Harnessing Human ADAR2 for Site-Directed RNA Editing

Dissertation

der Mathematisch-Naturwissenschaftlichen Fakultät
der Eberhard Karls Universität Tübingen
zur Erlangung des Grades eines
Doktors der Naturwissenschaften
(Dr. rer. nat.)

vorgelegt von
Jacqueline Wettengel

Tübingen
2016

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

Tag der mündlichen Qualifikation:	19.04.2016
Dekan:	Prof. Dr. Wolfgang Rosenstiel
1. Berichterstatter:	Prof. Dr. Thorsten Stafforst
2. Berichterstatter:	Prof. Dr. Ralf-Peter Jansen

Selbstständigkeitserklärung

Hiermit erkläre ich,

- dass ich diese Arbeit selbst verfasst und keine anderen als die angegebenen Quellen und Hilfsmittel benutzt habe und alle wörtlich oder sinngemäß aus anderen Werken übernommenen Aussagen als solche gekennzeichnet habe,
- dass ich die Richtlinien zur Sicherung guter wissenschaftlicher Praxis und zum Umgang mit wissenschaftlichen Fehlverhalten an der Eberhard-Karls-Universität beachtet habe,
- dass die Arbeit weder vollständig noch in wesentlichen Teilen Gegenstand eines anderen Prüfungsverfahrens gewesen und die Arbeit weder vollständig noch in wesentlichen Teilen bereits veröffentlicht ist,
- dass das in Dateiform eingereichte Exemplar mit eingereichten und gebundenen Exemplaren übereinstimmt.

Tübingen, den 17.Februar 2016

Danksagung

Meinem Chef und Mentor Dr. Thorsten Stafforst möchte ich für die großzügige Bereitstellung des Arbeitsplatzes und der Überlassung des hochinteressanten Promotionsthemas ganz herzlich danken. In jeder Phase meiner Arbeit wurde ich von ihm intensiv und warmherzig unterstützt. Außerdem möchte ich mich für die vielen hilfreichen und anregenden Diskussionen bei ihm bedanken und ganz besonders auch für die Freiheit, die mir während des gesamten Forschungsprojektes von ihm gewährt wurden. Auch Professor Dr. Ralf-Peter Jansen gilt ein besonderer Dank. Jederzeit konnte ich mich auf seinen kompetenten Rat und zielgerichtete Hilfe bei der Planung und Auswertung meiner Arbeit verlassen. Seine Erfahrung gab mir wertvolle Unterstützung. Für die kreativen Ideen und die uneingeschränkte Bereitschaft danke ich Dr. Frank Essmann. Dr. Gáspár Jékely möchte ich für die Möglichkeit danken, in seiner Arbeitsgruppe mit *Platynereis Dumerilii* innerhalb des Kooperationsprojekts zu arbeiten. In diesem Zusammenhang danke ich besonders Aurora Panzera für die zeitintensive und professionelle Einführung in die Mikroinjektionstechnik, sowie Luis Bezares und Nobuo Ueda für ihre zusätzliche Unterstützung.

An dieser Stelle möchte ich mich für die immer ausgesprochen angenehme Atmosphäre und die produktive wissenschaftliche Zusammenarbeit bei allen Kollegen aus dem Team Stafforst bedanken. Vor allem möchte ich mich bei Tahsin Kuzdere und Pia Mach für die tatkräftige Unterstützung bedanken. Philipp Reautschnig und Paul Vogel danke ich für die zahlreichen Tipps, die damit die Fertigstellung meiner Dissertation unendlich erleichterten.

Professor Dr. David Wharam danke ich sehr für das Korrekturlesen.

Meiner Mutter danke ich herzlich für ihre fortwährende Unterstützung und ihr Interesse an meiner Arbeit.

Allen meinen Freunden danke ich für die Ausdauer, Ruhe und Geduld, womit sie mir stets zur Seite standen und mich damit unterstützt haben.

Thomas danke ich von ganzem Herzen für seine unermüdliche Unterstützung, seine Liebe und Motivation. Ihm gilt mein besonderer Dank.

Table of Contents

List of Abbreviations

5-HT_3	5-Hydroxytryptamine$_3$
AAV vectors	adeno-associated viral vectors
ADAR	adenosine deaminase acting on RNA
ADAT	adenosine deaminase that acts on tRNA
ALS	amyotrophic lateral sclerosis
AMPA	α-amino-3-hydroxl-5-methyl-4-isoxazolepropionic acid
ASO	antisense oligonucleotides
ATP	adenosine triphosphate
bp	base pairs
BG-gRNA	5' O-benzylguanine-modified gRNA
BSA	bovine serum albumin
bw	backward
Cas	CRISPR-associated-endonuclease
CDA	cytidine deaminase
cDNA	complementary DNA
CMV promoter	cytomegalovirus immediate-early promoter
CRISPR	clustered regularly interspaced short palindromic repeats
cRNA	crispr RNA
DMEM	Dulbecco's Modified Eagle Medium
DNA	deoxyribonucleic acid
DNase	deoxyribonuclease
dNTP	deoxynucleotide
DSB	double stranded break
dsRBD	double-stranded RNA binding domain
dsRNA	double-stranded RNA
eCFP	enhanced cyan fluorescent protein
EDTA	ethylenediaminetetraacetic acid
eGFP	enhanced green fluorescent protein
EPAP	*E. coli* Poly(A) Polymerase
FBS	fetal bovine serum
FSC	forward scatter

FUS	fused in sarcoma
GluR	glutamate receptor
gRNA	guideRNA
HDR	homology directed repair
HEK293T	human embryonic kidney 293T
HH	hammerhead
hpf	hours post fertilization
ivT	in vitro transcription
IP_6	inositol hexakisphosphate
kb	kilo base pairs
KCl	potassium chloride
KH_2PO_4	potassium dihydrogen phosphate
mRNA	messenger RNA
miRNA	microRNA
Na_2HPO_4	sodium hydrogen phosphate
NaCl	sodium chloride
NEB	New England Biolabs
NES	nuclear export signal
NHEJ	non-homologous end joining
no.	number
NSW	natural sea water
OE PCR	overlap extension polymerase chain reaction
ORF	open reading frame
PBS	phosphate buffered saline
PCR	polymerase chain reaction
PDB	Protein Data Bank
PINK1	PTEN-induced putative kinase 1
pre-mRNA	precursor mRNA
pri-miRNA	precursor miRNAs
RISC	RNA-induced silencing complex
RLU	relative light units
RNA	ribonucleic acid
RNAi	RNA interference

RNase H	ribonuclease H
RT-qPCR	real-time quantitive PCR
SDS	sodium dodecyl sulfate
shRNA	short hairpin RNA
siRNA	small interfering RNA
ss-gRNA	single-stranded gRNA
SSC	side scatter
SSO	splice-switching oligonucleotides
TAE	Tris-acetate-EDTA
TALEN	transcription activator-like effector nuclease
TBE	Tris-borate-EDTA
TE	Tris-EDTA
TLS	translocated in sarcoma
tracrRNA	trans-activating crRNA
tRNA	transfer RNA
UTR	untranslated region
ZBD	Z-DNA-binding domain
ZFN	zinc finger nuclease

1 Abstract

Gene therapy has made significant progress in recent years. The first gene therapy application has been approved, novel technologies like TALEN and CRISPR/Cas9 are under development, and RNAi-based therapeutics behave well in phase 3 and will probably enter the clinics soon. Oligonucleotide-based drugs are very promising since they tackle genetic diseases by their roots. Furthermore, when targeting the genetic level (DNA or RNA) rather than the protein level, other interventions can be carried out since the proteins themselves can be adjusted and modified in their properties. Besides medical applications, DNA- and RNA-based tools are valuable to understand biological functions or to reengineer them (synthetic biology). One possibility of such an intervention is represented by A-to-I RNA editing, an RNA modification that allows the reprogramming of the genetic content at the level of codon sense and RNA processing. Recently, two strategies have demonstrated the recoding of mRNAs by site-directed A-to-I RNA editing. Both strategies apply reengineered deaminase enzymes for this purpose.

In this PhD-project we present a new strategy to overcome some limitations of these two recently presented RNA editing systems: the necessity to apply chemically modified guideRNAs and the need to express artificial proteins. For this, we have been engineering novel guideRNAs based on a natural substrate (neuronal AMPA receptor) of the endogenous editing enzyme hADAR2 (adenosine deaminases acting on RNA). Such guideRNAs join two functions, the site-specific binding of any arbitrary mRNA via duplex formation and the recruitment of the naturally occurring hADAR2 to elicit deamination of a specific single adenosine base. We could verify that our *trans*-acting guideRNAs work efficiently to direct hADAR2's editing activity to specific sites in user-defined target mRNAs, both, in the PCR tube as well as in cell culture. For the preferred 5´-UAG codon, high editing yields have been achieved in the PCR tube (up to 100%) and in cell culture (up to 57%). Also for the less preferred codons, such as 5'-CAG, 5'-CAC, and 5'-GAG, editing was feasible with >50% yield in the PCR tube and, for instance, 37% yield for the 5'-CAG triplet in cell culture, after optimization of the guideRNA sequences. Overall, we accumulated extensive knowledge about the optimal guideRNA architecture, the specific requirements to achieve efficient editing

with various codons, strategies to control off-target editing, and we generated some initial molecular understanding of the recognition of the mRNA/guideRNA duplex by the hADAR2 enzyme.
Furthermore, we could demonstrate our editing system not to be limited to target a single point mutation or a single gene only, but it rather allows the addressing of (two reporter) genes simultaneously, while maintaining high editing yields and specificities.
Finally, we generated initial data that show that our strategy functions in a living organism (*Platynereis dumerilii)*. We found some promising signs that suggest the recruitment of endogenous editing activity, and we performed early experiments that indicate the possibility to chemically stabilize the guideRNA.
In summary, we could demonstrate that our novel, genetically encodable approach is able to manipulate various codons at the RNA-level in order to restore protein function including the repair of disease-promoting genetic mutations.

Site-directed RNA editing represents an attractive alternative to DNA-based gene therapy, as it circumvents some ethical and safety issues related to the (irreversible) manipulation of the genome. Similar to genome editing, RNA editing preserves the endogenous gene regulation via transcription, splicing and transport. Additionally, site-directed RNA editing may complement DNA-based strategies due to its reversible and tunable nature that allows the creation of protein isoform mixtures with differing properties in a tunable stoichiometry and that allows the perturbation of a protein function suddenly and for a limited amount of time.

This PhD-thesis breaks the ground for the development of new tools to study basic biology and may well find translation into medicine. With respect to the latter, site-directed RNA editing, once fully developed, might allow diversifying protein function inside a living organism by administration or ectopic expression of a small RNA only. To achieve this goal we still have to better understand the recruitment of the (tissue-specific) endogenous ADAR proteins, ADAR1 and ADAR2, and its various isoforms and we still have to take the hurdle of delivery that currently limits most oligonucleotide-derived drugs in vivo.

Zusammenfassung

Die Gentherapie hat in den vergangenen Jahren beachtliche Fortschritte erzielt: die erste Gentherapie wurde zugelassen, neue Technologien wie TALEN und CRISPR/Cas9 sind in der Entwicklung, und RNAi-basierte Therapeutika sind erfolgreich in der klinischen Erprobung und werden wohl bald in die Kliniken gelangen. Solche Oligonukleotid-basierten Medikamente sind äußerst vielversprechend, da sie die genetischen Krankheiten an ihrem Ursprung angehen. Auch können auf der gene-tischen Ebene (DNA oder RNA), im Unterschied zur Proteinebene, ganz andere Eingriffe vorgenommen werden, denn die Proteine selbst können in ihren Eigenschaften angepasst werden. Neben der medizinischen Anwendung sind DNA- und RNA-basierte Werkzeuge äußerst nützlich, um biologische Funktionen zu verstehen oder um diese neu zu gestalten (synthetische Biologie). Eine Möglichkeit um verändernd einzugreifen ist die A-zu-I Editierung, eine RNA-Modifizierung die es erlaubt, den genetischen Inhalt auf Ebene der Kodierung und der RNA-Prozessierung umzuprogrammieren. Kürzlich wurden zwei Strategien vorgestellt, die mRNAs mittels zielgerichteter A-zu-I RNA Editierung umprogrammieren und beide verwenden dafür künstliche Deaminasen.

In dieser Dissertation stellen wir eine neue Strategie vor, die einige der Limitationen der beiden kürzlich präsentierten RNA Editierungssysteme überwindet: die Notwendigkeit, chemisch modifizierte guideRNAs zu verwenden und der Bedarf künstliche Proteine zu exprimieren. Dafür haben wir neue guideRNAs entworfen, die auf dem natürlichen Substrat (neuronaler AMPA Rezeptor) des endogenen hADAR2 (adenosine deaminases acting on RNA) beruhen. Diese guideRNAs verbinden zwei Funktionen; das zielgerichtete Binden jeder beliebigen mRNA durch Duplexbildung und die Rekrutierung des natürlich vorkommenden hADAR2, um die Deaminierung einer einzigen spezifischen Adenosinbase hervorzurufen. Wir konnten zeigen, dass unsere guideRNA sowohl im PCR-Gefäß, als auch in Zellkultur äußerst effizient die enzymatische Aktivität von hADAR2 an spezifische Stellen von ausgewählten mRNAs dirigiert. Für das bevorzugte 5´-UAG Triplett wurden hohe Editierungs-ausbeuten im PCR-Gefäß (> 100%) und in der Zellkultur (> 57%) erreicht. Nach Optimierung der guideRNA-Sequenz war die Editierung von schwierigeren Tripletts,

wie 5'-CAG, 5'-CAC, und 5'-GAG im PCR-Gefäß mit bis zu 50% Ausbeute und in der Zellkultur mit 37% Ausbeute für das 5'-CAG Triplett möglich.

Zudem konnten wir umfangreiches Wissen über die optimale guideRNA-Architektur und die spezifischen Voraussetzungen für die effiziente Editierung verschiedener Tripletts sammeln. Wir haben Strategien gefunden, um ungewünschte Editierungen zu kontrollieren und wir konnten erste Erkenntnisse über die molekulare Erkennung des mRNA/guideRNA Duplexes durch das hADAR2 Enzym gewinnen. Unser Editierungs-system ist nicht etwa auf eine einzelne Punktmutation oder einzelne Gene begrenzt, vielmehr können gleichzeitig zwei Gene (Reportergene) bei gleich-bleibend hoher Editierungsausbeute und -spezifität adressiert werden. Am Ende konnten wir erste Daten erzeugen, die zeigen, dass unsere Strategie auch im le-benden Organismus (*Platynereis dumerilii)* funktioniert. Wir haben vielversprechende Hinweise auf die Rekrutierung endogener Editierungsaktivität und erste Experimente weisen auf die Möglichkeit zur chemischen Stabilisierung der guideRNA hin.
Unsere neue Strategie ist in der Lage, verschiedene Tripletts auf RNA-Ebene zu verändern, z.B. krankheitsauslösenden genetische Mutationen, um die ursprüngliche Proteinfunktion wiederherzustellen.
Die zielgerichtete RNA-Editierung stellt eine attraktive Alternative zur DNA-basierten Gentherapie dar, da das Genom nicht (irreversibel) verändert wird und damit ethische und sicherheitsbezogene Bedenken wegfallen. Ähnlich zur Genomeditierung, bewahrt die RNA-Editierung die endogene Genregulation per Transkription, Spleißen und Transport. Die zielgerichtete RNA-Editierung könnte aufgrund seines reversiblen und einstellbaren Charakters die DNA-basierten Strategien ergänzen, indem mehrere Proteinisoformen mit unterschiedlichen Eigenschaften hergestellt werden können und die Funktion von Proteinen für eine begrenzte Zeit schlagartig gestört werden kann.

Diese Dissertation legt den Grundstein zur Entwicklung neuer Werkzeuge für die Grundlagenbiologie und sollte Anwendung in der Medizin finden. Hinsichtlich letzterem könnte die zielgerichtete RNA-Editierung, durch Gabe oder ektopische Expression kleiner RNA-Moleküle, die Proteinfunktion innerhalb eines lebenden Organismus verändern. Um dies zu erreichen müssen wir jedoch die Rekrutierung der (gewebespezifischen) endogenen ADARs und deren Isoformen besser verstehen und die ungünstige Pharmakokinetik Oligonukleotid-basierter Medikamente überwinden.

2 Introduction

2.1 RNA editing by adenosine deaminases acting on RNA

Post-transcriptional RNA modifications, such as RNA splicing, polyadenylation, 5'-capping and RNA editing, contribute to the control of gene expression and increase the complexity of an organism's proteome. RNA editing is one import mechanism that can change the sense of a transcript by insertion, deletion, or substitution of codons in the open reading frame, or by alteration of splicing (1,2). Beside pseudo-uridinylation, 2'-O-methylation and cytosine deamination, the modification of adenosine is one major RNA editing mechanism. This work focuses on adenosine deaminases acting on RNA (ADAR), which catalyze the hydrolytic deamination at the C6 position of adenosine, resulting in inosine, a process called A-to-I RNA editing (Figure 2-1). During translation, inosine will be interpreted as a guanosine (3). Therefore, A-to-I editing has the potential to change not only the primary sequence information of RNA but also the resulting amino acid composition of a protein (4).

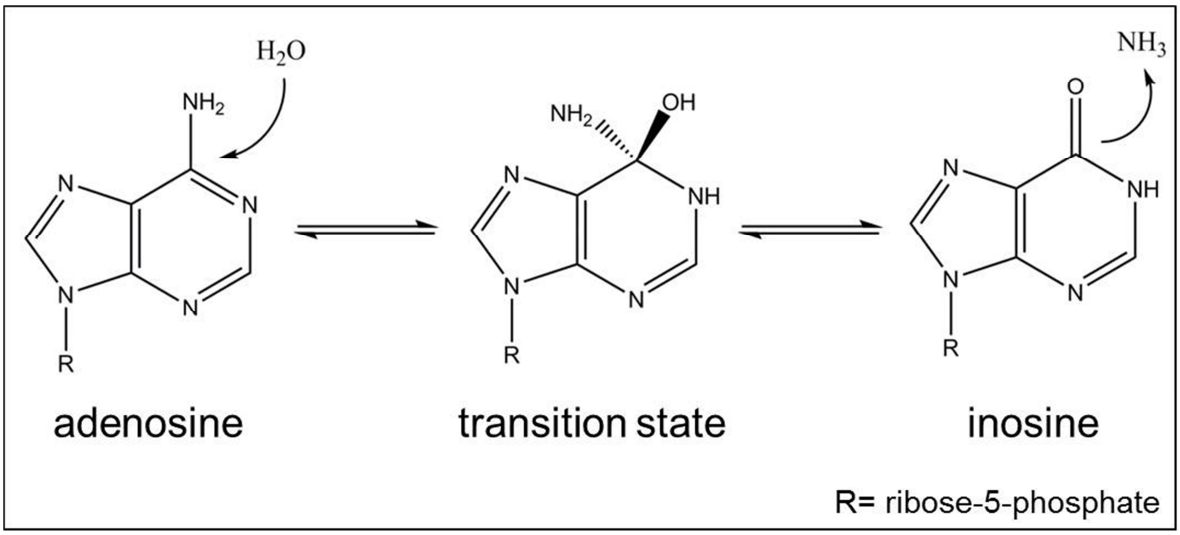

Figure 2-1: A-to-I editing. ADAR catalyzes the hydrolytic deamination at the C6 position of adenosine. R= ribose-5-phosphate.

However, RNA editing occurs more often in non-coding regions like introns, 3' untranslated regions (UTR) or in Alu repetitive elements (5,6). Levanon et al. found by a computational algorithm search for ADAR editing sites that 92% of sites are located within an Alu repeat (6). The effect of editing in Alu repeats, which are unique

to primates, can be hypothesized as an antitransposition mechanism that inhibits the integration of transcribed Alu sequences back into the genome (6). Furthermore, they could support the prediction that most pre-mRNA editing in the brain is located in noncoding regions: 12% were found in the 5'-UTR, 54% in the 3'-UTR and 33% in introns. Since the essential secondary structure for RNA editing is created by fold back structures of introns (or Alu repeats), RNA editing and splicing are postulated to be synchronized (7) (p.153). Therefore, editing of intronic sequences can alter splice sites by creation or deletion, whereas editing in 3'-UTRs is supposed to modulate stability and transport of RNA (6). The change of splice sites can also lead to a frame shift upon translation, which subsequently terminates the translation at the premature stop codon. The presence of a premature stop codon created by alternative splicing can also trigger the nonsense-mediated decay, with the result of lacking protein (8). Hyper-edited dsRNA is specifically cleaved (7,9). Thereby, editing has the potential to regulate gene expression by controlling transcript stability. A further influence that editing in the 3'-UTR can have is the modulation of the innate immune response (10). Liddicoat et al. (10) postulate that ADAR1 has the function to hyperedit endogenous dsRNA to prevent classification of endogenous dsRNA as nonself. Possible effects and the potential of A-to-I editing on pre-mRNA are summarized in Figure 2-2.

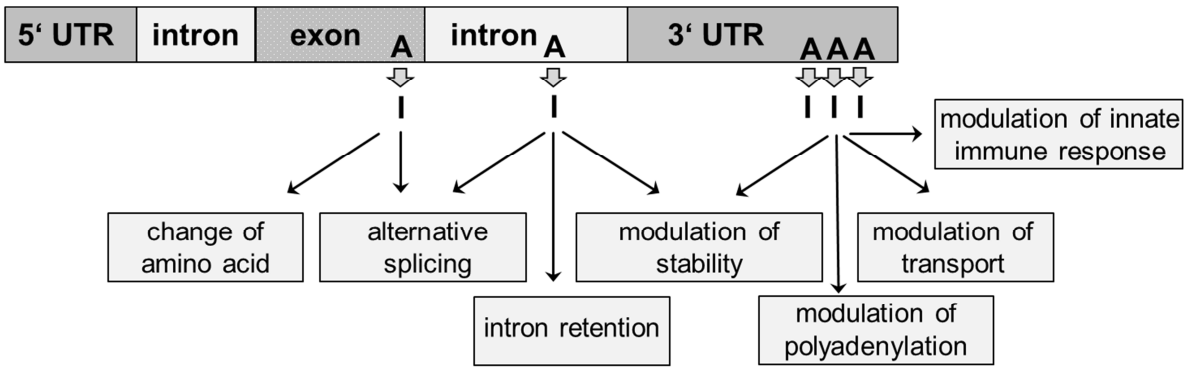

Figure 2-2: Potential of A-to-I editing of pre-mRNAs in the nucleus. The possible editing sites on the pre-mRNA are indicated with an A for adenosine and the potential consequences are shown in boxes. Picture similar to Öhmann in (8).

Although the function of editing in noncoding regions like in the 3'-UTR still remains unclear (5), one might assume that the effect of editing in coding regions (exons) is clearer. Many functional consequences under physiological conditions are known to be caused by aberrant RNA editing. For example the G-protein-coupled serotonin

receptor 5-HT_{2C} demonstrates connections to cognitive brain function, since its dysregulation was observed to be involved in depression, suicide and schizophrenia (11). The serotonin receptor 5-HT_{2C} transcript has five possible editing sites and by their combinatorial editing 24 receptor variants are producible (7) (p.151). Among these, 12 receptor isoforms were found in human brains demonstrating the possible isoform diversity from editing. About the complexity of different isoforms being necessary for cognitive brain function and how editing is involved in disease pathogenesis, little is known (11). This is due to the experimental gold-standard method that uses knock-in or knock-out animals for an abolished or constitutive RNA editing (11). Therefore, only dysregulations are observed, whereas it remains unclear when and where editing is regulated.

Since dsRNA represent targets for both RNA editing and RNA interference-mediated silencing (RNAi), crosstalk between these two pathways has been observed. Findings from Yang et al. proved that RNA editing suppresses RNAi in mammalian cells if I:U mismatches disrupt the structure of RNA duplexes to such an extent that Drosha cannot process the RNA duplex (7) (p.196) (12). It was demonstrated that even short siRNA, which are not long enough to induce ADAR activity, can be bound by ADAR with high affinity and are therefore not available for the RNAi pathway (7) (p.196) (13,14). Further studies support the idea that ADAR1-p150 antagonizes RNAi independent of editing by high-affinity binding to the siRNA and miRNA (15). In contrast, the findings from Ota et al. (16) postulate a different role of ADAR1 in miRNA processing and RNAi: The complex of ADAR1 and Dicer is supposed to increase the cleavage rate of pre-miRNAs by Dicer and loading of the miRNA into RISC (RNAi silencing complex) is facilitated. Until now it remains unclear how both pathways are coordinated in detail.

Another discovery revealed that precursor molecules of miRNAs (pri-miRNAs) that form a hairpin structure are subject of A-to-I RNA editing, too (12,17,18). These studies estimated that 6-16% of human pri-miRNAs are target for ADAR editing. Due to editing of the pri-miRNA outside the mature miRNA region, maturation can be blocked by alteration of the structure of pri-miRNA so that Drosha/Dicer cannot process the RNA further (Figure 2-3). Editing of the mature miRNA region might change the target specificity of the miRNA. Therefore, RNA editing increases the diversity of the “miRNome” (7) (p.22).

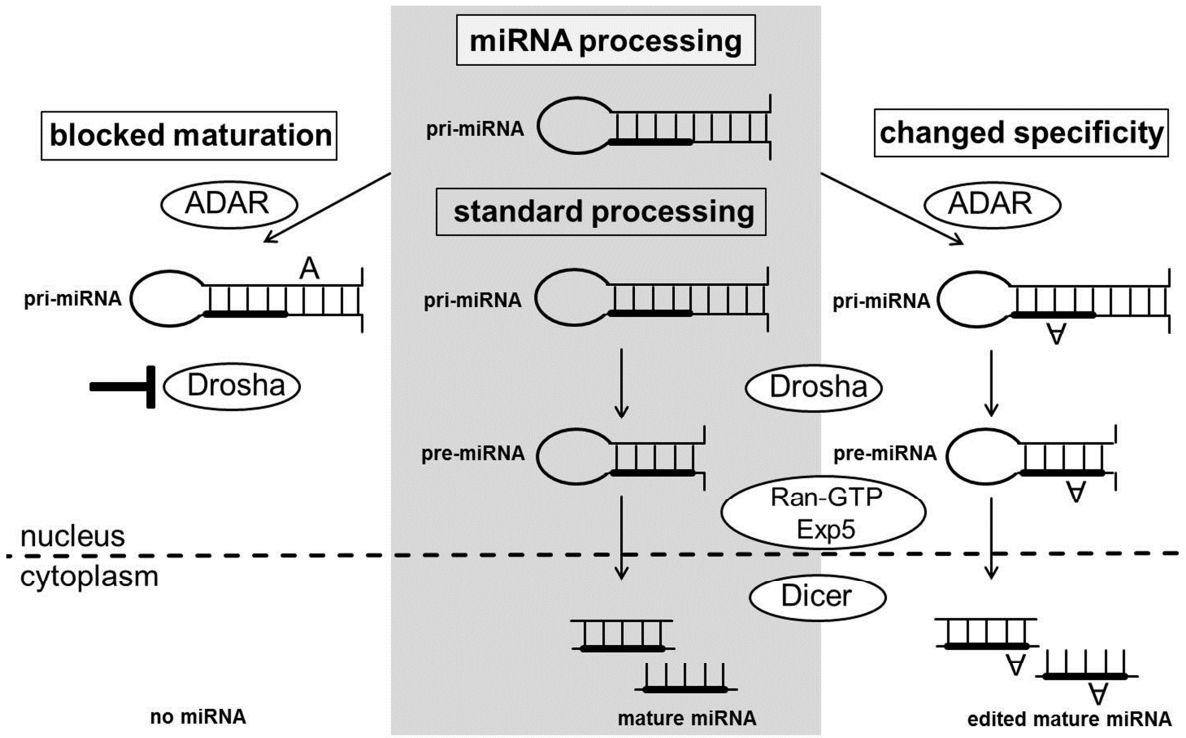

Figure 2-3: miRNA processing and influence of pri-miRNA editing on the mature miRNA. The central panel demonstrates the maturation pathway for miRNAs. The pri-miRNA is cleaved by Drosha, the resulting pre-miRNA is exported to the cytoplasm and processed by Dicer. The left panel demonstrates the possible consequences of pri-miRNA editing outside the mature miRNA region. The secondary structure might be changed by editing in a way that Drosha is not able to process this RNA duplex. The second option of RNA editing in the region of the mature miRNA (on the right) leads to a variable target specificity of the mature miRNA. Picture similar to (7) page 22.

2.2 The ADAR protein family

ADAR enzymes exist in all metazoans and are highly conserved (4). The number of genes and isoforms are very variable between species. In mammals the ADAR protein family includes three members: ADAR1, ADAR2 and ADAR3. All ADARs consist of a catalytic deaminase domain, located at the C-terminus of the protein and the N-terminus is variable for each family member (Figure 2-4) (4).

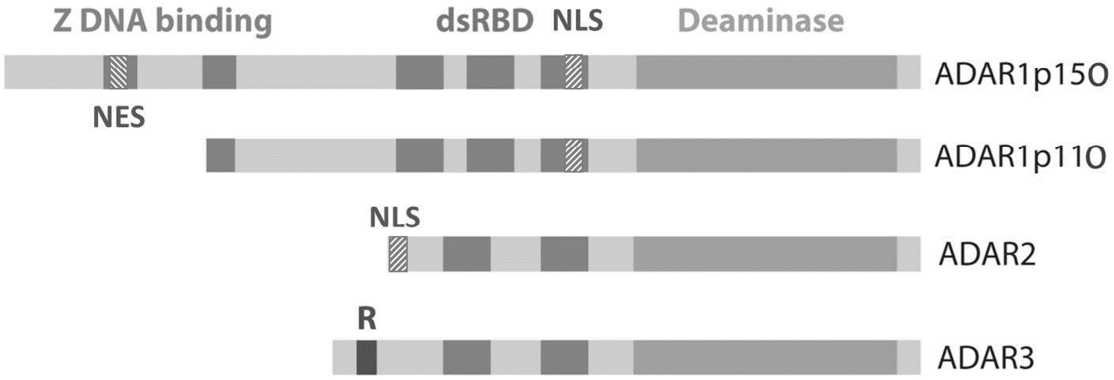

Figure 2-4: Mammalian ADAR family members. Three members of mammalian ADAR are known. ADAR1 exists in two isoforms (p150 and p110). The N-terminus of ADAR1 consists of three dsRBDs and two Z-DNA binding domains in the longer isoform and one Z-DNA binding domain in the short form. ADAR2 and ADAR3 consist both of two dsRBDs, whereas ADAR3 harbors a R-type RNA-binding domain. A nuclear localization signal is abbreviated by NLS and a nuclear export signal by NES. Figure changed from (19).

ADAR2 contains the shortest N-terminus harboring two double-strand RNA binding domains (dsRBMs) and a nuclear localization signal (NLS) located at the very end. ADAR1 demonstrates the longest N-terminus, including three dsRBMs, a NLS and nuclear export signal (NES), as well as two Z-DNA-binding domains (ZBDs). The ADAR3 protein harbors two dsRBMs and a single-stranded R-type RNA-binding domain and is found to be catalytically inactive, even though all essential amino acids are found in the active site (20). All ADARs are typically nuclear localized, whereas the long isoform of ADAR1 (ADAR1p150) is also found in the cytoplasm (4). The two isoforms of ADAR1 are transcribed by alternative promoters; the full-length isoform (ADAR1p150) by an interferon-inducible promoter and the amino-terminally shortened isoform (ADAR1p110) by a constitutive promoter (21). The long form of ADAR1 is thought to be involved in viral defense, by introducing hyper mutations in the viral open reading frames (22). The localization of the ADAR proteins is considered to be dependent on their biological function and their substrates (8). The ADAR1 enzyme is more ubiquitously expressed, whereas ADAR2 seems to be mainly limited to the brain (8,23,24). Since the intensity of RNA editing in humans is 35-fold higher than in mice it is suggested that RNA editing has brain-related functions and has led to an increased cognition and driven neuronal evolution (11). Also the majority of editing substrates are found in the central nervous system (24), which further suggests an important role for these enzymes in normal development

and life. An inactivation of ADARs results in neuronal and behavioral phenotypes: ADAR1 was demonstrated to be required for embryonic development and its lack leads to embryonic lethality, whereas the lack of ADAR2 results in severe neurological diseases, such as depression, schizophrenia and Prader-Willi syndrome (11,25-29). In which manner RNA editing is involved in disease pathogenesis still remains unclear and has to be the topic of further studies.

ADAT (Adenosine deaminases that act on tRNA) are structurally and functionally related to ADAR, thus presenting an extended member of the ADAR protein family (4,8).

2.3 ADAR substrates

The substrates of ADAR proteins are highly base-paired, usually consisting of one long, unbranched double helix (4). Other helix-forming substrates are interrupted by mismatches, bulges and loops. One of the shortest double-stranded RNA duplexes targeted by ADAR in vivo is the R/G editing site of the glutamate receptor transcript (4) and is, as the majority of all editing substrates, primarily expressed in the central nervous system (24).

The glutamate receptor channel (GluR) subtype AMPA, is one of the first ADAR substrates discovered. The GluR-B pre-mRNA of the AMPA receptor is the best studied substrate and contains two editing sites (8): Editing at the Q/R site changes a 5'-CAG codon (glutamine) to 5'-CIG, which is interpreted as arginine. A-to-I editing of the R/G-site alters the 5'-AGG codon (arginine) to 5'-IGG, coding for glycine. Both editing events of the GluR-B pre-mRNA are mediated by an RNA secondary structure that is formed via an intronic part, which folds back to the exonic part, creating an intramolecular hairpin (Figure 2-5). The intronic parts of the R/G- and Q/R-site of the glutamate receptor transcript are removed during splicing. Editing of these sites is essential for changing the ion permeability of the channel (4,30). The Q/R site is edited up to 99% and is absolutely necessary, since mice with a lack of ADAR2 editing at the Q/R-site, suffer from epileptic seizures and die within three weeks of birth (8,31). But ADAR2$^{-/-}$ animals can be rescued by changing the wild-type GluR-B allele with one that codes for the edited R codon (4). Interestingly, these rescued animals show a significantly reduced editing of the R/G-site, which is typically edited

up to 50% (32). By a variation of assembled subunits of the R-form or the G-form (R/G-site), the receptors have different recovery rates from desensitization (8,33).
The Q/R site of the gluR-B pre-mRNA is mainly edited by ADAR2, whereas both enzymes can edit the R/G site (8), however ADAR2 is the preferred enzyme.

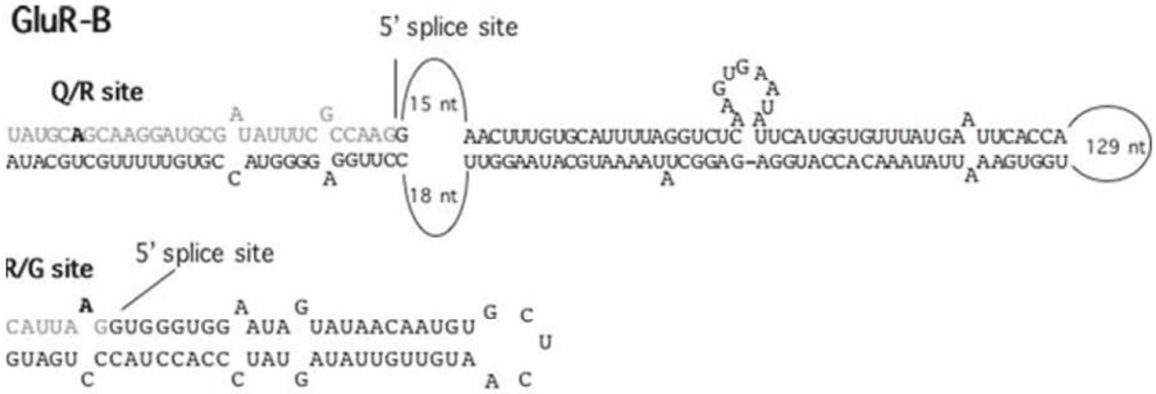

Figure 2-5: Predicted secondary structure of the GluR-B receptor transcript. The intronic part (black) folds back to the exonic part (grey) of the transcript, thereby creating the secondary structure with bulges and loops. The site-selective editing sites are indicated in bold. Picture from (8).

2.3.1 The role of dsRBDs in substrate recognition

The ADAR proteins vary in their number of dsRNA binding domains (dsRBDs) (2.2). The dsRBDs represent the second most abundant family of RNA recognition motifs (34) and even though they are not essential for the catalyzed reaction they play an important role in substrate recognition and determine the specificity of the ADAR enzymes. Besides RNA editing dsRBDs are well known from other mechanisms including post-transcriptional regulatory processes, RNAi, as well as RNA export and transport (35,36).
The dsRBDs are very abundant. In contrast to ssRNA binding domains, which often recognize specific nucleotide sequences, dsRBDs are believed to recognize shape rather than sequence (37). This is due to morphological characteristics of the A-form RNA double helix: The dsRNA helix has a wide and shallow minor groove and a deep and narrow major groove that can be interrupted by internal bulges. The edge of the minor groove is easily accessible, whereas access to the bases of the major groove is hampered (37). However, the major groove accommodates the Hoogsteen site of the nucleobase pairs, which are best suited to read out sequence information. Thus, sequence-specific protein readers are hardly accesible. Also binding analysis with

synthetic RNA duplexes supports the idea that the dsRBMs recognize the A-form helix of dsRNA in a sequence-independent manner (35,38). The weak sequence-dependency of dsRBDs fits well to the high promiscuity of ADARS. Nevertheless, NMR analysis of ADAR2's dsRBD with the R/G-motif reveals, that a few sequence specific contacts are formed in the minor groove (35): dsRBM1 binds to the RNA stem-loop by a hydrogen bond to the amino group of G22 in the mismatch G22:G50 and a hydrophobic contact to adenine A32. dsRBM2 forms a hydrogen bond to guanosine G9 and a hydrophobic contact to adenine A18. These bases are important for the binding affinity, since mutations of these lowered the affinity. Replacement of the G22:G50 mismatch by Watson-Crick G22-C50 hardly affects the binding affinity, confirming that dsRBM1 specifically recognizes the G22 base rather than the shape of the RNA helix (35). Mutagenesis supports this finding and sets the basis for the specific recognition of the R/G-motif by ADAR2 and explains the preference for guanosine as the 3'-neighbor of the edited adenosine.

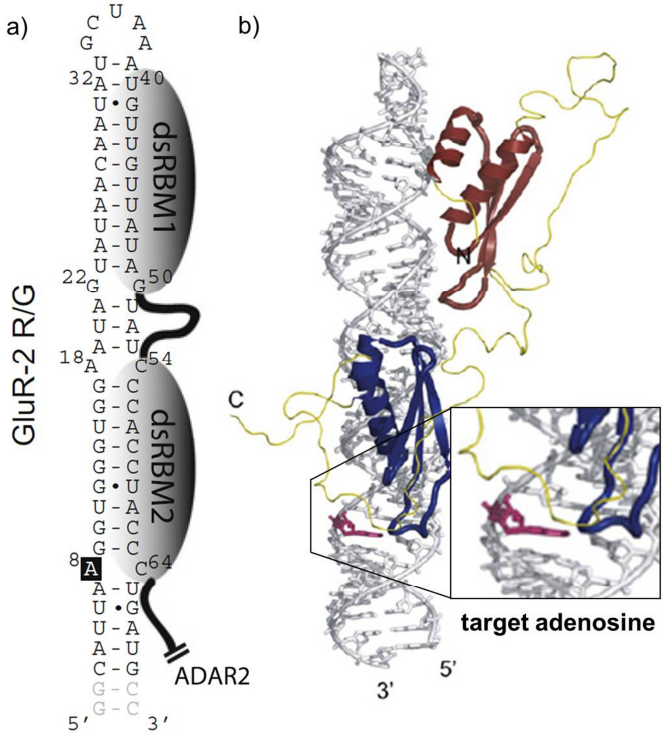

Figure 2-6: Secondary structure of GluR-B R/G RNA and structure of ADAR2 bound to GluR-B R/G.

a) Secondary structure of GluR-B R/G RNA

b) Model of the most representative RDC-reconstructed structure of ADAR2 bound to GluR-B R/G. The RNA is shown as a stick model in grey, edited adenosine is coloured in pink, dsRBM1 is shown in

red and dsRBM2 in blue. The linker between the two dsRBMs is coloured in yellow. Pictures adapted from (35).

2.4 Catalysis of A to I editing

The close relation of the deaminase domain to other nucleotide deaminases, such as adenosine deaminases (ADAs) and cytosine deaminases (CDAs) (4,39,40) led to the suggestion that zinc is responsible for the coordination of the water molecule acting as a nucleophile in the catalytic reaction (23). Indeed, the crystal structure of the deaminase domain revealed that a zinc ion is localized in the active site (41). The zinc ion is coordinated by one histidine (H394) and two cysteine residues (C451 and C516) in the active site (Figure 2-7). Unexpectedly, the crystallization of hADAR2 catalytic domain also revealed the tight binding of IP_6 in a conserved binding site. This cofactor was demonstrated to be essential for ADAR activity in vitro as well as in vivo, being involved in protein folding and formation of the catalytic site. The necessity for IP_6 also explains an old mystery why ADARs cannot be produced from *E. coli*. The amino acid E396 of the active site, which is thought to act as a proton-shuttle, coordinates together with zinc the nucleophilic water. A point mutation of this residue (E396A) results in a complete loss of editing activity (42).

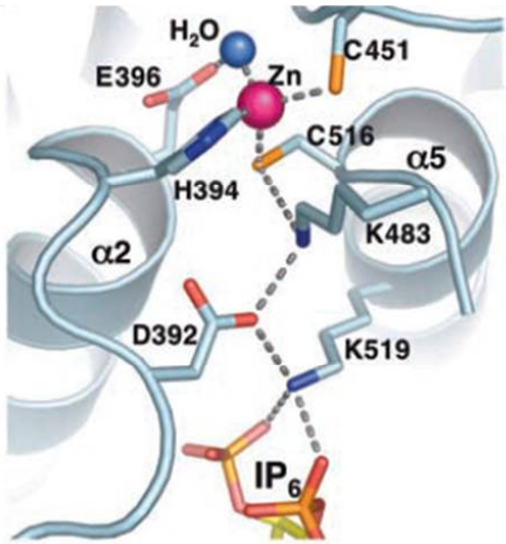

Figure 2-7: Residue interactions of ADAR2s active site. The zinc ion (pink) is coordinated by the residues H394, C451 and C516. The residue E396 coordinates together with zinc the nucleophilic water (blue). The coordination of IP_6 by hydrogen bonds to the catalytic site is indicated, as well. Picture adapted from (41).

It is proposed that the target adenosine for deamination is flipped out of the RNA-helix, a mechanism similar to the methyltransferases and the E488 residue contacts the orphaned counter base (43,44). This idea is supported by an experiment using the 2-aminopurine nucleotide analogue that replaces the substrate adenosine and suggests that the target adenosine is deeply buried inside the active site of the protein by flipping out of the RNA helix (45). Until now, there is no crystal structure of a RNA substrate bound to the deaminase to prove this theory.

There are ongoing discussions if the ADAR proteins are acting as dimers or monomers. According to the findings from Macbeth et al. 2005 (41), there exists no evidence for dimer formation in catalysis, whereas Nishikura postulates a dimer formation (19,46).

2.5 Specificity of ADAR enzymes

ADAR enzymes have a broad range of specificities. It has been observed that substrates of long (>50 bp) perfect dsRNA are non-selectively edited at numerous sites, which means that editing occurs at various adenosines at different frequency (43,47). In contrast, adenosines of RNA duplexes that contain shorter base-paired regions, bulges, mismatches and loops are more selectively edited (4), meaning that only a few or even a single adenosine is edited. Terms have been coined for this: preference and selectivity. Preference is the level of deamination at a specific editing site. The number of adenosines that are deaminated in a certain region of dsRNA, before the reaction stops, is described as selectivity. The generally accepted view is that preference is determined by the deaminase domain and selectivity by the dsRBDs.

The 5'-neighbor preference of hADAR1 and hADAR2, as well as for the deaminase domains alone is U>A>C>>G, whereas 5'-G<u>A</u>N is almost inactive (48). Also a week 3'-neighbor preference of G>C was observed. It is suggested that certain nearest neighbors facilitate the base flipping of the target adenosine over others, determining the editing preferences (4). The idea that the catalytic domain mainly determines the editing preference is supported by an experiments with ADAR1 and ADAR2 lacking the N-terminus (dsRBDs): Both enzymes are able to edit the same adenosines as the full-length proteins *in vitro*, but at different yields (48). Also strategies that use only the deaminase domain for site-selective RNA editing confirm that the editing

preference is mainly determined by the catalytic domain (49,50) (for details see 2.5.1).

The dsRBMs are thought to anchor the protein to the RNA (51) and thereby modulate the selectivity (52). It was observed that the editing reaction stops after editing 50-60% of adenosines in an adenosine-rich context (selectivity) due to instability of the duplex structure. The instability is caused by the introduction of I-U mismatches in the editing process and therefore not being recognized as a substrate (4).

2.5.1 E488Q ADAR2 and E1008Q ADAR1

Kuttan and Bass (53) set up a screening system, which allows the selection of evolved ADAR2 mutants obtained by random mutagenesis of the deaminase domain. Looking specifically for mutants that enable the editing of the difficult codon 5'-GAC they identified amino acid 488 as a hotspot in ADAR2. In particular, changing E488 (glutamic acid) into Q (glutamine) gave rise to a more reactive protein. Amino acid 488 is located on a loop proximal to the active site (the catalytic domain) and is supposed to form salt bridges with basic amino acids of neighboring loops (Figure 2-8).

