGATGGCCACGGCTGCTTCC
241        K  E  K  L  C  Y  V  A  L  D  F  E  Q  E  M  A  T  A  A  S
                790       800       810       820       830       840
781       AGCTCCTCCCTGGAGAAGAGCTACGAGCTGCCTGACGGCCAGGTCATCACCATTGGCAAT
261        S  S  S  L  E  K  S  Y  E  L  P  D  G  Q  V  I  T  I  G  N
                850       860       870       880       890       900
841       GAGCGGTTCCGCTGCCCTGAGGCACTCTTCCAGCCTTCCTTCCTGGGCATGGAGTCCTGT
281        E  R  F  R  C  P  E  A  L  F  Q  P  S  F  L  G  M  E  S  C
                910       920       930       940       950       960
901       GGCATCCACGAAACTACCTTCAACTCCATCATGAAGTGTGACGTGGACATCCGCAAAGAC
301        G  I  H  E  T  T  F  N  S  I  M  K  C  D  V  D  I  R  K  D
                970       980       990      1000      1010      1020
961       CTGTACGCCAACACAGTGCTGTCTGGCGGCACCACCATGTACCCTGGCATTGCCGACAGG
321        L  Y  A  N  T  V  L  S  G  G  T  T  M  Y  P  G  I  A  D  R
               1030      1040      1050      1060      1070      1080
1021      ATGCAGAAGGAGATCACTGCCCTGGCACCCAGCACAATGAAGATCAAGATCATTGCTCCT
341        M  Q  K  E  I  T  A  L  A  P  S  T  M  K  I  K  I  I  A  P
               1090      1100      1110      1120      1130      1140
1081      CCTGAGCGCAAGTACTCCGTGTGGATCGGCGGCTCCATCCTGGCCTCGCTGTCCACCTTC
361        P  E  R  K  Y  S  V  W  I  G  G  S  I  L  A  S  L  S  T  F
               1150      1160      1170      1180      1190      1200
1141      CAGCAGATGTGGATCAGCAAGCAGGAGTATGACGAGTCCGGCCCCTCCATCGTCCACCGC
381        Q  Q  M  W  I  S  K  Q  E  Y  D  E  S  G  P  S  I  V  H  R
               1210      1220      1230      1240      1250      1260
1201      AAATGCTTCTAGGCGGACTATGACTTAGTTGCGTTACACCCTTTCTTGACAAAACCTAAC
401        K  C  F  *
               1270      1280      1290      1300      1310      1320
1261      TTGCGCAGAAAACAAGATGAGATTGGCATGGCTTTATTTGTTTTTTTTGTTTTGTTTTGG
421
               1330      1340      1350      1360      1370      1380
1321      TTTTTTTTTTTTTTTTGGCTTGACTCAGGATTTAAAAACTGGAACGGTGAAGGTGACAGC
441
               1390      1400      1410      1420      1430      1440
1381      AGTCGGTTGGAGCGAGCATCCCCCAAAGTTCACAATGTGGCCGAGGACTTTGATTGCACA
461
               1450      1460      1470      1480      1490      1500
1441      TTGTTGTTTTTTTAATAGTCATTCCAAATATGAGATGCGTTGTTACAGGAAGTCCCTTGC
481                                                                   C
               1510      1520      1530      1540      1550      1560
1501      CATCCTAAAAGCCACCCCACTTCTCTCTAAGGAGAATGGCCCAGTCCTCTCCCAAGTCCA
501
               1570      3´-UTR site         1600      1610      1620
1561      CACAGGGGAGGTGATAGCATTGCTTTCGTGTAAATTATGTAATGCAAAATTTTTTTAATC
521
               1630      1640      1650      1660      1670      1680
1621      TTCGCCTTAATACTTTTTTATTTTGTTTTATTTTGAATGATGAGCCTTCGTGCCCCCCCT
541                                                                   P
               1690      1700      1710      1720      1730      1740
1681      TCCCCCTTTTTTTGTCCCCCAACTTGAGATGTATGAAGGCTTTTGGTCTCCCTGGGAGTGG
561
               1750      1760      1770      1780      1790      1800
1741      GTGGAGGCAGCCAGGGCTTACCTGTACACTGACTTGAGACCAGTTGAATAAAAGTGCACA
581
               1810      1820      1830      1840      1850
1801      CCTTAAAAATGAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA
601
```

Sequence of STAT1 mRNA (NM_007315.3) with chosen editing site Y701 (yellow).