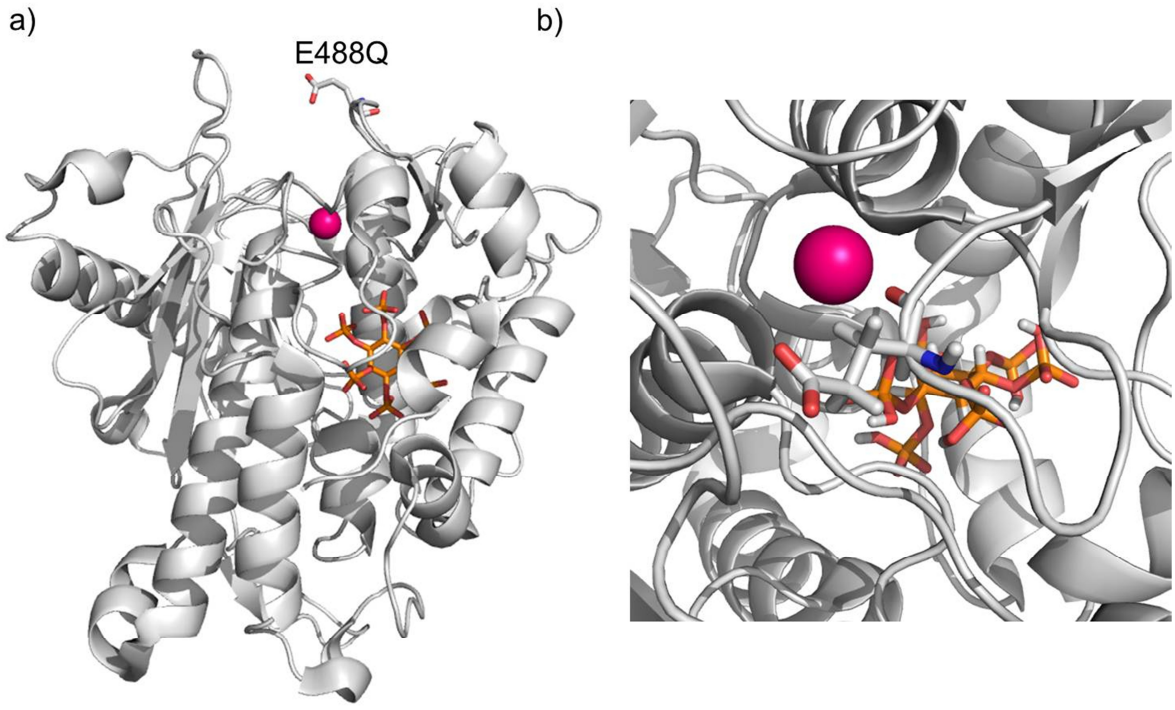

Figure 2-8: Ribbon model of the hADAR2 catalytic domain. Crystal structure of hADAR2 deaminase domain. The zinc ion is represented as a pink sphere. The IP6 is shown in orange. The glutamate residue of position 488 (E488) is located on the proximal loop above the active site. E488

forms salt bridges with basic amino acid of the adjacent loops. The entrance of the mRNA substrate into the RNA binding pocket is influenced by E488. a) General view of the hADAR2 deaminase domain. b) Closer view from above to the RNA binding pocket. PDB ID 1ZY7.

One aspect of this amino acid change from E488 to Q488 is that the entrance of the RNA substrate to the active site might be facilitated by preventing these salt bridges. Another aspect concerns the target adenosine and its orphaned counter base. As mentioned before, it is assumed that the target adenosine is flipped out of the RNA helix to be buried deeply inside the active site of the protein (45), similar to the mechanism of the cytosine-specific DNA methyltransferase M.HhaI (53,54). The glutamine Q237 of M.HhaI, which is surrounded by two glycine residues, is known to penetrate deep into the helix by pushing the target cytosine out of the helix. Furthermore, Q237 is known to contact the orphaned counter base guanine (54-56). Since the residues E488 of ADAR2 and the respective residue E1008 of ADAR1 are both flanked by two glycine residues, as is Q237 of M.HhaI, too, a similar role was hypothesized for E488/E1008 in ADAR2/1 (Figure 2-9).

Figure 2-9: Sequence context of E1008 hADAR1, E488 of hADAR2 and Q237 of M.HhaI.
Figure taken from (54).

E488Q/E1008 is supposed to increase the base flipping mechanism of adenosine, thus increasing the catalytic rate, and to influence the recognition of the orphan base (53). For both ideas indications have been found. The experiment using 2-aminopurine nucleotide analogue that replaces the substrate adenosine supports the base flipping mechanism of the target adenosine, which is deeply buried inside the active site of the protein (45). Wang et al. 2015 (54) investigated the ability of two further point mutations besides Q488 that were shown to be more active than the wild type ADAR, as well. This analysis was performed in the context of ADAR1 (E488Q corresponds to the position E1008 in ADAR1) to edit an adenosine mismatched with a guanosine (A•G), which is usually not accepted. They found the following pattern of preference for the 1008 site chain: E1008R> E1008Q> wt ~ E1008H. They assume

that the preference of the 1008 side chain (or 488 for ADAR2) is dependent on the identity of the counter base, since the amino acid at position 488 or 1008 is supposed to penetrate the RNA duplex and contacts the orphaned counter base.

2.6 Site-directed RNA editing in-vitro and in cell culture

Regarding the genetic code, 12 out of 20 canonical amino acids are potential targets for editing, namely Asp, Glu, Asn, Gln, His, Lys, Arg, Ser, Thr, Tyr, Ile, Met/Start and all three Stop codons (57). Most of these amino acids are essential for enzyme catalysis, posttranslational modification (protein signaling) and protein glycosylation. The potential of a tool that specifically edits certain mRNAs by (site-directed) RNA editing for basic research or medical application is evident. To date two independent approaches have been successfully established.

Since ADARs use protein domains (dsRBDs) to recognize their RNA substrates, it is not easily possible to re-address the ADAR enzymes. The application of a single guideRNA, similar to riboproteins that are addressable by external guideRNAs to a new desired target mRNA, would be most favorable. Such a guideRNA could be easily designed in order to complementary target the mRNA region for site-directed RNA editing.

The first method for site-directed RNA editing was demonstrated by our group in 2012 (57) and overcomes this obstacle. In order to turn hADAR into a guideRNA-dependent enzyme, a SNAP-tag domain (58) replaced the whole N-terminus, including the natural RNA-binding domains, by fusing it to the catalytic domain (Figure 2-10) (49,57). The SNAP-tag domain of the fusion protein reacts specifically with 5'-O-benzylguanine (BG) modified guideRNAs and generates 1:1 covalent conjugates. The guideRNA-deaminase conjugate enables site-directed RNA editing of arbitrary mRNA targets by using a guideRNA complementary to the target mRNA. This artificial editing system was shown to achieve highly selective and efficient repair of nonsense and missense mutations in reporter genes (59). With the SNAP-tag BG-system we could confirm the predicted 5'-neighbor preferences of ADAR2 *in vitro* (49). In this context we demonstrated that by optimization of the guideRNA, every codon is addressable and even less preferred codons, such as 5'-CAG and 5'-GAG, are editable to at least 50% *in vitro* (Figure 2-11). A further study of the SNAP-tag BG-system showed that chemical modifications (2'-O-methylation and PTO) of the

RNA backbone allows fine-tuning of the selectivity of the guideRNA-deaminase conjugate (59). We could demonstrate in this publication that editing in cell culture is possible to 27%. But by newer optimized experiments editing yields up to 50% in cell culture and 70% in *Platynereis* zygotes are obtainable (60).

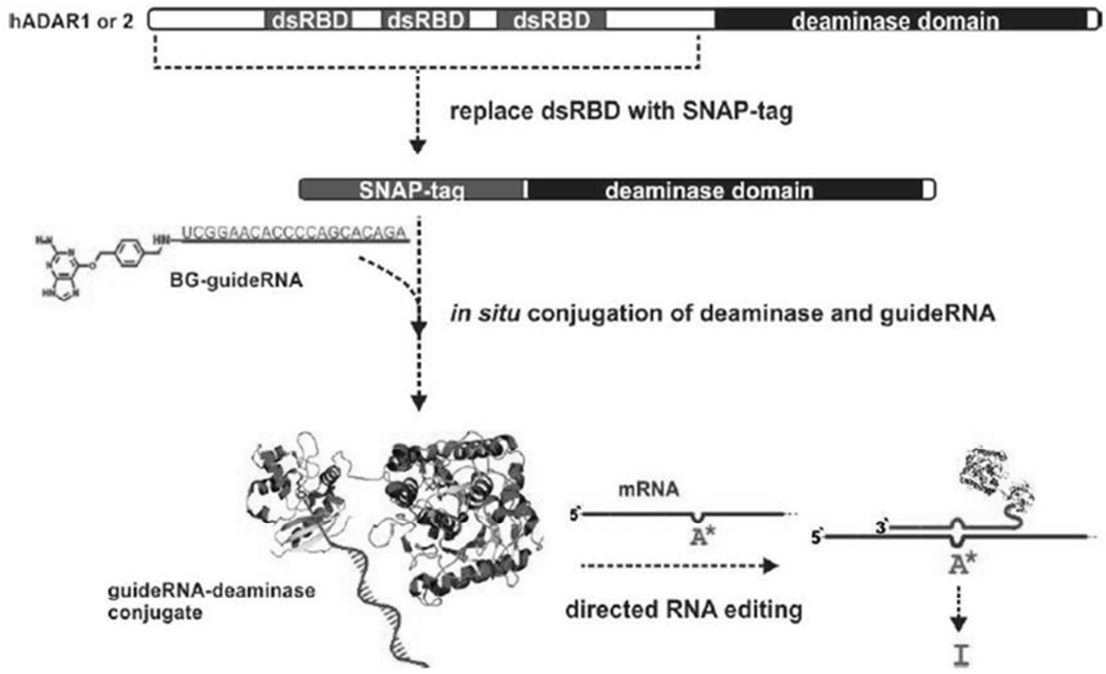

Figure 2-10: Turning hADAR into a guideRNA-deaminase conjugate. The SNAP-tag domain replaces the natural dsRBDs of the N-terminus by fusing it to the catalytic domain. Incubation of the produced and purified fusion protein with a 5'-O-benzyguanine (BG) modified guideRNA generates 1:1 conjugates. This covalent guideRNA-deaminase conjugate enables site-directed RNA editing of arbitrary mRNA targets by using a complementary guideRNA sequence to the target mRNA. Figure taken from (49).

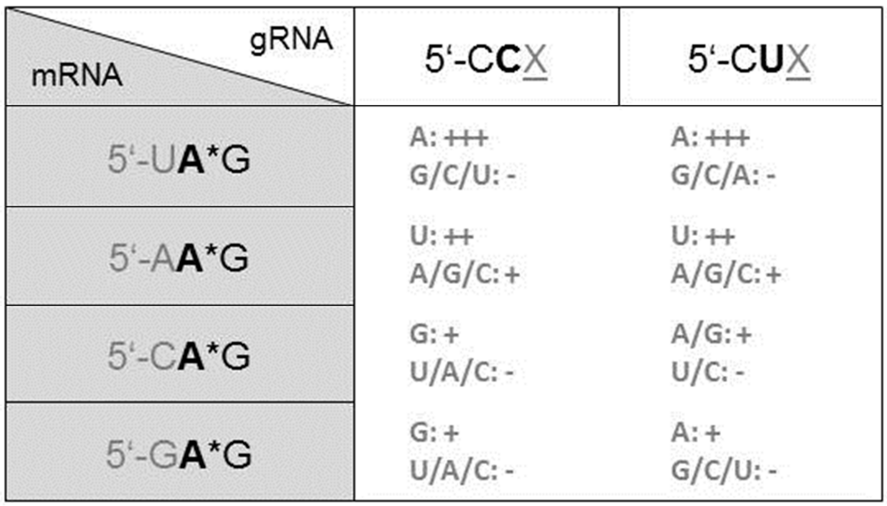

Figure 2-11: Overview of guideRNA preferences of SNAP-ADAR2 for 32 anticodons. SNAP-ADAR2 has for each of the four mRNA codons different guideRNA preferences. The editing level is pictured with four types: - (less than 15% yield), + (15-50% yield), ++ (>50-80% yield) and +++ (>80% yield). These editing yields were obtained at 0.75 mM magnesia. Table adapted from (49).

A second strategy for site-directed RNA editing was established by Montiel-Gonzalez et al. 2013 (50). The deaminase domain of ADAR was fused to a 22 amino acid long lN-peptide (61,62) that recruits the catalytic domain to a boxB-containing guideRNA by a specific RNA-peptide recognition with nanomolar affinity. In contrast to our established SNAP-tag BG-system, the lN-peptide deaminase domain boxB-system is entirely encodable. They achieved 20% editing yield in cell culture.

Both methods have been shown to be promising tools *in vitro*, as well as in cell culture to address different mRNA targets to be specifically edited with high editing yields with medical impact (49,50,57,59).

2.7 Gene therapy methods

Curing human diseases is the desired goal of medicine. While for many diseases their origin is still unknown, so that only the symptoms can be alleviated, however, many diseases for which the genetic background is now known cannot be treated due to a lack of medicinal methods that target the nucleic acid level rather than the protein level. But over the past years a few methods have been established to treat diseases that arise from mutations of the genome and make medical applications on the protein level, such as enzymes or receptors, redundant. Methods targeting the DNA or the RNA level for treating diseases are summarized under the term gene therapy.

The current gene therapy trials are mainly focused on treating cancer, followed by monogenic disorders, infectious diseases and cardiovascular diseases (63). Inheritable monogenic disorders represent, according to Ginn et al 2013 (64), 9% of the current trials of gene therapy. Table 2-1 lists a selection of monogenic disorders that are subject of gene therapy.

Table 2-1: Exemplary list of inheritable monogenic disorders

Monogenic disorders	
Adrenoleukodystrophy	Haemophilia A and B
β-Thalassaemia	Huntingtons's chorea
Cystic fibrosis	Leukocyte adherence deficiency
Galactosialidosis	Sickle cell disease

The applications of genetic engineering are now developing rapidly. In recent years new genome editing technologies to introduce precise changes in a certain genome locus have been developed. Besides these genome editing methods, approaches were established that successfully target the RNA level for intervention. Fomivirsen (trade name Vitravene) was the first antisense drug approved in 1998 to treat cytomegalovirus retinitis. In 2013 Mipomersen (trade name Kynamro) was the next antisense drug that was approved for treating homozygous familial hypercholesterolemia. In 2012 Glybera was the first approved gene therapy treatment in Europe that treats a rare inherited disorder by delivery of an intact copy of the human lipoprotein lipase (65).

2.7.1 Diseases caused by single point mutation

In gene therapy, diseases caused by single point mutations represent attractive study cases to restore normal gene function, since only one nucleobase has to be exchanged. In this PhD project three genes have been studied for the repair of single point mutations that are reported to give defined disease states. Furthermore, these disease phenotypes are directly detectable in cell culture, thus allowing the evaluation of the potential effect of therapeutic RNA editing at the cellular level.

The p53 gene has become one of the most studied tumor suppressor genes. The p53 protein regulates cellular stress responses for example cell cycle arrest and apoptosis

(66). Therefore, cellular transformation and tumorgenesis is prevented by excluding damaged cells from proliferation (66). Mutations of p53 that inactivate its normal function are observed in about 50% of human cancers (66). Essmann and Schulze-Osthoff highlighted in their review from 2012 (66) several mutational hotspots in the p53 gene that result in loss-of-function. These point mutations affect the DNA binding ability or sequence specificity of the transcription factor p53, leading to an ineffective function. Restoring p53 activity in human tumors is the strategy of many anti-tumor therapies (67).

Further examples that were chosen as diseases caused by a single point mutation, were FUS/TLS (Fused in sarcoma/translocated in sarcoma) and PINK1 (PTEN-induced putative kinase 1) gene. Mutations in these genes lead to a distinct phenotype in cell culture, which offers the possibility to prove a successful restored wild type protein not only on the RNA level, but also by visualization in cell culture. FUS/TLS is a DNA/RNA binding protein and was found to be commonly mutated in amyotrophic lateral sclerosis (ALS), a progressive degeneration of motor neurons. Most mutations affect the FUS nuclear localization and lead to an unusual cytoplasmic accumulation of FUS, which is characteristic for ALS (68). Single point mutations such as FUS R521H, which cause abnormal FUS localization, offer an attractive target for gene therapy methods and a simple phenotype read-out on the cellular level.

PINK1 mutations are beside *PARKIN* mutations the most common cause of recessive Parkinson's disease (69). Functional PINK1 kinase activity is a prerequisite for activation of Parkin, a multifunctional E3 ubiquitin ligase. Defective PINK1 kinase activity can be determined either by an alteration of the PINK1 and Parkin protein localization, or by an abnormal mitophagy. The molecular contribution to the Parkinson's disease pathogenesis is still unclear. Single point mutations, such as *PINK1 G309D*, *W437X* or *R407Q* have been identified in Parkinson patients.

2.7.2 Tools for genome editing

Today, several genome editing technologies including zinc finger nucleases (ZFNs), transcription activator-like effector nucleases (TALENs) and clustered regularly interspaced short palindromic repeats (CRISPR)/CRISPR-associated-endonuclease (Cas) are available. These technologies use restriction enzymes to create double

strand breaks (DSBs) in DNA at defined sites based on engineered proteins or guideRNAs that target and specifically bind to the designated genome sequence (63). However, the genome editing tools only provide the site-specific DSB, the knock-in or knock-out relies on endogenous DNA repair pathways: non-homologous end joining (NHEJ) or homology directed repair (HDR) (Figure 2-12). NHEJ repairs DSBs by degradation of 5'-ends and re-ligation of the strand without a homologous DNA template. This mechanism is harnessed to provide knock-out via indel mutation. In contrast, HDR is a precise repair mechanism that requires a homologous DNA template, which serves as a guide for the repair (70). By providing a donor template, which is homologous to the target sequence, the HDR pathway might be used to precisely edit the genome, change nucleotides, include deleted parts or delete inclusion, or even include completely new genes.

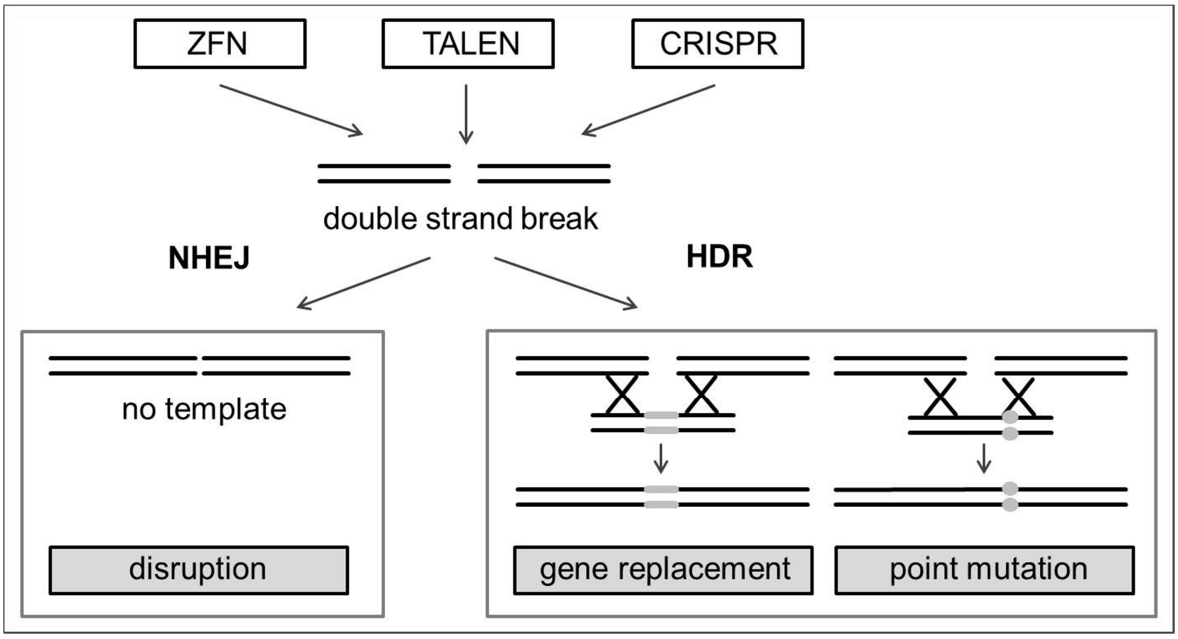

Figure 2-12: Three gene correction approaches exploiting endogenous mechanism of DSB repair.

ZFN, TALEN and CRISPR insert specific double strand breaks (DSBs) in DNA that are corrected by NHEJ (non-homologous end joining) or HDR (homology directed repair). Thereby, genes can either be disrupted (NHEJ), replaced by donor knock-in or equipped with a point mutation.

2.7.2.1 Zink finger nucleases (ZFNs)

The zinc finger nuclease was the first technique developed in 2002 to specifically edit the genome (71,72). The ZFN is formed by fusing two components: a sequence of 3 to 6 zinc finger proteins and the restriction enzyme *Fok*I. Since the *Fok*I only acts on

DNA as a dimer (73), two ZFNs are required that bind to the top and bottom strands of DNA to initiate a DSB (63) (Figure 2-13). The necessity of dimerization improves the specificity and reduces off-target effects. Each zinc finger protein binds only to one specific nucleotide triplet. Therefore, a sequence of up to 36 nucleotides is recognizable. The major advantage of ZFNs over the other two genome editing strategies is that the sequence encoding a ZFN monomer of ~1kb is with careful design not limited by vector capacities. Thinking of current delivery strategies such as adeno-associated viral (AAV) vectors, which restrict an expression cassette to ~4.5 kb, the transfer of ZFN with an AAV vector is possible. The biggest limitations for ZFNs are the limited number of sites that can be targeted and the reported cytotoxicity due to off-target effects by reason of the difficulty to specifically direct ZFNs (63). But a clear advantage of ZFNs in gene therapy lies in the huge experience in application, which has facilitated human clinical trials. The phase I trial in treating HIV-positive patients was successfully completed in 2014 (63).

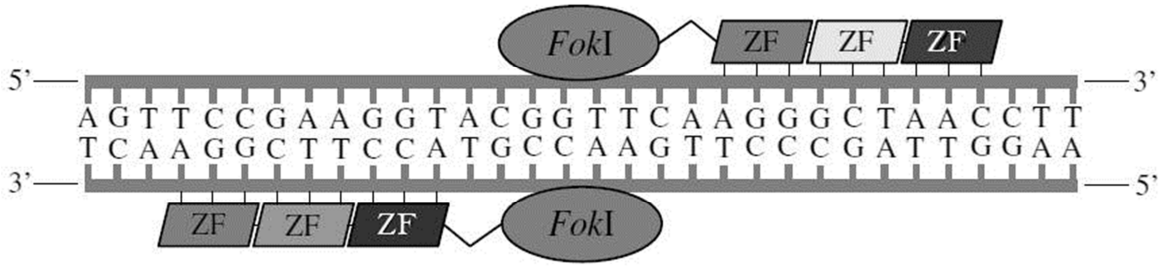

Figure 2-13: Illustration of a zinc finger nuclease (ZFN) pair. The ZFN pair consists of a left and right monomer of up to 6 zinc finger proteins and *Fok*I nuclease. *Fok*I cleaves DNA only when a dimer is formed. Each zinc finger protein recognizes 3 nucleotides. Figure taken from (74).

2.7.2.2 Transcription activator-like effector nucleases (TALENs)

Like the zinc finger proteins, TAL effectors are also DNA binding proteins. Four different TAL effector proteins were found that specifically bind each of the four individual DNA bases (75,76). This 1:1 binding affinity allows the construction of TALE arrays that can bind any DNA sequence with 30 – 40 bp. The TALE array is fused to the same endonuclease as used for ZFN, *Fok*I (Figure 2-14). Advantages of TALENs are reduced cytotoxicity compared to ZFNs (77), bigger flexibility for the design of the 1:1 binding affinity and the rational engineering. However, the size of one TALEN monomer of ~3 kb limits the delivery by AAV vectors, since usage of one pair TALEN monomers would require a big cassette with more than 6 kb. TALENs are

huge proteins with highly repetitive gene sequences. These repetitive gene sequences were reported to limit the choice of vectors that are able to successfully transfer intact TALEN genes into human cells, since highly repetitive DNA sequences are known for their instability during replication, repair and recombination (63). Since TALENs are not that established as ZFNs, no clinical trials have been started yet. But several *in vitro* experiments with the focus on treating inherited monogenic disorders, such as severe combined immunodeficiency or duchenne muscular dystrophy, demonstrate the progress of TALEN technologies (63).

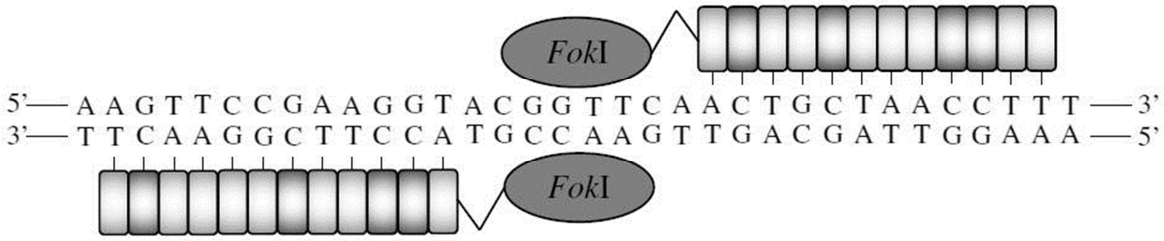

Figure 2-14: Illustration of transcription activator-like effector nucleases (TALENs). One TALE array consists of up to 40 TAL effector proteins, which bind one of the four DNA bases. The TALE array is fused to the *Fok*I nuclease. Simultaneous binding of two TALENs on each strand of the DNA (flanking the target site) leads to a DSB by *Fok*I. Figure taken from (74).

2.7.2.3 CRISPR/Cas9

The youngest technology for specific genome editing is CRISPR/Cas9. This system is derived from a natural defense mechanism against viruses and plasmids in prokaryotes. A ~20 bp fragment of invading DNA is copied at a locus of clustered regularly interspaced short palindromic repeats (CRISPR) (63). This locus is used as a genomic memory for future invasions. Upon subsequent invasion, a corresponding crispr RNA (crRNA) will be transcribed and processed from this locus. A complex of crRNA, endogenous CRISPR-associated endonuclease (Cas) and a separate trans-activating crRNA (tracrRNA) recognizes and binds to the foreign DNA by base pairing. The Cas nuclease creates with the crRNA a DSB in the pathogenic DNA. Thereby, the integration and replication of the pathogens genome is prevented (63) (Figure 2-15).

Researchers modified this natural system in such a way that the Cas9 endonuclease includes a nuclear localization signal for human cells and a single guideRNA can be

designed and applied to target any 20 bp DNA sequence to induce DSB (63,78) (Figure 2-15 b)).

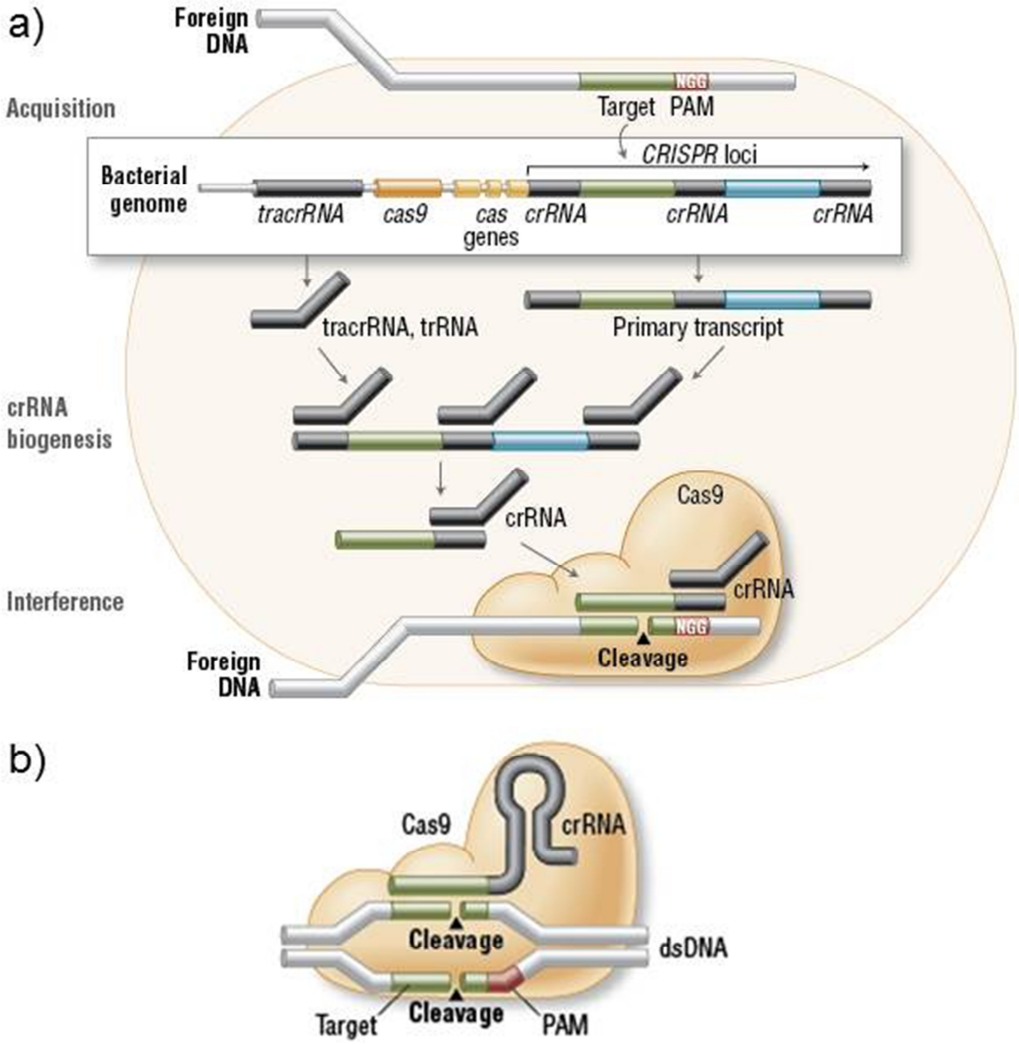

Figure 2-15: Bacterial defense mechanism by Cas9 *in vivo* and genome engineering with Cas9 nucleases. a) The bacterial adaptive immunity copies a ~20 bp fragment of invading DNA at the locus of clustered regularly interspaced short palindromic repeats (CRISP). The corresponding crispr RNA (crRNA) will be transcribed and processes from this locus during its biogenesis. The complex of Cas9 nuclease, crRNA and separate trans-activating crRNA (tracrRNA) recognizes and cleaves foreign DNA. b) Genome engineering with Cas9 nuclease, which specifically cleaves dsRNA with a designed guideRNA that is complementary to the target sequence. Figure from NEB (http://neb.com/tool-andresources/feature-articles/crispr-cas9-and-targeted-genome-editing-a-new-era-in-molecular-biology).

In contrast to ZFNs and TALENs, Cas9 cleaves DNA as a monomer. Thereby, the size of the Cas9 and a single guideRNA is ~4.3 kb, and is intermediate in size (63). A clear advantage of CRISPR/Cas9 is the reduction of costs and time for engineering.

But some cases reported off-target mutagenesis, leading to concerns of limitations in target-specific delivery (63). It was suggested that the CRISPR/Cas9 system may be less specific compared to ZFNs and TALENs due to a shorter targeting sequence. Future research is ongoing to optimize the guideRNA in order to limit off-target effects (79). Since CRISPR/Cas9 is a very young technology compared to the other two genome editing methods it is expected to take time before further progress for applications is achieved. CRISPR/Cas9 was demonstrated to be effective in correcting a mutation in the CFTR gene in patient-derived intestinal stem cells, which suffer from cystic fibrosis. Further proof-of-concept experiments *in vitro* demonstrated the therapeutical potential of the CRISPR/Cas9 technology (63).

2.7.3 Targeting the RNA for gene therapy

Three different approaches that exploit oligonucleotides by base pairing to a complementary target mRNA are used in gene therapy to address the RNA level: RNA interference (RNAi), antisense oligonucleotides (ASO) and splice-switching oligonucleotides (SSO). RNAi and antisense oligonucleotides downregulate the targeted mRNA by inducing enzymatic degradation, whereas SSO block the access of the cellular machinery to the pre-mRNA and mRNA (80) (Figure 2-16).

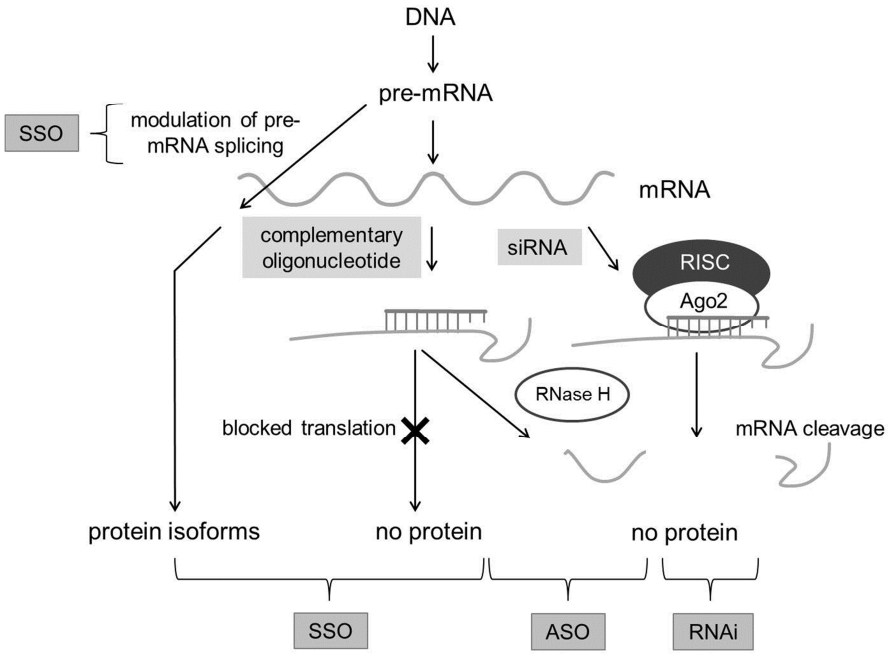

Figure 2-16: Overview of antisense- and RNA interference approaches on the mRNA level.

Antisense oligonucleotides (ASO) induce RNase H mediated cleavage of the target mRNA. RNA interference (RNAi) is induced by siRNA and results in mRNA cleavage. Splice-switching oligonucleotides (SSO) can lead to blocked translation, but their main ability for their application lies on the modulation of pre-mRNA splicing.

Antisense oligonucleotides are typically 20 nucleotides long and consist of a phosphorothioate backbone with flank modifications, which are usually 5 residues of 2'-O-methoxyethyl or 2'-O-methyl residues. These flank modifications improve the stability of the ASO towards exonuclease degradation and enhance its binding to the target mRNA (80). The 10 deoxynucleotide gap between the flank modifications in the gapmer-mRNA duplex is recognized by ribonuclease H (RNase H), which degrades DNA-RNA-duplex (Figure 2-17). The oldest approved antisense drug is Fomivirsen. It simply acts by sterically blocking the translation of viral mRNA and permits the treatment of cytomegalovirus retinitis. In 2013, the first antisense drug that harnesses RNAse H activity, Mipomersen (commercial name Kynamro), was approved for the treatment of homozygous familial hypercholesterolemia.

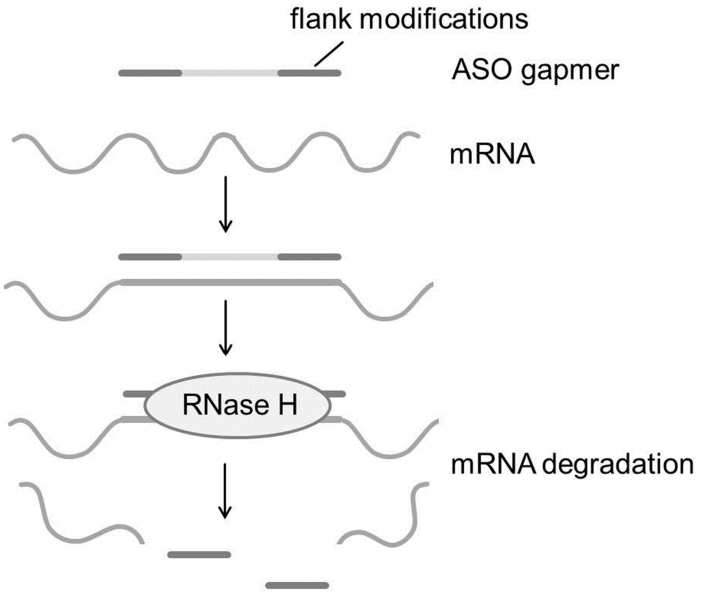

Figure 2-17: Degradation of mRNA by ASO gapmer mediated RNase H cleavage. The antisense oligonucleotide (ASO) consists of a phosphorothioate backbone and modified flanks (ASO gapmer). These flanks are usually 5 nt long 2'-O-methoxyethyl or 2'-O-methyl residues and improve the stability of the ASO towards exonuclease degradation and enhance its binding to the target mRNA. The 10 nucleotide gap between the flank modifications in the gapmer-mRNA duplex is recognized and cleaved by ribonuclease H (RNase H). Picture similar to (80).

Splice-switching oligonucleotides are oligonucleotides that are similar to ASOs but that do not induce RNase H-mediated cleavage of the target mRNA. Instead of mRNA degradation they block the translation of the targeted mRNA. But their main ability is the modulation of pre-mRNA splicing with several outcomes: Redirection of alternative splicing to repair defective RNA and restore protein function or generation of novel proteins with new desired properties (80). The therapeutic application of SSO is currently being tested in clinical trials (phase 3) to modulate splicing for treating duchenne muscular dystrophy and spinal muscular atrophy (Nusiversen). For current state information see http://www.wiley.com/legacy/wileychi/genmed/clinical/.

The RNAi pathway is a natural antiviral defense mechanism present in the cytoplasm. Since cellular RNA is single-stranded, double stranded RNA is interpreted to originate from a virus. This defense mechanism degrades dsRNA that is complementary to single stranded “guide”RNAs. Besides being part of the immune response, RNAi also has the function to regulate gene expression by endogenous miRNAs and was first described in *C. elegans* (81).

The therapeutic potential of RNA interference (RNAi) is that RNAs can be destabilized or that their translation can be down regulated in a sequence-specific manner. The sequence-specificity results from a 21-23 nucleotide dsRNA that is from now on referred to as siRNA. The siRNA is either directly transfected to the cell or derives from a short hairpin RNA (shRNA) coded on a plasmid. This siRNA is incorporated into the RNA-induced silencing complex (RISC), consisting of Dicer and TRBP (the human immunodeficiency virus trans-activating response RNA-binding protein) in mammalian cells (Figure 2-18). Subsequently, RISC forms a complex with Argonaute 2 (Ago2) protein and transfers the siRNA. The incorporated siRNA unwinds; the sense strand is cleaved and released from the complex. The antisense strand of the siRNA directs RISC to the target mRNA. Ago2 cleaves the phosphodiester bond of the target mRNA and releases the resulting fragments, which will be finally degraded resulting in gene silencing (82). The gene silencing is a post transcriptional down regulation of gene expression (knockdown) and does not eliminate the genetic function completely. Therefore, the functional RNA or protein is reduced to lower levels.

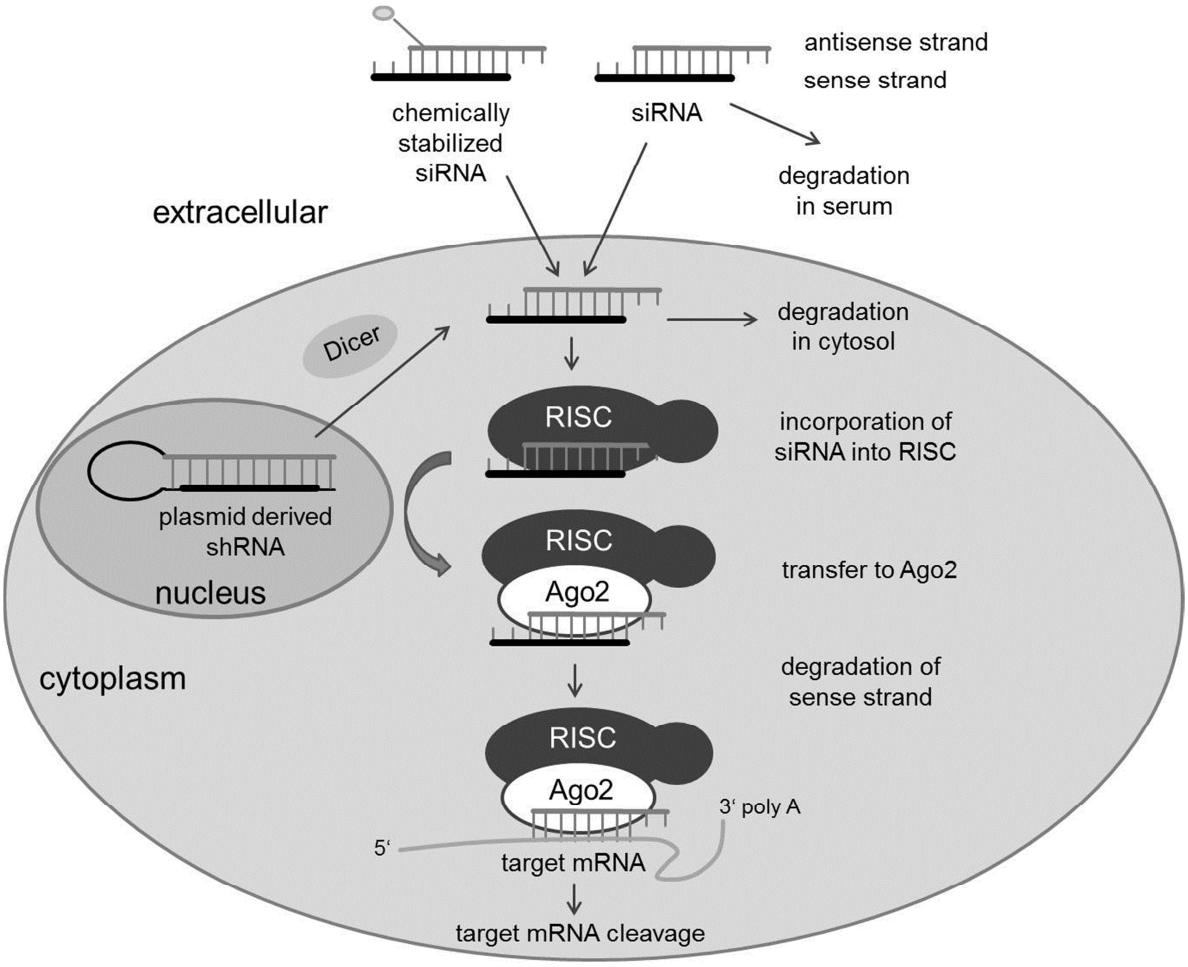

Figure 2-18: Therapeutic potential of the RNA interference (RNAi) pathway. The RNAi pathway is induced by a 21-23 nt siRNA (double stranded RNA) that is incorporated into the RNA-induced silencing complex (RISC). RISC consists of Dicer and TRBP (the human immunodeficiency virus trans-activating response RNA-binding protein) in mammals. RISC forms a complex with Argonaute 2 (Ago2) protein and transfers the siRNA. Ago2 cleaves the sense strand of the siRNA. The antisense strand directs the RISC Ago2 complex to the target mRNA, which will be cleaved. The chemical modification of siRNAs increases the stability in serum and cytoplasm, leading to a more efficient gene silencing.

According to various *in vitro* and *in vivo* studies, nearly every human disease that is based on a toxic gain-of-function genetic lesion or an overexpression of disease causing genes represents a potential target for RNAi-mediated therapy (82). Cancer is in the focus of RNAi-based therapy, since oncogenes, mutated tumor suppressor genes and other genes are potential important targets for gene silencing. RNAi-based therapeutics offer both a low risk of side-effects, and the possibility to simultaneously target multiple genes. This is extremely valuable for cancer therapy, since different cellular pathways are involved in tumor progression. Also the fact that the siRNA-

loaded RISC can be used for multiple rounds of gene silencing is very attractive for therapy (83). A further application of RNAi is the determination of loss-of-function phenotypes in functional genomics.

In recent years a lot of studies to improve the siRNA stability and its design have been performed to achieve higher target efficiencies and reduced unspecific effects of siRNAs and interferon-responses (84). Several RNAi-based drugs are in clinical trials. Due to the poor intracellular delivery of siRNA, the majority of clinical trials are locally applied. One of them targets the vascular endothelial growth factor to treat age-related macular degeneration. Other diseases, such as hypercholesterolemia, Huntington's disease or cancer were demonstrated to be effectively treated *in vivo* (82) and are undergoing different phases of clinical trials, as well.

2.7.4 Genome editing or antisense-strategies?

The choice of method to be used, either genome editing knockout, RNAi-mediated knockdown, or SSO depends on the therapy aims. With all strategies the expression level can be changed. If changes of the genetic code are undesirable, RNAi-mediated gene silencing or SSO are the methods of choice. Those strategies are more preferable over genome editing, if a genetic function has to be temporarily reduced. Either by transfection of antisense oligonucleotides, siRNAs or a stable integration of shRNAs into the genome for siRNA delivery, restoring of the normal genetic function remains possible. Usage of inducible promoters for shRNA expression, such as tetracycline inducible U6 promoters (85), enables the gene expression to be knocked down at desired periods and to switch it back. The need to administer or express a single guideRNA only to knock down a specific gene product, is another advantage compared to the genome editing methods. A further reason to choose RNAi-mediated knockdown is given if a full suppression of the gene function would harm the organism. Usually, RNAi does not completely switch off gene function. Many gene defects or diseases are simply caused by an overexpression of certain RNAs or proteins. But the complete loss of its function would lead to cell death. With SSO alternative splicing of pre-mRNA can be modulated and thereby repair defective RNA to restore normal protein function, if the defective protein or disease orgins from incorrect splice variants. The biggest limitation of RNA therapeutics is the poor

intracellular uptake. Furthermore, the knock-in of single nucleotides or genes is not possible with RNAi or antisense-strategies.

Genome editing technologies are preferable for the creation of genetic null alleles or a permanent elimination of genetic function, shortly described as knock-out tools. Since it is not possible to control the repair pathway of a DSB (NHEJ or HDR) and NHEJ occurs much more frequently than HDR. These technologies require the isolation of correct single clones due to the inhomogeneity of cells and can be very time consuming. Also off-target editing cannot be completely excluded. Off-target editing displays a huge risk, if other genes are affected and not repaired for example by providing a homologous template.

2.8 Aims of this study

Significant progress has been made in gene therapy over the past years: the first gene replacement therapy has been approved in 2014, and more and more antisense drugs have finished clinical phase 3 studies successfully. In particular, toxicity and delivery of nucleic acid-based drugs has improved and is still under development. However, the need to repair disease causing point mutations efficiently, rationally and safely remains unmet. The RNAi-based knockdown offers an attractive platform for gene therapy applications, such as down-regulation of mutated tumor suppressor genes. However, RNAi and other antisense technologies other than SSO are limited to down-regulation of proteins. For many diseases the restauration of protein function is required, which might get lost by point mutations, deletions or insertions. Genome editing or gene replacements represent potential approaches, which are under development. Correcting gene function at the RNA level is attractive, since wild type function could be restored without manipulation of the genome and thus avoiding ethical issues (germline intervention) and safety concerns. With respect to the latter, correction of the RNA would make off-target effects less probable and less severe compared to the correction of the genome, because the manipulation is reversible and tunable, and this should also apply for potential adverse effects, which anyway can be expected to be less dramatic in phenotype for off-target editing than for off-target DNA scission. Furthermore, at the RNA level gene function can be corrected without interfering with the endogenous regulation via transcription, splicing and transport. Current gene therapy is limited to non-toxic alterations, since only an

irreversible shift between two isoforms can be performed. Therefore, manipulation at the genome level typically gives an irreversible and complete switch to a new protein isoform. However, both, the completeness or the irreversibility of the intervention can cause severe adverse effects or lethality and may thus be inaccessible at the genome level. In contrast, by manipulation at the RNA level it is well conceivable to actively fine-tune the stoichiometry of protein isoforms and to perturb the balance of isoforms temporarily. Thus, manipulation at the RNA level will complement genome editing strategies and make interventions possible that are inaccessible with the latter.

Two systems are under development that address the RNA level based on A-to-I RNA editing (2.6). Our group has already successfully established the SNAP-tag BG-system for site-specific RNA deamination with a guideRNA that steers the deaminase activity to a specific target at the mRNA (49,57,59,60). The other strategy developed by Montiel et al. (50), uses a genetically encodable editing system with an engineered ADAR deaminase. The first strategy is limited by the need to repetitively administer the chemically synthesized guideRNA, which is not genetically encodable by principle. Both strategies, the first and the latter, require the ectopic expression of artificial, engineered proteins.

The aim of this thesis was to develop a fully genetically encodable editing system that harnesses the naturally occurring wild type hADARs for site-specific RNA editing in order to overcome the limitations of the two editing strategies described above. For this, a guideRNA had to be constructed that directs endogenous hADARs to any target mRNA in *trans*. This guideRNA has to fulfill two tasks: a) specific binding of the target mRNA region via RNA duplex formation, and b) recruitment of hADARs to elicit the deamination reaction. Previous experiences in designing an optimal anticodon to achieve high editing yields of the target adenosine of choice have been taken into account (49).

The new system was supposed to be tested and analyzed at different levels. Firstly, the developed guideRNA working in *trans* should be tested in a PCR reaction tube to understand the molecular prerequisites for efficient RNA editing. This included various guideRNA architectures and various ADAR variants that differ in their dsRNA binding domains. Secondly, the editing system should be optimized in cell culture. Issues were the choice of promotor for expression of the guideRNA and the other components, stabilization of the guideRNA, variation of guideRNA architecture,

editing times and ADAR variants. Third, RNA editing should be applied for the proof-of-concept repair of a disease causing point mutation. And finally, first experiments for site-directed RNA editing were performed in a living animal, *Platynereis dumerilii*.

3 Material and methods

3.1 Material

3.1.1 Chemicals

All chemicals and solutions used in the following experiments are listed in Table 3-1 with their source of supply. The chemicals had an analytically pure quality.

Table 3-1: Used chemicals and solutions

Chemicals	Supplier
Agarose NEEO	Roth
Agarose HR Plus	Roth
Ampicillin	Life Technologies
Ammoniumsulfat	Roth
Borsäure	Roth
Brillant Blau G 250	Roth
Bromphenolblau Na-Salz	Roth
Dextran, Rhodamine B	Life Technologies
Dimethylsulfoxid	Roth
1,4-Dithiothreit	Roth
DMEM	Life Technologies
dNTP solutation mix 40x	NEB
EDTA (Ethylendiamintetraacetat)	Roth
Ethanol	Sigma-Aldrich
FBS	Life Technologies
Ficoll®400	Roth
Formaldehydlösung 37%	Th.Geyer
Glucose Monohydrat	Roth
Glycerol 86%	Roth
Kaliumchlorid	Roth
LB-Agar (Lucia/Miller)	Roth
LB-Medium (Lucia/Miller)	Roth
Lipofectamine 2000	Life Technologies
L-Trypophan	Roth
LMW-SDS marker	LMW-SDS marker
Opti-MEM®	Life Technologies

Chemicals	Supplier
SDS (Sodiumdodecylsulfat)	Roth
Spermidine	Roth
Luciferase Assay Reagent	Promega
Luciferase Cell Culture Lysis 5x Reagent	Promega
Penicillin-Streptomycin 100x	Life Technologies
Phusion HF Buffer 5x	NEB
Potassium acetate	Roth
rNTPs mix (25 mM each)	NEB
Roti-load	Roth
Rotiphorese SDS-PAGE 10x	Roth
Roti-Safe	Roth
Rotiphorese® Sequenziergel Konzentrat	Roth
Rotiphorese® Sequenziergel Puffer-Konzentrat	Roth
Rotiphorese® Sequenziergel Verdünner	Roth
Rhodamine-labeled Dextran (10 000 MW)	Sigma-Aldrich
Spermidin	Roth
ThermoPol Reaction Buffer 10x	NEB
TRIS PUFFERAN®	Roth
Trypanblau 5x	Biochrom
Trypsin/EDTA	Biochrom
Water (Nuclease free)	VWR
Yeast N2 base	Roth
2-Log DNA Ladder (0.1-10.0 kb)	NEB
6x Loading Dye	NEB

3.1.2 Consumption items

All consumption items used in the following experiments are listed in Table 3-2 with their source of supply.

Table 3-2: Consumption items

Name	Supplier
15 ml 10 kDa MWCO Amicon® Ultra centrifugal filter	EMD Millipore
0.5 ml 30 kDa MWCO Amicon® Ultra centrifugal filter	EMD Millipore
HiTrap™ Heparin columns	GE Healthcare
Femtotips II microcapillarie	Eppendorf
Roti®-PAGE Gradient (4-20 %)	Roth
SDS-PAGE gradient gel (4-20%)	Roth

3.1.3 Equipment

The equipment, which was used in the following experiments, is listed in Table 3-3.

Table 3-3: Equipment

Name	Supplier
ARKTIK Thermal Cycler	Thermo Scientific
Axiovert 200M microscope	Zeiss
AxioCam HRc microscope camera	Zeiss
Axiovert 40 CL inverted microscope	Zeiss
Axio Observer.Z1	Zeiss
Axio Imager Z1 widefield fluorescence microscope	Zeiss
Cyclo 2 water pump	Roth
Femtojet express microinjector	Eppendorf
FLUOstar OPTIMA	BMG LABTECH
French press	-
LSRII	BD (Becton, Dickinson and Company)
Neubauer counting chamber	Roth
micromanipulator	Luigs and Neumann
Badcontroller V cooling system	Luigs and Neumann
MIKRO 120	Hettich Zentrifugen
MIKRO 220 R	Hettich Zentrifugen
POWER PAC 300	BIO-RAD

Name	Supplier
POWER PAC Basic	BIO-RAD
ThermoMixer C	Eppendorf
Vortexer	neoLab

3.1.4 Media, buffers and solutions

The composition of media, buffers and solutions used for the experiments of this work can be seen in Table 3-4, Table 3-5 and Table 3-6.

Table 3-4: Media and their composition

Media	Component	Concentration
DMEM+FBS+PS	DMEM	1x
	FBS	10%
	Penicillin/streptomycin	1%
LB-Media	LB-Media	25g/L
LB-Agar	LB-Agar	40g/L
Yeast growth medium (SD-CAA+Trp)	Glucose	22 g/L
	Yeast N2 base	6.7 g/L
	Na_2HPO_4	5.4 g/L
	$NaH_2PO_4 \cdot 2H_2O$	9.7 g/L
	CAA	5 g/L
	Tryptophan	0.04 g/L
	Inositol	0.02 g/L
Yeast induction medium (SG-CAA+Trp)	Galactose	20 g/L
	Raffinose	5 g/L
	Yeast N2 base	6.7 g/L
	Na_2HPO_4	5.4 g/L
	$NaH_2PO_4 \cdot 2H_2O$	9.7 g/L
	CAA	5 g/L
	Tryptophan	0.04 g/L
	Inositol	0.02 g/L

Table 3-5: Buffers and their composition

Buffer	Component	Concentration
Elution buffer (for protein production)	Tris-HCl	20mM
	NaCl	100mM
	Glycerol	5%
	Imidazole	400mM
		pH 8.0

Buffer	Component	Concentration
5x isothermal reaction buffer (ISO)	PEG-8000	25%
	Tris-HCl pH 7.5	500 mM
	$MgCl_2$	50 mM
	DTT	50 mM
	dNTPs	1 mM each
	NAD	5 mM
Gibson-mix	5x ISO buffer	320 µl
	T5 Exonuclease (10 U/µl)	0.64 µl
	Q5® Hot Start High-Fidelity DNA Polymerase (2 U/µl)	20 µl
	Taq DNA Ligase (40 U/µl)	160 µl
		Ad 1.2 mL store at -20°C
Yeast Lysis buffer	Tris-HCl	20 mM
	NaCl	750 mM
	Glycerol	5%
	Imidazole	10 mM
	b -ME	1 mM
		pH 8.0
10x transcription buffer	Tris-HCl pH 8.1@25°C	400 mM
	Spermidine	10 mM
	$MgCl_2$	220 mM
	Triton-X-100	0.1 mM
10x RNAPol Reaction buffer NEB	Tris-HCl	400 mM
	$MgCl_2$	60 mM
	Spermidine	20 mM
	Dithiothreitol	10 mM
	pH 7.9 @ 25°C	
10x RT editing buffer	Tris	125 mM
	Tris-HCl	125 mM
	KCl	750 mM
	DTT	20 mM
RNA loading buffer 5x	TBE	5x
	Urea	8.3 M
	Ficoll without bromphenolblue	small spatula
RNA loading buffer	Formamide	95%
	Bromphenol blue	0.025%
Washing buffer 1 (W1)	Tris-HCl	20mM
	NaCl	750mM
	Glycerol	5%
	Imidazole	10mM
		pH 8.0
Washing buffer 2 (W2)	Tris-HCl	20mM
	NaCl	300mM

Buffer	Component	Concentration
	Glycerol	5%
	Imidazole	10mM
		pH 8.0
Washing buffer 3 (W3)	Tris-HCl	20 mM
	NaCl	100mM
	Glycerol	5%
	Imidazole	10mM
		pH 8.0

Table 3-6: Solutions and their composition

Solution	Component	Concentration
Coomassie staining solution	Coomassie Brilliant Blue G-250	0.2 mg/mL
	$Al_2(SO_3)_4$	50 mg/mL
	100% Ethanol	10%
	85% H_3PO_4	2%
PBS 10x	NaCl	8 g
	KCl	0.2 g
	Disodium hydrogen phosphate (Na2HPO4)	1.42 g
	Potassium dihydrogen phosphate (KH_2PO_4)	0.27 g
		Ad to 1 L
1x TAE	TRIS-base	4.84g/L
	acetic acid	1.14 mL/L
	EDTA, pH 8.3	1mM
1x TBE	TRIS-base	10.8 g/L
	boric acid	5.58 g/L
	EDTA, pH 8.0	2 mM

3.1.5 Commercially available kits

Table 3-7 lists all applied kit-systems that were used in the following experiments.

Table 3-7: Kit-systems and source of supply

Kit-system	Supplier
Frozen-EZ Yeast Transformation II Kit™	Zymo Research
NucleoBond® Xtra Midi	Machery-Nagel
Nuclespin® Plasmid	Machery-Nagel
RNeasy® MinElute® Clean-up Kit	QIAGEN
Rnase-Free Dnase Set	QIAGEN

Kit-system	Supplier
NucleoSpin® RNA Plus	Machery-Nagel
RNeasy® Mini Kit	QIAGEN

3.1.6 Enzymes

Table 3-8: Enzymes and their source of supply

Enzyme	Supplier
Antarctic Phosphatase	NEB
DNase I	NEB
Taq DNA Polymerase	NEB
Q5® Hot Start High-Fidelity DNA Polymerase	NEB
Phusion® High-Fidelity DNA Polymerase	NEB
T7 RNA Polymerase	NEB
Murine RNase Inhibitor	NEB
M-MuLV Reverse Transcriptase	NEB
Taq DNA Ligase	NEB
AgeI-HF	NEB
AscI	NEB
BamHI-HF	NEB
BglII	NEB
Eco53KI	NEB
HindIII-HF	NEB
SalI-HF	NEB
SapI	NEB
XbaI-HF	NEB
XhoI	NEB
NdeI	NEB

3.1.7 Plasmids and constructs

Table 3-9: Plasmids

Internal plasmid name	Vector backbone	Insert	Reference
pRJ114	pGAL1_416	Empty vector	Professor Jansen
pTS7	pMG211	Stop417 luciferase	This study
pTS8	pMG211	Stop66 eCFP	Provided by group
pTS9.1	pGAL1_416	hADAR2	Provided by group
pTS11.2	pMG211	HH-R/G-stop66gRNA-CCA	Provided by group
FUS	pCMV-HA	FUS wt	Professor Kahle
Pink	pcDNA6v5 Hismyc	PINK W437X	Professor Kahle
pTS32	pMG211	FUS wt	This study
pTS39	pMG211	R521H FUS	This study
pTS49.2	p*Silencer*	R/G-gRNA stop eCFP CCA 16 nt	This study
pTS55.2	pGAL1_416	short-hADAR2	This study
pTS56	Biosystvector	hADAR1	Obiosystem
pTS57	pcDNA3.1	hADAR2	This study
pTS58	pcDNA3.1	WT-eGFP	Provided by group
pTS59	pcDNA3.1	W58X-eGFP	Provided by group
pTS63.1	pcDNA3.1	E396A hADAR2	This study
pTS67	p*Silencer*	R/G-gRNA W437X PINK CCA 16 nt	This study
pTS68	p*Silencer*	R/G-gRNA W58X eGFP CCA 16 nt	This study
pTS71	pMG211	mRNA+R/G-CCA	This study
pTS72	pcDNA3.1	SNAPf-ADAR2	Provided by group
pUC57	pUC57	Empty vector	Dr. Gáspár Jékely
pTS82	pUC57	eGFP	This study
pTS83	pUC57	W58X eGFP	This study
pTS86	p*Silencer*	R/G-gRNA W58X eGFP CCA 16 nt with BoxB	This study
pTS87	p*Silencer*	R/G-gRNA W58X eGFP CCA 18 nt with BoxB	This study
pTS88	p*Silencer*	R/G-gRNA W58X eGFP CCA 20 nt with BoxB	This study
pTS89	pMG211	E488Q ADAR2	This study
pTS90	pcDNA3.1	E488Q ADAR2	This study
pTS93.1	p*Silencer*	R/G-gRNA W58X eGFP CCA 25 nt with BoxB	This study
pTS94.1	p*Silencer*	R/G-gRNA W58X eGFP CCA 29 nt with BoxB	This study

Internal plasmid name	Vector backbone	Insert	Reference
pTS101.1	pUC57	ADAR2	This study
pTS115	pMG211	Luciferase	Provided by group
pTS116.8	pcDNA3.1	W417X luciferase	This study
pTS121.1	p*Silencer*	R/G-gRNA W417X luciferase CCA 16 nt with BoxB	This study
pTS137.2	pcDNA3.1puro	ADAR2 and 4x U6+R/G-gRNA W58X eGFP 16 nt	This study
pTS139	pMG211	HH-R/G-W58X-eGFP-gRNA-CCA	This study
pTS155.2	p*Silencer*	R/G-gRNA W58X eGFP CCA 16 nt with BoxB position 3	Provided by group
pTS156.1	p*Silencer*	R/G-gRNA W58X eGFP CCA 16 nt with BoxB position 4	Provided by group
pTS157.1	p*Silencer*	R/G-gRNA W58X eGFP CCA 16 nt with BoxB position 5	Provided by group
pTS164.1	p*Silencer*	RG gRNA PINK R407Q CUA editing position 6	Provided by group
pTS169	p*Silencer*	R/G-gRNA W417 luciferase CCA 16 nt with BoxB position 5	This study
pTS170	p*Silencer*	R/G-gRNA W417 luciferase CCA 16 nt with BoxB position 8	This study
pTS171	p*Silencer*	R/G-gRNA W417 luciferase CCA 16 nt with BoxB position 9	This study
pTS181	p*Silencer*	R/G-gRNA R175H p53 GCG 16 nt with BoxB position 8	This study
pTS182	p*Silencer*	R/G-gRNA R175H p53 GUA 16 nt with BoxB position 8	This study
pTS185	p*Silencer*	R/G-gRNA W58X eGFP CCA 16 nt with BoxB position 7	Provided by group
pTS186	p*Silencer*	R/G-gRNA W58X eGFP CCA 16 nt with BoxB position 8	Provided by group
pTS187	p*Silencer*	R/G-gRNA W58X eGFP CCA 16 nt with BoxB position 9	Provided by group
pTS188	p*Silencer*	R/G-gRNA W58X eGFP CCA 16 nt with BoxB position 10	Provided by group
pTS193	p*Silencer*	R/G-gRNA W417 luciferase CCA 16 nt with BoxB position 7	Provided by group
pTS194	p*Silencer*	R/G-gRNA W417 luciferase CCA 16 nt position 6	Provided by group
pTS195	p*Silencer*	R/G-gRNA W417 luciferase CCA 16 nt position 8	Provided by group
pTS203	p*Silencer*	R/G-gRNA W58X eGFP CCA 16 nt position 8	Provided by group
No. 289 pEntry	pEntry	wt p53	Frank Essmann

Internal plasmid name	Vector backbone	Insert	Reference
p53α			
pTS217	pcDNA3.1	R175H p53	This study
pTS241	p*Silencer*	R/G-gRNA R282Q p53 CCG 16 nt position 8	Provided by group
pTS269.4	pcDNA3.1	R175H & R282Q p53	Provided by group
pTS284	pcDNA3.1	ADAR1	This study

3.1.8 Strains

Table 3-10: Strains and source of supply

Strain	Supplier
E.coli XL1 Blue	Gift from ETH Zürich
YVH10	Gift from ETH Zürich
HEK293T	DSMZ
HeLa ACC-57	DSMZ
Flp-In T-REx	AG Jung

3.1.9 Oligonucleotides

The following table lists oligonucleotides used for plasmid generation, PCR and klenow reaction.

Table 3-11: Oligonucleotides and their sequences

Internal number	Name	5´-3´sequence
1	seq_egfp_end_b	CAGCGGTGGCAGCAGCCAAC
5	Stop417_Luci_fw	CAAGGACGGCTAGCTGCACAGC
6	Stop417_Luci_bw	GCTGTGCAGCTAGCCGTCCTTG
13	ADAR2_bamDES_fw	GGCATCCAGGGTTCCCTGCTCAG
14	ADAR2_bamDES_bw	CTGAGCAGGGAACCCTGGATGCC
52	antiStop66_ssDNA	ATCTGTGCTGGGGTGTTCCGAT
83	Xba1_ADAR2_fl_fw	GCTCTAGAATGGATATAGAAGATGAAGAAAACATG
84	ADAR2_H6_Sal1_b	CGGCGTCGACTATTAATGGTGATGGTGATGGTGGGGCGTGAGTGAGAACTGGTC
87	seq_post_T7p_fw	GCGGATAACAATTCCCCTCTAG
92	Stop66c-Xho1_bw	GCCTCGAGACTCTGTGCTGGGGTGGTGGG

Internal number	Name	5´-3´sequence
94	p416Gal_Galp_fw	GCGCAATTAACCCTCACTAAAGG
118	ADAR2-_Xba1_bw	CCTCTAGAGGGCGTGAGTGAGAACTGGTC
121	BamH1_Koz_GFP_fw	GCGGATCCACCATGGCTAGCAAAGGAGAAGAACTC
122	GFP_Xba1_bw	CCTCTAGAGCCGGATTTGTATAGTTCATCCATGCC
125	PCR_EagI_pMG211	GGTCAGGCCCAGGTTCTCCG
126	T7p_clon_fw	CGCGAAATTAATACGACTCACTATAGG
127	Fus_A_G_Xho1_bw	GGCTCGAGTTAATACGGCCTCTCCCTGYGATCCTG
128	Nde1_Fus_fw	CCGGCATATGGCCTCAAACGATTATACCC
133	FUS_H521_Xho1_bw	GGCTCGAGTTAATACGGCCTCTCCCTGTGATCCTG
139	RG15nt_H521-c_b	GACAGGATCGCAGGGGTGGGATACTATAACAAC
140	RG15nt_H521-u_b	GACAGGATCACAGGGGTGGGATACTATAACAAC
141	RG16nt_H521-c_b	AGACAGGATCGCAGGGGTGGGATACTATAACAAC
142	RG16nt_H521-u_b	AGACAGGATCACAGGGGTGGGATACTATAACAAC
144	BGH backward	CTAGAAGGCACAGTCGAGGC
152	RG16nt_FusAUG_b	AGACAGGATTACAGGGGTGG
157	RGLuci16nt_CCA_b	AAGGACGGCTGGCTGCGTGGGATACTATAACAAC
169	RG_E66_GCCbw	ACTCTGTGCCGGGGTGGTGG
170	RG_E66_AUCbw	ACTCTGTGCTAGGGTGGTGG
171	RG_E66_CCCbw	ACTCTGTGCGGGGGTGGTGG
176	pSilencer_Sap1_fw	CCGCCTTTGAGTGAGCTGATACC
177	RG_Stop58_Hind3_bw	GCCAAGCTTTCCAAAAAACCTGTTCCGTGGCCGAGTGGGAT ACTATAACAAC
178	XbaI_flA2_dsR2_fw	CTCTAGAATGGATCCGAGTGGGAAGAATC
187	BamHI_Koz_flA2fw	GCGGATCCACCATGGATATAGAAGATGAAGAAAACATG
201	RG_PINKW437_TAGb	GCTGATGCCTGGGCAGGTGGGATACTATAACAAC
202	ADAR2_E396A_fw	GACTGCCATGCAGCAATAATATCTC
203	ADAR2_E396A_bw	GAGATATTATTGCTGCATGGCAGTC
204	SilHind3PinkAmb_bw	GCCAAGCTTTCCAAAAAAGGCTGATGCCTGGGCAGGTGGGA TACTATAACAAC
214	W58xgRNATeil1	GCCCTCTAGACCTGTTCCGTGGCCGAGTG
215	W58xgRNABoxBT2	CCAAGCTTTCCAAAAAAGGGCCCTCTTCAGGGCCCTCTAGA CCTGTTCCGTGG
216	Asc1_GFP_fw	ATGGCGCGCCTAGCTAGCAAAGGAGAAGAACTC
217	Age1_GFP_bw	TAACCGGTTTTGTATAGTTCATCCATGCCATG
229	18nt_W58X	GTCTAGATGCCTGTTCCGTGGCCGAGTGGG
230	20nt_W58X	GTCTAGAACTGCCTGTTCCGTGGCCGAGTGG
232	flADAR2_E/Qfw	GAGTCTGGTCAGGGGACGATTCCAGTGC

Internal number	Name	5´-3´sequence
233	flADAR2_E/Qbw	GCACTGGAATCGTCCCCTGACCAGACTC
234	25nt_W58X	GTCTAGAGGCAAACTGCCTGTTCCGTGGCCGAGTGG
235	29nt_W58X	GTCTAGATACTGGCAAACTGCCTGTTCCGTGGCCGAGTGG
258	AscI_fullADAR2fw	ATGGCGCGCCTAGATATAGAAGATGAAGAAAACATGAG
259	AgeI_fullADAR2bw	TAACCGGTGGGCGTGAGTGAGAACTGG
264	RG_W58X_16cca	CCTGTTCCGTGGCCGAGTGGGATACTATAACAAC
299	BamHI_Koz_Luci_fw	GCGGATCCACCATGGAAGATGCCAAAAACATTAAGAAG
300	XbaI_Luci_bw	CCTCTAGATTACACGGCGATCTTGCCGC
304	XbaI_RGLuci_16nt	CCCTCTAGAAGGACGGCTGGCTGCG
359	1.Gibson_RGfw	ATTTCCCCGAAAAGTGCCACCTGACGTCGACGGATCGGGAC GACGGCCAGTGCCAAGCTTTCC
360	2.Gibson_RGbw	CCTCGTAAATATAGGTCACTCGTGCCCAGGATAATTAAGGCC CAGTGGAAAGACGCGCAGGC
361	3.Gibson_RGfw	CCTTAATTATCCTGGGCACGAGTGACCTATATTTACGAGGCG ACGGCCAGTGCCAAGCTTTCC
362	4.Gibson_RGbw	GCATGGAGGAACATTGTACACATGAACAAAGGTGCTTGCGC CCAGTGGAAAGACGCGCAGGC
363	5.Gibson_RGfw	CGCAAGCACCTTTGTTCATGTGTACAATGTTCCTCCATGCCG ACGGCCAGTGCCAAGCTTTCC
364	6.Gibson_RGbw	CGTAAATGGCACCGTCGAATTATGAGGCCCAGATATGGCGC CCAGTGGAAAGACGCGCAGGC
365	7.Gibson_RGfw	CGCCATATCTGGGCCTCATAATTCGACGGTGCCATTTACGC GACGGCCAGTGCCAAGCTTTCC
366	8.Gibson_RGbw	CAGATTGTACTGAGAGTGCACCATAGGGGATCGGGAGATCC CCAGTGGAAAGACGCGCAGGC
431	LuciRG_P5Primer1	CAAGGACGGCTGGCTGGTGGGATACTATAACAACATTTAGC
432	LuciRG_P8Primer1	GGACGGCTGGCTGCACGTGGGATACTATAACAACATTTAGC
433	LuciRG_P9Primer1	GACGGCTGGCTGCACAGTGGGATACTATAACAACATTTAGC
435	LuciRG_Primer2_P5	CCCTCTAGACAAGGACGGCTGGCTGGTGG
436	LuciRG_Primer2_P8	CCCTCTAGAGGACGGCTGGCTGCACGTGG
437	LuciRG_Primer2_P9	CCCTCTAGAGACGGCTGGCTGCACAGTGG
449	RG_R175H_pos8_GCG_ Pr1	GGCGCTGCCCCGTGGGATACTATAACAACATTTAGC
450	RG_R175H_pos8_GUA_P r1	GAGGTACTGCCCCGTGGGATACTATAACAACATTTAGC
455	R175H_pos8_GCG_Pr2	CCCTCTAGATTGTGAGGTACTGCCCGTGGG
456	R175H_pos8_GUA_Pr2	CCCTCTAGATGTGAGGCGCTGCCCCGTGG
521	Luci_Pos6gRNA	CAAGGACGGCTGGCTGCGTGGG
522	Luci_Pos8gRNA	GGACGGCTGGCTGCACGTGGG
523	Luci_HindIII_RGpos6	CCAAGCTTTCCAAAAAACAAGGACGGCTGGCTGCGTGGG

Internal number	Name	5´-3´sequence
524	Luci_HindIII_RGpos8	CCAAGCTTTCCAAAAAAGGACGGCTGGCTGCACGTGGG
537	p53R175H_fw	GGAGGTTGTGAGGCACTGCCCCCACCATG
538	p53R175H_bw	CATGGTGGGGGCAGTGCCTCACAACCTCC
561	BamHI_p53fw	CTCGGATCCACCATGGAGGAGCCGCAGTCAGATCC
562	XbaI_p53bw	CCTCTAGAGCCGTCAGTCTGAGTCAGGCCCTTCTG
566	GFPanalyse327fw	GACACGTGCTGAAGTCAAGTTTGAAGGTG
586	p53 taq PCR 382fw	CAGCCAAGTCTGTGACTTGCACGTACTCCC

3.2 Molecular biology methods

3.2.1 PCR with Phusion DNA polymerase

Phusion DNA polymerase was used for the amplification of DNA fragments that were ligated into vectors and PCR products for downstream applications, especially OE PCR (3.2.2) for generating blunt end products with fewer errors. An exemplary 50 µl size PCR reaction mixture is given in Table 3-12. The corresponding program of amplification is shown in Table 3-13. The applied annealing temperature of the primer pair was calculated by the 'Oligonucleotide Properties Calculator' (86).

Table 3-12: Phusion PCR reaction mix

Component	Volume [µl]
Water	31.95
5x High-fidelity-Buffer	10
40x dNTPs	1.25
10 µM fw primer	2.5
10 µM bw primer	2.5
Template (ca. 100 ng)	1
Phusion polymerase	0.8

Table 3-13: Phusion PCR program

Step	Temperature	Time	Cycles
Denaturation	98°C	30 sec	1
Denaturation	98°C	10 sec	36
Annealing	52-58°C	30 sec	
Elongation	72°C	30 sec/1000bp	
Elongation	72°C	5-10 min	1
Hold	4°C	∞	

3.2.2 Overlap extension PCR

Overlap extension PCR (OE-PCR) was used for the introduction of point mutations into a desired sequence/gene and is explained in (87). The same components were used for OE- and Phusion-PCR (Table 3-12). Two separate PCRs generate the 5' PCR-product and 3' PCR-product. The overlapping complementary sequence of the 5'- and 3' PCR-product serve as primers in the first 5 cycles of a third PCR. For the following 36 cycles of the PCR two primers flanking the sequence of interest are added.

3.2.3 PCR with Taq DNA polymerase

PCR amplification with Taq DNA polymerase was used for cDNA-amplification and colony PCR (3.2.4). The reaction mixture of a 50 µl PCR is given in Table 3-14. The corresponding program of amplification shows Table 3-15.

Table 3-14: Taq DNA polymerase PCR reaction mix

Component	Volume [µl]
Water	36.25
10x ThermoPol-Buffer	5
40xdNTPs	1.25
10 µM fw primer	2.5
10 µM bw primer	2.5
Template (ca. 25ng cDNA)	2
Taq DNA polymerase	0.5

Table 3-15: Taq DNA polymerase PCR program

Step	Temperature	Time	Cycles
Denaturation	95°C	30 sec	1
Denaturation Annealing Elongation	95°C 52-58°C 68°C	15 sec 30 sec 60 sec/1000bp	30
Elongation Hold	68°C 4°C	5-10 min ∞	1

3.2.4 Colony PCR

Colony PCR is a fast screening method to test yeast or *E.coli* colonies grown on the selective LB-Agar plates for the correct plasmid or ligation of vector and insert. The PCR reaction mixture for Taq DNA polymerase is used (Table 3-14) with an increased volume of 38.25 µl purified water. The PCR reaction mixture is provided in the PCR tube and the colony is touched with a sterile toothpick and dissolved in the solution. The first 95°C step of the PCR program is prolonged up to 5 min for cell lysis.

3.2.5 Klenow reaction

The Klenow fragment without 5´-3´and without 3´-5´ exonuclease activity was used to fill-up hybridized oligonucleotide-fragments with dNTPs. These fragments were digested in the ongoing process for direct ligation into vectors. Table 3-16 shows the components of a Klenow fill-up reaction mix.

Table 3-16: Klenow reaction mixture

Component	Volume [µl]
Water	64
Neb4 buffer	10
40x dNTPs	2,5
100 µM fw oligonucleotide	10
100 µM bw oligonucleotide	10

The reaction was done by the following reaction steps: 2 min 60°C, 5 min 50°C, 5 min 37°C, 5 min 30°C, addition of 6 µl Klenow fragment; 30 min 30°C, 60 min 37°C.

For the subsequent digestion reaction the Klenow fragment was inactivated by incubation at 75°C for 20 min. The whole volume of the Klenow reaction mix was increased with 41 µl of water, 5 µl of Neb4 buffer and 2 µl of each restriction enzyme. The digestion mix was incubated overnight at 37°C. For the subsequent ligation protocol see 3.2.8.

3.2.6 Agarose gel electrophoresis

PCR-fragments, RNA fragments and plasmids were separated by agarose gel electrophoresis. The standard concentration for analysis of PCR-fragments and plasmids was 1.4% agarose in 1x TAE. RNA fragments were separated in 1% agarose gel in 1x TBE. For in-gel staining of nucleic acids 5 µl of Rotisafe was added to a 100 ml agarose-solution. The DNA samples were mixed with 6x Gel Loading Dye for loading. RNA was denatured before loading by addition of RNA loading buffer and incubation of the probes for 5 min at 70°C to avoid formation of secondary structures. As a length standard 5 µl 2-Log DNA Ladder was used. The separation was carried out with 10 volts/cm gel length and the visualization followed by UV-illumination.

3.2.7 Gel extraction

DNA was extracted from the agarose gel with the NucleoSpin® Gel and PCR Clean-up kit (Machery-Nagel) according to the manufacturer's protocol. The drying time was prolonged by 5 min.

3.2.8 Restriction and ligation for subcloning

Plasmids and PCR-fragments were restricted with the respective enzyme according to manufacturer's protocol (NEB). The appropriate buffer condition for a double digest was determined with the Double Digest Finder from NEB (https://www.neb.com/tools-and-resources/interactive-tools/double-digest-finder).

Table 3-17 shows an exemplary restriction mixture. Samples were incubated for at least 2 h at 37°C.

Table 3-17: Restriction mixture

Component	Plasmid	Insert
Restriction enzyme 1	1 µl	0,5 µl
Restriction enzyme 2	1 µl	0,5 µl
Recommended buffer	2 µl	2 µl
BSA 10x	If necessary	If necessary
Water	Ad to 20 µl	Ad to 20 µl

Restricted DNA was purified via agarose gel electrophoresis and gel extraction or, if a separation of fragments was not necessary, directly purified with the NucleoSpin® Gel and PCR Clean-up kit (cf. 3.2.6; 3.2.7). For an optimal ligation reaction 250 ng vector and a molar ratio of 1: 3 (vector: insert) were chosen. In a total reaction volume of 20 µl, 2 µl T4 DNA reaction buffer and 2 µl T4 DNA ligase were used. The ligation reaction was performed either for 2-4 h at room temperature or at 16 °C overnight.

3.2.9 Dephosphorylation

For many ligation reactions a previous dephosphorylation of the vector is necessary to prevent self-ligation. A standard reaction mixture includes 1 µl of Antarctic Phosphatase (NEB) and adjusted 10x Antarctic Phosphatase buffer. Incubation was carried out at 37°C for 30 min. The subsequent 5 min 70°C heat inactivation step allows the direct usage of the vector for the ligation reaction.

3.2.10 Preparation of $CaCl_2$ competent *E.coli* cells and transformation

For the production of $CaCl_2$ competent cells XL1 Blue *E. coli* was inoculated into 5 ml of LB and grown O/N. With 1 ml of the fresh O/N culture 200 mL LB media was inoculated and grown for 2-4 h at 37°C, 230 rpm, until an OD_{600} of 0.4 was reached. The culture was transferred to sterile 50 mL falcon tubes and spun for 10 min at 3000 rpm and 4°C. From this point on the cells were constantly kept on ice. The supernatant was removed and the pellets were resuspended in 10 ml cold 0.1 M $CaCl_2$ solution. After a second centrifugation step (10 min, 3000 rpm, 4°C) and removal of the supernatant the pellet was carefully resuspended with 2.5 ml cold

0.1 M $CaCl_2$/ 20% glycerin solution. The cells were incubated for at least 30 min on ice and subsequently frozen with liquid nitrogen in 100 µl aliquots.
For transformation 100 µl aliquots of $CaCl_2$ competent *E. coli* cells were thawed on ice. Subsequently 100 µl 1x TE buffer and 1 µl of diluted miniprep plasmid DNA or 5 µl of ligation reaction mix was added to the cells. Cells were kept on ice for 30 min and heat shocked for 1 min at 42°C. The cells were kept for 5 min at RT followed by a recovery step with the addition of 800 µl LB media for 1 h at 37°C, without shaking. Appropriate dilutions were distributed on selective agar-plates.

3.2.11 Plasmid isolation of *E.coli*

Plasmids were isolated from overnight grown *E. coli* culture with the Nucleospin® Plasmid kit (Machery-Nagel) according to manufacturer's protocol. For bigger amounts of purified plasmids the NucleoBond® Xtra Midi kit (Machery-Nagel) was used to isolate plasmids from a 200 ml *E. coli* culture grown overnight. Eluates of plasmid DNA were concentrated with the NucleoBond® Finalizers provided in the kit.

3.2.12 DNA sequencing

DNA sequences of plasmids or PCR fragments were sequenced with the Sanger method by Eurofins MWG operon. For purified plasmid DNA 1200 ng had to be mixed with 1.5 µl 10 µM primers in a total volume of 15 µl. For purified PCR fragments 120 ng of DNA was used.

3.2.13 Template generation for mRNA and R/G-gRNA transcription

Different mRNAs were used for in vitro studies. Every gene was cloned into pMG211 via XhoI/XbaI restriction sites. An overview of the produced plasmids for mRNA transcription of this study and their cloning information is provided in Table 3-18. These plasmids served as a template for Phusion PCR with primer pair no.126 and no.1 to generate the mRNA transcription templates.

Table 3-18: Produced plasmids of this work for mRNA transcription

Internal plasmid name	Vector backbone	Insert	Used oligonucleotides	PCR template	Restriction sites
pTS7	pMG211	W417X luciferase	5'-fragment: no. 126 & 6 3'-fragment: no. 1 & 5	pRJ114	XbaI/XhoI
pTS32	pMG211	FUS wt	no. 127 & 128	FUS	NdeI/XhoI
pTS39	pMG211	R521H FUS	no. 128 & 133	pTS32	NdeI/XhoI
pTS89	pMG211	E488Q ADAR2	5'-fragment: no. 83 & 233 3'-fragment: no. 84 & 232	pTS9.1	SalI/XbaI

The plasmid pTS11.2 contains a T7 polymerase promoter followed by HH-cassette and R/G-motif and served for R/G-gRNA templates generation. The Phusion PCR was performed with primer no. 125 and the respective primer, containing the gRNA sequence. The sequence of all produced R/G-gRNAs and the sequence of the corresponding bw primer for template generation are listed in Table 6-3.

3.2.14 Plasmid generation for cell culture

Each gene that was analyzed for editing was cloned into pcDNA3.1 vector using the restriction sites BamHI/XbaI. The R/G-gRNAs were inserted into a p*Silencer*™ 2.1-U6 Hygro vector. The initial cloning of a R/G-gRNA into the p*Silencer* vector was performed according to manufacturer's protocol containing the following steps: hybridization of two oligonucleotides that are coding for the R/G-gRNA, phosphorylation and ligation of the annealed inserts into the linearized p*Silencer* vector (pTS49.2).

For further insertions of R/G-gRNAs only two stepwise Phusion PCRs had to be performed, always using fw primer no. 176 and two respective PCR primers, which anneal to the R/G-motif and add the gRNA, poly-U-signal and HindIII restriction site to the fragment. This R/G-gRNA fragment was inserted into the p*Silencer* vector by the SalI and HindIII restriction sites.

A second option for cloning R/G-gRNAs into p*Silencer* was enabled: The first insertion method for cloning R/G-gRNAs was used to add a XbaI restriction site followed by a boxB-motif after the gRNA sequence. Thereby, the insertion of the R/G-gRNA fragment via XbaI and SapI restriction sites was enabled.

An overview of the plasmids used and an exemplary design of the p*Silencer* vector for R/G-gRNA expression are given in Figure 3-1.

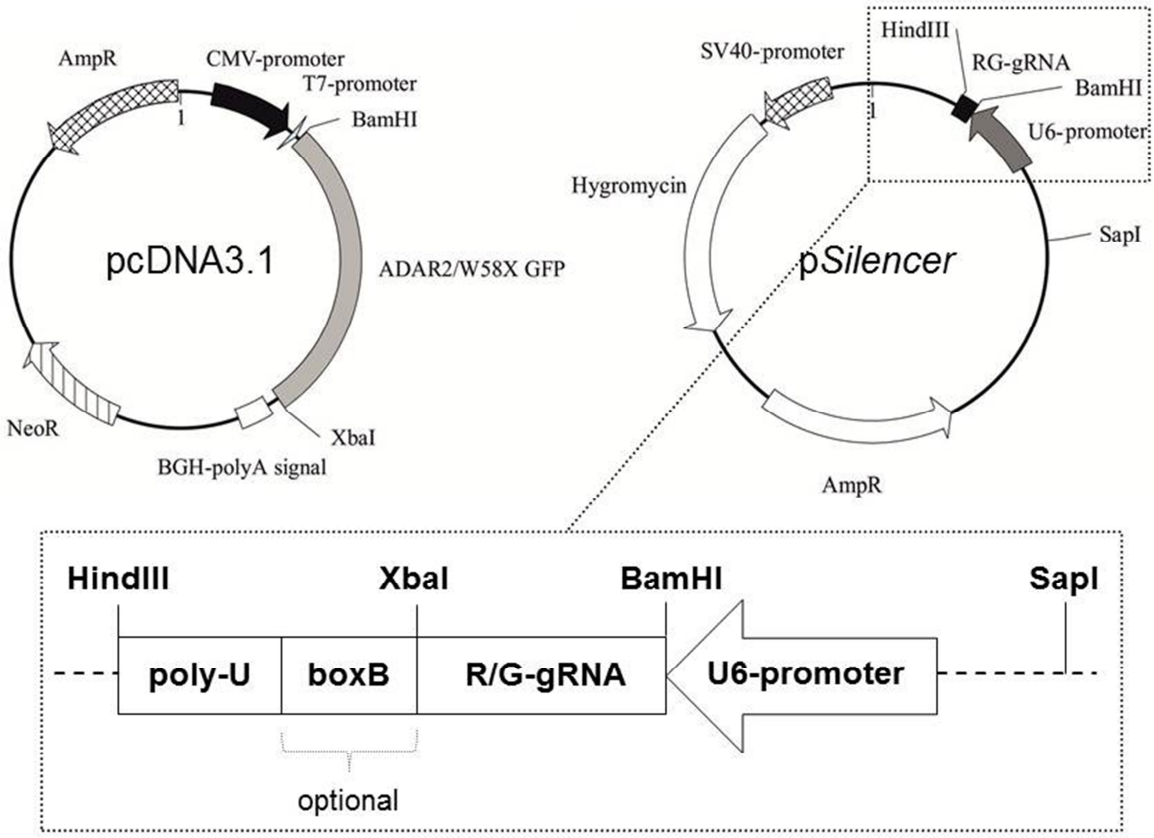

Figure 3-1: Design of the expression plasmids for cell culture application.

The ADAR2 and the gene for RNA editing analysis, here W58X eGFP, are cloned into pcDNA3.1 using the BamHI/XbaI restriction sites. The p*Silencer*™ 2.1-U6 Hygro vector was used for R/G-gRNA expression. Without the optional boxB-motif attached to the R/G-gRNA, the cloning is performed by the HindIII/SalI restriction sites. The XbaI restriction site is used for a construct expressing the R/G-gRNA with a boxB-motif attached to the 3'-end.

The following Table 3-19 summarizes the plasmids that were cloned during this study and provides information about the primers, templates, plasmid backbone and restriction sites used.

Table 3-19: Produced plasmids of this work for cell culture use and cloning information

Internal plasmid name	Vector backbone	Insert	Used oligonucleotides	PCR template	Restriction sites
pTS57	pcDNA3.1	ADAR2	5'-fragment: no. 14 & 187 3'-fragment: no. 13 & 118	pTS9.1	BamHI/XbaI
pTS63.1	pcDNA3.1	E396A ADAR2	5'-fragment: no. 126 & 203 3'-fragment: no. 144 & 202	pTS57	BamHI/XbaI
pTS90	pcDNA3.1	E488Q ADAR2	5'-fragment: no. 187 & 233 3'-fragment: no. 118 & 232	pTS57	BamHI/XbaI
pTS116.8	pcDNA3.1	W417X luciferase	no. 299 & no. 300	pTS7	BamHI/XbaI
pTS217	pcDNA3.1	R175H p53	5'-fragment: no. 538 & 561 3'-fragment: no. 537 & 562	No 289	BamHI/XbaI
pTS67	p*Silencer*	R/G-gRNA W437X PINK 16nt	no. 176 & no. 204	pTS49.2	HindIII/SapI
pTS68	p*Silencer*	R/G-gRNA W58X eGFP 16nt	no. 176 & no. 177	pTS49.2	HindIII/SapI
pTS86	p*Silencer*	R/G-gRNA W58X GFP 16nt + boxB	1. no. 176 & no. 214 2. no. 176 & no. 215	pTS68 1. PCR product	HindIII/SapI
pTS87	p*Silencer*	R/G-gRNA W58X eGFP 18nt + boxB	no. 176 & no. 229	pTS86	SapI/XbaI
pTS88	p*Silencer*	R/G-gRNA W58X eGFP 20nt + boxB	no. 176 & no. 230	pTS86	SapI/XbaI
pTS93.1	p*Silencer*	R/G-gRNA W58X eGFP 25nt + boxB	no. 176 & no. 234	pTS88	SapI/XbaI
pTS94.1	p*Silencer*	R/G-gRNA W58X eGFP 29nt + boxB	no. 176 & no. 235	pTS88	SapI/XbaI
pTS121	p*Silencer*	R/G-gRNA W417X luciferase 16nt P6 +boxB	1. no. 176 & no. 157 2. no. 176 & no. 304	pTS68 1. PCR product	SapI/XbaI
pTS169	p*Silencer*	R/G-gRNA W417X luciferase 16nt P5 +boxB	1. no. 176 & no. 431 2. no. 176 & no. 435	pTS121 1. PCR product	SapI/XbaI
pTS170	p*Silencer*	R/G-gRNA W417X luciferase 16nt P8 +boxB	1. no. 176 & no. 432 2. no. 176 & no. 436	pTS121 1. PCR product	SapI/XbaI
pTS171	p*Silencer*	R/G-gRNA W417X luciferase 16nt P9 +boxB	1. no. 176 & no. 433 2. no. 176 & no. 437	pTS121 1. PCR product	SapI/XbaI
pTS181	p*Silencer*	R/G-gRNA R175H p53 16nt P8 +boxB 5'-GCG	1. no. 176 & no. 449 2. no. 176 & no. 455	pTS68 1. PCR product	SapI/XbaI
pTS182	p*Silencer*	R/G-gRNA R175H p53 16nt P8 +boxB	1. no. 176 & no. 450 2. no. 176 & no. 456	pTS68 1. PCR	SapI/XbaI

Internal plasmid name	Vector backbone	Insert	Used oligonucleotides	PCR template	Restriction sites
		5'-GUA		product	
pTS194	p*Silencer*	R/G-gRNA W417X luciferase 16nt P6	1. no. 176 & no. 521 2. no. 176 & no. 523	pTS121 1. PCR product	HindIII/SapI
pTS195	p*Silencer*	R/G-gRNA W417X luciferase 16nt P8	1. no. 176 & no. 522 2. no. 176 & no. 524	pTS170 1. PCR product	HindIII/SapI

3.2.15 Gibson cloning

The method used to assemble multiple overlapping DNA molecules into one plasmid was described in 2009 by Gibson et al. (88) and uses 3 enzymes: 5' exonuclease, DNA polymerase and DNA ligase.