```
                10        20        30        40        50        60
1           GCTGAGCGCGGAGCCGCCCGGTGATTGGTGGGGGCGGAAGGGGGCCGGGCGCCAGCGCTG
1
                 70        80        90       100       110       120
61          CCTTTTCTCCTGCCGGGTAGTTTCGCTTTCCTGCGCAGAGTCTGCGGAGGGGCTCGGCTG
21
                130       140       150       160       170       180
121         CACCGGGGGGATCGCGCCTGGCAGACCCCAGACCGAGCAGAGGCGACCCAGCGCGCTCGG
41
                190       200       210       220       230       240
181         GAGAGGCTGCACCGCCGCGCCCCCGCCTAGCCCTTCCGGATCCTGCGCGCAGAAAAGTTT
61
                250       260       270       280       290       300
241         CATTTGCTGTATGCCATCCTCGAGAGCTGTCTAGGTTAACGTTCGCACTCTGTGTATATA
81
                310       320       330       340       350       360
301         ACCTCGACAGTCTTGGCACCTAACGTGCTGTGCGTAGCTGCTCCTTTGGTTGAATCCCCA
101
                370       380       390       400       410       420
361         GGCCCTTGTTGGGGCACAAGGTGGCAGGATGTCTCAGTGGTACGAACTTCAGCAGCTTGA
121                                   M   S   Q   W   Y   E   L   Q   Q   L   D
                430       440       450       460       470       480
421         CTCAAAATTCCTGGAGCAGGTTCACCAGCTTTATGATGACAGTTTTCCCATGGAAATCAG
141           S   K   F   L   E   Q   V   H   Q   L   Y   D   D   S   F   P   M   E   I   R
                490       500       510       520       530       540
481         ACAGTACCTGGCACAGTGGTTAGAAAAGCAAGACTGGGAGCACGCTGCCAATGATGTTTC
161           Q   Y   L   A   Q   W   L   E   K   Q   D   W   E   H   A   A   N   D   V   S
                550       560       570       580       590       600
541         ATTTGCCACCATCCGTTTTCATGACCTCCTGTCACAGCTGGATGATCAATATAGTCGCTT
181           F   A   T   I   R   F   H   D   L   L   S   Q   L   D   D   Q   Y   S   R   F
                610       620       630       640       650       660
601         TTCTTTGGAGAATAACTTCTTGCTACAGCATAACATAAGGAAAAGCAAGCGTAATCTTCA
201           S   L   E   N   N   F   L   L   Q   H   N   I   R   K   S   K   R   N   L   Q
                670       680       690       700       710       720
661         GGATAATTTTCAGGAAGACCCAATCCAGATGTCTATGATCATTTACAGCTGTCTGAAGGA
221           D   N   F   Q   E   D   P   I   Q   M   S   M   I   I   Y   S   C   L   K   E
                730       740       750       760       770       780
721         AGAAAGGAAAATTCTGGAAAACGCCCAGAGATTTAATCAGGCTCAGTCGGGGAATATTCA
241           E   R   K   I   L   E   N   A   Q   R   F   N   Q   A   Q   S   G   N   I   Q
                790       800       810       820       830       840
781         GAGCACAGTGATGTTAGACAAACAGAAAGAGCTTGACAGTAAAGTCAGAAATGTGAAGGA
261           S   T   V   M   L   D   K   Q   K   E   L   D   S   K   V   R   N   V   K   D
                850       860       870       880       890       900
841         CAAGGTTATGTGTATAGAGCATGAAATCAAGAGCCTGGAAGATTTACAAGATGAATATGA
281           K   V   M   C   I   E   H   E   I   K   S   L   E   D   L   Q   D   E   Y   D
                910       920       930       940       950       960
901         CTTCAAATGCAAAACCTTGCAGAACAGAGAACACGAGACCAATGGTGTGGCAAAGAGTGA
301           F   K   C   K   T   L   Q   N   R   E   H   E   T   N   G   V   A   K   S   D
                970       980       990      1000      1010      1020
961         TCAGAAACAAGAACAGCTGTTACTCAAGAAGATGTATTTAATGCTTGACAATAAGAGAAA
321           Q   K   Q   E   Q   L   L   L   K   K   M   Y   L   M   L   D   N   K   R   K
               1030      1040      1050      1060      1070      1080
1021        GGAAGTAGTTCACAAAATAATAGAGTTGCTGAATGTCACTGAACTTACCCAGAATGCCCT
341           E   V   V   H   K   I   I   E   L   L   N   V   T   E   L   T   Q   N   A   L
               1090      1100      1110      1120      1130      1140
1081        GATTAATGATGAACTAGTGGAGTGGAAGCGGAGACAGCAGAGCGCCTGTATTGGGGGGCC
361           I   N   D   E   L   V   E   W   K   R   R   Q   Q   S   A   C   I   G   G   P
               1150      1160      1170      1180      1190      1200
1141        GCCCAATGCTTGCTTGGATCAGCTGCAGAACTGGTTCACTATAGTTGCGGAGAGTCTGCA
381           P   N   A   C   L   D   Q   L   Q   N   W   F   T   I   V   A   E   S   L   Q
               1210      1220      1230      1240      1250      1260
1201        GCAAGTTCGGCAGCAGCTTAAAAAGTTGGAGGAATTGGAACAGAAATACACCTACGAACA
```