Using this method four copies of the U6 promoter and R/G-gRNA were cloned into a pcDNA3.1 vector via the restriction site BglII. Overlapping DNA molecules are generated by Phusion PCR with the following primer pairs and pTS68 as a template (Figure 3-2): 1) no. 359/360, 2) no. 361/362, 3) no. 363/364 and 4) no. 365/366.

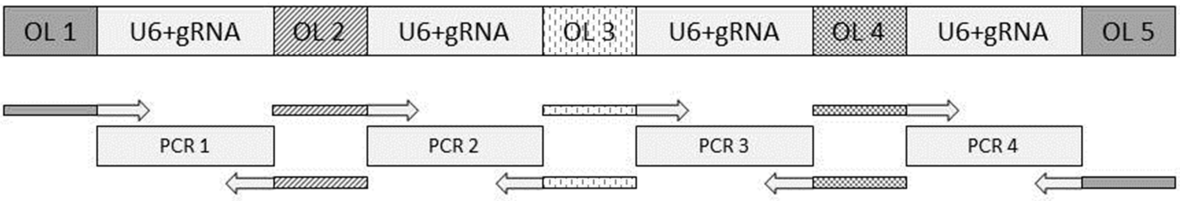

Figure 3-2: Overlapping DNA molecules for Gibson reaction. The overlap (OL) is attached to the U6-promotor+R/G-gRNA part by Phusion PCR. The first and the last OL are identical in their sequence to the 5' respective 3' part of the restriction site of the vector.

For the Gibson reaction the DNA fragments including the restricted vector were mixed together in 5 µl and kept on ice. The DNA should be in equimolar amounts. A vector amount of 100 ng was used. The thermo cycler was preheated to 50°C (70°C lid) allowing a quick start to the reaction. The DNA-mix was added to 15 µl of the Gibson-mix, briefly vortexed and directly placed into the preheated cycler. After one hour of reaction 5 µl of the Gibson-mix were used for transformation into XL1Blue *E.coli* cells (3.2.10).

The insertion of the ADAR1 gene into the pcDNA3.1 vector was performed with Gibson cloning, too. The primer pair no.276/277 and the template pTS56 were used to generate the overlapping molecule. The generated molecule was inserted via BamHI and XbaI digested pcDNA3.1 vector resulting in pTS284.

3.2.16 In vitro transcription of RNA

3.2.16.1 mRNA synthesis and DNase I digestion

All mRNAs were transcribed by T7 RNA Polymerase from 200 ng purified Phusion PCR-template containing the T7 promoter sequence. The synthesis was set up according to the following table at 37°C for 3 h:

Table 3-20: Transcription mixture

Component	Volume [µl]
Purified PCR template	X
10x transcription buffer	7,5
DTT [100 mM]	7,5
BSA [1 mg/ml]	3
rNTP mix [25 mM each]	12
T7 RNA-polymerase	4,5
Nanopure RNase free water	Ad to 75 µl

The QIAGEN RNeasy® MinElute® Clean-up Kit was used to purify mRNA. The protocol of the manufacturer was modified by drying the silica membrane for 5 min at 14000 rpm and eluting 2 times with 30 µl of RNAse free water. The mRNA was then treated with DNase I (QIAGEN RNase-Free DNase Set, 30 min at 37°C) to remove the DNA template and was again cleaned by RNeasy spin column work-up. The removal of DNA template was proven by Taq PCR with respective DNA template primers. For all sequences and varieties of mRNA that were synthesized see Table 6-2.

3.2.16.2 guideRNA synthesis

The in vitro transcription (ivT) is performed in the same way as mRNA transcription (3.2.16.1) with about 100 ng DNA template. For template generation of the R/G-

gRNAs see 3.2.13. The in vitro transcription is followed by the ethanol precipitation and purification of the guideRNA with Urea PAGE gel (3.2.17, 3.2.18).

3.2.17 Ethanol precipitation

When the ivT was finished, the volume was increased up to 300 µl with nanopure water, and then mixed with 33 µl of 3 M NaOAc. For precipitation of guideRNAs that are used for microinjection KOAc was used instead of NaOAc. Finally, the solution was mixed with 1000 µl of 100% EtOH. The samples were left overnight at -20°C or for 1 h at -80°C.

3.2.18 Urea PAGE

For the purification of transcribed guideRNAs an 8%-PAGE with 7.5N urea was always used. The following components were carefully mixed together for a 25 ml scale:

Table 3-21: 8%-PAGE gel Mixture with 7.5 N urea

Component	Volume [ml]
Nanopure water	2.4
Gel-sequencing-concentrate	8
Diluter	12.1
Gel-sequencing-buffer concentrate	2.5

Subsequently 125 µl of 10% ammonium persulfate solution was added and carefully mixed by inverting the tube twice, followed by adding 12.5 µl TEMED. The gel was casted with a serological pipette between the glass slides before the polymerization started. After one hour the gels are ready. The precipitated gRNAs were spun at -4°C, 14000 rpm for 30 min while the gels were polymerizing. The liquid phase was carefully removed and the pellet was washed with 500 µl precooled 70% EtOH. After a second spinning step the liquid was discarded. The RNA pellet was redissolved into 40 µl loading buffer (5x TBE, 8.3 M urea, ficoll without bromophenolblue). To remove the excess ammonium persulfate, the gel was run for 20 min (const. 100V). The pockets were washed before loading the RNA. A 60mer oligonucleotide was used as a length standard and buffer with bromophenolblue served as a control for the

duration of the gel run (2.5-3 h at 130V). The RNA products were cut out on a TLC-plate that was covered with a wrapping film under low-intensity 254 nm UV-light exposure (Figure 3-3). gRNA was isolated from the gel by shaking in 600 µl RNAse-free water at 4°C overnight. To remove the buffer and urea the gRNA was precipitated with ethanol (3.2.17). The RNA pellet was dissolved in 40 µl nanopure water. The RNA concentration was determined with UV absorbance at 260 nm (Nanodrop) and calculated with the respective extinction coefficient of the gRNA.

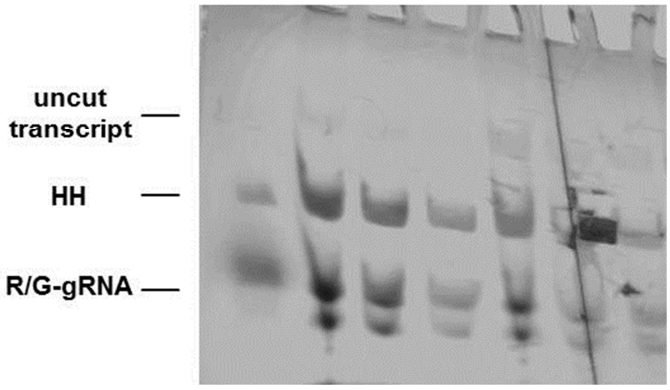

Figure 3-3: Urea PAGE gel of in vitro transcribed guideRNA. The in vitro transcribed hammerhead cassette, which cleaves itself out of the transcript and the R/G-gRNA are separated by urea PAGE gel.

3.3 Protein expression in S. cerevisae

Each ADAR protein was cloned into the pRS426 protein production plasmid for expression in yeast. The plasmids for hADAR2 and SNAP-ADAR2 were provided by the group. The hADAR2 mutant with glutamine instead of glutamic acid at position 488 (named as E488Q) (53), was cloned by overlap extension PCR (3.2.2) with primer no. 233 and no. 232 for insertion of the point mutation and flanking primers no. 83 and no. 84 (Table 3-11). E488Q was inserted into pRS426 at the SalI/XbaI restriction sites generating pTS89. The dsRBM1 deletion mutant of hADAR2, called short-ADAR2, was constructed as follows. The shortened gene of ADAR2 was amplified by Phusion PCR from pTS9 with primer pair no. 178 and no. 84. After digestion with SalI/XbaI the restriction fragment replaced the hADAR2 coding DNA fragment in pTS9 generating pTS55.2.

3.3.1 Transformation of *S. cerevisae*

The yeast strain YVH10 was used for protein expression. Yeast cells were cultivated in YPD-media at 30°C. Preparation of competent yeast cells and following transformation was performed with the Frozen-EZ Yeast Transformation II Kit (Zymo Research). All steps were performed according to the manufacturer protocol. The selection of positive transformants was done by spreading out 100 µl transformation mix on SD -CAA+Trp plates and incubating for 3 days at 30 °C.

3.3.2 Protein expression

With one colony of interest, 5 ml of SD-CAA+Trp media was inoculated and incubated overnight at 30°C. The next day, 1 ml of the culture grown overnight was put into two 500 ml Erlenmeyer flasks containing SD-CAA+Trp media (2 L size). The cultures were incubated at 30°C and 200 rpm for one day and at 28°C for another two days. After about one day the lid of the flask was removed to ensure a constant oxygen supply inside the flasks. The cells were harvested (10 min, 4 °C, 3000 rpm) and resuspended in 500 ml of SG-CAA+Trp to start protein expression. Incubation at 20°C and 200 rpm followed for the next five days. The cells were pelleted (10 min, 4 °C, 3000 rpm) for subsequent cell lysis (3.3.3) and protein purification (3.3.4).

3.3.3 Cell lysis

The cell pellet was dissolved in 90 ml yeast lysis buffer (Table 3-5). The cell walls were broken by the french press (1000 psi 3 times per suspension). The rough lysate was separated by centrifugation (1 h, 4 °C, 18000 g) into 3 fractions. The middle light phase of the fractions was used for subsequent purification steps of the protein.

3.3.4 Protein purification

ADAR proteins were purified by using a Ni-NTA column. The Ni-NTA column was equilibrated with lysis buffer. The lysate (middle light phase) was filtered by a 0.4 µm PES syringe filter and directly loaded on the Ni-NTA column. The column was washed with 40 ml W1, W2 and W3 each and eluted 3 times with 5 ml elution buffer (Table 3-5). The fractions obtained were further purified with a cation exchanger. The HiTrap^{TM} Heparin column (GE-Healthcare) was equilibrated with 100 mM NaCl. The

protein was eluted by gradually increasing the salt concentration, from 100 mM to 1000 mM NaCl (5 ml per 100 mM concentration shift).

3.3.5 SDS-PAGE

Each of the fractions of purified protein was run on a gradient SDS-PAGE to determine the main fractions of eluted protein and the efficiency of each purification step. Hence, 8 µl LMW-SDS marker and 16 µl of fraction were added to the respective amount of 4x Roti-load and incubated for 3 min at 95 °C. After centrifugation of the samples (3 min, 10000 g), 16 µl of the supernatant were loaded on a 4-20% Roti®-PAGE gradient gel (Roth). The gel was run in 1x SDS running buffer at 120 V for 90 min. The gel was washed three times with Millipore water and stained for at least 2 h with coumassie-staining solution (Table 3-6). For destaining, the gel was incubated in 20% ethanol overnight and washed with water. The protein containing fractions were identified and pooled together (3.3.6)

3.3.6 Protein concentration

The fractions identified containing protein were concentrated with two centrifugal filters at 6 °C and 3500 g to a final volume of 200 µl (15 ml 10 kDa and 0.5 ml 30 kDa MWCO Amicon® Ultra centrifugal filter). During protein concentration, the salt concentration was adjusted to 200 mM NaCl. The final protein concentration was calculated using NanoDrop measurement and extinction coefficient of the protein (Table 3-22). The storage conditions at -20°C (20% glycerol and 2 mM DTT) were included into the calculation of the protein concentration. All SDS-PAGE scans of purified ADAR proteins produced in this work, as well as their yields after concentration are provided in 6.3.

Table 3-22: Extinction coefficients of ADAR proteins

ADAR protein	Extinction coefficient [$M^{-1}cm^{-1}$]
ADAR2	130.000 (A_{260})
E488Q ADAR2	130.000 (A_{260})
Short-ADAR2	621.000 (A_{230})
SNAP-ADAR2	120.000 (A_{260})

3.4 In vitro editing:

3.4.1 In vitro editing reaction

The standard editing reaction was carried out in 0.2 ml PCR tubes on a 25 µl scale with the following components: mRNA, ADAR protein, R/G-gRNA, 10x RT editing buffer and Magnesia. The concentrations of each component were varied during experiments. Table 3-23 shows an example for the concentrations of each component in standard editing reaction mix. For lower concentrations of mRNA the scale was increased up to 100 µl to facilitate the following reverse transcription of mRNA to cDNA.

Table 3-23: Standard editing reaction mixture

Component	Concentration
mRNA	25 nM
ADAR protein	350 nM
R/G-gRNA	125 nM
$MgCl_2$ [50 mM]	3 mM
10x RT buffer	1x
Rnase Murine Inhibitor	0,4 µl

The annealing of mRNA and RG-gRNA, as well as the secondary structure formation of the R/G-gRNA was facilitated by heating steps (70°C for 2 min, 50°C for 2 min, 37°C 5 for min) in the absence of enzymes. This was followed by the addition of 0.4 µl RNase Inhibitor and ADAR protein to the mixture. Editing reactions were carried out for 3 h cycling every 30 min from 30°C to 37°C in a PCR machine. Variations of the standard editing reaction included the supplement of spermidine and heparin to the reaction mixture as well as reduced reaction times. The editing reaction was stopped by the addition of 12.25 µl of reverse transcription mix (Table 3-24).

3.4.2 Reverse transcription

To generate cDNA from the mRNA of the editing reaction mix, the reverse transcription mix (Table 3-24) was added, incubated for 2 min at 70°C, and directly put on ice. Subsequently 0.5 µl M-MulV RT was added. An excess of a short fully reverse complementary ssDNA to the corresponding/respective gRNA is applied to

keep the gRNA blocked from binding to the mRNA in order to facilitate reverse transcription of mRNA. The heating at 70°C was used to break secondary structures and to enable binding of the primer.

Table 3-24: Reverse transcription mixture for a 25 µl editing mix

Component	Volume [µl]	Final concentration of editing reaction + reverse transcription mix
10x M-MuLV Reverse transcriptase reaction buffer	1.25	1x (reverse transcription mix)
DTT [100 mM]	2	5.33 mM
dNTPs [40 mM]	0,94	1 mM
$MgCl_2$ [50 mM]	x	2 mM
Antisense ssDNA [1 mM]	0.75	20 µM
Bw primer [10 µM]	1.87	0.5 µM
Nanopure water	Fill up to 12.25 µl	37.5 µl in total

This reverse transcription mixture was incubated for 2 h at 42°C. The synthesized cDNA was purified, eluted with 30 µl and 2 µl of the eluate and were amplified with Taq polymerase as described in 3.2.3 with the corresponding primers.

To create cDNA from the isolated RNA of human cells about 80-100 ng RNA were mixed with 2 µl reverse transcription reaction buffer, 1 µl dNTPs, 1 µl bw primer and filled up to a final volume of 19.25 µl and heated at 70°C for 3 min. The mixture was directly shifted to ice for 2 min. After the addition of 0.25 µl RNase Inhibitor and 0.5 µl of M-MulV RT the reverse transcription mix was incubated for 2 h at 42°C. Subsequent purification of cDNA and PCR amplification followed. The editing efficiencies were estimated by Sanger sequencing (3.2.12).

3.5 Cell culture techniques for human cell lines

3.5.1 Cell cultivation, sub cultivation and transfection

HEK (human embryonic kidney) 293T cells were cultivated with DMEM+10%FBS+1%P/S at 37°C with 5% CO_2. Cells were detached from culture flasks with trypsin: the medium was removed, cells were washed with PBS (5 ml for 25 cm^2 culture flask), trypsin/EDTA solution (500 µl for small flask) was added, and after 2 min of incubation, the cells were resuspended with DMEM+10%FBS+1%PS

(4.5 ml for 25 cm^2 culture flask). The number of cells per ml was determined with the Neubauer counting chamber.
Human cell line specific plasmids were transiently transfected using Lipofectamine™ 2000 (Invitrogen). For the 24-well plate format $1.75x10^5$ cells and $3x10^4$ cells for the 96-well plate respectively were seeded out for the following day to reach a cell density of about 80-90%. The DNA to Lipofectamine ratio used for transfection was 1:3. Cells were imaged at 200x or 100x magnification by the use of Axio Observer.Z1 microscope (Zeiss).

3.5.2 Cell storage

Cells were stored with $3x10^6$ cells/ml storage medium (70% DMEM, 20% FBS, 10% DMSO) in liquid nitrogen.

3.5.3 Cell thawing

One vial of cells was put briefly into a warm water bath to thaw the cells frozen in liquid nitrogen. The cell suspension was transferred to 5 ml of culture medium and centrifuged for 5 min at 300 g to separate DMSO from the cells. The cells were resuspended in culture medium and grown at 37°C with 5% CO_2.

3.5.4 Total RNA isolation

To analyze the editing yield in cells, RNA was isolated with NucleoSpin® RNA Plus (Machery-Nagel) and RNeasy® Mini Kit (QIAGEN). Cells were detached from the well plate with DMEM and centrifuged for 5 min at 300 g. The supernatant was discarded and the cell pellet washed with cold PBS. After a final centrifugation step (5 min, 300 g) cells were immediately flash frozen by liquid N_2 and either stored at -80°C or directly thawed on ice for RNA isolation. The cell membranes were stripped by resuspending the pellet in lysis buffer and pulling the solution through a 0.6 mm size syringe. The RNA isolation procedure was performed according to the manufacturer`s protocol. RNA was eluted with 30 µl nanopure water. DNase I digestion was necessary for the use of the RNeasy® Mini Kit to isolate RNA, the NucleoSpin® RNA Plus kit includes removal of genomic DNA.
The genomic DNA was removed by the addition of 3 µl RDD-buffer, 1 µl DNase I to a final volume of 30 µl and incubated for 30 min at 37°C. The DNase I was heat

inactivated (65°C 10 min) by the addition of 3 µl 25 nM EDTA solution. Successful removal of genomic DNA was tested parallel to the cDNA amplification with Taq polymerase (3.4.2).

3.5.5 Luciferase assay

Luciferase was used as a reporter system for positive editing reaction events in HEK293T cells. The Luciferase Assay System from Promega provides luciferin as a substrate, which is oxidized by the luciferase protein. The resulting chemical energy of the luciferin oxidation is used by luciferase to emit light at a wavelength of 562 nM. The bioluminescent reaction needs elemental oxygen (O_2), ATP and magnesium ions (Mg^{2+}) as cofactors (89) and is detected with the FLUOstar OPTIMA plate reader from BMG LABTECH.

The luciferase assay was performed at least 48 hours after transfection. For the performance of the assay culture medium of the cells was removed, the cells were washed once with PBS and 20 µl of 1x Lysis Reagent (4 volumes water, 1 volume 5x Lysis Reagent) was added to detach the cells. For the solid luciferase activity measurement, dilutions of the cell Lysis suspension were made. The luminescence measurement was carried out in a 96 well plate format. Each well contained 7 µl of the cell suspensions that were provided and 35 µl Luciferase Assay Reagent added by the plate reader. The total measurement time per well was 30 s at room temperature (20-25°C), including 60 intervals of 0.5 s duration. The value obtained is the average of 20 intervals, excluding the first 10 intervals (4.5 s delay).

3.5.6 FACS analysis

Fluorescent activated cell sorting (FACS) was used to measure the amount of cells that expressed eGFP. The cells were fixed with formaldehyde for FACS analysis. Therefore, cells were washed with 500 µl PBS, harvested (300 g, 5 min), resuspended in PBS containing 4% of formaldehyde and incubated for 30 min on ice. After a final centrifugation step (300 g, 5 min), cells were dissolved in PBS and stored in the fridge at 4°C.

The FACS measurment was performed on LSRII (BD) and the analysis of FACS data was done with FACSDiva software from BD. More than 10 000 events (cells) were measured and analyzed by FACS for each sample.

3.6 Microinjecion of *Platynereis Dumerilii*

3.6.1 Preparation of mRNAs coding for reporter gene and ADAR2 variants

Plasmid pUC57 was used for an enhanced protein translation of the desired transcripts in *Playtynereis Dumerilii* since it contains a 169 bp 5'-UTR from the Platynereis 60 S acidic ribosomal protein P2 downstream of the T7 promoter, which enhances the protein expression of the transcripts. Reporter gene W58X eGFP and ADAR2 were subcloned into pUC57 via AscI and AgeI restriction sites. In order to attach the desired restriction sites to W58X and ADAR2 the following primer pairs were used in a Phusion PCR mixture: No. 216/217 for W58X and no. 258/259 for ADAR2. The plasmids pTS59 for W58X and pTS57 for ADAR2 served as templates.

3.6.2 In vitro transcription

In vitro transcription of W58X mRNA and ADAR2 mRNA was performed using the mMACHINE T7 Ultra Kit from Life Technologies. One specialty is the novel cap analog ARCA (Anti-Reverse Cap Analog), in which one of the 3' OH groups is eliminated and replaced by methoxy group ($-OCH_3$). According to the manufacturer this leads to 100% functional transcripts (mRNA transcripts with the correct orientation), since T7 RNA polymerase can only initiate transcription from the other remaining hydroxyl group of the cap analog.

According to the manufacturer's protocol the plasmids were linearized using Eco53KI. This linearization facilitates T7 RNA polymerase to bind properly, prevents formation of long, heterogeneous mRNA transcripts since the processive T7 RNA polymerase solely runs off the template by reaching the 3' end. Circular DNA on the other hand tends to lower mRNA yields because of supercoiling.

The whole transcription process including enzymatic poly(A)-tailing and m_7G capping was performed according to the manufacture's protocol. The purification of synthesized mRNA was done with RNeasy MinElute Cleanup Kit from QIAGEN.

The polyadenylation state and quality of the synthesized mRNA was controlled by 1% TBE agarose gel electrophoresis (3.2.6). An aliquot of the transcript without *E.coli* Poly(A) Polymerase (EPAP) treatment served as a control. A mass shift after EPAP treatment indicates successful polyadenylation.

3.6.3 Preparation of Platynereis zygotes

Mature male and female Platynereis worms were received from the laboratory of Dr. Gáspár Jékely (Max Planck Institute for Developmental Biology, 72076 Tübingen, Germany). In order to initiate the reproduction of the worms, two female and male adults were collected into one beaker containing natural sea water (NSW). The release of sperm and oocytes had to be checked by eye. After fertilization the animals were removed cautiously with a plastic Pasteur pipette and half of the NSW is replaced by fresh NSW to decrease the concentration of sperm as much as possible, since high concentrations of sperm are harmful for the eggs. The eggs are incubated for 55 min at 14.8°C until the egg jelly is removed. The removal of the jelly is necessary to soften the egg envelope in order to facilitate the microinjection. For this the eggs were washed with 500 ml of NSW in a 100 µM sieve, treated with proteinase K for 1 minute with a final concentration of 60 µg/ml and finally washed with 500 ml NSW. About 150 to 200 zygotes were put in the channel of the 2% agarose injection stage (dissolved in NSW). The zygotes must be covered always with NSW during washing and microinjection to prevent draining.

3.6.4 Microinjection

All injection samples consisted of 150 ng/µl reporter mRNA, 500 ng/µl ADAR mRNA and 40 µM of gRNA. The samples were mixed with 1.125 µg/µl Rhodamine-labeled dextran (10 000 MW), which was used as an indicator for positive injected zygotes. Thereby, a separation of injected and non-injected eggs is possible. The injection samples were spun down (40 min, 4 °C, and 14000 g) and 4.5 µl of the supernatant was loaded into Femtotips II microcapillarie (Eppendorf).

The microinjection was performed at the laboratory of Dr. Gáspár Jékely, which provided a Femtojet express microinjector (Eppendorf) on a Zeiss Axiovert 40 CL inverted microscope combined with a Luigs and Neumann micromanipulator. With the Luigs and Neumann Badcontroller V cooling system and a Cyclo 2 water pump (Roth) the temperature was controlled and kept at 14.8°C. The injection was started after one hour post fertilization (hpf) and ended by the first cleavage of the zygotes (around 2 hpf). The injection session started with an injection pressure of 700 hPa, injection time of 0.1 s and a compensation pressure of 35 hPa and was adjusted during the injection according to the conditions of the microcapillarie. The injected zygotes were

bred in a 6 well plate, filled with 6 ml NSW at 19°C until microscopy analysis was performed.

3.6.5 Imaging of Platynereis larvae

Microscopy analysis of Platynereis larvae was performed with an AxioCam HRc microscope camera connected to an Axio Imager Z1 wide field fluorescence microscope from Zeiss. For immobilization of the living larvae, glass slides were taped with three layers of adhesive tape on both sides. The larvae were placed in the middle of the glass slide (10 µl drop) and fixed with a coverslip. Using three layers of adhesive tape provides enough space for the larvae to stay alive and fixed in one spot.

4 Results and discussion

The aim of this work was to establish a fully genetically encodable editing system with the ability to steer the wild type ADAR2 protein to a specific adenosine in diverse mRNA targets. This encodable editing system should be addressable to different mRNA targets and 5'-codons in a highly specific way. The deamination reaction is supposed to achieve high editing yields. In order to harness ADAR2`s activity towards defined mRNA targets a specific guideRNA had to be designed that works exclusively and efficiently in combination with the ADAR2 protein. The concept of the guideRNA design was firstly analyzed and proved for its applicability in vitro with variable mRNA transcripts under variable reaction conditions. The valuable information obtained from the in vitro editing experiments was transferred to cell culture, benefiting from the possibility to encode both components of the editing system on plasmids. Many optimizations took place to achieve high editing yields, demonstrating the exclusiveness and specificity of the genetically encodable editing system. In a last part of this work the established editing system had to be successfully applied in a living organism, demonstrating the proof-of-principle for the encodable editing system in the zygotes environment of Platynereis dumerilii.

4.1 In vitro editing

ADAR proteins, target genes, and guide-RNAs had to be designed and produced to study a new site-directed RNA editing strategy with the ability to steer hADAR2 to the respective target mRNA. For ADAR protein production, as well as purification and mRNA synthesis, see chapter 3.3 and 3.2.16.1.

4.1.1 Construction of the R/G-gRNA

The challenge to design a fully genetically encodable guideRNA with the aim to steer ADAR2 to an arbitrary mRNA in *trans* and to achieve a specific and selective deamination reaction is demonstrated and explained in the following section. Furthermore, the concept of synthesis will be illustrated as well.

The design of the *trans*-acting guideRNA was inspired by the natural R/G-site motif of a glutamate receptor transcript (Figure 4-1).

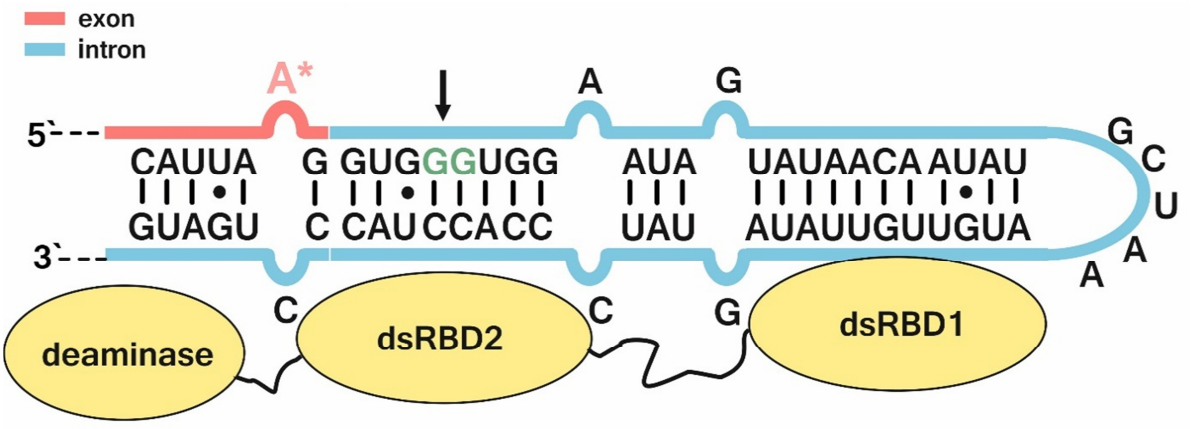

Figure 4-1: Natural *cis*-acting R/G-site of the glutamate receptor transcript. The intronic part (blue) of the transcript folds back to the exonic part (red) and is removed during splicing. This *cis*-acting motif of the glutamate receptor transcript recruits hADAR2, which edits the target adenosine of the exon (highlighted by asterisk) in a highly specific and quantitative manner, before splicing takes place.

The *cis*-acting element of the glutamate receptor, which is formed by folding back of the intronic sequence (blue) to the exonic part (red) of the transcript, recruits ADAR2 and allows a highly specific and quantitative editing of the target adenosine in the exon. Figure 4-1 illustrates the approximate positioning of the two double strand RNA binding domains (dsRBMs) of hADAR2 targeting the R/G-motif according to the findings from Stefl et al. and Masliah et al. (34,35,90). The idea was to use this R/G-loop motif for hADAR2 recruitment and binding, and to combine it with a RNA part, which is reverse complementary to the mRNA target sequence. The complementary RNA part is supposed to guide the whole *trans*-acting element to its target mRNA by Watson Crick base pairing.

Based on NMR structure analysis of the two dsRBMs in complex with the glutamate receptor transcript (35), the RNA secondary structure of the intact R/G-motif was cut between the two guanosine bases five and six nucleotides downstream of the edited adenosine to design one part of the engineered guideRNA (Figure 4-2).

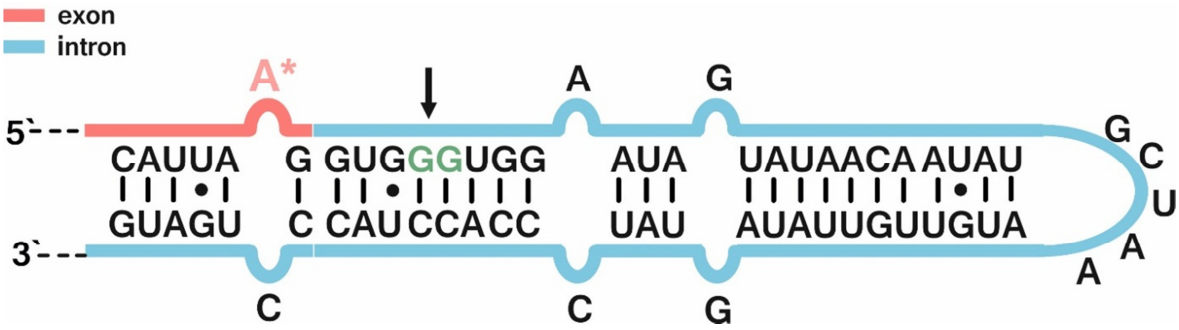

Figure 4-2: Position to cut the natural *cis*-acting R/G-site of the glutamate receptor transcript. The black arrow and the two highlighted guanosines mark the position, where the glutamate receptor transcript is cut to use the R/G-motif as a *trans*-acting element.

Thereby, an emersion point of the mRNA for the in *trans* applicated guideRNA is created. This position seemed to be most suitable for low interference of the mRNA with the dsRBD2 recognition, since the dsRBM2 is positioned at the opposite site (Figure 4-3).

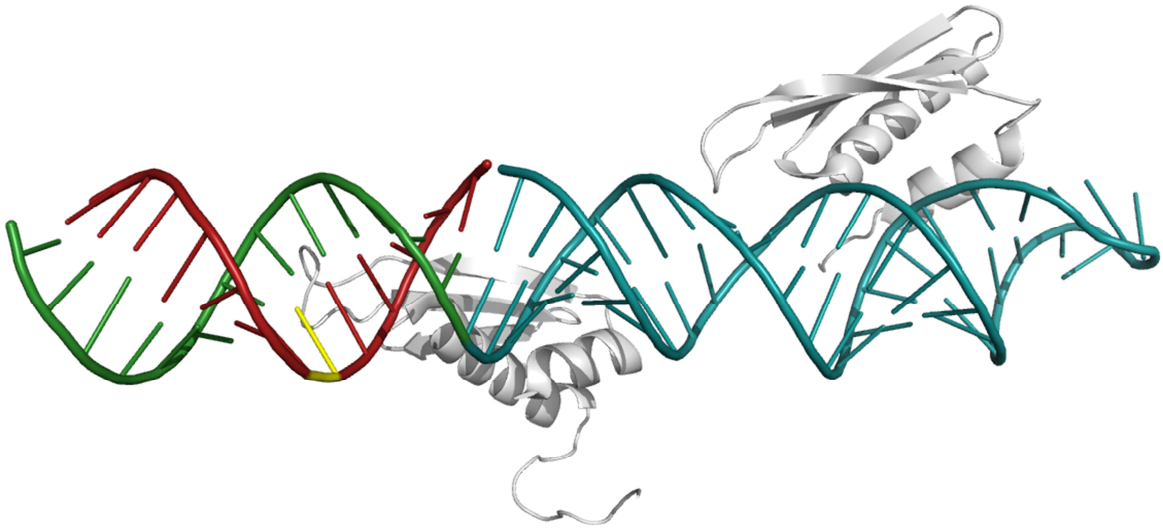

Figure 4-3: ADAR2 bound to the R/G-gRNA and mRNA. The mRNA is colored in red and the R/G-gRNA in green and blue. The target adenosine is highlighted in yellow. Each dsRBMs of ADAR2 binds the R/G-gRNA at two distinct positions. The dsRBM1 binds the stem-loop region, whereas dsRBM2 binds to the opposite of the created mRNA exit. Picture modified from PDB ID 2L3J.

Consequently, the engineered *trans*-acting guideRNA consists of two parts (Figure 4-4): a conserved part of the intronic section of the natural R/G-motif (blue), which folds back, and a flexible part for mRNA binding (green). The flexible part of the guideRNA can bind to any target of the mRNA by exchanging it with the reverse complementary sequence of the target mRNA (3.2.16.2).

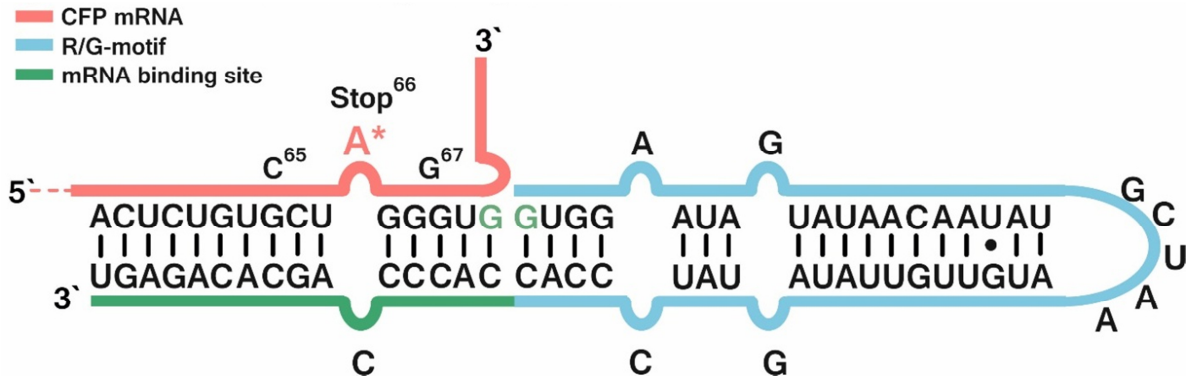

Figure 4-4: Engineered trans-acting R/G-guideRNA. The engineered R/G-guideRNA is build up from two parts: a conserved part from the natural R/G-motif (blue) and a flexible part (green). The target mRNA (eCFP) is bound by the flexible part of the R/G-guideRNA.

To achieve high levels of A-to-I conversion for each tested codon by ADAR2, the design of the anticodon of the guideRNA was adopted from our previous reported global analysis of editing preferences (49). There, we report that for high editing yields of the 5'-UAG codon, the 5'-CCA and 5'-CUA anticodons were most successful for SNAP-ADAR2 editing. In the case of a 5'-CAG codon, the most efficient editing was obtained using 5'-CCG and 5'-CUA anticodons in the gRNA (Figure 2-11). These editing preferences of SNAP-ADAR2 for all the four analyzed codons were considered for the design of the anticodons of the R/G-gRNAs to harness ADAR2.

The different R/G-gRNAs were supposed to be transcribed by a T7 RNA polymerase in vitro. Using a class III promoter, T7 RNA polymerase achieves highest product yields for the presence of two guanosine nucleotides at position +1 and +2 of the initiation domain (91). Also the T7 RNA polymerase tends to produce heterogeneous 5'-ends during the transcription depending on the 5'-terminal sequence (92,93). In order to achieve high transcription yields and to reduce crowding of negative charge at the 5'-end of the R/G-motif, a hammerhead (HH) ribozyme cassette was put 5´ to the guideRNA, which cleaves itself out of the transcript during synthesis (94), generating a strictly, homogenous 5´ r(GUGG) end of the R/G-gRNA.
The cloning strategy to obtain various R/G-gRNA templates for in vitro transcription is given in 3.2.13. Figure 4-5 demonstrates exemplarily the construct of hammerhead cassette and R/G-gRNA. For precise sequence information see the attached sequence details in chapter 6.3. An example of R/G-gRNA in vitro transcription and hammerhead ribozyme cleavage separated by urea PAGE gel is illustrated in Figure 4-6.

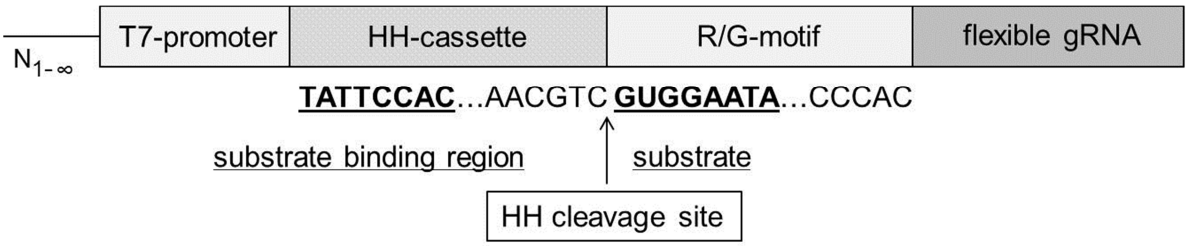

Figure 4-5: Illustration of the R/G-gRNA template construction. The R/G-gRNA template consists, depending on the chosen fw primer of $N_{1-\infty}$ nucleotides before the T7-promoter starts. The T7-promoter follows the hammerhead (HH) - cassette, which has complementary substrate binding region in the beginning to the substrate sequence. The R/G-motif and gRNA (flexible mRNA binding site) is cut by the HH-ribozyme, as indicated by the arrow between a cytosine and guanosine.

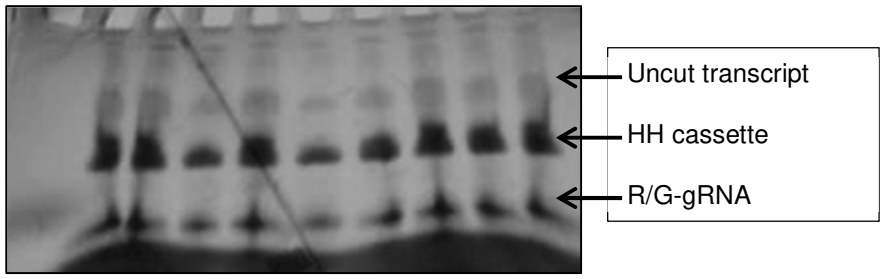

Figure 4-6: Urea PAGE gel of in vitro transcribed R/G-gRNA. The in vitro transcribed hammerhead (HH) cassette, which cleaves itself out of the transcript and the R/G-gRNA are separated by urea PAGE gel.

4.1.2 In vitro editing with R/G-gRNA – efficient conditions

The ability of the engineered guideRNAs to steer hADAR2 editing activity towards W66X eCFP mRNA had to be tested in a next step in correspondence to the previously established SNAP-tag BG-system (57). Also different conditions have been tested to find the most efficient conditions for this editing system in the PCR reaction tube to highly edit the target adenosine. Therefore, ADAR2 was produced and purified from yeast, as described in section 3.3. ADAR2 turned out to be highly active. Incubation of purified ADAR2 with in vitro transcribed W66X eCFP mRNA in a PCR tube resulted in editing at two positions 99 and 127 to 45% and 55% yield, even though no guideRNA was present. This is in contrast to the SNAP-ADAR enzyme, which shows much less off-target editing under the same conditions (500 nM ADAR protein, 250 nM guideRNA, 50 nM mRNA at 0.75 mM Mg^+ in a 25 µl scale). We expect the guideRNA independent editing to result from the binding of ADAR2 to

RNA duplex structures formed from the mRNA itself. It is known from the literature that ADAR2 edits in a sequence independent way (35,95). The two dsRBDs of ADAR2 that are lacking in case of SNAP-ADAR2 would explain the increased tendency of ADAR2 to edit RNA guideRNA independently. We then tested to elicit editing at our target codon W66X by adding a short single-stranded guideRNA (ss-gRNA) reverse complementary to the target site, but lacking the designed R/G-motif to the editing mix. Editing at the target site of 95% was found, whereas guideRNA independent editing remained unaltered (Figure 4-7). Then the R/G-motif containing guideRNA was introduced for in vitro editing. Unexpectedly, we found a decrease in the editing yield at the targeted codon.

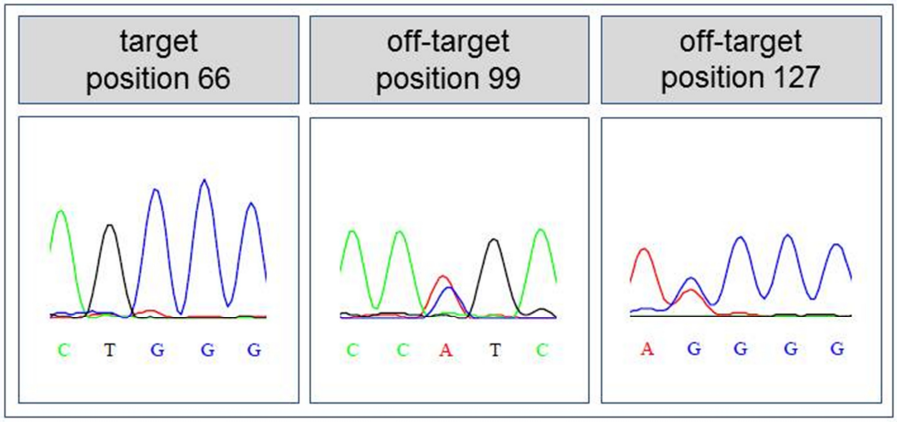

Figure 4-7: Editing of stop66 eCFP mRNA with ss-gRNA and ADAR2. The in vitro editing reaction was performed with 50 nM stop66 eCFP mRNA, 250 nM ss-gRNA and 500 nM ADAR2 at 0.75 mM Mg^{2+} and stopped after 3 hours of incubation. The target adenosine at position 66 is edited up to 95%, whereas other off-targets at positions 99 and 127 are edited as well.

Therefore, the editing reaction conditions had to be reconsidered. The first consideration was successful: previous heating steps (70°C 2min; 50°C 2 min; 37°C 5 min) before adding ADAR2 to the reaction mix improved the editing efficiency in all samples up to 50-90% depending on the chosen magnesia concentration (data not shown). Thus, we assumed that heating steps are facilitating the R/G-gRNA folding. Consequently, all experiments were performed with previous heating steps.

The decrease of enzyme, R/G-gRNA and mRNA to a ratio of 350 nM: 125 nM: 25 nM did not change the editing yields compared to the conditions that were chosen before. In order to save protein and RNA resources the following editing reactions were mostly performed at these decreased concentrations.

From the results of our previous studies with the SNAP-tag BG-system analyzing the editing efficiency, it was known that concentrations of magnesia higher than 0.75 mM are not well tolerated by SNAP-ADAR proteins. In the SNAP-tag BG-system, the addition of 0.75 mM magnesia prevents the enzyme from unspecific over editing, while the editing yield at the target adenosine is unaltered (49,57). Increasing magnesia concentrations were used for the ADAR2 R/G-gRNA system to test the tolerance of the enzyme, as well as the ability of magnesia addition to prevent unspecific over editing. The ability to edit the desired adenosine at position 66 of the eCFP mRNA was neither decreased nor improved by increasing the magnesia concentration (Figure 4-8).

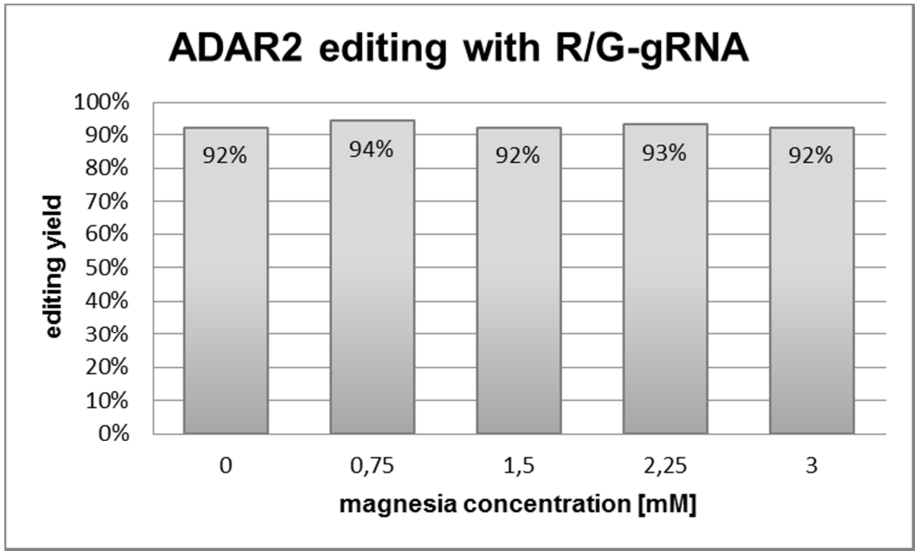

Figure 4-8: Editing yields of ADAR2 and R/G-gRNA at varying magnesia concentration. The editing yields of ADAR2 with the R/G-gRNA stay almost constant around 92% to 94% with increasing magnesia concentrations ranging from zero to three millimolar.

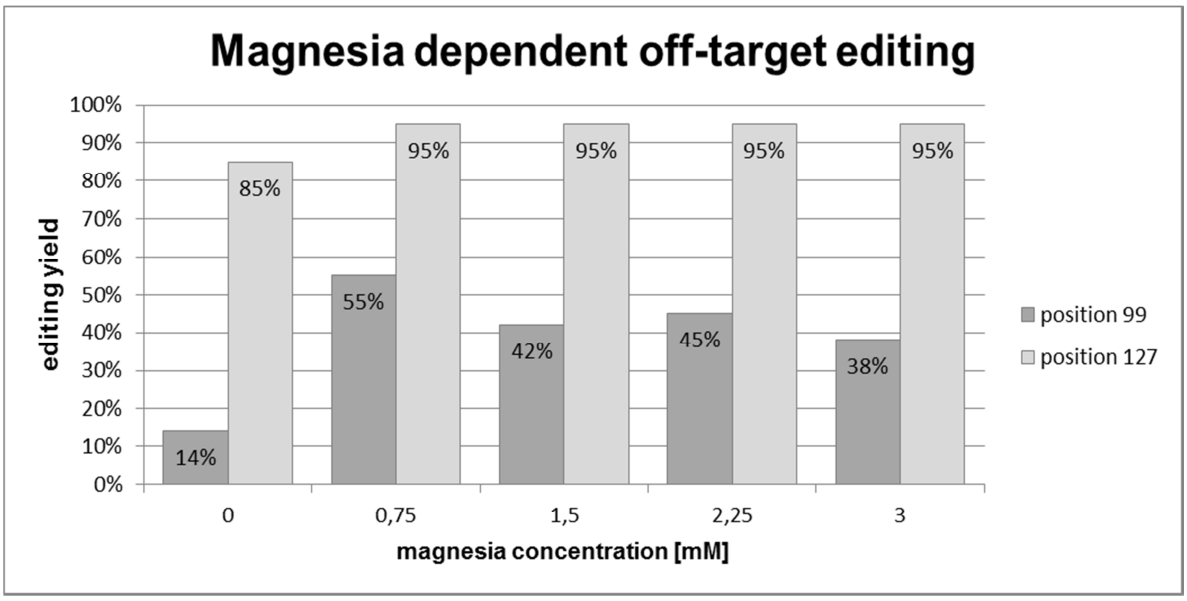

Figure 4-9: Magnesia dependent off-target editing at position 99 and 127. Absence of magnesia in the editing mix resulted in the lowest off-target editing yield. The off-target editing at position 127 is not influenced by magnesia addition, whereas the off-target editing of position 99 is slightly reduced by higher magnesia concentrations.

In former experiments SNAP-ADAR2 turned out to be sensitive towards high magnesia concentration. In absence of magnesia SNAP-ADAR2 was most active, but a magnesia concentration of 0.75 mM was tolerated (data not shown). The ADAR2 protein seems to be much more tolerant towards high magnesia concentrations. This effect was unexpected. Unfortunately, the undesired and unspecific over editing could not be prevented by its addition. The off-target editing at position 127 was stable at ~95% from 0.75 mM until 3 mM Mg^{2+} (Figure 4-9), whereas the off-target editing at position 99 was slightly reduced. Obviously, off-target editing turned out to be magnesia dependent. But the absence of magnesia in the editing mix resulted in slightly lower off-target editing yields than in the presence of magnesia. Therefore, other additional components were tested for their ability to maintain high editing yields at the target adenosine while preventing other adenosines from being edited.

4.1.3 Increasing stringency

Heparin has been shown to efficiently suppress guideRNA-independent off-target editing in the SNAP-tag BG-system. Since heparin is a negatively charged polysaccharide, it is assumed to compete with RNA at the ADAR binding sites.

The addition of 0.5 µM or 1 µM heparin to the reaction mixture decreased the editing yield of the targeted adenosine dramatically at both chosen magnesia concentrations (2.25 mM Mg^{2+} and 3 mM Mg^{2+}) (Figure 4-10). Not only was the A-to-I editing at the target adenosine position affected, the unspecific off-target editing at positions 99 and 127 were reduced and abolished, as well (Figure 4-11, Figure 4-12).

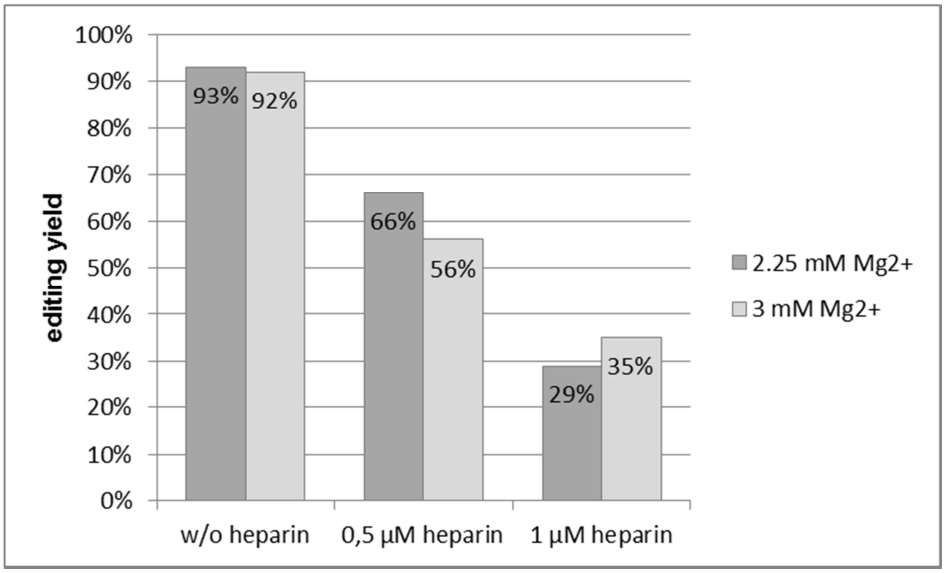

Figure 4-10: Influence of heparin to the editing yield of ADAR2 with R/G-gRNA. An increasing heparin concentration dramatically decreases the editing yield of ADAR2 with R/G-gRNA at 2.25 mM Mg^{2+}, as well as at 3 mM Mg^{2+}.

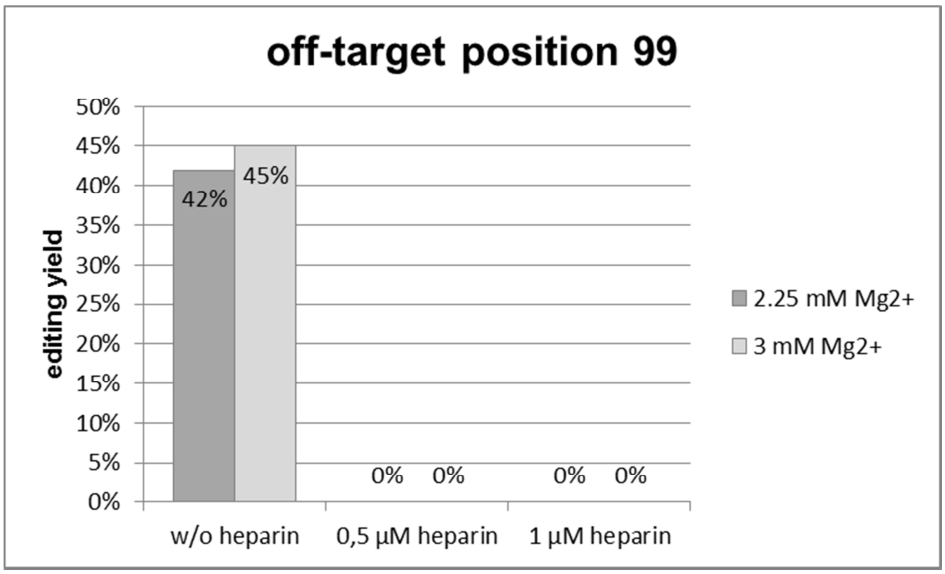

Figure 4-11: Influence of heparin on the editing yield of ADAR2 for the off-target position 99. Both heparin concentrations (0.5 µM and 1 µM) abolish completely the editing yield of the off-target adenosine position 99 at 2.25 mM Mg^{2+}, as well as at 3 mM Mg^{2+}.

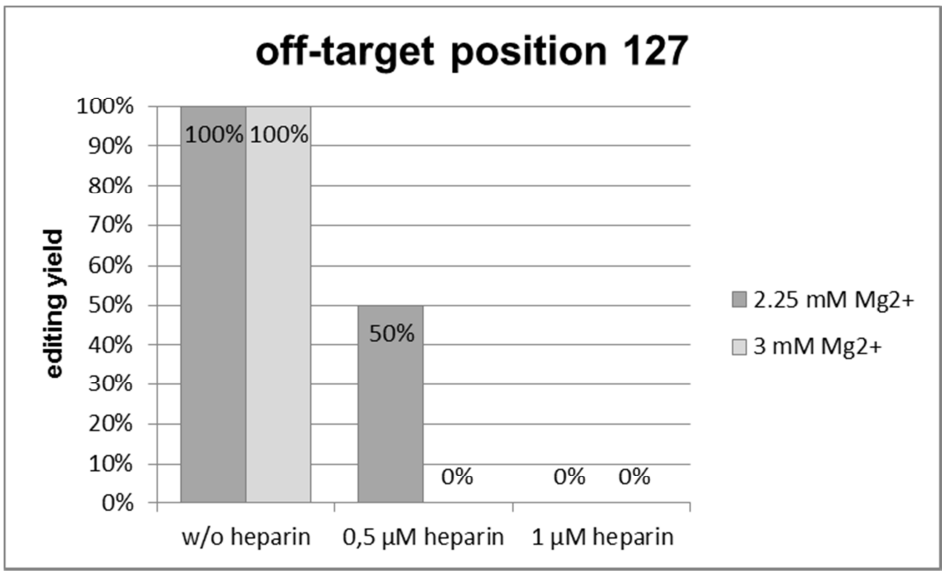

Figure 4-12: Influence of heparin on the editing yield of ADAR2 for the off-target position 127. A heparin concentration of 0.5 µM completely decreases the editing yield of the off-target adenosine position 127 at 2.25 mM Mg^{2+} about 50% and at 3 mM Mg^{2+}. The heparin concentration of 1 µM completely abolishes the editing at position 127.

The reduction of the editing yields at the off-target positions and the decreased editing at position 66 shows the sensitivity of ADAR2 towards heparin. In contrast SNAP-ADAR2 is much less sensitive towards heparin for two reasons. The first reason is that the two dsRBMs of ADAR2 are replaced by the SNAP-tag and are not present for heparin binding. Secondly, the covalent conjugate of SNAP-ADAR2 and the guideRNA enables the deaminase domain to stay close to the mRNA/gRNA-duplex and the recruitment of the deaminase remains unaffected (Figure 4-13). Therefore, SNAP-ADAR2 tolerates 2 µM heparin at physiological magnesia concentration (0.75 mM) and still edits its target adenosine with high yields (71%). ADAR2 is bound by heparin at three sites (Figure 4-13): The two dsRBMs and the deaminase domain. Since ADAR2 binds reversible to the R/G-gRNA, ADAR2 will diffuse away from the RNA substrate duplex upon heparin complexation. Thus, heparin addition reduces editing at off-target positions, but also at the target position. Heparin is not suitable for the R/G-gRNA ADAR2 system to prevent off-target editing and parallel to maintain high editing yields at the target adenosine position.

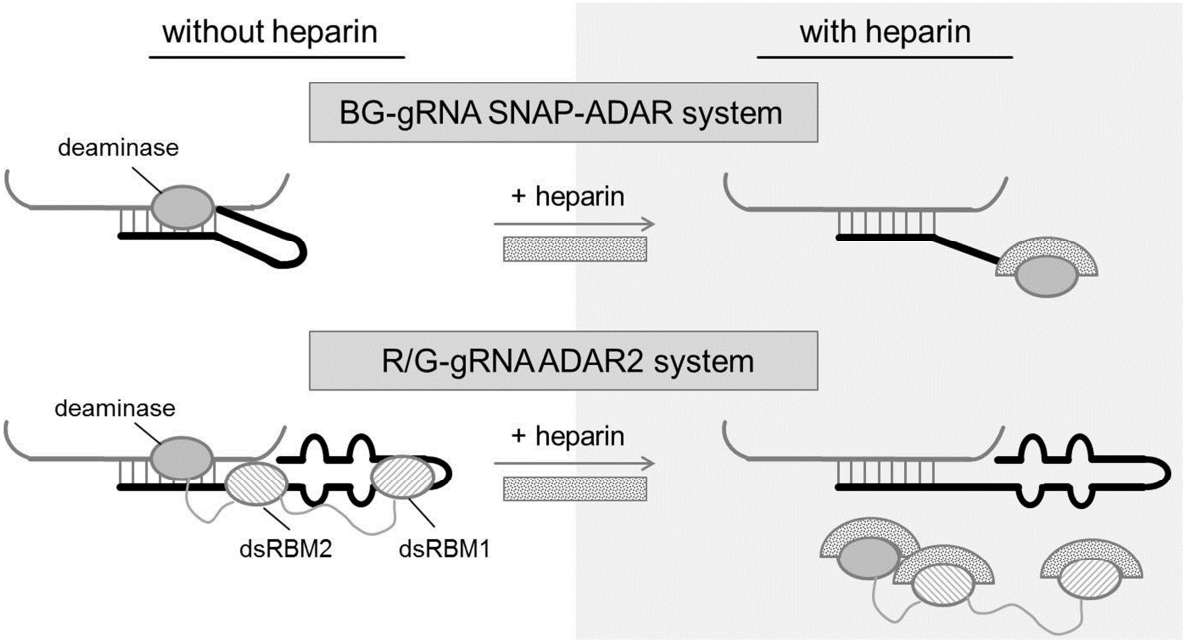

Figure 4-13: Illustration of the heparin influence at the BG-gRNA SNAP-ADAR system and the R/G-gRNA ADAR2 system. Heparin binds beside to the deaminase domain also to the dsRBMs of ADAR2. The editing efficiency of SNAP-ADAR2 is less affected, since SNAP-ADAR2 lacks of the two dsRBMs of ADAR2 and is covalently linked to the guideRNA. Thus, the binding of the target mRNA is not influenced by heparin at the SNAP-ADAR2 BG-gRNA system. In contrast, the ADAR2 binds reversible to the RNA substrate duplex. Therefore, ADAR2 will diffuse away from the RNA substrate duplex upon heparin complexation.

The second additive tested was spermidine. Since spermidine is a positively charged polyamine, it can bind to RNA and has the potential to compete with ADAR2 for RNA binding.

A concentration of 2 mM spermidine in combination with varying magnesia concentrations in the reaction mixture only slightly reduces the editing efficiency at the target adenosine, whereas common off-target editing is drastically reduced, and at position 99 completely prevented (Figure 4-14, Figure 4-15, Figure 4-16). Since SNAP-ADAR turned out to be very sensitive towards spermidine (data not shown), this tolerance of ADAR2 towards high spermidine concentrations was not expected. Spermidine accomplishes the desired conditions of maintaining high editing yields at the target position 66 while reducing off-target editing of other adenosines.

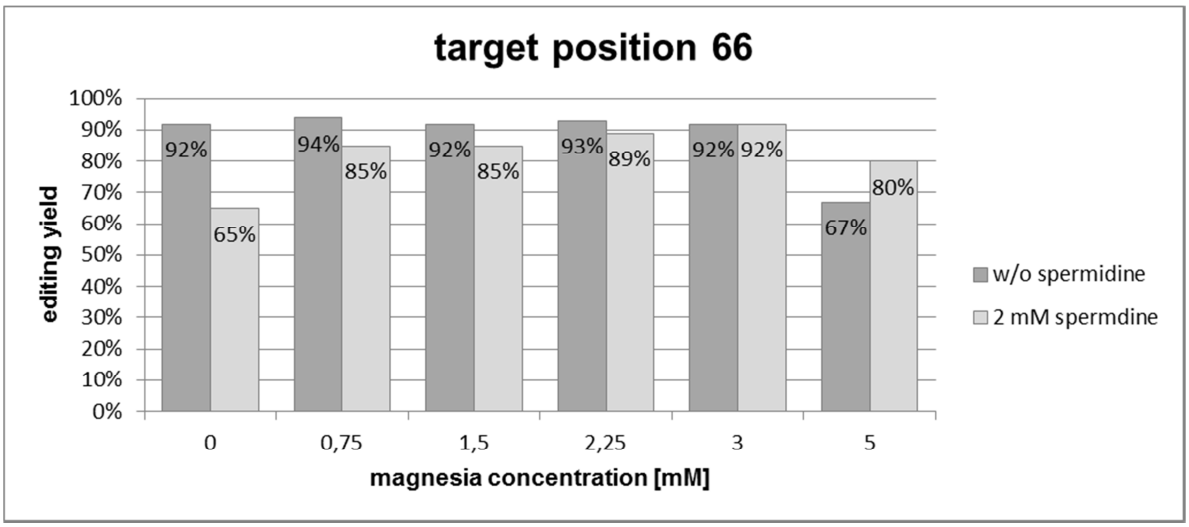

Figure 4-14: Influence of spermidine on the editing yield of the target adenosine position 66. The addition of 2 mM spermidine to the reaction mixture by varying magnesia concentrations only slightly reduces the editing yields of the target adenosine at position 66. The reduction of the editing yields decreases with higher magnesia concentrations and at 3 mM Mg^{2+} the editing yield stays the same. At 5 mM Mg^{2+} the concentration of 2 mM spermidine was even beneficial regarding the editing yield.

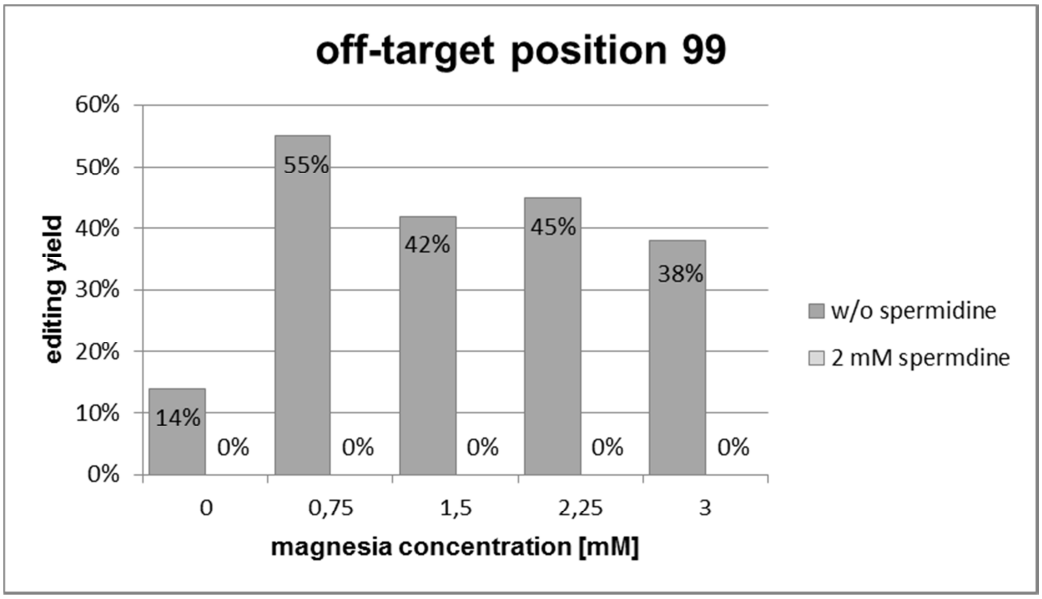

Figure 4-15: Influence of spermidine on the editing yield of the off-target position 99. The addition of 2 mM spermidine to the reaction mixture by varying magnesia concentrations completely abolishes the editing yields of the off-target adenosine at position 99.

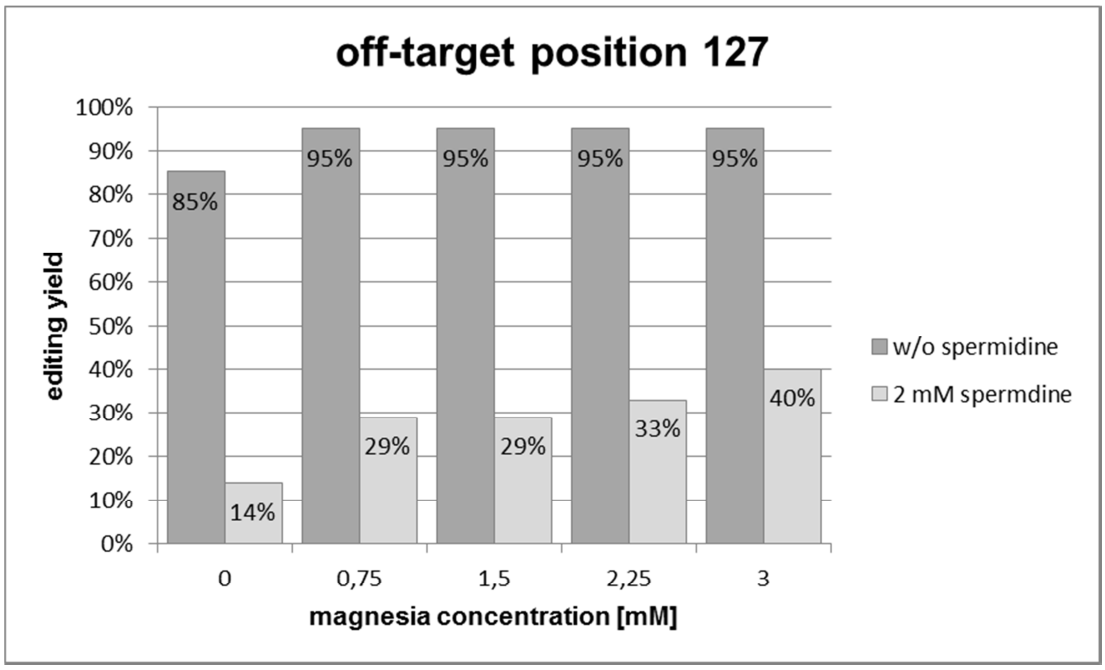

Figure 4-16: Influence of spermidine on the editing yield of the off-target position 127. The addition of 2 mM spermidine to the reaction mixture by varying magnesia concentrations dramatically reduces the editing yields of the off-target adenosine at position 127. With increasing magnesia concentrations, the reductive effect of spermidine to the editing is less distinct.

With increasing magnesia concentration and at constant 2 mM spermidine, the off-target editing at position 127 was rising (Figure 4-16). The assumed reason for a reduced off-target editing effect of spermidine towards off-target position 127 by increasing magnesia concentrations is that magnesia stabilizes the secondary structure or even facilitates the secondary structure formation of the mRNA. Therefore, the addition of spermidine at 3 mM Mg^{2+} in the reaction mixture has fewer positive effect on preventing ADAR2 from off-target editing.
It was also observed, that the absolute editing yield is dependent on the R/G-gRNA production and its quality, since different batches lead to different absolute editing levels.

Summarized, the conditions for a highly efficient and specific in vitro editing reaction at the target adenosine position and strongest off-target editing reduction with ADAR2 and R/G-gRNA is as follows:

Table 4-1: Suggested optimal in vitro editing reaction mixture for ADAR2 and R/G-gRNA.

Component	Final concentration
ADAR2	350 nM
R/G-gRNA	125 nM
mRNA	25 nM
$MgCl_2$ [50 mM]	3 mM
Spermidine	2 mM

4.1.3.1 Combination of transcription and editing

The tolerance of ADAR2 towards high magnesia and spermidine concentrations was unexpected. This tolerance raised up the idea to combine the transcription of the reporter mRNA and their R/G-gRNA with the editing of this transcribed mRNA in the same buffer. A successful combination of transcription and editing would offer the possibility to obtain high amounts of mRNA containing inosine at a specific position. A chemical synthesis of inosine containing mRNA could be substituted.

The transcription buffer from NEB (10x RNAPol Reaction buffer) was chosen for this experiment, since it has a lower magnesia and spermidine concentration compared to the self-made transcription mix (Table 3-5): 6 mM magnesia and 2 mM spermidine (NEB) vs. 22 mM magnesia and 1 mM spermidine (self-made).

At first the ability of ADAR2 to edit 5'-UAG at position 66 of the eCFP mRNA in the NEB 10x RNAPol Reaction buffer was tested. The reaction mixture contained 10x RNAPol Reaction buffer, 350 nM ADAR2, 125 nM R/G-gRNA and 25 nM mRNA at 25 µl scale. ADAR2 was able to edit the target adenosine at these buffer conditions up to 78% (Figure 4-17).

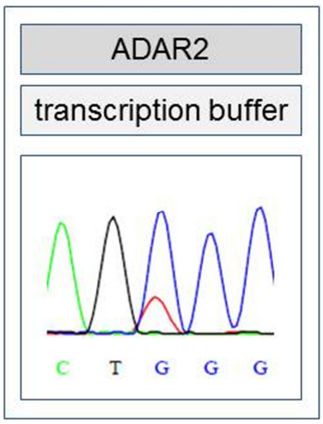

Figure 4-17: Test editing of ADAR2 in the transcription buffer from NEB. ADAR2 edits the target adenosine at position 66 up to 78% in a transcription buffer containing 6 mM Mg^{2+} and 2 mM spermidine.

The following diagram shows the protocol combining RNA transcription and RNA editing (Figure 4-18):

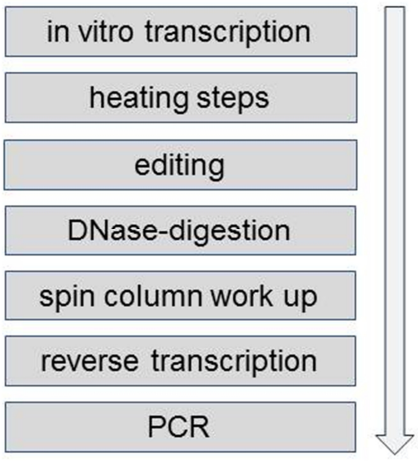

Figure 4-18: Diagram of the protocol combining RNA transcription and RNA editing.

Both RNA transcription templates - eCFP mRNA and R/G-gRNA - were used for a 2 h long transcription (3.2.13). In order to achieve approximately the same ratios of the R/G-gRNA and mRNA like in other in vitro experiments, five times as much R/G-gRNA template was used as eCFP mRNA templates for the transcription. The heating steps served for the parallel inactivation of the RNA polymerase and the folding of the R/G-gRNA. After the heating steps, 650 nM of ADAR2 were added to the solution. After 3 h of editing, RNA purification with the column and DNase digestion followed (3.2.16.1). The reverse transcription, cDNA purification (3.4.2), Taq

PCR (3.2.3) and DNA sequencing (3.2.12) were performed according to the standard protocols.

The usage of both RNA transcription templates for the transcription and the editing combination protocol reproducibly resulted in very low RT-PCR yields. In order to check if the presence of the R/G-gRNA transcription template was the reason, the R/G-gRNA transcription was omitted in the experimental protocol. Instead, 125 nM purified R/G-gRNA were added to the reaction mixture after mRNA transcription was performed and before the heating steps followed. This experimental setup was successful in generating high amounts of PCR product, suggesting that the PCR problems were caused by the presence of the R/G-gRNA transcription template. The editing efficiency with the directly added R/G-gRNA was 31 % (Figure 4-19).

Since the direct application of the R/G-gRNA was not hampering the eCFP mRNA RT-PCR reaction, it was assumed that the template of the R/G-gRNA created the problems to obtain bigger amounts of eCFP PCR product. Analyzing the hybridization of the primers that were used for the mRNA reverse transcription revealed that indeed both primers can bind to the R/G-gRNA template: the fw primer no. 87 binds to the 5'-UTR (before the HH-cassette) and the bw primer no. 1 binds a sequence that follows the gRNA. Since the DNase digest is not always achieving full destruction of the DNA templates, a shorter R/G-gRNA template was used at the same 1:5 stoichiometry (mRNA template: R/G-gRNA template), without changing the R/G-gRNA itself. The yield of the eCFP cDNA and PCR product was improved by this shorter template, but the editing yield of 27% for the transcription and editing reaction experiment could still not be improved.

The idea of combining mRNA transcription and editing in the same buffer has not achieved high editing yields but it provides a starting point for further optimizations like variation of the template amounts and ADAR2 concentration.

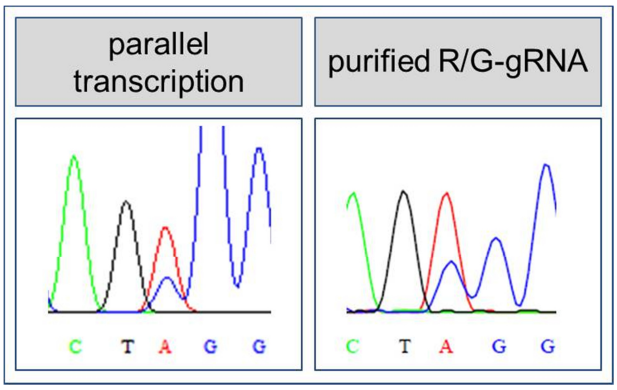

Figure 4-19: Combination of mRNA transcription and editing. The combination of mRNA and R/G-gRNA transcription with editing in the same buffer achieved 27% yield. The direct addition of 125 nM purified R/G-gRNA resulted in 31% editing yield.

4.1.4 Variation of ADAR2 concentration

In the beginning of analyzing the editing ability of the R/G-gRNA ADAR2 system, the amount of 350nM enzyme to 125 nM R/G-gRNA and 25 nM reporter mRNA was predetermined. We were wondering if at 25 nM reporter mRNA and 125 nM R/G-gRNA the amount of ADAR2 could be reduced for efficient editing or if the ratio of 14:5:1 (enzyme: guideRNA: mRNA) was most suitable for our experiments.

By lowering the ADAR2 concentration from 350 nM down to 10 nM at standard 3 mM Mg^{2+} editing reaction mixture, the editing level at position 58 of W58X eGFP mRNA was steadily reduced (Figure 4-20). But still, with an ADAR2 concentration of only 10 nM, which corresponds to a ratio of 0.4:5:1 (enzyme: guideRNA: mRNA), editing at 32% yield was achieved. Lowering the ADAR2 concentration also decreased the off-target editing (Figure 4-21, Figure 4-22). Starting from 75 nM or lower the off-target editing at position 99 was completely avoided. To abolish off-target editing at position 127 completely, an ADAR2 concentration of less than 10 nM is needed. Obviously, the ratio of enzyme to guideRNA and mRNA is determining the editing efficiency. Also the diluting effect of ADAR2 in the solution influences the editing efficiency. Since every protein needs a special environment, we assume that a drastic reduction of the protein concentration is not tolerated by ADAR2. The addition of BSA might be an option to balance this dilution effect.

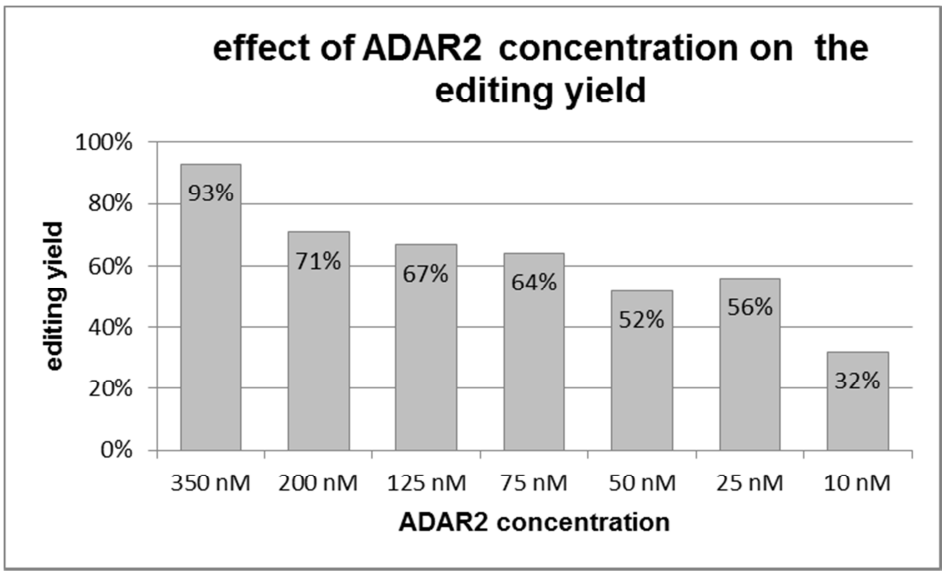

Figure 4-20: Effect of decreasing ADAR2 concentrations on the editing level. The editing yield decreases with lower ADAR2 concentrations at 3 mM Mg^{2+} added to the reaction mixture.

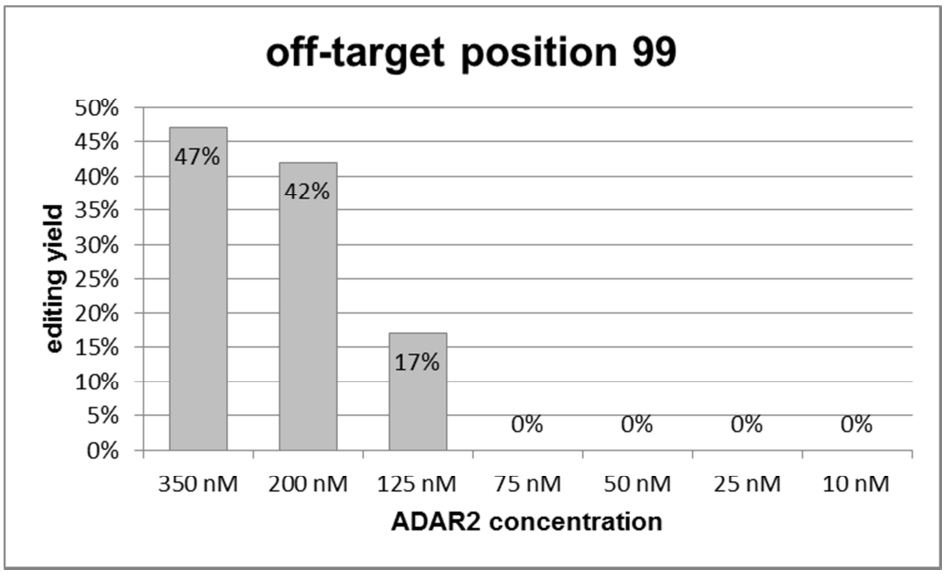

Figure 4-21: Effect of decreasing ADAR2 concentrations on the editing at position 99. The editing of the off-target position 99 is reduced by lower ADAR2 concentrations. Starting from 75 nM the off-target editing is completely avoided.

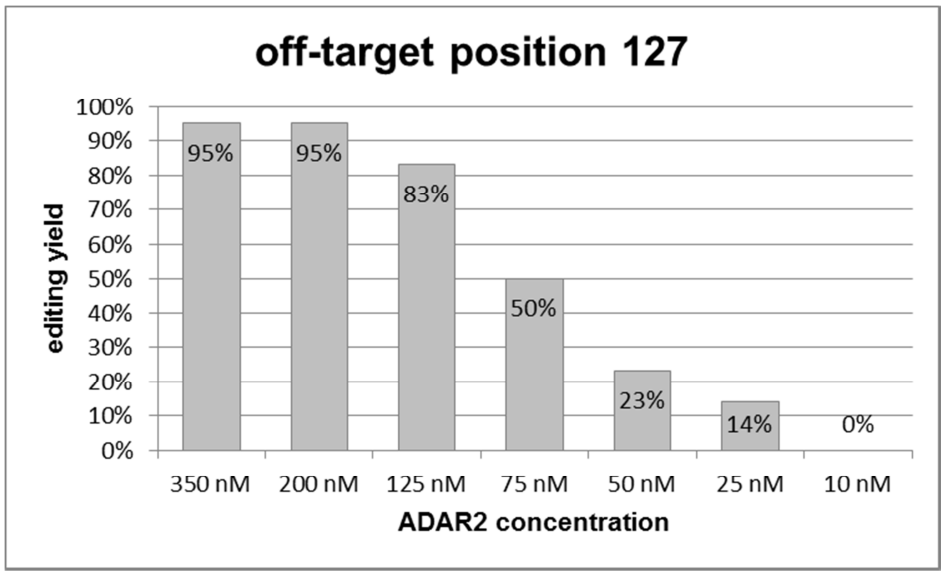

Figure 4-22: Effect of decreasing ADAR2 concentrations on the editing at position 127. Reduced editing levels of the off-target position 127 are obtained by lowering the ADAR2 concentration. For a concentration of 10 nM ADAR2 the off-target editing is avoided.

4.1.5 Reduction of the editing time

Different time points of incubation were tested with and without spermidine addition to find out 1) how quickly the maximum editing yield of W58X eGFP mRNA is achieved, 2) which time-dependency of A-to-I conversion does exist and 3) how quickly proceeds the target versus off-target editing and what changes the spermidine addition on this proportion?

The longer the incubation time with ADAR2 was chosen, the higher the editing yields were obtained (Figure 4-23). The maximum editing yield was obtained after 3 hours of incubation time, with and without spermidine addition. For determing the time-dependeny of the A-to-I editing reaction, we only focus on the mRNA substrate. Since all other components (R/G-gRNA and ADAR protein) were present in excess, we use a single-turnover kinetic to plot the deamination product as a function of time (Figure 4-24). The rate constant was three-fold higher without spermidine addition (k_{deam}= 0.0545 min^{-1}) than with spermidine (k_{deam}= 0.0165 min^{-1}), indicating that the presence of spermidine slows down the deamination reaction of ADAR2. Compared to Stephens et al. (45), their calculated rate constant for a synthetic substrate (k_{deam}= 1.2 ± 0.1 min^{-1}) is 22-fold higher than our calculated rate constant of ADAR2. The deamination reaction is not proportional to the time: the major part of editing takes

place within the first 90 min of incubation. The extension of the incubation time by further 90 minutes (with a total of three hours editing time) increased the editing by 13% for the editing condition with spermidine. However, the editing yield of the reaction mix without spermidine has reached its maximum almost after 45 min and can only increase at 3% by extending the editing time beyond 90 min.

For high editing yields (≥50% conversion) the reaction mix should be incubated at least for 45 min. As expected, spermidine addition to the reaction mixture slows down the conversion of A-to-I in all chosen time points, since it competes with the ADAR2 enzyme for RNA binding. But after 3 h of editing time this delay is overcome. If spermidine is added to the reaction mix this delay for the maximum editing yield should be regarded.

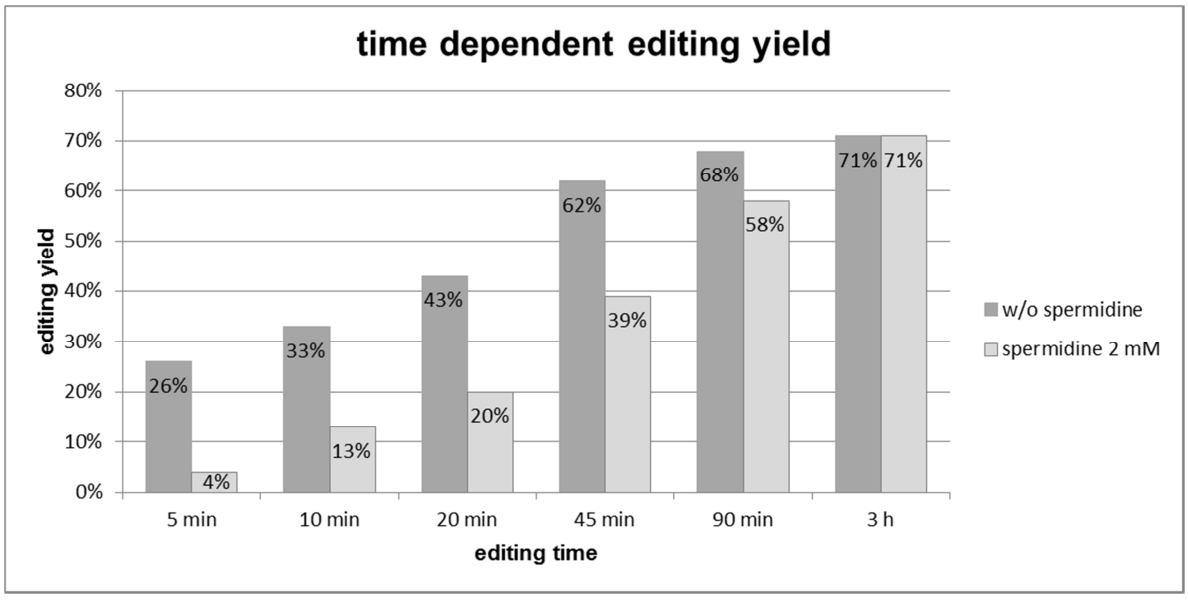

Figure 4-23: Time dependent editing yield of W58X eGFP mRNA. The editing of W58X eGFP mRNA at standard editing reaction mixture is time dependent. The longer the incubation time with ADAR2 the higher the editing yields for both conditions – with and without spermidine. After 90 min of incubation the editing yield of the condition lacking the spermidine is only increased about 3 % by extending the editing time until three hours. For the editing condition with spermidine the prolongation of editing time to further 90 min results in 13% more editing yield. The following amounts of the editing components were used: 350 nM ADAR2, 125 nM guideRNA and 25 nM mRNA.

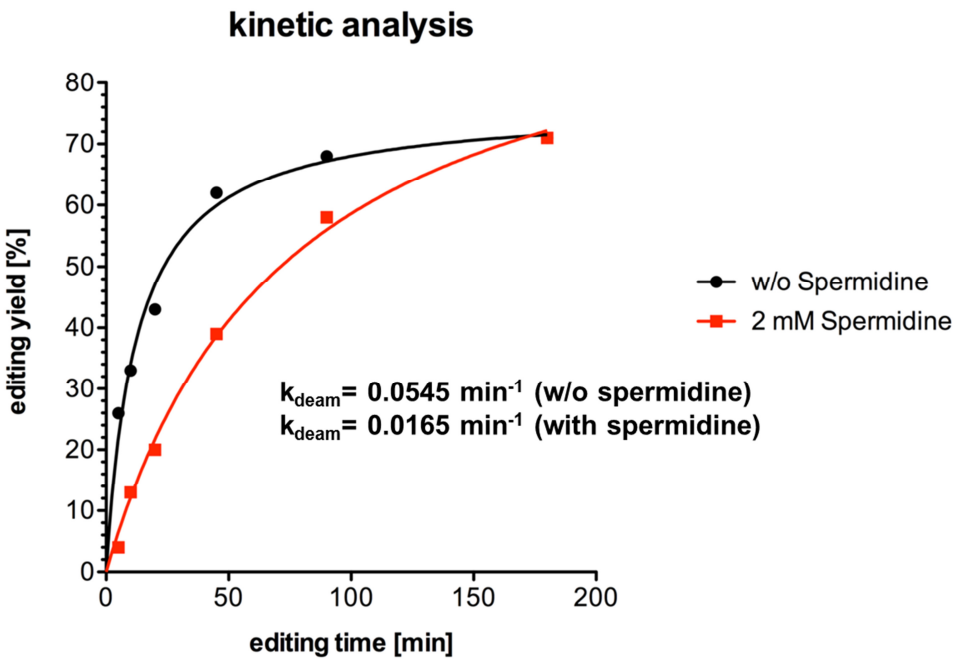

Figure 4-24: Plot of deamination product as a function of time. Reactions were carried out under single turnover conditions: 25 nM mRNA, 125 nM R/G-gRNA and 350 nM ADAR2 with and without spermidine (2 mM). Graph was generated using GraphPad Prism.

The off-target editing yields depend on the incubation time, as well (Figure 4-25). As obtained for the target editing position, the maximum editing yield for the off-target position 99 without spermidine was achieved within the first 90 min of incubation and was not increasable. Thus, the restriction of the reaction time is one possibility to control off-target editing. But the addition of spermidine to the reaction mixture abolishes off-target editing and was discussed in chapter 4.1.3.

The editing time can be reduced dependent on the question that is investigated, since the major part of the editing reaction takes place within the first 45 min without spermidine and 90 min with spermidine addition.

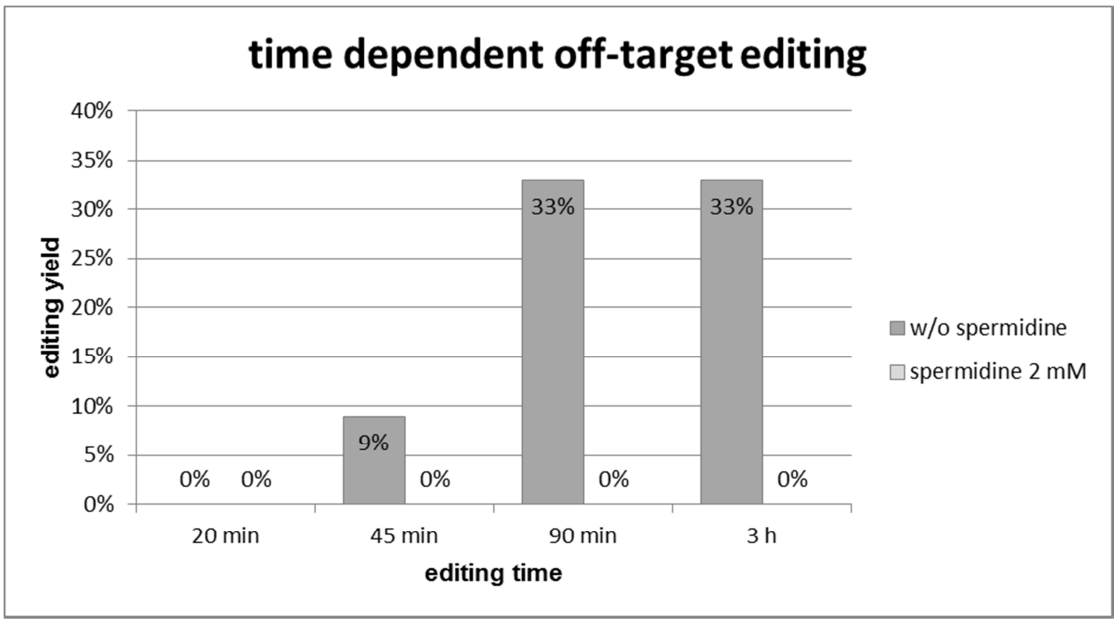

Figure 4-25: Time dependent off-target editing of W58X eGFP mRNA at position 99. The position 99 of W58X eGFP mRNA is edited time dependent with a maximum yield of 33% after 90 min without spermidine addition. In the first 20 min is no editing of the position 99 visible. The off-target editing is abolished completely by spermidine addition.