```
401          Q  V  R  Q  Q  L  K  K  L  E  E  L  E  Q  K  Y  T  Y  E  H
                 1270      1280      1290      1300      1310      1320
1261     TGACCCTATCACAAAAAACAAACAAGTGTTATGGGACCGCACCTTCAGTCTTTTCCAGCA
421          D  P  I  T  K  N  K  Q  V  L  W  D  R  T  F  S  L  F  Q  Q
                 1330      1340      1350      1360      1370      1380
1321     GCTCATTCAGAGCTCGTTTGTGGTGGAAAGACAGCCCTGCATGCCAACGCACCCTCAGAG
441          L  I  Q  S  S  F  V  V  E  R  Q  P  C  M  P  T  H  P  Q  R
                 1390      1400      1410      1420      1430      1440
1381     GCCGCTGGTCTTGAAGACAGGGGTCCAGTTCACTGTGAAGTTGAGACTGTTGGTGAAATT
461          P  L  V  L  K  T  G  V  Q  F  T  V  K  L  R  L  L  V  K  L
                 1450      1460      1470      1480      1490      1500
1441     GCAAGAGCTGAATTATAATTTGAAAGTCAAAGTCTTATTTGATAAAGATGTGAATGAGAG
481          Q  E  L  N  Y  N  L  K  V  K  V  L  F  D  K  D  V  N  E  R
                 1510      1520      1530      1540      1550      1560
1501     AAATACAGTAAAAGGATTTAGGAAGTTCAACATTTTGGGCACGCACACAAAAGTGATGAA
501          N  T  V  K  G  F  R  K  F  N  I  L  G  T  H  T  K  V  M  N
                 1570      1580      1590      1600      1610      1620
1561     CATGGAGGAGTCCACCAATGGCAGTCTGGCGGCTGAATTTCGGCACCTGCAATTGAAAGA
521          M  E  E  S  T  N  G  S  L  A  A  E  F  R  H  L  Q  L  K  E
                 1630      1640      1650      1660      1670      1680
1621     ACAGAAAAATGCTGGCACCAGAACGAATGAGGGTCCTCTCATCGTTACTGAAGAGCTTCA
541          Q  K  N  A  G  T  R  T  N  E  G  P  L  I  V  T  E  E  L  H
                 1690      1700      1710      1720      1730      1740
1681     CTCCCTTAGTTTTGAAACCCAATTGTGCCAGCCTGGTTTGGTAATTGACCTCGAGACGAC
561          S  L  S  F  E  T  Q  L  C  Q  P  G  L  V  I  D  L  E  T  T
                 1750      1760      1770      1780      1790      1800
1741     CTCTCTGCCCGTTGTGGTGATCTCCAACGTCAGCCAGCTCCCGAGCGGTTGGGCCTCCAT
581          S  L  P  V  V  V  I  S  N  V  S  Q  L  P  S  G  W  A  S  I
                 1810      1820      1830      1840      1850      1860
1801     CCTTTGGTACAACATGCTGGTGGCGGAACCCAGGAATCTGTCCTTCTTCCTGACTCCACC
601          L  W  Y  N  M  L  V  A  E  P  R  N  L  S  F  F  L  T  P  P
                 1870      1880      1890      1900      1910      1920
1861     ATGTGCACGATGGGCTCAGCTTTCAGAAGTGCTGAGTTGGCAGTTTTCTTCTGTCACCAA
621          C  A  R  W  A  Q  L  S  E  V  L  S  W  Q  F  S  S  V  T  K
                 1930      1940      1950      1960      1970      1980
1921     AAGAGGTCTCAATGTGGACCAGCTGAACATGTTGGGAGAGAAGCTTCTTGGTCCTAACGC
641          R  G  L  N  V  D  Q  L  N  M  L  G  E  K  L  L  G  P  N  A