4.1.6 Is the R/G-motif necessary for steering ADAR2 to its target?

One of the main questions to be tested with the ADAR2 R/G-gRNA system is whether the R/G-motif is in fact necessary for guiding ADAR2 to its target adenosine and if the system of ADAR2 and R/G-gRNA works exclusively in combination or if one component can be replaced by another guideRNA or ADAR protein variant. To figure out if the R/G-gRNA can be replaced by another guideRNA, a ss-gRNA lacking the R/G-motif (Figure 4-26) was tested for the editing reaction. For the following experiments the standard editing reaction mix with varying magnesia concentrations was used but without spermidine (Table 3-23).

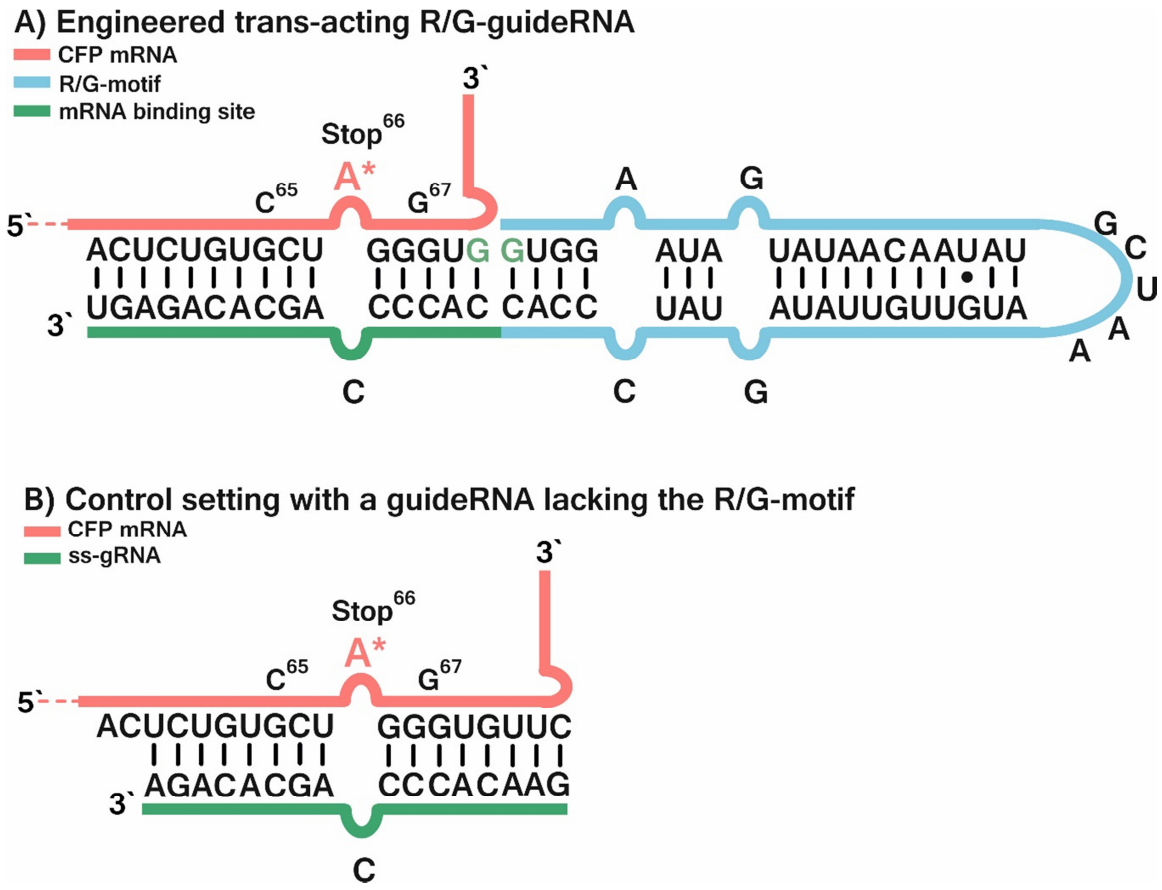

Figure 4-26: Engineered *trans*-acting R/G-gRNA and control ss-gRNA lacking the R/G-motif. The ss-gRNA is lacking the R/G-motif and the target adenosine is placed in the middle of the ss-gRNA sequence.

In the range from zero to 2.25 mM Mg^{2+}, there is no advantage of the R/G-motif to edit stop66 eCFP mRNA visible – the editing yields for both gRNAs stay more or less above 90% yield (Figure 4-27). But for the highest tested magnesia concentration (3 mM) some evidence is obtained that the R/G-motif is beneficial for steering ADAR2 to its target: the editing yield of the R/G-gRNA was 21% higher than that of the ss-gRNA. We then increased the stringency of the editing reaction by the addition of spermidine to the reaction mixture. As already mentioned in section 4.1.3 the ADAR2 R/G-gRNA system tolerates high magnesia and spermidine concentrations. The maximum editing yield for the R/G-gRNA was obtained at 3 mM Mg^{2+} with a 92% conversion. In contrast the editing yields at all chosen magnesia concentrations of the ss-gRNA were significant lower than the editing yields of the R/G-gRNA (Figure 4-28). The maximum editing level of 15% was achieved at 3 mM Mg^{2+} by the ss-gRNA. These reduced editing yields of ADAR2 with the ss-gRNA demonstrate the

definite advantage of the R/G-motif and highlights its necessity in steering ADAR2 to the target adenosine.

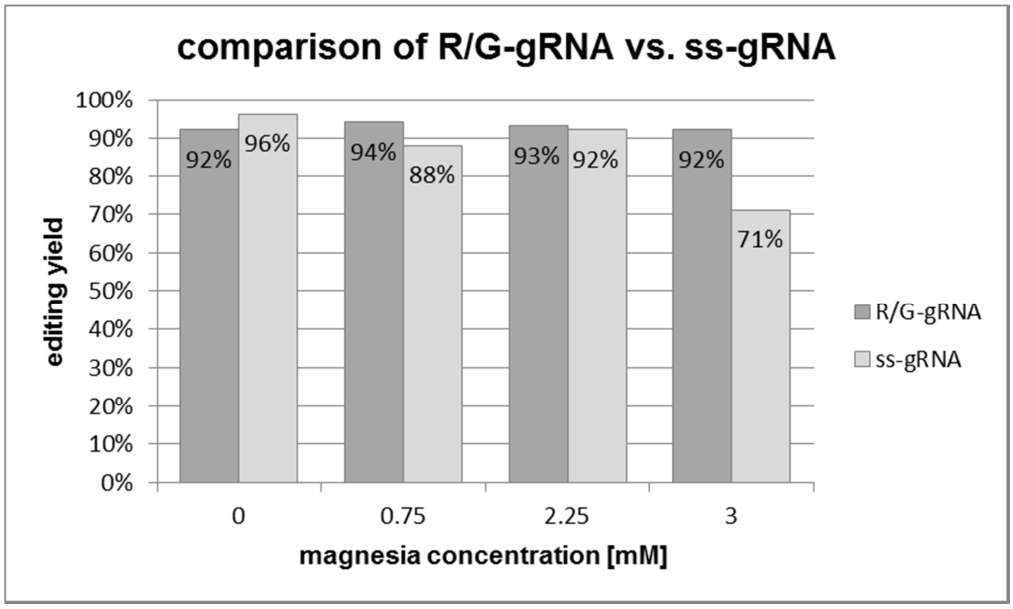

Figure 4-27: Comparison of ADAR2 editing with R/G-gRNA and ss-gRNA. In the range from zero to 2.25 mM Mg^{2+} is no significant advantage of R/G-gRNA about the ss-gRNA visible, if spermidine is absent in the standard reaction mix. At 3 mM Mg^{2+} the editing yield of ADAR2 with the ss-gRNA is decreased in comparison to the one of the R/G-gRNA.

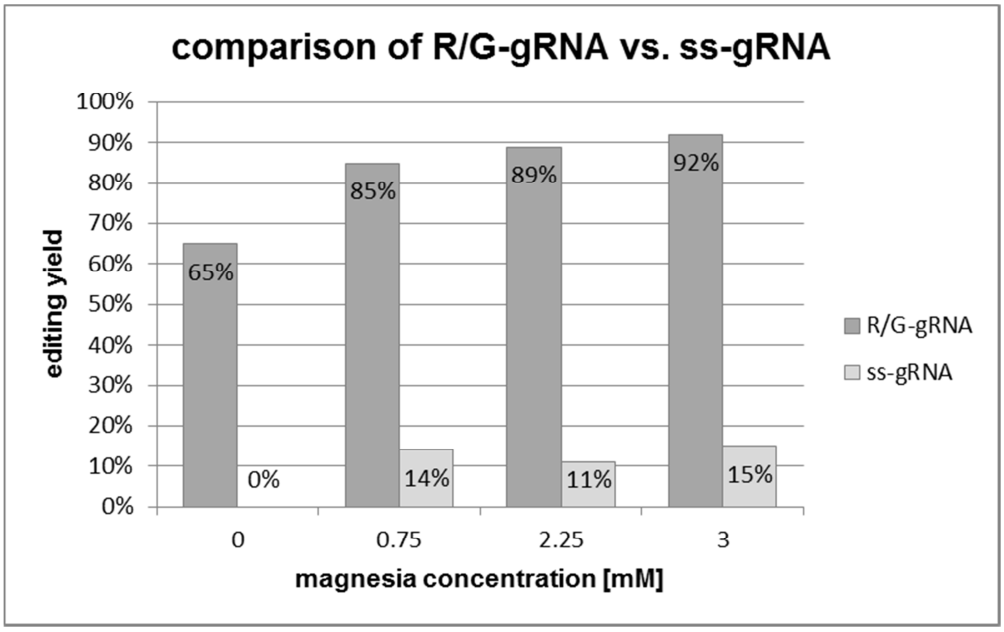

Figure 4-28: Comparison of ADAR2 editing with R/G-gRNA and ss-gRNA at 2 mM spermidine. The editing yield of ADAR2 with R/G-gRNA increases with rising magnesia concentration up to 92% at 3 mM Mg^{2+} and 2 mM spermidine in the reaction mix. In comparison the maximum editing level of ADAR2 with ss-gRNA is obtained with 15% at 3 mM Mg^{2+}.

Not only should the R/G-gRNA be specific and exclusive for the ADAR2 protein, ADAR2 should be specific and exclusive for the R/G-gRNA, as well. To figure out if ADAR2 can be replaced by another protein variant that can also be directed to deaminate the target adenosine, an ADAR2 variant from our previous studies – the so-called SNAP-ADAR2 - was used (49). The two dsRBDs of ADAR2 are replaced by a SNAP-tag (58) domain. A 5'-O-benzylguanine (BG)-modified gRNA can form a covalent bond with the SNAP-tag of the enzyme and thus the covalent gRNA-deaminase conjugate is directed to the complementary target sequence of the gRNA. For the first comparison of ADAR2 and SNAP-ADAR2 editing with the R/G-gRNA, the standard reaction conditions of ADAR2 where chosen: 2.25 mM Mg^{2+} and 3 mM Mg^{2+} with 2 mM spermidine addition. ADAR2 edited up to high yields (93%/92%) at both editing conditions, while SNAP-ADAR2 was only able to edit up to 10% at 2.25 mM Mg^{2+} (Figure 4-29). Editing at 3 mM Mg^{2+} and 2 mM spermidine was not possible with SNAP-ADAR2.

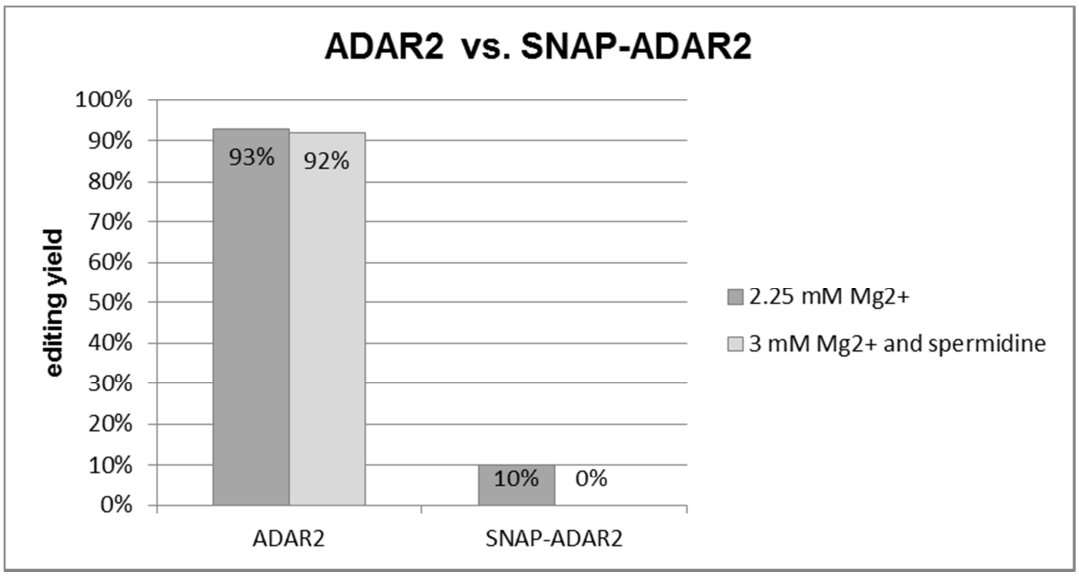

Figure 4-29: Comparison of ADAR2 versus SNAP-ADAR2 editing with R/G-gRNA. The editing yields of ADAR2 are at both reaction conditions (2.25 mM Mg^{2+} and 3 mM Mg^{2+} with 2 mM spermidine) higher than the editing yields of SNAP-ADAR2.

Since the ideal editing reaction conditions of SNAP-ADAR2 and ADAR2 are different, the preferences of SNAP-ADAR2 with the BG-gRNA and R/G-gRNA were compared to each other (Figure 4-30). SNAP-ADAR2 was able to edit the target adenosine guided by BG-gRNA at 0 mM Mg^{2+} and 0.5 µM heparin addition up to 100%. In comparison to the R/G-gRNA, editing of only 37.5% was possible at these preferred

editing conditions of SNAP-ADAR2. Interestingly, by the increase of the magnesia concentration up to 2.25 mM at same heparin concentrations, SNAP-ADAR2 still edits up to 71% with the BG-gRNA, whereas editing with the R/G-gRNA is not possible anymore.

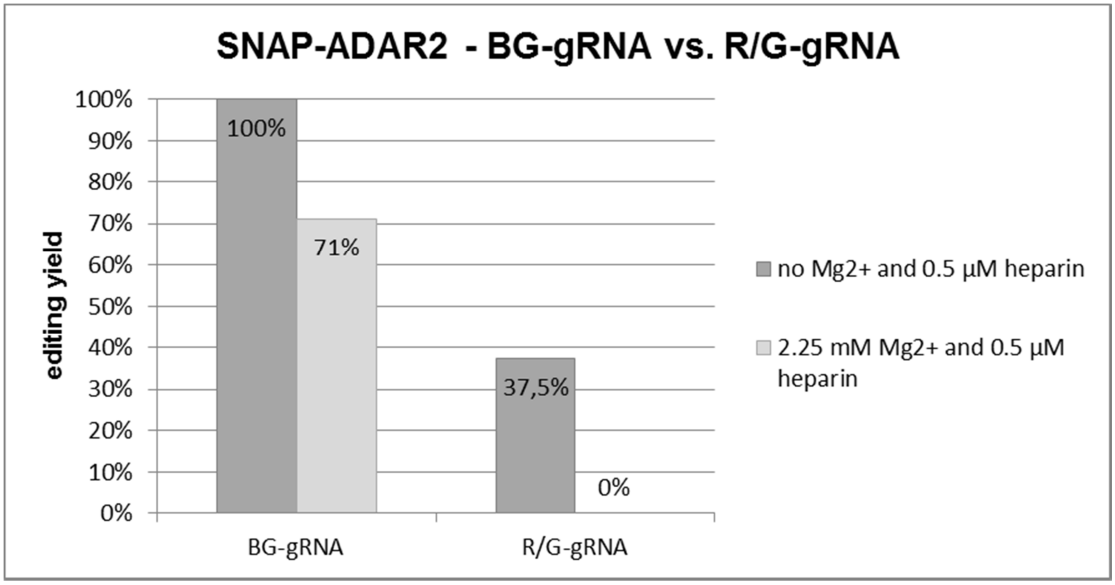

Figure 4-30: Comparison of BG-gRNA versus R/G-gRNA editing of SNAP-ADAR2. SNAP-ADAR2 achieves with BG-gRNA at both reaction conditions (no Mg^{2+} and 2.25 mM Mg^{2+} with 0.5 µM heparin) higher editing yields than with the R/G-gRNA.

Both, ADAR2 and R/G-gRNA can be replaced by another variant to some extent but only under loosing editing efficiency. On the other hand, it is easy to find conditions under which each protein, ADAR2 and SNAP-ADAR2, works exclusively with its respective engineered guideRNA. To summarize, the experiments of the comparison of R/G-gRNA with ss-gRNA and the comparison of ADAR2 with SNAP-ADAR2 show that the ADAR2 R/G-gRNA system works specifically and exclusively together. The two editing strategies are mutually orthogonal to each other.

4.1.7 Challenging the R/G-gRNA ADAR2 system to edit other codons and mRNAs targets

After demonstrating that the R/G-gRNA ADAR2 editing system highly edits 5'-UAG codons in eGFP and eCFP sequence contexts, we wanted to challenge the system further. Therefore, we addressed the system to edit 5'-UAG in a luciferase and PINK sequence context and also less preferred codons. The following Table 4-2 shows all

tested codons of the eCFP mRNA, as well as other gene sequence contexts and their codons that were analyzed as targets for in vitro editing.

Table 4-2: Target genes and their codons for in vitro editing analysis.

Reporter gene	Codon	Position
CFP	5‘-CAG	W66Q
CFP	5‘-GAG	W66E
luciferase	5‘-UAG	W417X
PINK	5‘-UAG	W437X
FUS	5‘-CAC	R521H

4.1.7.1 Editing of 5’-UAG codons in luciferase and PINK mRNA

The first challenge was to achieve high editing yields of 5’-UAG codons, as demonstrated for the eGFP and eCFP mRNA, but now in another sequence context of PINK and luciferase. Like for the 5’-UAG codon of eGFP and eCFP mRNA, the same 5’-CCA anticodon of the R/G-gRNA was used for editing the 5’-UAG codon in both mRNAs. For precise sequence details of the R/G-gRNAs see Table 6-3 in section 6.3. The 5’-UAG codon was editable in both sequence contexts, as well, but to a different extent (Figure 4-31): At standard editing reaction conditions (Table 3-23) and a magnesia concentration of 2 mM, the W417X luciferase mRNA was edited up to 93%, whereas the editing level of W437X PINK mRNA was 67%. The editing yield of W437X PINK mRNA could be improved up to 78% by decreasing the magnesia concentration down to 1 mM. A lower magnesia concentration might lead to a less strong secondary structure of the PINK mRNA and thus facilitates the editing reaction compared to higher magnesia concentrations.

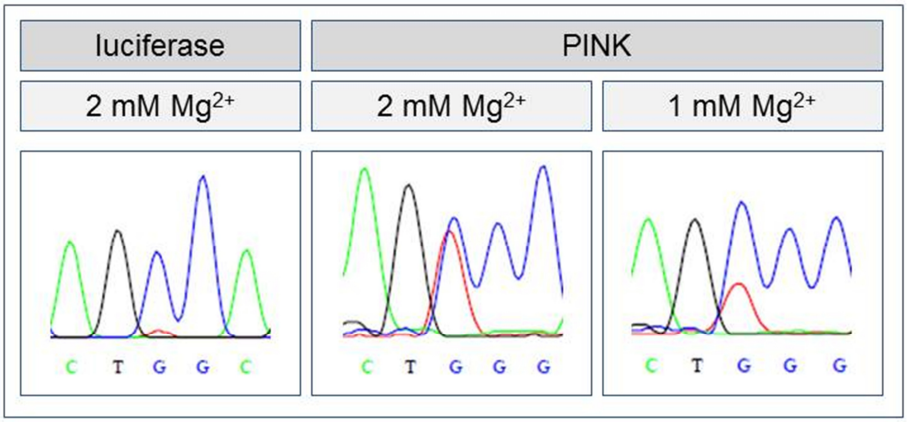

Figure 4-31: Editing of the 'UAG' codon of W417X luciferase and W437X PINK mRNA. W417X luciferase mRNA is edited with 93% at 2 mM Mg^{2+}. W437X PINK mRNA is edited at 2 mM Mg^{2+} with 67% yield. The editing yield of W437X PINK mRNA is increased up to 78% by a reduced magnesia concentration of 1 mM.

In both genes off-target editing was obtained with a maximum of 10% yield for luciferase and 40% yield for PINK. As already mentioned in section 4.1.3 off-target editing can be reduced by the addition of spermidine, but this was not tested.

The editing level of the target adenosine of the PINK mRNA is obviously lower than expected. Other codons were tested to be edited in the PINK mRNA, as well and it was found that all of the target adenosines were not converted to very high levels like in the eCFP or eGFP mRNA context. Different guideRNAs were constructed and it was reasoned, that the PINK mRNA is less addressable in general than the other mRNAs, since the PINK mRNA has a lot of self-pairing sequence parts with a high GC content (62,8 % vs. 54,9 % eCFP mRNA), which presumably forms more secondary structures than other mRNAs. Thereby, some parts of the PINK mRNA might not be addressable, but also the R/G-gRNA itself tends to hybridize with itself or with each other dependent on the sequence context and the magnesia concentration. This issue could turn out as a major limitation of an editing system applied in *trans*.

4.1.7.2 Editing of 5'-CAG and 5'-GAG codon in eCFP mRNA

The next challenge was to edit a 5'-CAG codon, which is known to be much more difficult than a 5'-UAG codon for deamination reaction, and was analyzed in the context of eCFP mRNA. In order to assess the ADAR2 protein's ability to edit the 5'-CAG codon of Q66 eCFP mRNA with high editing yields, two different anticodons

of R/G-gRNAs were tested, selected on our findings of the global editing analysis (49): 5'-CCG and 5'-CUA. The editing reaction was performed using the standard editing reaction mix (Table 3-23). The application of the 5'-CUA and 5'-CCG anticodon was expected to result in high editing yields, since both anticodons led in our previous studies to the highest editing levels for SNAP-ADAR2 (Figure 2-11) (49). However, the editing yields of the 5'-CCG anticodon in the R/G-gRNA were slightly higher (8%) than for the 5'-CUA anticodon at both conditions - with and without the addition of spermidine to the reaction mixture (Figure 4-32).

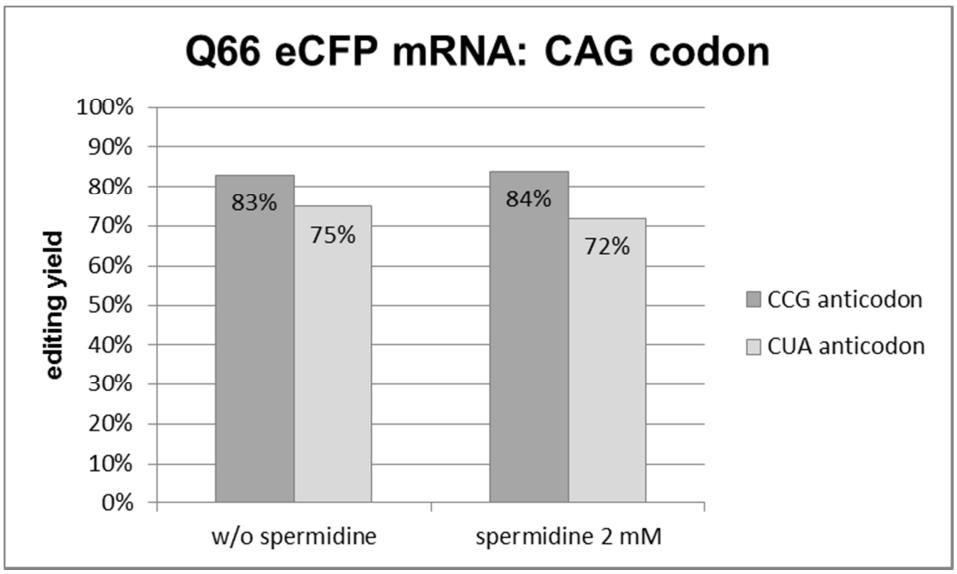

Figure 4-32: Editing of the Q66 eCFP mRNA CAG codon. Two different anticodons were tested for the ability to achieve high editing yields at the position Q66 eCFP mRNA. The 5'-CCG anticodon leads at both conditions, with and without the addition of spermidine to the reaction mixture, to slightly higher editing yields than the CUA anticodon.

Since the 5'-GAG codon is one of the most difficult codons to be edited (49), the challenge was to achieve high editing yields for its deamination. Therefore, a set of three different anticodons of R/G-gRNAs was tested: 5'-CCG, 5'-CUA and 5'-CCC. According to our findings from 2014 (49) the anticodon 5'-CCG or 5'-CUA was expected to be the most advantageous one. As expected, the anticodon 5'-CCG leads to the highest editing yields of E66 eCFP mRNA with 55% (Figure 4-33). Correlating to the results of the previous study of the SNAP-ADAR BG-gRNA system the editing yield of about 55% seems to have its maximum at this level with both editing systems. The other tested anticodons of the R/G-gRNAs resulted in much lower editing levels. The 5'-CUA anticodon mediated 29% editing, whereas the

5’-CCC anticodon only led to 5% editing. Contrary to the findings from 2014, 5’-CUA was not that suitable for mediating high editing yields of the 5’-GAG codon with ADAR2. As expected, the use of the 5’-CCC anticodon resulted in lowest editing yields. Off-target effects for eCFP mRNA are discussed in chapter 4.1.3.

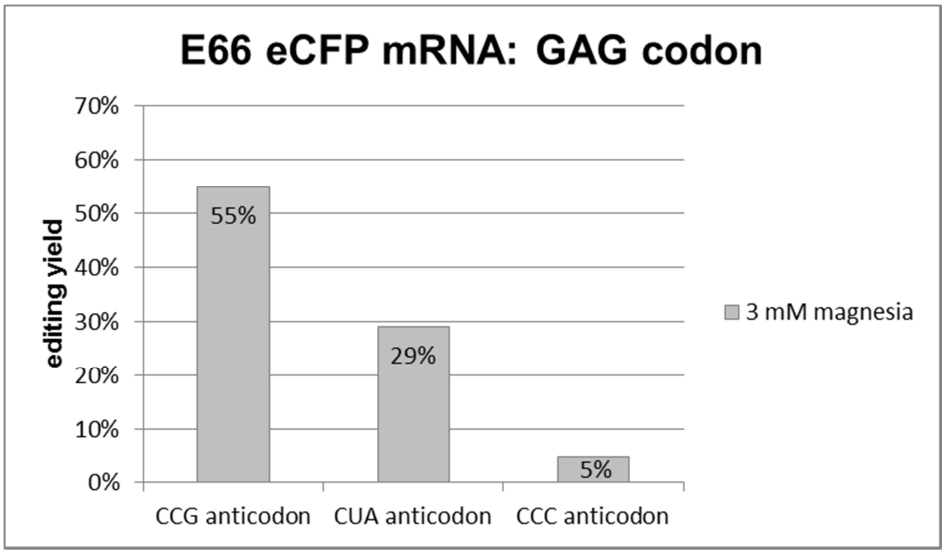

Figure 4-33: Editing of the E66 eCFP mRNA GAG codon. Three different anticodons were tested for the ability to achieve high editing yields at the position E66 eCFP mRNA. The editing with the 5’-CCG anticodon achieved the highest yields with 55%. The editing with the 5’-CUA anticodon lead only to 29% editing whereas the 5’-CCC mediated only 5% editing at the 5’-GAG codon.

4.1.7.3 Editing of 5’-CAC in FUS mRNA

As mentioned, the future aim is to address the genetically encodable editing system towards disease causing single point mutations (2.7.1). Therefore, the R521H FUS mutation was chosen. We wanted to demonstrate that the editing system is able to edit less preferred codons such as 5’-CAG or 5’-CAC codons not only in a eGFP or eCFP sequence context, but also in the R521H FUS sequence context (5’-CAC). The experiments were performed using the standard editing reaction mix (Table 3-23) except for the magnesia concentration, which was adjusted to 1.5 mM. The R521H FUS mRNA with the codon 5’-CAC was editable by the 5’-GCG anticodon gRNA of 15 nt or 16 nt length up to 50% (Figure 4-34). The editing yields of both lengths of the R/G-gRNAs were similar. The replacement of the counter base in the anticodon from cytosine (GCG) to uridine 5’-GUG lead to a complete loss of editing at the target position R521H FUS. But the change of the 3‘-neighbor of the counter base from

guanosine to adenosine 5'-GUA restored the editing ability again (Figure 4-35). The usage of this 16 nt R/G-gRNA allowed editing up to 51% at 3 mM Mg^{2+}.

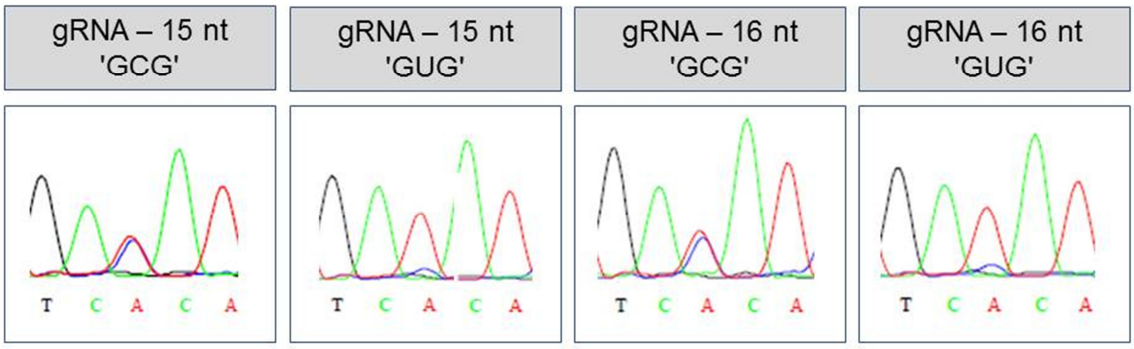

Figure 4-34: Editing of R521H FUS mRNA with different anticodons and length of R/G-gRNAs. R521H FUS mRNA was editable with the anticodon 5'-GCG up to 45%, whereas editing with the anticodon 5'-GUG was not successful. The yields for both lengths of the gRNA (15 nt or 16 nt) were similar.

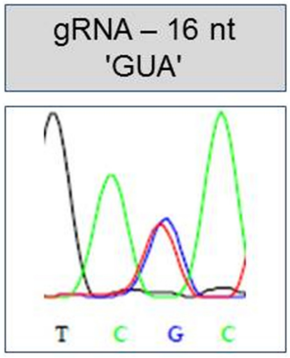

Figure 4-35: Editing of R521H FUS mRNA with 'GUA' anticodon. The R/G-gRNA with a 5'-GUA anticodon and 16 nt length allowed editing of the FUS R521H mRNA up to 51% at 3 mM Mg^{2+}.

Unfortunately, the FUS mRNA is difficult to reverse transcribe – probably because of some secondary structure formations that hamper the reverse transcriptase – and the following Taq PCR of this generated cDNA always results in a shortened product. Different optimizations took place from varying the annealing temperature and the usage of a new set of primers for amplification, but none of these trials resulted in an improved cDNA and PCR product amplification. This is the reason why no statements for the off-target editing of this mRNA can be made, since the full range of mRNA could not be analyzed. But a valuable result from the FUS R521H in vitro editing was obtained: editing of the 5'-CAC codon is possible with the R/G-gRNA comprising a 5'-GCG or 5'-GUA anticodon.

To summarize, the following codons are addressable with an engineered R/G-gRNA steering ADAR2 to its target mRNA: 5'-UAG, 5'-CAG, 5'-CAC and 5'-GAG. These codons are editable to at least 50% depending on the chosen anticodon of the R/G-gRNA.

4.1.8 Steering E488Q for site-specific RNA editing - a more active mutant of ADAR2

The following experiments of this chapter were performed together with Tahsin Kuzdere during the supervision of his bachelor thesis. Like in the previous chapter demonstrated, we wanted to address also less preferred codons, such as 5'-GAG or 5'-CAG codons. A further solution beside optimization of the editing conditions and anticodons of the guideRNA is the application of a more active editing mutant of ADAR2, named E488Q. This mutant was described by Kuttan and Bass to target 5'-GAC codons at high editing yields (53) (2.5.1). We analyzed if this mutant of ADAR2 is also steerable with the engineered R/G-gRNA to edit adenosines of a 5'-GAG and 5'-CAG at high levels and we compared the results with wild type ADAR2.

The experiments were performed using the standard editing reaction mix (Table 3-23) and the same anticodons for the guideRNAs, as for the wild type ADAR2. The 5'-CCG, 5'-CUA and 5'-CCC anticodons were selected according to our findings from 2014 (49) to achieve high editing yields for the 5'-GAG codon. E488Q and ADAR2 differ in their preference of R/G-gRNAs to edit the 5'-GAG triplet at high editing levels (Figure 4-36). Regarding the 5'-GAG codon, wild type ADAR2 deaminates adenosine to inosine best with a 5'-CCG anticodon (55%) and less effective with 5'-CUA (29%). E488Q shows a reverse preference for those anticodons: Using the 5'-CCG anticodon resulted in 44% editing, whereas the 5'-CUA anticodon achieved 56% yield. But surprisingly, the 5'-CCC anticodon, which was not successful with wt ADAR2 to convert adenosine to inosine, achieved now in combination with E488Q the highest editing yields of 89% at the 5'-GAG codon. Referring to our previous studies of anticodon preferences, the 5'-CCC anticodon, which completely base pairs except for the target adenosine, was not expected to be the most suitable anticodon to mediate 5'-GAG editing. The reason of the changed anticodon preference and the fact that E488Q achieves higher editing yield must be caused by changes during the

base flipping mechanism. This mechanism is assumed to be facilitated by the mutation of E488Q (2.5.1). According to further studies that analyzed mutations in the side chain of E488 (or E1008 in ADAR1) and proposed that ADAR uses a similar mechanism as the M.HhaI Mtase (53,55,56), this residue promotes not only base flipping, it contacts and stabilizes the orphaned counter base, as well. Since there is no difference in the counter base of the 5'-C**C**G or 5'-C**C**C anticodon itself, it is very likely that this residue is not only involved in contacting the counter base, but also interacting or been influenced by the 3'-neighbor of the counter base, as well. Therefore, the E488Q has a different preference of the anticodon as the wild type ADAR2.

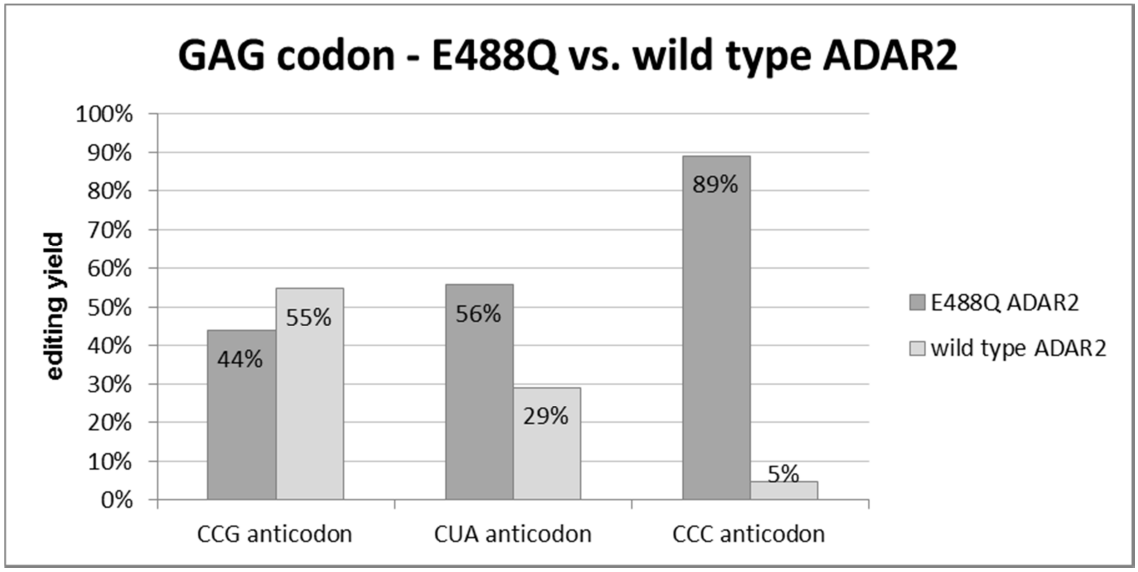

Figure 4-36: Editing of the 5'-GAG codon – E488Q in comparison to the wild type ADAR2. Both enzymes have different preferences of R/G-gRNAs for efficient editing of the 5'-GAG codon. Wild type ADAR2 prefers the 5'-CCG anticodon most, whereas E488Q converts adenosine to inosine best with the 5'-CCC anticodon.

The enhanced editing activity of E488Q for 5'-GAG codons has a downside of more off-targeting positions (Figure 4-37) compared to the ADAR2, whose common off-target positions are 99 and 127. Some of the additional and already known off-target positions are edited up to very high levels and are independent of the chosen R/G-gRNA (Figure 4-38).

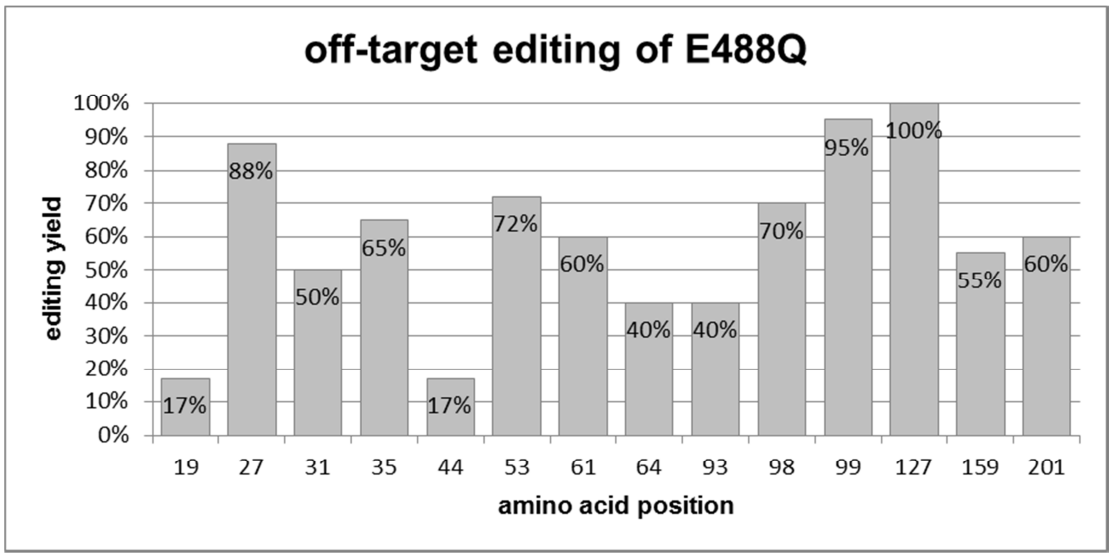

Figure 4-37: Off-target editing positions of E488Q in eCFP mRNA. Many off-target editing sites appear for the standard editing reaction mix of E488Q with eCFP mRNA and the 5'-CCC anticodon of the R/G-gRNA. Some of these off-target editing positions achieve high editing yields. The numbering of position 1 starts with the 5'-AUG start codon of eCFP mRNA and all amino acid positions refer to the consecutive numbering.

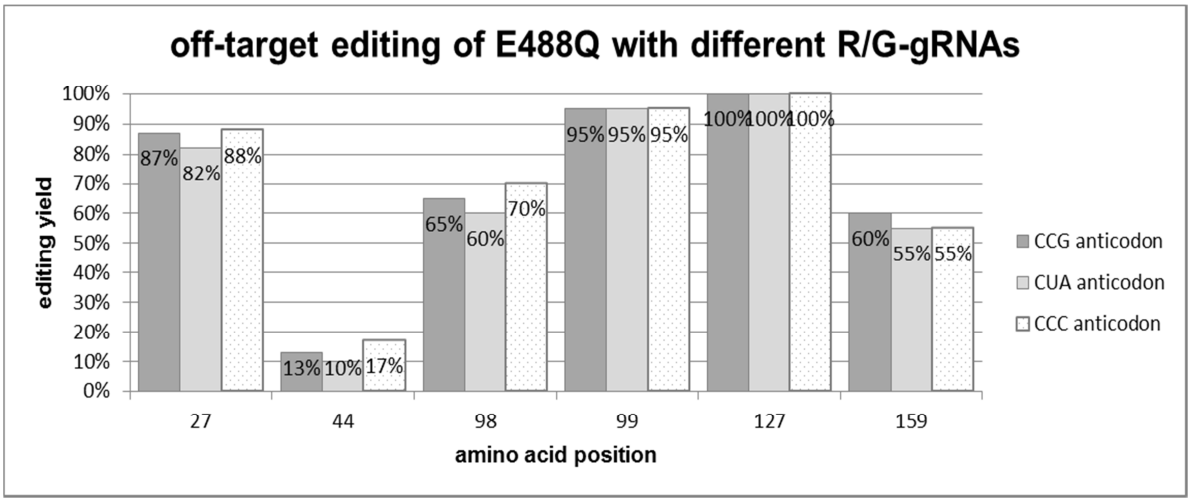

Figure 4-38: Off-target editing positions of E488Q in eCFP mRNA with different R/G-gRNAs. The extent to which the off-target editing sites are edited is independent from the anticodon of the R/G-gRNA. The numbering of position 1 starts with the 5'-AUG start codon of eCFP mRNA and all amino acid positions refer to the consecutive numbering.

The addition of spermidine to prevent the E488Q from over editing was investigated in the context of a 5'-CAG triplet, which has an identical mRNA sequence except for the target codon.

In the work of Kuttan and Bass was not investigated whether the ADAR2 mutant is more active in editing a 5'-CAG codon, as well, but we assumed. Therefore, the same

experimental setup was used as for the 5'-GAG editing with two different R/G-gRNAs, to compare the editing yields of E488Q with the ones of ADAR2. The analysis included the anticodons 5'-CCG and 5'-CUA.

Both ADAR2 enzymes show the same preference for the anticodon of the R/G-gRNA; the highest editing yields were obtained for the 5'-CCG anticodon (Figure 4-39), which matches to our previous studies (49). The editing levels of E488Q were higher with both chosen R/G-gRNAs than of ADAR2, demonstrating that the point mutation of glutamic acid to glutamine is not only improving 5'-GAG and 5'-GAC triplet editing, but also 5'-CAG codons. The use of spermidine in the reaction mixture is not affecting the editing yield of the target editing position, but also hardly influencing the off-target editing (Figure 4-40).

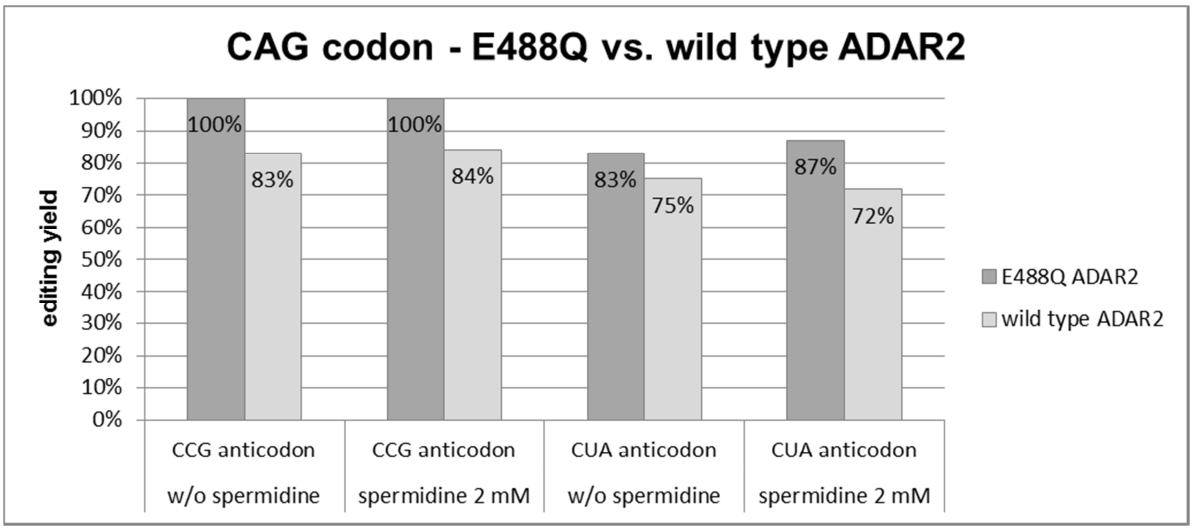

Figure 4-39: Editing of the 5'-CAG codon – E488Q in comparison to wild type ADAR2. The two enzymes have the same preferences of R/G-gRNAs for efficient editing of the 5'-CAG codon. Both achieved highest editing yields with a 5'-CCG anticodon containing R/G-gRNA. E488Q edited the target position to a higher extend than the wild type ADAR2. The editing yield of the target position is not influenced by spermidine.

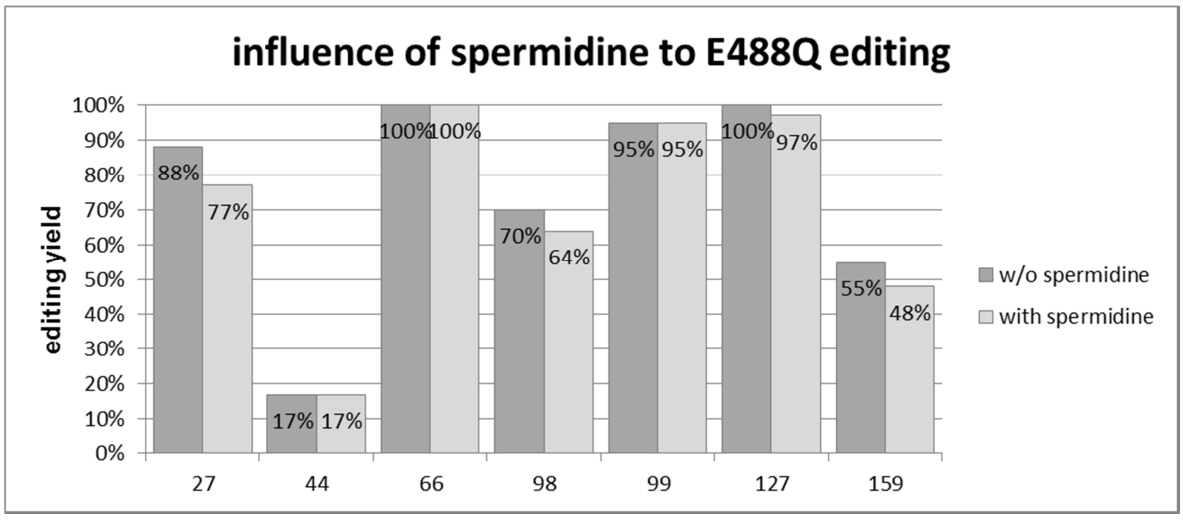

Figure 4-40: Influence of spermidine to the editing yield caused by E488Q. The presence of spermidine in the reaction mixture was not affecting the editing yield at the target position 66. But also the off-target editing remained less influenced by the spermidine.

E488Q seems to be much more active than the wt ADAR2. To prove the assumption that E488Q is catalytically more active and faster in editing than the wild type ADAR2, experiments with reduced editing times were performed. The editing times were reduced to 45 min and 1.5 hours with a standard editing reaction mix without spermidine. E488Q edits the 5'-CAG codon with a 5'-CCG anticodon R/G-gRNA much faster than the wild type ADAR2 (Figure 4-41), confirming that E488Q is more active and faster in editing than the wild type ADAR2. The same effect of activity was obtained at off-target sites, with longer incubation time, the editing yields of the off-target sites of E488Q increase (data not shown). As illustrated in chapter 2.5.1 the glutamate residue 488 is located on the loop above the catalytical site of ADAR2. The change of glutamate to glutamine is supposed to facilitate the entrance of the mRNA substrate and thus substrate binding. By this amino acid change the enzyme kinetics might be increased.

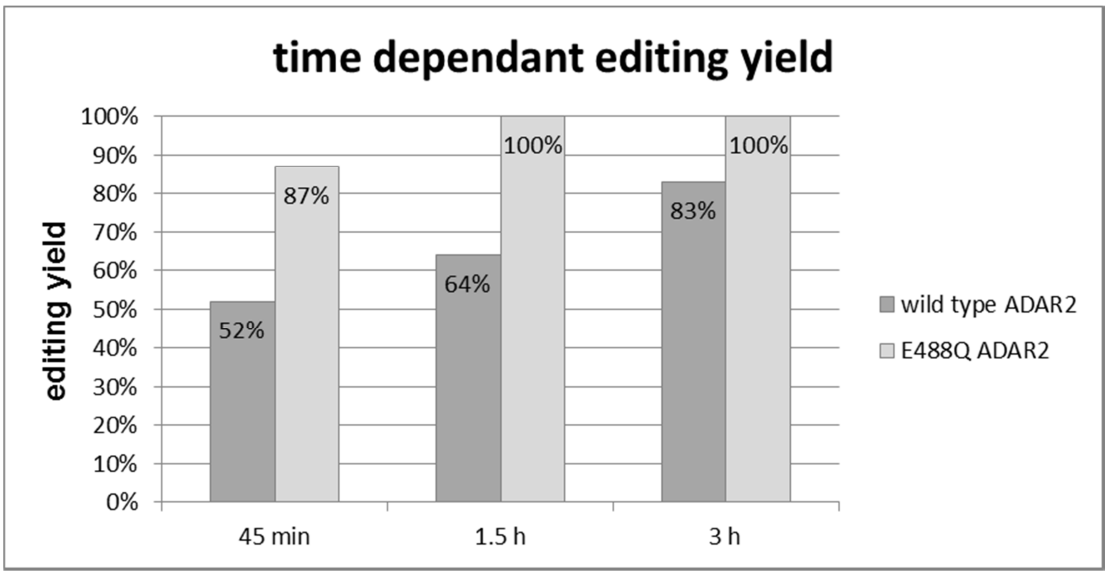

Figure 4-41: Time dependant editing yield of wild type and E488Q ADAR2. The editing time to convert adenosine of a 5'-CAG codon with a 5'-CCG anticodon R/G-gRNA was reduced to 45 min and 1.5 h. E488Q is faster and more active than the wild type ADAR2, indicated by higher obtained editing yields at every incubation time.

To summarize, E488Q can be steered by our engineered guideRNA, such as wild type ADAR2, and can be used for the deamination of less preferred codons. The point mutation of glutamic acid to glutamine leads to an increased catalytic rate for editing in general, since not only 5'-GAC codons are edited to higher levels, but also 5'-GAG and 5'-CAG codons and various off-target codons. The codon 5'-UAG was tested, as well and resulted in full conversion of adenosine to inosine, as did in the work of Kuttan and Bass (data not shown). The preferences of the two ADAR2 enzymes for most efficient editing of the target position 66 eCFP differ for the 5'-GAG codon; ADAR2 edited best with the 5'-CCG anticodon, whereas E488Q achieved highest editing levels with the 5'-CCC anticodon. The 5'-CAG codon was edited by both enzymes best in combination with a 5'-CCG anticodon R/G-gRNA. Spermidine has not been found to be useful as additive in the editing reaction mix for controlling the off-target editing sites of E488Q.

4.1.9 Is ADAR´s substrate binding via the dsRBD2 bothered in the trans-acting system?

Naturally, the glutamate receptor transcript is edited by ADAR2 in *cis*, which means that the double stranded RNA substrate is formed from only one RNA molecule that

folds back on itself. Since our designed R/G-gRNA is acting in *trans* to target any arbitrary mRNA, the protruding 3'-part of the target mRNA might disturb the recognition and binding of ADAR2. From literature we know that the dsRBM2 of ADAR2 recognizes and binds to that region of the natural R/G-motif, where the exit of the 3'-part of the mRNA in the *trans* editing situation is created (35). In order to analyze the effect that the mRNA exit has on the recognition and binding of ADAR2 and to figure out the contribution of the dsRBM2 of ADAR2 in the *trans* setting, a deletion mutant lacking the dsRBM1 domain was constructed, called short ADAR2. For plasmid construction and production of short ADAR2 see chapter 3.3. Details of protein expression and isolation can be found in section 6.3. By comparing short ADAR2 with wt ADAR2, we wanted to evaluate differences in the yields for editing in *trans* versus *cis* (Figure 4-42). Therefore, an eCFP mRNA was created that includes the R/G-gRNA, mimicking the editing situation in *cis*. This so called mRNA+R/G construct was obtained by two stepwise overlap extension PCRs (Figure 4-43).

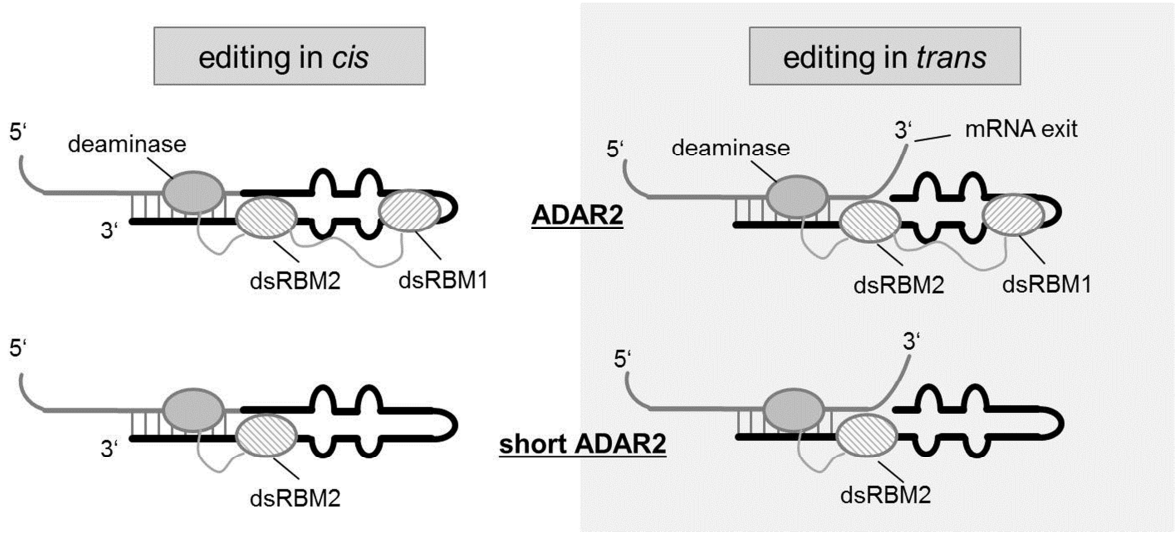

Figure 4-42: Editing in *cis* and *trans* with ADAR2 and short ADAR2. The editing situation in *trans* differs from the editing situation in *cis* by having a 3'-mRNA exit at the binding site of dsRBM2. Short ADAR2 lacks the dsRBM1 for recognition and binding.

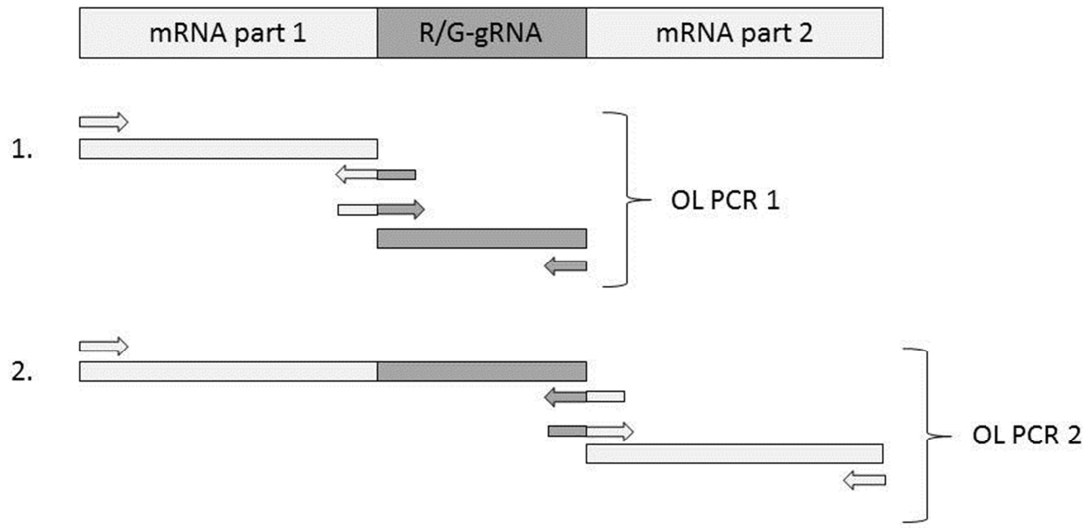

Figure 4-43: Illustration of stepwise overlap extension PCRs for generating the mRNA+R/G construct. The mRNA+R/G construct consists of three parts: the first part of the mRNA, ending 4 nucleotides after the stop66 codon, the R/G-gRNA and the second part of the residual following part of the eCFP mRNA. The mRNA+R/G construct was cloned by two stepwise overlap extension PCRs.

The first overlap PCR fused the first part of the eCFP gene, which ends 4 nucleotides after the stop66 codon, together with the R/G-motif. The 5'-part of this overlap was created by a Phusion PCR on pTS8 with the primer pair no. 126 and 186 that contains 1/3 part of the R/G-gRNA motif and the 3'-part was amplified by the primer pair no. 183 and 148 from pTS11.2 to map the R/G-motif and a part of the gRNA. The overlap PCR was performed with the flanking primers no. 126 and 148. In the second overlap PCR the first generated overlap fragment served as a template for a Phusion PCR with primers no. 126 and 184 to generate the 5'-product, containing the first part of the eCFP gene, the whole part of the R/G-gRNA and a small section of the second part of the eCFP mRNA. The 3' product was generated by a Phusion PCR using the primer pair no. 185 and 1 with pTS8 as a template, to fuse the missing part of the eCFP gene to the 5'-construct. The overlap PCR was done with the primer pair no. 126 and 1. In every Phusion PCR DMSO (5%) and the special GC-buffer were used that are provided by NEB. The mRNA+R/G construct was inserted by the restriction sites XbaI/XhoI into pMG211 vector, creating pTS71. The mRNA+R/G template for editing in *cis* was in vitro transcribed as described in chapter 3.2.13.

At first the editing in *cis* with the mRNA+R/G construct was tested with both ADAR2 protein variants, the wild type and the short version, applying standard editing conditions (Table 3-23). An mRNA concentration of 25 nM was chosen. Both

enzymes are able to fully convert adenosine to inosine in the mRNA+R/G molecule (Figure 4-44).

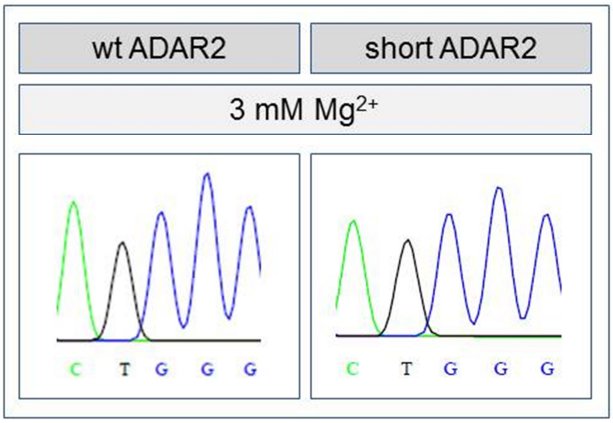

Figure 4-44: Editing of the mRNA+R/G construct by wild type and short ADAR2. Both enzymes fully edit the mRNA+RG construct at position 66 at 3 mM magnesia.

The use of the mRNA+R/G construct and the sequence analysis of it allowed catching more details about the fate of the R/G-motif during the editing reaction. In the *trans* editing situation this was never observed, since the R/G-gRNA was never reverse transcribed and analyzed. Nine sites in the R/G-motif were found to be edited by both enzymes (Figure 4-45). Interestingly, short ADAR2 edits these nine positions at higher levels than wt ADAR2. Presumably, short ADAR2 is more promiscuous for editing in general, since the dsRBM1 is lacking. Another explanation can be assumed as well: Since the dsRBM1 of the ADAR2 binds naturally the stem-loop region of the R/G-motif (34,35,90), the lack of dsRBM1 opens the possibility that the deaminase domain gets more access for editing this region, resulting in higher editing yields at these nine positions.

In the literature (96) it was mentioned that the R/G-motif itself is strongly edited by ADAR2, as well and this could be confirmed by our findings. Since the highly edited part of the R/G-motif of the natural glutamate receptor is removed during splicing, this hyper-editing is not affecting the sequence of the mRNA transcript of the glutamate receptor.

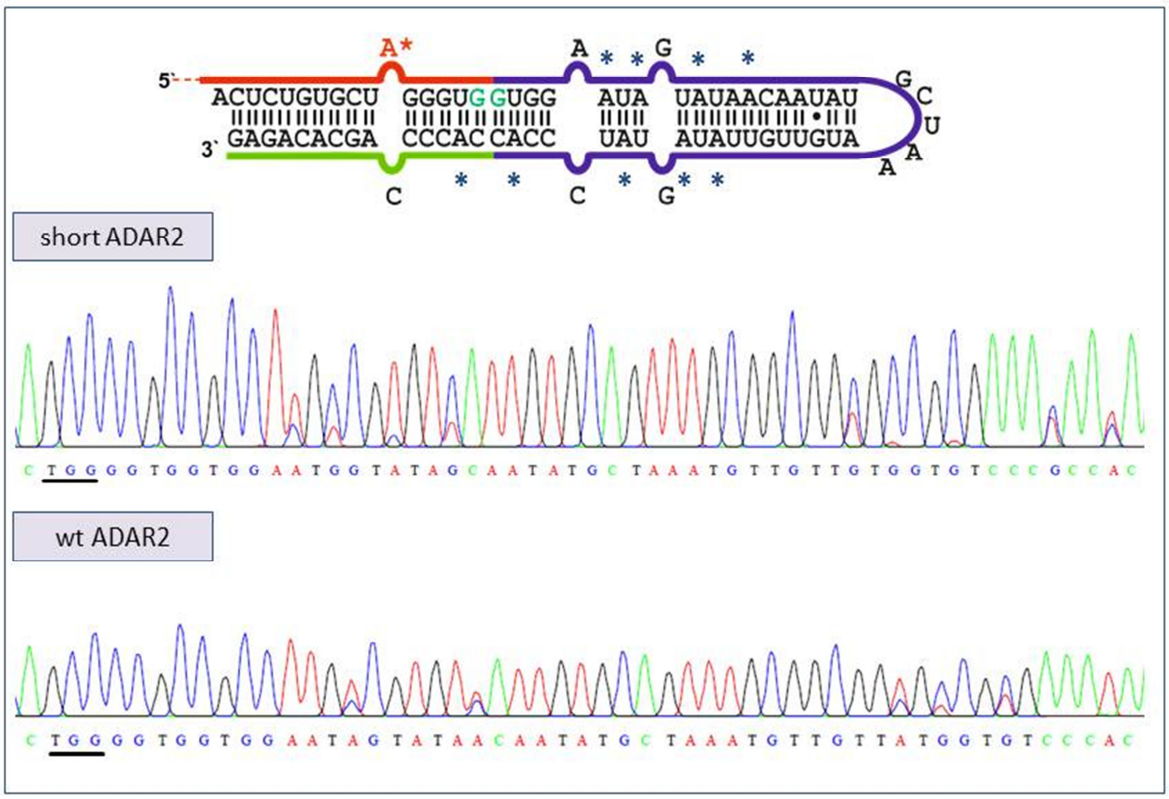

Figure 4-45: The R/G-motif editing in the mRNA+R/G construct. There are 9 more editing sites for wt and short ADAR2 in the mRNA+R/G construct. Short ADAR2 edits these 9 positions at higher yields than wt ADAR2.

Regarding the target position stop66 eCFP, neither the addition of 2 mM spermidine nor high heparin concentrations of 2 µM hampered editing of both enzymes (Figure 4-46). The observation that both enzymes are insensitive towards spermidine and heparin addition is only valid for the target adenosine and not for the two main off-target sites 99 and 127. The off-target editing of both positions is completely prevented by the addition of 2 mM spermidine or 2 µM heparin (data not shown). The hyper-editing of the R/G-motif was reduced, but for both enzymes not completely prevented (Figure 4-47). In the *trans* situation neither wt ADAR2 nor short ADAR2 showed this tolerance at the target editing site, also not for low heparin concentration (Figure 4-48). The editing yields achieved approximately 50% for both enzymes in the *trans* setting.

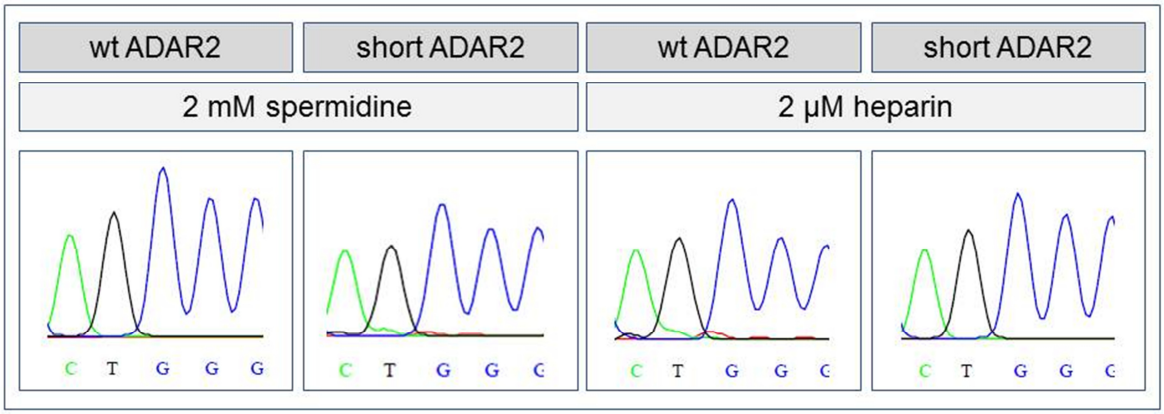

Figure 4-46: Editing of mRNA+R/G construct at increased stringency. At 2 mM spermidine or 2 µM heparin both enzyme still fully convert adenosine to inosine at the target position.

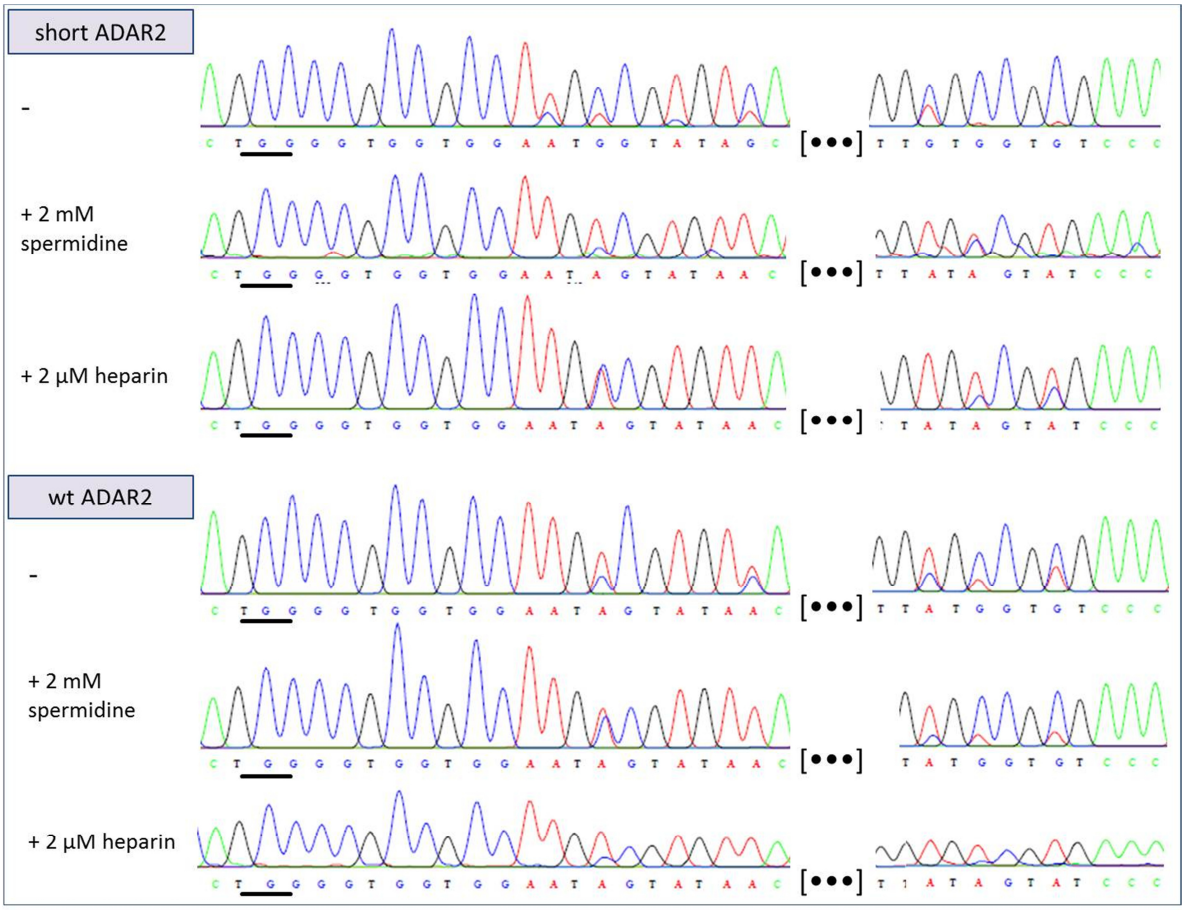

Figure 4-47: The R/G-motif editing in the mRNA+R/G construct in presence of spermidine and heparin. In presence of 2 mM spermidine or 2 µM heparin the hyper-editing of both enzymes – wt ADAR2 and short ADAR2 – at the R/G-motif is reduced, but not completely prevented. The missing middle part of the R/G-motif is devoid of editing sites and is indicated as dots.

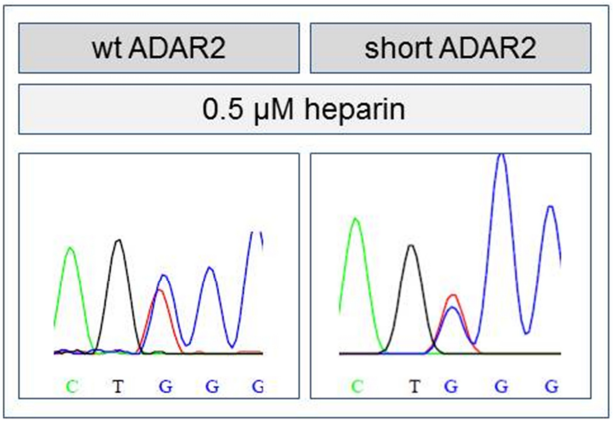

Figure 4-48: Editing in *trans* of wt ADAR2 and short ADAR2 at 0.5 µM heparin. Wt ADAR2 and short ADAR2 are both affected by heparin addition to the reaction mixture for editing in *trans* – both editing yields resulted in approximately 50%.

Surprisingly, short ADAR2 was still able to edit adenosine at full conversion in the *trans* editing situation, if only the standard editing reaction mix was used, containing 3 mM magnesia (Figure 4-49). As soon as spermidine or heparin was added to the reaction mixture, short ADAR2 lost its ability to edit the target adenosine at high levels. The inhibiting effect of spermidine for editing was stronger than the influence of the heparin concentration. This general increased sensitivity towards spermidine was not confirmed by analyzing the off-target editing of position 99 and 127: heparin fully prevented off-target editing, whereas spermidine prevented off-target editing at position 99, while only reducing editing at position 127 to 45% (data not shown).

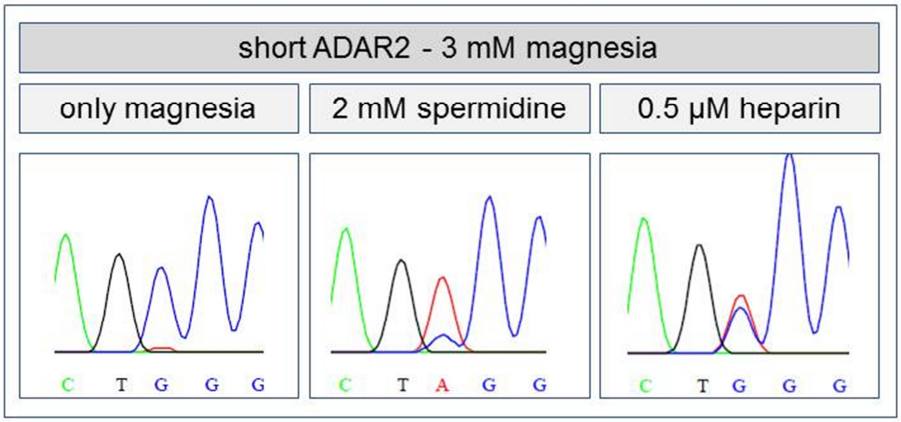

Figure 4-49: Editing yields of short ADAR2 at standard and stressful editing conditions. Short ADAR2 edits at standard editing conditions with 100% whereas the addition of spermidine or heparin decreases drastically the editing level. Spermidine affects the editing yield stronger than heparin.

The effect that short ADAR2 is more sensitive towards spermidine than wt ADAR2 is caused by the lack of dsRBM1. Therefore, wt ADAR2 showed a less extensive reduction of the editing yields at spermidine addition (Figure 4-50). The presence of heparin decreased the editing yield at the target adenosine of both enzymes similarly.

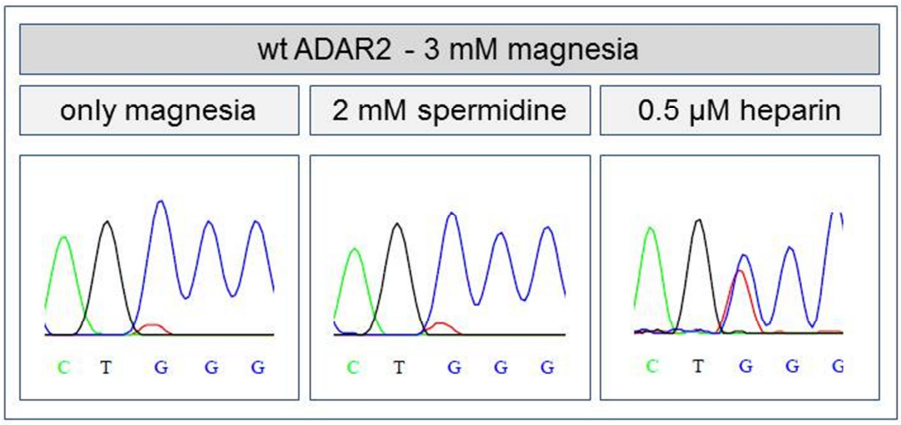

Figure 4-50: Editing yields of wt ADAR2 at standard and stressful editing conditions. Wt ADAR2 edits at standard editing conditions and spermidine addition up to 92% whereas the addition of heparin decreases the editing level to 57%. Only heparin addition affects the editing yield of wt ADAR2.

These experiments demonstrate that editing in *cis* with the R/G-motif is very insensitive towards disturbing additives, such as heparin or spermidine, no matter if the full-length or shortened enzyme is used. Obviously, the binding sites of the mRNA are well formed and one of the dsRBM is sufficient for binding and catalysis. Also editing in *trans* is efficient with both enzyme variants. However, editing in *trans* is very sensitive towards heparin addition for both enzymes. Since heparin binds to the deaminase domain and the dsRBMs, the hampering effect is similar for both enzymes (Figure 4-51). The editing in *trans* with spermidine addition shows a drastic difference for both enzymes in their editing efficiency: While full-length ADAR2 is completely unperturbed, short ADAR2 loses its ability to efficiently edit the target adenosine. Spermidine binds competitively to the mRNA (Figure 4-51). Now, there is a huge difference if ADAR2 has both dsRBMs, or only one. The multivalent binding of the full-length ADAR2 is more difficult to disturb, demonstrating the importance of dsRBM1. The observation that spermidine reduces the editing ability of short ADAR2 only at editing in *trans* and not in *cis,* indicates that the exit of the mRNA at the editing in *trans* setting affects and hampers the recognition and binding of the dsRBM2, in comparison to the *cis* editing situation.

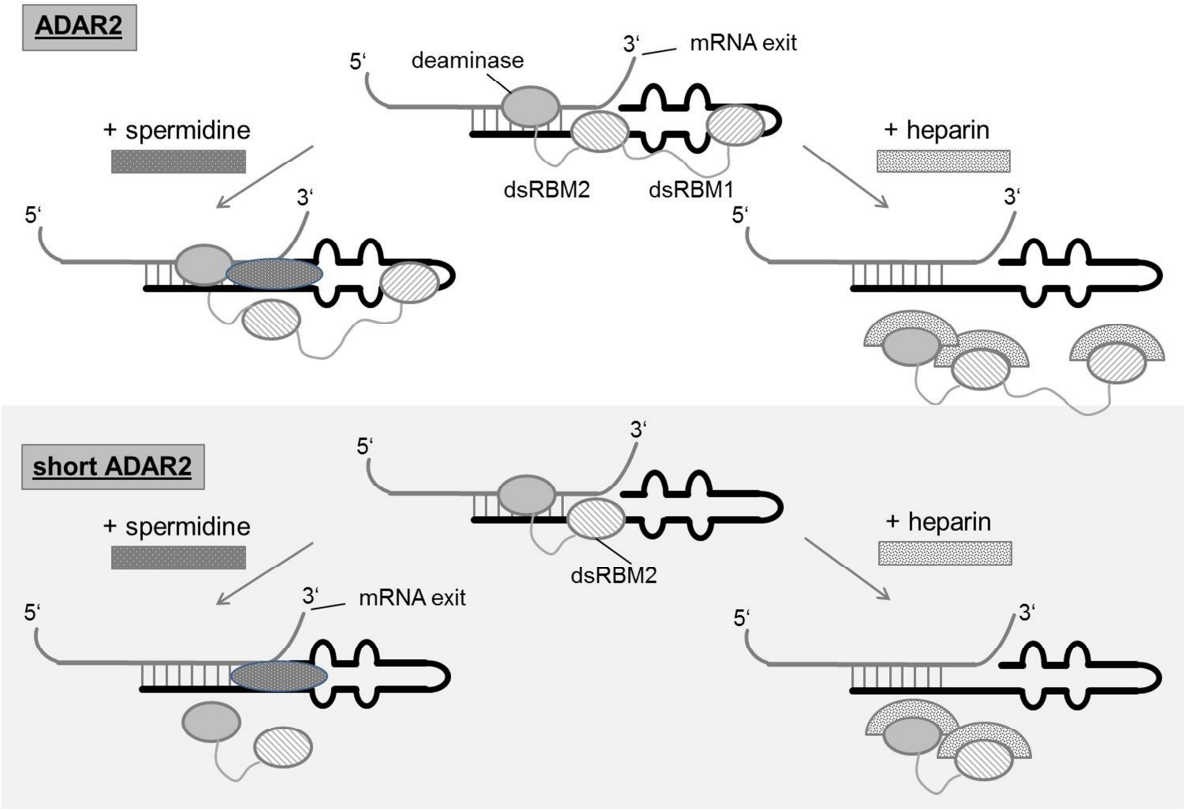

Figure 4-51: Editing in *trans* of both ADAR2 variants with heparin or spermidine present in the reaction mix. The addition of heparin to the reaction mix disturbs both enzyme variants similar, since heparin binds the deaminase domain and the dsRBMs. In contrast, spermidine addition hampers only the editing efficiency of short ADAR2, indicating that the exit of the mRNA in the *trans* editing situation hampers the recognition and binding of the dsRBM2.

4.1.10 Lowering the amounts of the editing components

All previously presented experiments were performed on a 25 µl scale with high nanomolar amounts of enzyme, guideRNA and reporter mRNA. This scale was ideal to analyze different conditions, ratios and peculiarities of the editing system. But to get an idea whether the R/G editing system is able to be used for editing in living cells, the system should be tested at low nanomolar concentrations, since living cells or organism produce proteins and RNAs in much lower concentrations, than they were used in the 25 µl scale. In a rough estimation, 25 nM mRNA as used so far, relates to about 25 000 copies per cell. Typical copy numbers of mRNA are more in the range of 50-500 copies per cell. This suggests to carry out experiments at mRNA concentrations in the range of 50 to 500 pM. To perform experiments at such low concentrations, we had to scale up the reaction volume from 25 µl to 100 µl in order to get enough mRNA for reverse transcription.

Different concentrations of enzyme and R/G-gRNA were tested to edit a constant amount of 0.1 nM eCFP mRNA at 3 mM magnesia at a 100 µl scale. A concentration of 187 nM ADAR2 (ratio 1870:1 enzyme:mRNA) and 5 nM R/G-gRNA (ratio 50:1 gRNA:mRNA) achieved 91% conversion of adenosine to inosine in 0.1 nM stop66 eCFP mRNA (Figure 4-52). A reduction of the ratio R/G-gRNA to mRNA from 50:1 to 5:1, as used before for the 25 µl scale, was half as effective in editing the target adenosine. Lowering the ADAR2 concentration from 187 nM to 30 nM at a R/G-gRNA concentration of 5 nM resulted in 39% editing. This is in accordance with our previous finding that ADAR2 cannot be diluted below 100 nM in a PCR tube without significant loss in enzyme activity. Therefore, a protein concentration of 187 nM and a R/G-gRNA of 5 nM were chosen for further experiments at the 100 µl scale.

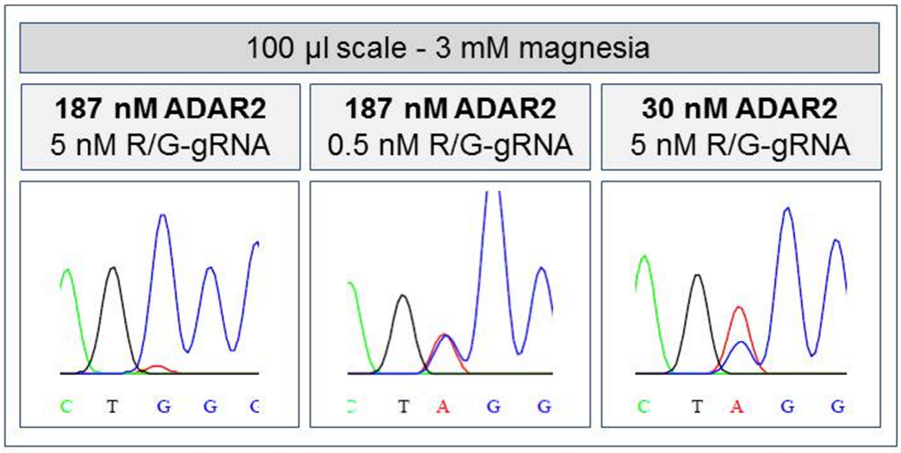

Figure 4-52: Variations of enzyme and R/G-gRNA concentrations at 100 µl scale. The editing with drastically reduced mRNA concentration (0.1 nM stop66 eCFP mRNA) was tested in presence of 3 mM magnesia and various ratios of ADAR2 and R/G-gRNA. The full conversion of adenosine to inosine was achieved with 187 nM ADAR2 and 5 nM R/G-gRNA. A reduction of either the protein concentration or R/G-gRNA concentration decreased the editing yield drastically.

For the evaluation of the ideal magnesia concentration at nanomolar concentrations of the components of the editing system, three different magnesia concentrations were tested: no Mg^{2+}, 1.5 mM Mg^{2+} and 3 mM Mg^{2+}. No addition of magnesia led to the lowest editing yield with 73% (Figure 4-53). The standard magnesia concentration of the 25 µl scale (3 mM Mg^{2+}) resulted for the 100 µl in 91% editing, whereas a lower magnesia concentration of 1.5 mM achieved full conversion of adenosine to inosine.

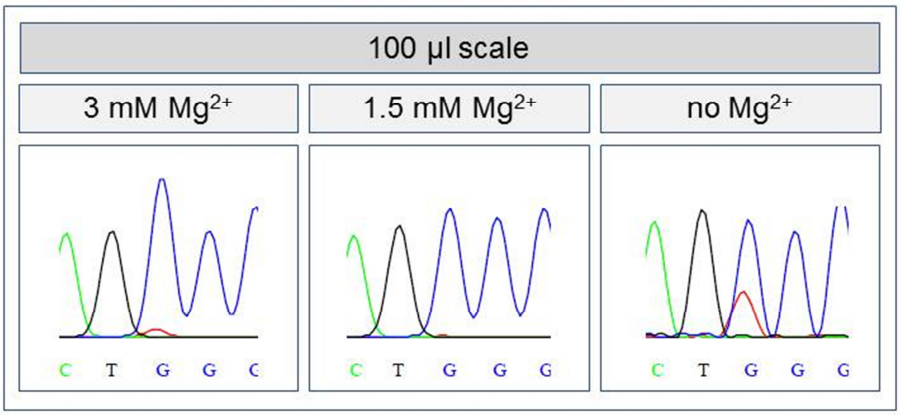

Figure 4-53: Variation of the magnesia concentration at 100 µl scale. Three different magnesia concentrations were tested for their influence on the editing yield at a100 µl scale: no magnesia, 1.5 mM Mg^{2+} and 3 mM Mg^{2+}. No addition of magnesia resulted in the lowest editing yield with 73%. The standard editing mix concentration of 3 mM Mg^{2+} resulted in 91% editing yield and with a magnesia concentration of 1.5 mM a full conversion from adenosine to inosine was achieved.

Since the reverse transcription resulted many times in a low cDNA amount, a higher mRNA concentration of 0.5 nM was chosen to improve the downstream analysis. The increase of the mRNA concentration from 0.1 nM to 0.5 nM resulted in 94% editing yield instead of full conversion of adenosine to inosine, assuming that the ratios in the editing system were now optimal.

Next we compared the lack of spermidine versus the addition of 0.5 mM and 1 mM spermidine in order to analyze the possibility to prevent off-target editing in the sub-nanomolar scale. Interestingly, only the off-target position 127 was affected by over editing of the editing machinery at nanomolar concentrations, whereas off-target editing at position 99 was not present.

The positive influence of spermidine towards blocking off-target editing is less intense, since the concentrations of 0.5 mM and 1 mM spermidine were pretty low compared to 2 mM spermidine of the 25 µl scale. If only 1.5 mM Mg^{2+} was added to the reaction mix this resulted in 67% over editing at position 127 (Figure 4-54). Usage of 0.5 mM spermidine slightly reduces the off-target editing down to 54%, whereas the target adenosine was less affected, as well. But the increase of spermidine up to 1 mM decreases not only the editing of the off-target adenosine down to 36%, but also reduces the editing of the target adenosine at position 66 to 54%.

Lower concentrations of the editing components are much more sensitive towards higher spermidine amounts than higher concentrations: At a 25 µl scale the addition

of 2 mM spermidine at 1.5 mM Mg^{2+} only slightly reduced the editing yield at the target adenosine (7%) (Figure 4-14), whereas already the addition of 1 mM spermidine to the editing mix halved the editing yield at 1.5 mM Mg^{2+} at low nanomolar concentrations (Figure 4-54).

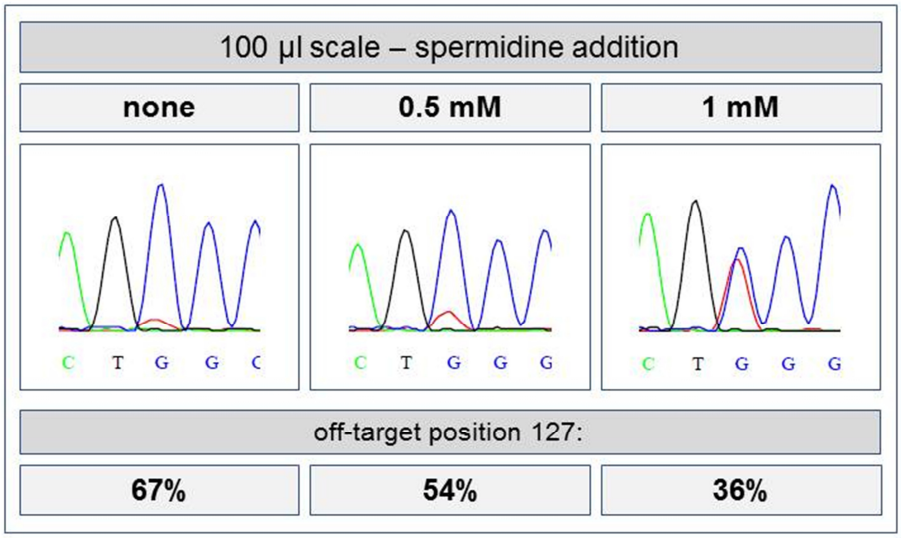

Figure 4-54: Addition of different spermidine concentrations at a 100 µl scale. Three different spermidine concentrations were tested at 3 mM Mg^{2+} in a 100 µl scale editing reaction mix to edit 0.5 nM stop66 eCFP mRNA with 187 nM ADAR2 and 5 nM R/G-gRNA: no addition, 0.5 mM and 1 mM. The addition of 0.5 mM spermidine slightly reduced the editing yield at the target adenosine position compared to non-addition, as well as for the off-target position 127. The increase of the spermidine concentration up to 1 mM drastically reduced the editing level of the target position down to 54% and the off-target editing was reduced down to 36%.

In order to confirm the conclusions been made from the 25 µl scale that the R/G-motif is necessary for the editing system and the editing system is specific and exclusively working together, some of the experiments of the 25 µl were repeated at sub-nanomolar concentrations.

At first, the comparison of the editing yields obtained from the use of a ss-gRNA versus a R/G-gRNA with and without spermidine addition at 1.5 mM Mg^{2+} were performed. The editing yields are in both cases – with 0.5 mM spermidine and without – much higher for the use of the R/G-gRNA than for the ss-gRNA (Figure 4-55). The use of the ss-gRNA achieves 65% editing yield without spermidine addition. The spermidine addition drastically reduced the editing level of the ss-gRNA, whereas the editing level of the R/G-gRNA was only slightly reduced. This confirms again the contribution of the R/G-motif for harnessing ADAR2.

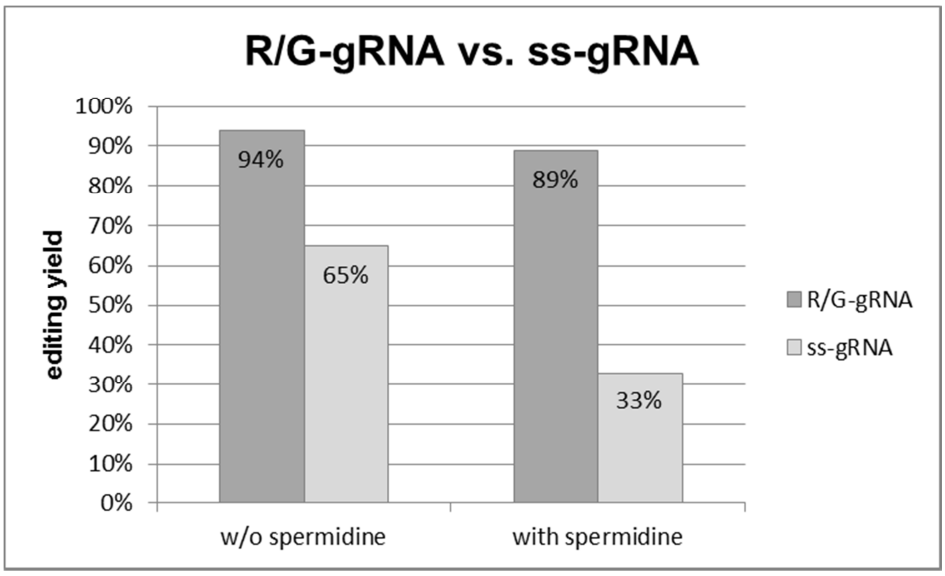

Figure 4-55: Comparison of R/G-gRNA versus ss-gRNA at nanomolar concentrations. The editing yields with 0.5 mM spermidine and without were highest for the use of the R/G-gRNA.

Also the exclusiveness of the editing system could be verified again: The use of 187 nM of SNAP-ADAR2 achieved in combination with 5 nM R/G-gRNA at 1.5 mM Mg^{2+} only 21% editing yield (Figure 4-56). The addition of 0.5 mM spermidine to the reaction mix completely inhibited SNAP-ADAR2 to edit the target adenosine at position 66 eCFP mRNA.