                 1990      2000      2010      2020      2030      2040
1981     CAGCCCCGATGGTCTCATTCCGTGGACGAGGTTTTGTAAGGAAAATATAAATGATAAAAA
661          S  P  D  G  L  I  P  W  T  R  F  C  K  E  N  I  N  D  K  N
                 2050      2060      2070      2080      2090      2100
2041     TTTTCCCTTCTGGCTTTGGATTGAAAGCATCCTAGAACTCATTAAAAAACACCTGCTCCC
681          F  P  F  W  L  W  I  E  S  I  L  E  L  I  K  K  H  L  L  P
                 2110      2120      2130      2140      2150      2160
2101     TCTCTGGAATGATGGGTGCATCATGGGCTTCATCAGCAAGGAGCGAGAGCGTGCCCTGTT
701          L  W  N  D  G  C  I  M  G  F  I  S  K  E  R  E  R  A  L  L
                 2170      2180      2190      2200      2210      2220
2161     GAAGGACCAGCAGCCGGGGACCTTCCTGCTGCGGTTCAGTGAGAGCTCCCGGGAAGGGGC
721          K  D  Q  Q  P  G  T  F  L  L  R  F  S  E  S  S  R  E  G  A
                 2230      2240      2250      2260      2270      2280
2221     CATCACATTCACATGGGTGGAGCGGTCCCAGAACGGAGGCGAACCTGACTTCCATGCGGT
741          I  T  F  T  W  V  E  R  S  Q  N  G  G  E  P  D  F  H  A  V
                 2290      2300      2310      2320      2330      2340
2281     TGAACCCTACACGAAGAAAGAACTTTCTGCTGTTACTTTCCCTGACATCATTCGCAATTA
761          E  P  Y  T  K  K  E  L  S  A  V  T  F  P  D  I  I  R  N  Y
                 2350      2360      2370      2380      2390      2400
2341     CAAAGTCATGGCTGCTGAGAATATTCCTGAGAATCCCCTGAAGTATCTGTATCCAAATAT
781          K  V  M  A  A  E  N  I  P  E  N  P  L  K  Y  L  Y  P  N  I
                 2410      2420      2430      2440      2450      2460
2401     TGACAAAGACCATGCCTTTGGAAAGTATTACTCCAGGCCAAAGGAAGCACCAGAGCCAAT
801          D  K  D  H  A  F  G  K  Y  Y  S  R  P  K  E  A  P  E  P  M
                 2470      2480      2490      2500      2510      2520
2461     GGAACTTGATGGCCCTAAAGGAACTGGATATATCAAGACTGAGTTGATTTCTGTGTCTGA
821          E  L  D  G  P  K  G  T  G  Y  I  K  T  E  L  I  S  V  S  E
                 2530      2540      2550      2560      2570      2580
2521     AGTTCACCCTTCTAGACTTCAGACCACAGACAACCTGCTCCCCATGTCTCCTGAGGAGTT
841          V  H  P  S  R  L  Q  T  T  D  N  L  L  P  M  S  P  E  E  F
                 2590      2600      2610      2620      2630      2640
2581     TGACGAGGTGTCTCGGATAGTGGGCTCTGTAGAATTCGACAGTATGATGAACACAGTATA
861          D  E  V  S  R  I  V  G  S  V  E  F  D  S  M  M  N  T  V  *
                 2650      2660      2670      2680      2690      2700
2641     GAGCATGAATTTTTTTCATCTTCTCTGGCGACAGTTTTCCTTCTCATCTGTGATTCCCTC
881
```

Sequence SERPINA1 mature peptide cDNA (NM_001127707.1) with chosen editing site (PiZZ E342K, yellow).

```
              10        20        30        40        50        60
1         GAGGATCCCCAGGGAGATGCTGCCCAGAAGACAGATACATCCCACCATGATCAGGATCAC
1          E  D  P  Q  G  D  A  A  Q  K  T  D  T  S  H  H  D  Q  D  H

               70        80        90       100       110       120
61        CCAACCTTCAACAAGATCACCCCCAACCTGGCTGAGTTCGCCTTCAGCCTATACCGCCAG
21         P  T  F  N  K  I  T  P  N  L  A  E  F  A  F  S  L  Y  R  Q

              130       140       150       160       170       180
121       CTGGCACACCAGTCCAACAGCACCAATATCTTCTTCTCCCCAGTGAGCATCGCTACAGCC
41         L  A  H  Q  S  N  S  T  N  I  F  F  S  P  V  S  I  A  T  A

              190       200       210       220       230       240
181       TTTGCAATGCTCTCCCTGGGGACCAAGGCTGACACTCACGATGAAATCCTGGAGGGCCTG
61         F  A  M  L  S  L  G  T  K  A  D  T  H  D  E  I  L  E  G  L

              250       260       270       280       290       300
241       AATTTCAACCTCACGGAGATTCCGGAGGCTCAGATCCATGAAGGCTTCCAGGAACTCCTC
81         N  F  N  L  T  E  I  P  E  A  Q  I  H  E  G  F  Q  E  L  L

              310       320       330       340       350       360
301       CGTACCCTCAACCAGCCAGACAGCCAGCTCCAGCTGACCACCGGCAATGGCCTGTTCCTC
101        R  T  L  N  Q  P  D  S  Q  L  Q  L  T  T  G  N  G  L  F  L

              370       380       390       400       410       420
361       AGCGAGGGCCTGAAGCTAGTGGATAAGTTTTTGGAGGATGTTAAAAAGTTGTACCACTCA
121        S  E  G  L  K  L  V  D  K  F  L  E  D  V  K  K  L  Y  H  S

              430       440       450       460       470       480
421       GAAGCCTTCACTGTCAACTTCGGGGACACCGAAGAGGCCAAGAAACAGATCAACGATTAC
141        E  A  F  T  V  N  F  G  D  T  E  E  A  K  K  Q  I  N  D  Y

              490       500       510       520       530       540
481       GTGGAGAAGGGTACTCAAGGGAAAATTGTGGATTTGGTCAAGGAGCTTGACAGAGACACA
161        V  E  K  G  T  Q  G  K  I  V  D  L  V  K  E  L  D  R  D  T

              550       560       570       580       590       600
541       GTTTTTGCTCTGGTGAATTACATCTTCTTTAAAGGCAAATGGGAGAGACCCTTTGAAGTC
181        V  F  A  L  V  N  Y  I  F  F  K  G  K  W  E  R  P  F  E  V

              610       620       630       640       650       660
601       AAGGACACCGAGGAAGAGGACTTCCACGTGGACCAGGTGACCACCGTGAAGGTGCCTATG
201        K  D  T  E  E  E  D  F  H  V  D  Q  V  T  T  V  K  V  P  M

              670       680       690       700       710       720
661       ATGAAGCGTTTAGGCATGTTTAACATCCAGCACTGTAAGAAGCTGTCCAGCTGGGTGCTG
221        M  K  R  L  G  M  F  N  I  Q  H  C  K  K  L  S  S  W  V  L

              730       740       750       760       770       780
721       CTGATGAAATACCTGGGCAATGCCACCGCCATCTTCTTCCTGCCTGATGAGGGGAAACTA
241        L  M  K  Y  L  G  N  A  T  A  I  F  F  L  P  D  E  G  K  L

              790       800       810       820       830       840
781       CAGCACCTGGAAAATGAACTCACCCACGATATCATCACCAAGTTCCTGGAAAATGAAGAC
261        Q  H  L  E  N  E  L  T  H  D  I  I  T  K  F  L  E  N  E  D

              850       860       870       880       890       900
841       AGAAGGTCTGCCAGCTTACATTTACCCAAACTGTCCATTACTGGAACCTATGATCTGAAG
281        R  R  S  A  S  L  H  L  P  K  L  S  I  T  G  T  Y  D  L  K

              910       920       930       940       950       960
901       AGCGTCCTGGGTCAACTGGGCATCACTAAGGTCTTCAGCAATGGGGCTGACCTCTCCGGG
301        S  V  L  G  Q  L  G  I  T  K  V  F  S  N  G  A  D  L  S  G
```

```
                970       980       990      1000      1010      1020
961       GTCACAGAGGAGGCACCCCTGAAGCTCTCCAAGGCCGTGCATAAGGCTGTGCTGACCATC
321        V  T  E  E  A  P  L  K  L  S  K  A  V  H  K  A  V  L  T  I

               1030      1040      1050      1060      1070      1080
1021      GACAAGAAAGGGACTGAAGCTGCTGGGGCCATGTTTTTAGAGGCCATACCCATGTCTATC
341        D  K  K  G  T  E  A  A  G  A  M  F  L  E  A  I  P  M  S  I

               1090      1100      1110      1120      1130      1140
1081      CCCCCCGAGGTCAAGTTCAACAAACCCTTTGTCTTCTTAATGATTGAACAAAATACCAAG
361        P  P  E  V  K  F  N  K  P  F  V  F  L  M  I  E  Q  N  T  K

               1150      1160      1170      1180
1141      TCTCCCCTCTTCATGGGAAAAGTGGTGAATCCCACCCAAAAATAA
381        S  P  L  F  M  G  K  V  V  N  P  T  Q  K  *
```