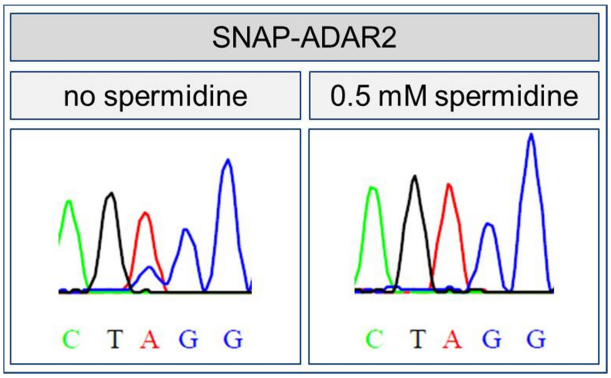

Figure 4-56: Editing efficiency of SNAP-ADAR2 in the presence and absence of spermidine at the nanomolar scale. SNAP-ADAR2 (187 nM) achieved with 5 nM R/G-gRNA 21% editing yield of position 66 eCFP mRNA at 1.5 mM Mg^{2+}. The addition of 0.5 mM spermidine to the editing mix resulted in a complete loss of editing.

After demonstrating that editing with the engineered guideRNA working in *trans* is even possible at a low nanomolar scale, we wanted to verify that a reduction of the editing components has also no influence on the editing ability of short ADAR2 and ADAR2 in comparison to the high nanomolar 25 µl scale.

Therefore, short ADAR2 and wt ADAR2 were tested for their deamination efficiency of 0.1 nM mRNA+R/G construct at 3 mM Mg^{2+}, in order to compare it again to the editing in *trans* level. Both, short ADAR2 and wt ADAR2 edited the target adenosine at full conversion in the *cis* editing situation (data not shown). In contrast, in the *trans* situation, short ADAR2 was able to edit the targeted adenosine to 75% at 1.5 mM Mg^{2+} (Figure 4-57). The addition of 0.5 mM spermidine decreased the editing yield of short ADAR2 down to 52%. This decrease was expected from the results of the 25 µl scale.

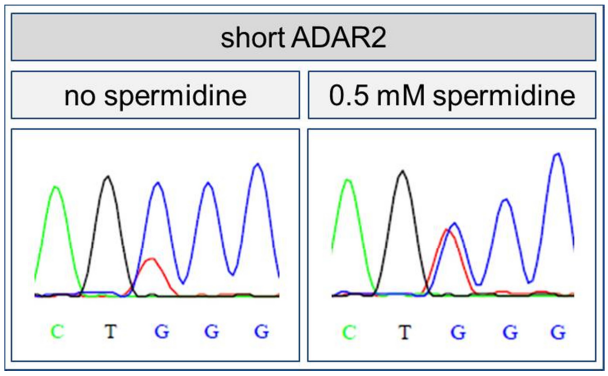

Figure 4-57: Editing efficiency of short ADAR2 in the presence and absence of spermidine at the sub-nanomolar scale in *trans*. Short ADAR2 (187 nM) achieved with 5 nM R/G-gRNA 75% editing yield of position 66 eCFP mRNA at 1.5 mM Mg^{2+}. The addition of 0.5 mM spermidine to the editing mix resulted in a reduction down to 52% editing yield.

These experiments for testing the influence of the concentrations of the editing system towards the exit of the mRNA demonstrate, that the exit of the mRNA in the *trans* situation affects the recognition and binding of the dsRBM2, confirming the results obtained for the 25 µl with higher nanomolar concentrations.

The reduction of the components of the editing system down to a sub-nanomolar concentration is not affecting the peculiarities and characteristics of the editing system. The editing system is still efficiently, specifically and exclusively working. The exit of the mRNA is shown to have no decreasing effect at the deamination efficiency of ADAR2 with the editing system. Furthermore, the editing system showed to edit the target adenosines at high inosine yields at low magnesia concentrations. An efficiency of the editing system at low magnesia concentration is important for an in vivo application, since a cellular magnesia concentration of around 0.75 mM is expected (97).

The in vitro established editing system is expected to work successfully in cell culture, since the system showed to be specific, efficient and robust towards stringent conditions and to be efficient on a broad range of component concentrations & stoichiometries.

4.2 Editing in cell culture

The ability of the genetically encodable editing system to specifically edit a certain adenosine was proven on an in vitro basis before and should now be transferred to cell culture, as well. The in vitro results for editing at low nanomolar concentrations were pretty promising in steering ADAR2 for site-specific RNA editing in cell culture. In order to set up the cell culture experiments, different plasmids had to be provided to analyze the ability of the editing system in HEK293T cells, coding for the reporter genes, the ADAR2 protein and the R/G-gRNA.
The transfection experiments of the following chapter represent only a section of all transfection experiments that were performed during the dissertation project. Each chapter will present a selection of these pretests to reveal the optimal experimental design to achieve high editing yields in cell culture.

4.2.1 Plasmid construction

The plasmid pcDNA3.1 was used for reporter gene and ADAR2 transcription in HEK293T cells. All genes were put under CMV promoter to ensure a strong mRNA expression. To preclude the R/G-gRNA from being poly(A)-tailed or capped at its 3 prime or 5 prime ends respectively, it was transcribed by the RNA polymerase III, using a U6 promoter and poly-U signal for termination. The commercial p*Silencer*™ 2.1-U6 Hygro vector was used for these matters. The cloning strategies and an overview of the design of the pcDNA3.1 and *pSilencer* vectors are described in chapter 3.2.14.

4.2.2 Editing of W58X eGFP in HEK293T cells

In order to prove the editing system of ADAR2 and the engineered R/G-gRNA in a living cell, the eGFP reporter gene was chosen as a target enabling an easy read out by fluorescence. The 5'-UAG point mutation at position 58 was chosen and the same R/G-gRNA as already used for in vitro experiments. This 5'-UAG mutation allows a

straight microscopy analysis of positive editing events, since the deamination of the stop codon results in a functional tryptophan restoring eGFP fluorescence. The usage of eCFP as a reporter gene is possible, but leads in the positive controls to a visibly weaker fluorescence signal than eGFP and was, therefore, not further investigated.

In order to evaluate the editing yield for each individual experiment, microscopy analysis was complemented by RNA extraction and sequencing (3.5.4). For this DNaseI digestion, cDNA preparation by reverse transcription (3.4.2), amplification by Taq-PCR and the subsequent sequence analysis of the amplified PCR product was performed. The primer no. 144 was used as bw primer for reverse transcription, primer no. 144 and no. 121 for Taq-PCR and primer no. 121 for sequence analysis. The experimental design is shown in Figure 4-58.

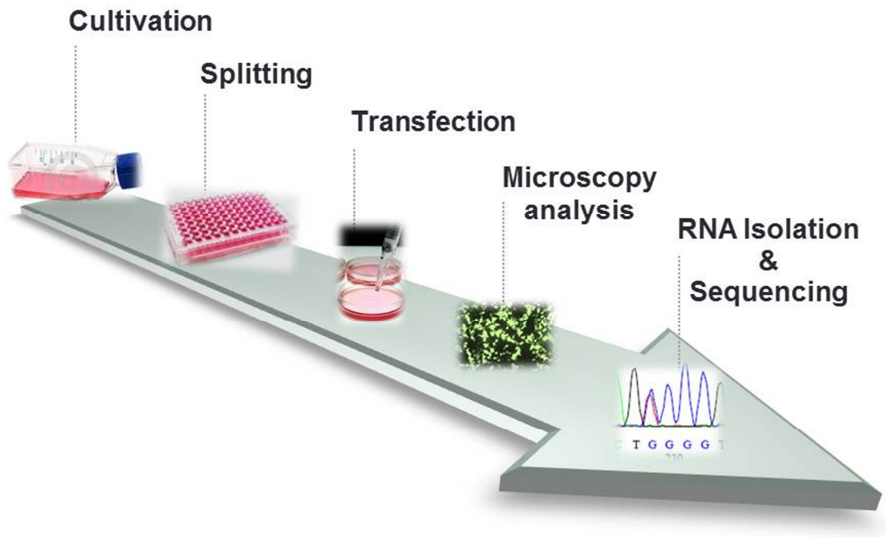

Figure 4-58: Experimental setup for the analysis of editing in cell culture. HEK293T cells were seeded in the 24-well or 96-well format and transfected on the following day. The success of the editing reaction was verified by microscopy analysis and RNA isolation depending on the chosen incubation time of the editing reaction.

The cell culture experiments were started with an optimization of the transfection protocol with Lipofectamin™ 2000 as the chosen transfection reagent. Different ratios of Lipofectamin to DNA were transfected, using an eGFP expressing pcDNA3.1 plasmid (pTS58) and the ratio of 1 µl Lipofectamin for 3 µg of DNA plasmid was found to be ideal according to microscopy analysis. The amount of eGFP plasmid was adjusted for the following editing experiments to a level, which assured a good fluorescence signal but the lowest plasmid amount, in order to keep the amount of

transfection reagent possibly low. The amount that fulfilled this task was 300 ng of eGFP plasmid.

Corresponding to the amount of the eGFP plasmid or W58X eGFP plasmid (pTS59) the amount of ADAR2 plasmid (pTS57) had to be set for the first analysis. The same amount of 300 ng as for the W58X eGFP plasmid was chosen for the transfection of the ADAR2 plasmid. The design of the R/G-gRNA was the same as already used for the in vitro editing studies: 16 nt long mRNA binding site (green) and a 5'-CCA anticodon (Figure 4-59).

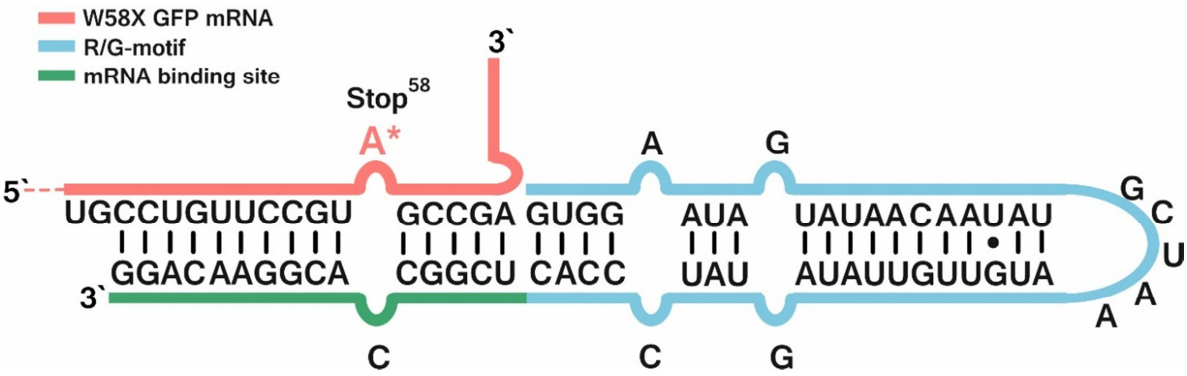

Figure 4-59: Design of the engineered R/G-gRNA for editing W58X eGFP mRNA. The W58X eGFP mRNA (red) is bound by a 16 nt long mRNA binding site (green) of the R/G-gRNA (green+blue). The target adenosine is highlighted with an asterisk.

The ability of the genetically encodable editing system to deaminate the adenosine at position 58 of a W58X eGFP mRNA in cell culture was analyzed with the following transfection experiment: 300 ng of W58X eGFP plasmid (pTS59), 300 ng ADAR2 plasmid (pTS57) and variable amounts of R/G-gRNA plasmid (pTS68) were transfected together in a 24-well format and the result of the editing reaction was analyzed 24 hours post transfection.

Each tested amount of the R/G-gRNA plasmid was able in combination with the ADAR2 plasmid to restore the fluorescent signal of W58X eGFP (Figure 4-60). No differences in the cell number and brightness of the fluorescent signal were distinguishable between the samples, except for the sample with 2000 ng of transfected R/G-gRNA plasmid. The 2000 ng R/G-gRNA plasmid sample showed slightly more and stronger fluorescence signals. But the sequencing results of the isolated eGFP mRNA demonstrated a slightly different ranking of the highest obtained editing yields: starting with an editing yield of 19% for transfecting 300 ng of

R/G-gRNA plasmid, the further increase of the R/G-gRNA plasmid amount leads to an increment of the editing yield until 30% for the transfection of 1600 ng R/G-gRNA plasmid. This increment demonstrates a distinct tendency that more transfected R/G-gRNA plasmid results in a higher editing yield. But for the sample with 2000 ng of R/G-gRNA plasmid a slightly lower editing yield of 28% was obtained.

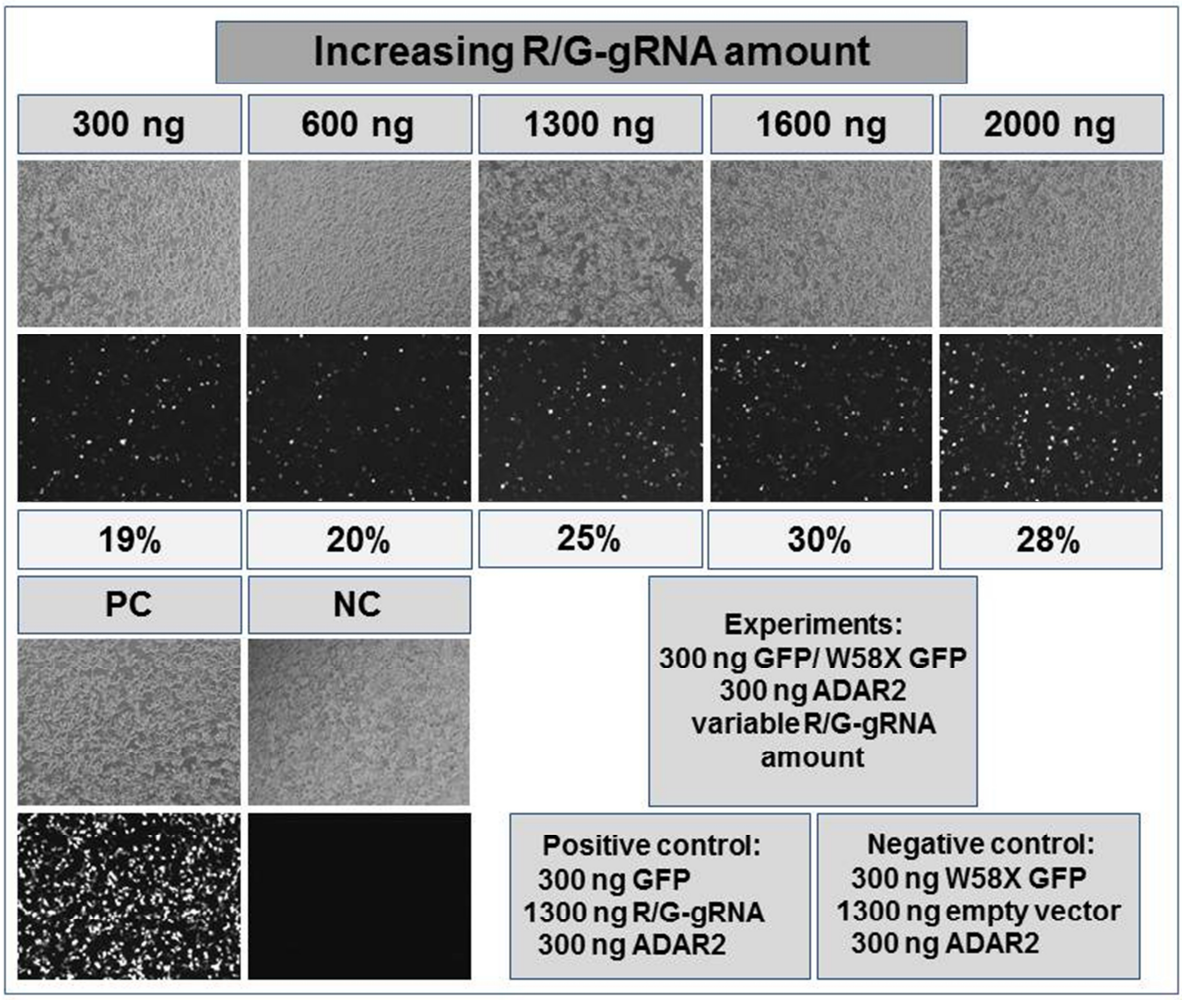

Figure 4-60: Analysis of increasing R/G-gRNA plasmid amounts affecting the editing yield in cell culture experiments. Variable amounts of R/G-gRNA plasmid were transfected together with 300 ng of W58X eGFP plasmid and 300 ng of ADAR2 plasmid in a 24-well plate format. Microscopy analysis and RNA isolation followed by sequence analysis were performed 24 hours post transfection. The microscopy analysis showed hardly differences in the amount and brightness of the fluorescent signal, except for the sample that was transfected with 2000 ng of R/G-gRNA plasmid. But the sequence analysis demonstrates that the 1600 ng of R/G-gRNA plasmid sample achieved with 30% the highest editing yields, followed by the 2000 ng of R/G-gRNA plasmid sample. Magnification: 100x, GFP exposure time: 50 ms.

This experiment showed that the deamination of a specific adenosine with the encodable editing system is possible and it was the starting point for further optimizations with the aim to increase the total editing yield.

4.2.2.1 Editing takes time

The analysis of positive editing events, as indicated by fluorescence, was initially performed 24 hours after transfection. These 24 hours of incubation were adopted from our previous findings with the SNAP-tag BG-system in cell culture (59). In contrast to the SNAP-tag BG-system, the guideRNA has to be produced and delivered by the cell, what raise up the idea of prolonging the editing time in order to increase the total editing yield for the isolated W58X eGFP mRNA.

The microscopy analysis and RNA isolation were additionally performed 48 hours after transfection. The following amounts of plasmids were chosen to analyze the effect of a longer incubation time for the editing reaction: 300 ng of W58X eGFP, 300 ng of ADAR2 and 1300 ng or 1600 ng of R/G-gRNA. Such high amounts of R/G-gRNA plasmid were chosen based on their positive editing results of the editing experiment with varying amounts of R/G-gRNA plasmids and an incubation time of 24 hours.

The prolongation of the incubation time from 24 hours to 48 hours increased the fluorescent signal and the number of eGFP expressing cells in the positive controls (Figure 4-61). Also the editing samples (1300 ng or 1600 ng of R/G-gRNA) showed at 48 hours of incubation an increased fluorescent signal, compared to the editing samples at 24 hours of incubation. The number of eGFP expressing cells rose as well, indicating that more cells edited W58X eGFP mRNA. The assumption that more editing took place, based on the microscopy analysis, is validated by the sequence results: Indeed, the total editing yield of the isolated total W58X eGFP mRNA was increased from 25% (24 hours) to 41% (48 hours) for the editing sample with 1300 ng of R/G-gRNA plasmid and from 30% (24 hours) up to 44% (48 hours) for the sample including 1600 ng of R/G-gRNA plasmid. The tendency that a higher R/G-gRNA plasmid amount leads to higher total editing yields was thereby demonstrated another time.

Also the question arises from the in vitro editing experiments, whether off-target editing is present in cell culture. The off-target editing of W58X eGFP mRNA could

completely excluded for experiments with an incubation time of 24 hours (Figure 4-62). After 48 hours of incubation the already from in vitro experiments well known off-target position 127 was edited up to 5-7%. The sequence analysis with a second primer (no. 566), in order to cover the whole length of the eGFP mRNA excluded further off-target sites. By off-target editing of position 127 the fluorescent phenotype was not limited, since both triplets (5'-AAA and 5'-AAG) code for lysine.

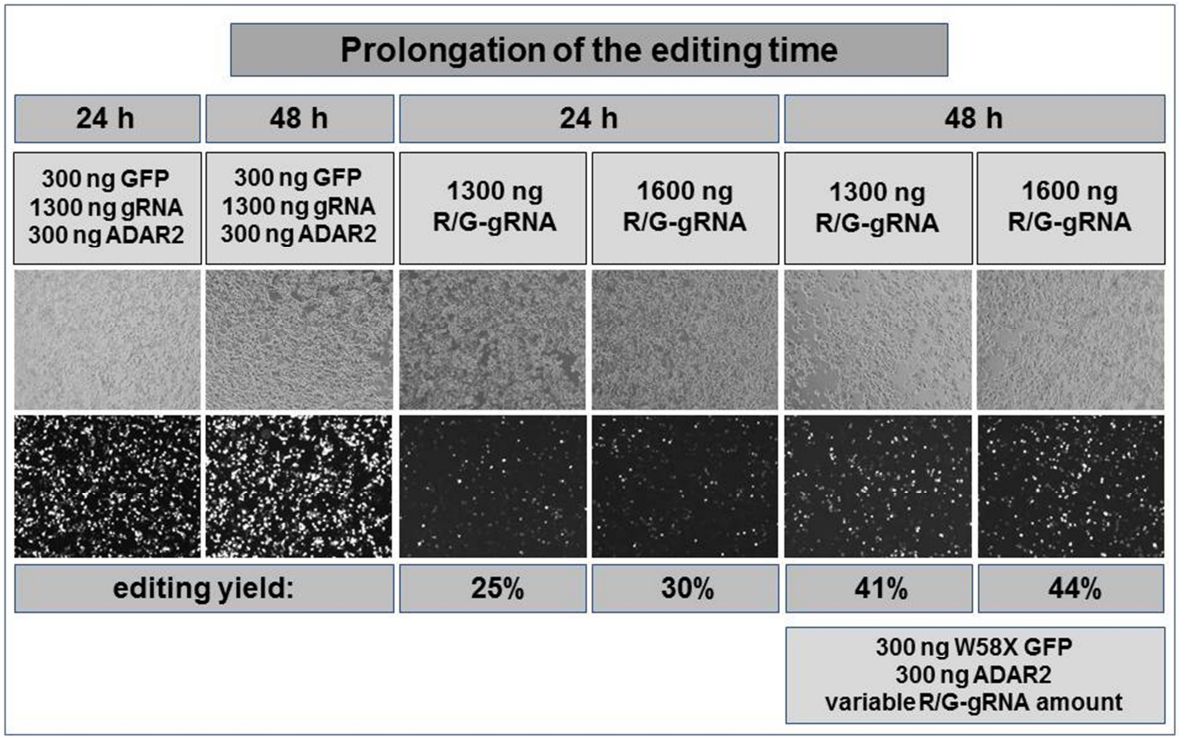

Figure 4-61: Prolongation of the editing time in cell culture from 24 hours to 48 hours. Each transfection sample included the amount of 300 ng W58X eGFP plasmid, 300 ng of ADAR2 plasmid and 1300 ng or 1600 of R/G-gRNA plasmid. The positive controls contained 300 ng eGFP plasmid, 300 ng of ADAR2 plasmid and 1300 ng of R/G-gRNA. An increment of the fluorescent signal is obtained for the positive controls for prolonging the incubation time up to 48 hours. The same increment of the fluorescent signal and the amount of eGFP containing cells is visible for the two editing samples with the double incubation time. The editing yields of 41% for 1300 ng of R/G-gRNA plasmid and 44% for 1600 ng of R/G-gRNA plasmid are increased for the longer incubation time, as well, compared to 24 hours (25% for 1300 ng R/G-gRNA and 30% for 1600 ng R/G-gRNA). Magnification: 100x, GFP exposure time: 50 ms.

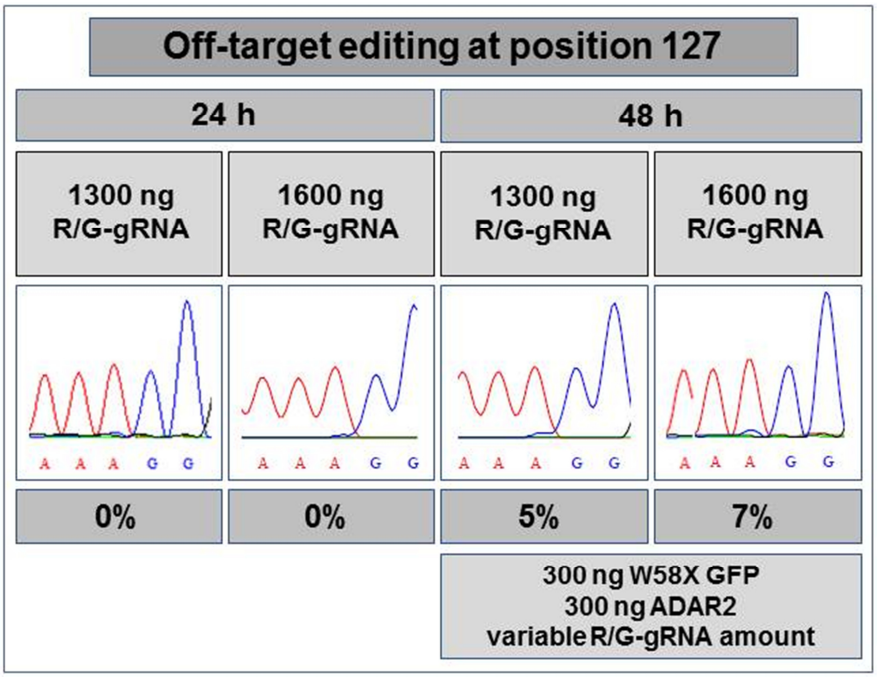

Figure 4-62: Off-target editing of position 127 at 24 hours and 48 hours of incubation.

The assumption was confirmed that longer incubation times increase the editing yields; obviously, editing with ADAR2 and the engineered R/G-gRNA takes time. The increase of the incubation time has a stronger impact on the maximum editing yield than the use of higher amounts of it: By increment of the incubation time the editing level could be increased up to 44%, whereas the maximum editing yield of various R/G-gRNA plasmid amounts was 30%.

This experiment also shows that the genetically encodable editing system and the SNAP-tag BG-system differ in the incubation time that is needed to achieve high editing yields, suggesting that the production and provision of a sufficiently high guideRNA amount may be a limiting factor. The SNAP-tag BG-system achieves 63 % editing yield after 24 hours (unpublished data from Paul Vogel), whereas the encodable editing system achieves 30% after 24 hours and 44% after 48 hours. This clearly demonstrates a difference in response for both editing systems.

A matter difference of the two editing systems is that the SNAP-tag BG-system forms a covalent and stable conjugate of enzyme and guideRNA as soon as they encounter each other, whereas the contact and binding of ADAR2 with the R/G-gRNA is reversible. From the stochastic point of view, this reversibility of the R/G-gRNA ADAR2 binding may result in a requirement of more time for the editing process, since the R/G-gRNA and ADAR2 protein have to meet each other each time at the

target mRNA (probability) and forming a tripartite binding complex (statistics). Contrary, the SNAP-tag BG-system benefits from the covalent bound, so that the conjugate has to form just once and only the process of hybridization to the target mRNA is time limiting.

The idea to increase the editing yield by a prolongation of the editing time was successful for both chosen R/G-gRNA plasmid amounts. The question if the editing yield of 44% can be further increased is subject of the following chapters. But one condition had to be fulfilled beforehand: the assumption that only the combination of R/G-gRNA and ADAR2 protein is responsible for editing W58X eGFP mRNA and nothing else.

4.2.2.2 R/G-gRNA & ADAR2 are responsible for W58X GFP mRNA editing

The question if both components of the editing machinery, R/G-gRNA and ADAR2, are required for the editing of W58X eGFP mRNA is subject of the following experiments.
The transfection experiments were performed in a 24-well format, always using 300 ng of W58X GFP plasmid for each well and individual plasmid combinations. After 48 hours the microscopy analysis and RNA isolation was carried out.

As indicated by fluorescence microscopy no restoration of the stop codon in W58X eGFP to tryptophan was visible in the absence of both, W58X eGFP R/G-gRNA and ADAR2 plasmid (Figure 4-64). The replacement of 1600 ng of W58X eGFP R/G-gRNA plasmid with 1600 ng of PINK R407Q R/G-gRNA plasmid (pTS164.1), co-transfected with 300 ng of ADAR2 plasmid resulted in no editing at all. This demonstrates that only a specific guideRNA, which is complementary to the mRNA target region, allows editing of the target adenosine by ADAR2.
We then tested whether a single-stranded-gRNA (ss-gRNA), which is complementary to the W58X eGFP region of the target adenosine as well, but lacking the R/G-motif (as shown for the in vitro editing in chapter 4.1.6 Figure 4-26), enables editing of the W58X eGFP mRNA with ADAR2. For this 300 ng W58X eGFP plasmid and ADAR2 plasmid were co-transfected in a 24-well format. After 24 hours the cells were detached by trypsin, splitted into a 96-well format ($4 \cdot 10^4$ cells per well) and rested for further 24 hours until transfection with 50 pmol of a chemically stabilized ss-gRNA (Figure 4-63).

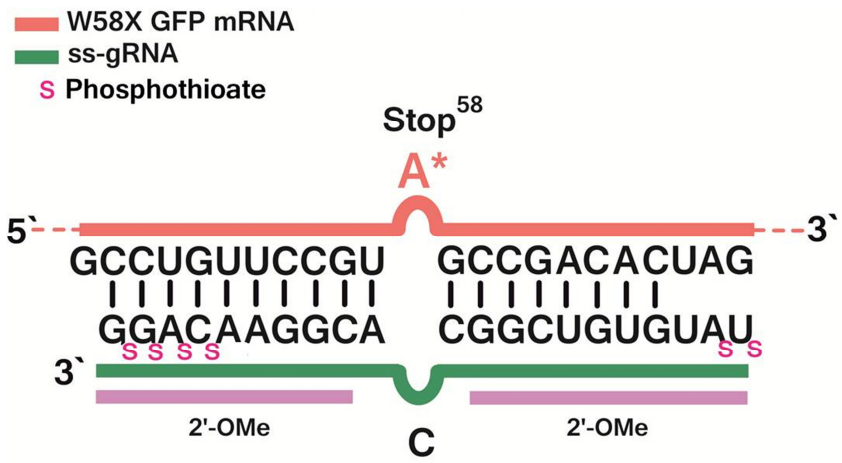

Figure 4-63: Chemically stabilized 19 nt single-stranded-gRNA binding the W58X GFP mRNA.

An amount of 50 pmol was typically used for the SNAP-tag BG-system to achieve high editing yields (data not shown). The microscopic analysis was performed 24 hours post transfection with the ss-gRNA. The amount and the intensity were markedly lower than using the R/G-gRNA to edit W58X eGFP (Figure 4-65). This shows that the R/G-gRNA and the ADAR2 protein are more specific for each other and more effective in editing the target adenosine than ADAR2 editing with a stabilized ss-gRNA, as already proven for in vitro experiments (4.1.6).

Also the replacement of ADAR2 by the SNAP-ADAR2 enzyme (pTS72) resulted in cell culture in lower and less fluorescent signals. The R/G-gRNA is able to steer SNAP-ADAR2 for editing the target adenosine of W58X eGFP mRNA, but the efficiency is clearly lower than for ADAR2. This demonstrates the importance of the recognition of the R/G-motif by the two dsRBMs of ADAR2, giving ADAR2 a distinct advantage over SNAP-ADAR2, lacking the two dsRBMs. This outcome was expected from the data obtained for the in vitro editing experiments with R/G-gRNA and SNAP-ADAR2, where ADAR2 had also a clear advantage over SNAP-ADAR2 at stringent editing conditions (4.1.3). The analysis whether ADAR1 (pTS284) is steerable by the R/G-gRNA to edit W58X eGFP mRNA, too, was not performed in vitro but now in cell culture. The microscopy analysis demonstrates that ADAR1 is steerable by the R/G-gRNA, but in a less effective way than ADAR2 (Figure 4-65). The exchange of ADAR2 by ADAR1 was anticipated to result in lower editing levels for the following reasons. One reason is that the glutamate receptor transcript, from which the R/G-motif of our guideRNA was derived, is naturally mainly edited by ADAR2 (8). Another indication gave the binding analysis of a phenotypic screen with a substrate derived

from GluR-B receptor, showing that this substrate is only effective for studying the binding properties of ADAR2 and not of ADAR1 (54). Both ADAR enzymes vary in their substrate specificities and preferences (2.5). Therefore, the R/G-motif was expected to be less effective in steering ADAR1`s activity to edit the target adenosine. It was rather expected that the editing result of W58X eGFP mRNA with R/G-gRNA and ADAR1 is similar to the one of a ss-gRNA and ADAR2; the R/G-motif being not optimal to be recognized by ADAR1 and the ss-gRNA lacking the recognition motif for ADAR2.

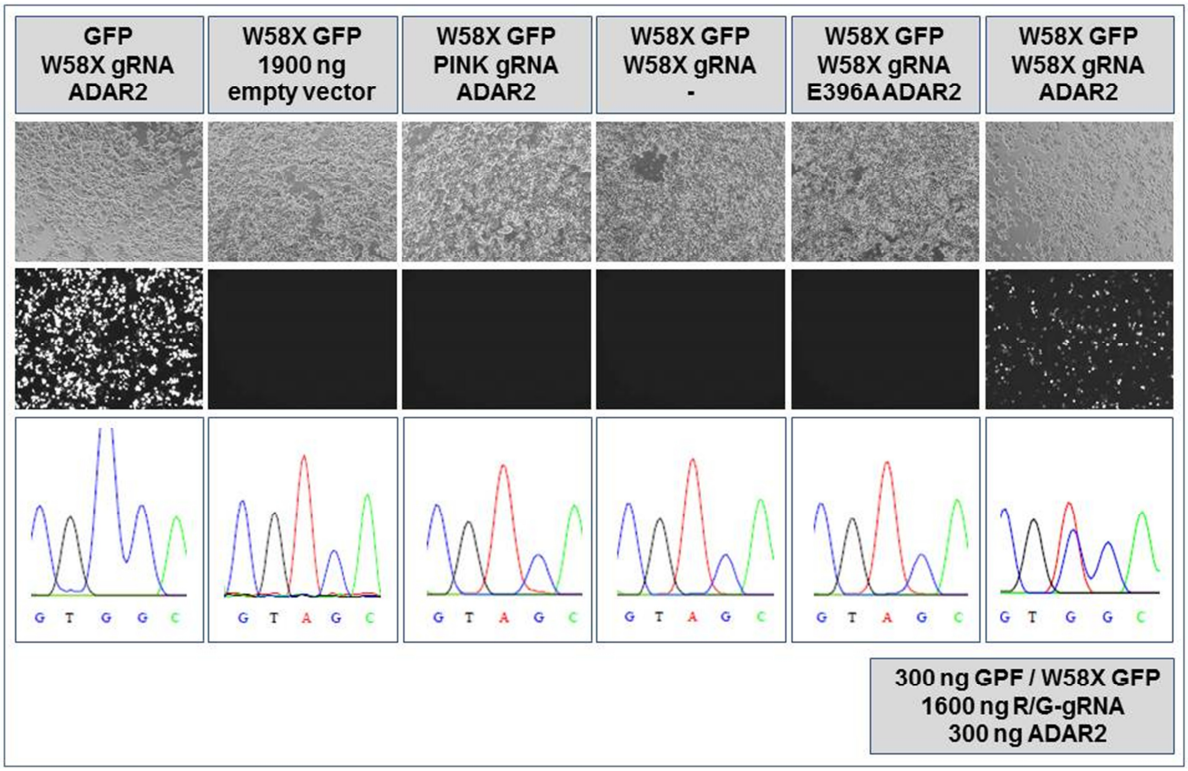

Figure 4-64: Evidence that only the combination of R/G-gRNA and ADAR2 mediates editing. Each 24-well was transfected with 300 ng of W58X eGFP plasmid, always lacking one or two components of the editing system (R/G-gRNA or ADAR2), in order to prove its necessity for each other to edit the W58X eGFP mRNA. The sample that co-transfected 1900 ng of empty vector shows no GFP signal. Also the co-transfection of PINK R407Q R/G-gRNA and ADAR2 does not lead to a restored fluorescent signal. If ADAR2 is replaced by a nonfunctional mutant called E396A or is completely left out for the transfection, this results in hardly any fluorescence signal. The sequencing traces of these samples confirm that no editing of W58X eGFP mRNA took place. In comparison the co-transfection of R/G-gRNA and ADAR2 plasmid enables editing of W58X eGFP mRNA, demonstrated by many bright fluorescing cells and high editing yields. Magnification: 100x, GFP exposure time: 50 ms.

If ADAR2 is replaced by an enzymatically nonfunctional mutant, called E396A (42) or missing in the co-transfection together with W58X eGFP and the corresponding W58X eGFP R/G-gRNA, only a few less fluorescing cells are visible (Figure 4-64, not visible on the picture print). This shows on the one hand, that truly ADAR2 is enabling the editing of W58X eGFP and is necessary for the editing reaction in cell culture and on the other hand, that the endogenous editing is steerable by the engineered R/G-gRNA, as well. This effect that the endogenous ADAR2 is addressable to edit W58X eGFP mRNA will be discussed more in detail in chapter 4.2.2.8.

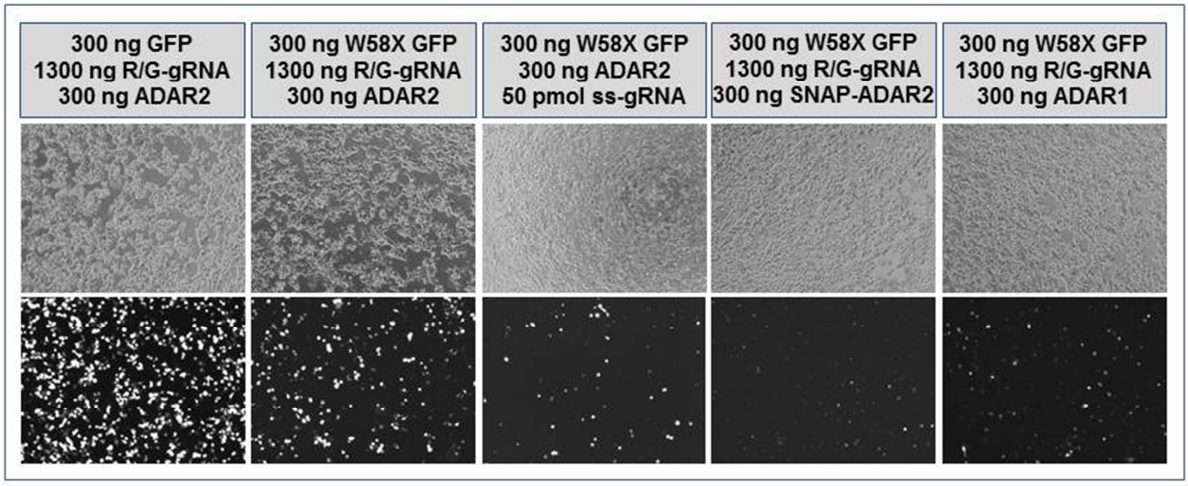

Figure 4-65: Comparison of the editing efficiency using SNAP-ADAR2, ADAR1 and a ss-gRNA. In order to replace the R/G-gRNA plasmid by a ss-gRNA, a co-transfection with 300 ng of W58X eGFP and ADAR2 plasmid was performed in 24-well format. After 24 hours the cells were detached and reseeded in a 96-well format and rested for 24 hours prior transfection with 50 pmol ss-gRNA. The combination of ss-gRNA and ADAR2 is able to restore the fluorescent signal but in a less effective way than the combination of R/G-gRNA and ADAR2. The replacement of ADAR2 by SNAP-ADAR2 or ADAR1 in a co-transfection experiment together with 300 ng W58X GFP and 1300 ng R/G-gRNA plasmid drastically reduced the amount and brightness of eGFP expressing cells, demonstrating that this replacement is less effective in editing W58X eGFP mRNA. Magnification: 100x, GFP exposure time: 50 ms.

In order to complete the tested variants of ADAR2 protein, an experiment with the E488Q mutant of ADAR2, which was introduced in the in vitro editing part (4.1.8), was performed to verify if this enzyme has the ability to increase the editing yield in cell culture, as well. The same experimental setup was used: transfection of 300 ng W58X eGFP plasmid, 1300 ng R/G-gRNA plasmid (pTS186) and 300 ng E488Q plasmid (pTS90) in a 24-well-format. The microscopic analysis and RNA isolation followed by sequence analysis were performed 24 and 48 hours post transfection.

Contrary to our expectation, the amount and brightness of fluorescing cells was distinctly lower than of the samples transfected wt ADAR2 (Figure 4-66). This result was confirmed by the sequence results and demonstrates that at both chosen incubation times the E488Q enzyme was clearly less effective in editing W58X eGFP mRNA. This outcome was unexpected since in vitro it was clearly shown that E488Q has not only an advantage in editing of less preferred codons, but also in general being faster in editing the target position compared to wt ADAR2. This may be caused by hyper-editing of the R/G-motif itself, destroying the originary form of the motif and thus diminishing the ability to steer E488Q. This phenomenon was already observed by in vitro editing of the mRNA+R/G transcript by ADAR2 and short ADAR2 (4.1.9).

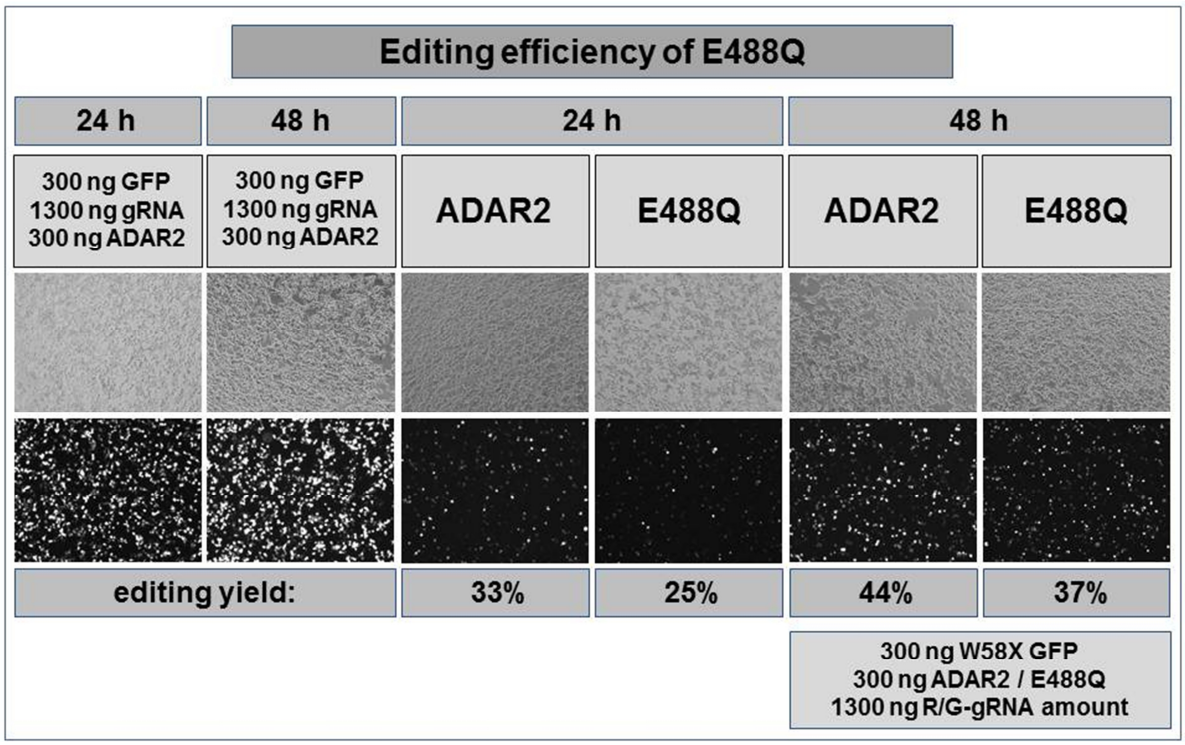

Figure 4-66: Editing efficiency of E488Q in cell culture compared to wt ADAR2. A transfection experiment was performed in a 24-well format, using 300 ng of W58X eGFP plasmid, 1300 ng of R/G-gRNA and 300 ng of E488Q. The microscopic analysis and RNA isolation were performed 24 and 48 hours after transfection and compared to the results obtained for the wt ADAR2. The microscopic images of E488Q demonstrate for both incubation times less bright fluorescent signal, as obtained for the wt ADAR2. The sequence analyses reflect the same trend. The E488Q was less effective in editing the W58X eGFP mRNA in cell culture. Magnification: 100x, GFP exposure time: 50 ms.

Contrary to our in vitro findings, the off-target editing (of the eGFP mRNA) of E488Q in cell culture was restricted to position 127. Off-target editing by E488Q at position 127 was obtained already after 24 hours of incubation at 25% (data not shown), whereas usage of ADAR2 shows not yet off-target editing. After 48 hours, the off-target editing of E488Q at postion 127 was increased up to 37%. This trend was also shown in vitro, that a longer incubation time yielded always in higher off-target editing levels (4.1.8).

It was proven that the W58X eGFP R/G-gRNA is successful in steering ADAR2's activity for editing W58X eGFP mRNA and that only a guideRNA that is complementary to the target site of W58X eGFP mediates editing together with ADAR2. As already shown for in vitro experiments, a ss-gRNA binding the mRNA target site of W58X eGFP is also successful in steering ADAR2 to edit W58X eGFP mRNA and to restore the 5'-UAG to a functional tryptophan, but in a less effective manner. Again the guideRNA benefits from having the R/G-motif in order to recruit ADAR2 and to steer it to the target adenosine. The R/G-gRNA also specifically steers ADAR2`s activity towards the target adenosine since other ADAR variants like SNAP ADAR1/2, but also ADAR1, are less addressable towards editing W58X eGFP mRNA. These experiments proved that the combination of both components - the R/G-gRNA and the ADAR2 protein – are specific for another and necessary to specifically edit a favored adenosine at high editing yields.

4.2.2.3 Which length of the mRNA binding site is optimal?

The necessity for the cognate R/G-gRNA and the presence of ADAR2 for editing W58X eGFP mRNA was proven. Now the aim was pursued to increase the editing yields by optimizing different details of the R/G-gRNA.

One interesting point that has not been investigated in the in vitro editing part in detail, is now subject in the cell culture experiments: the optimal length of the flexible part of the R/G-gRNA that hybridizes with the target mRNA. Starting point was a 16 nt long R/G-gRNA, which was now prolonged up to 18, 20, 25, and 29 nucleotides (Figure 4-67). Shorter constructs were not tested, but it was assumed, that according to our findings from 2012 (57) the length of the flexible part of the guideRNA can be shortened down to 11 nt, as well, as it was proven for the SNAP-tag BG-system. It was thought, that a longer guideRNA might lead to higher editing yields. On the other

hand, this prolongation could lead to off-target editing in the bound region of the mRNA for further adenosine bases that are present in the RNA duplex and has to be considered, as well. Also the longer RNA duplex could elicit unwanted antisense effects.

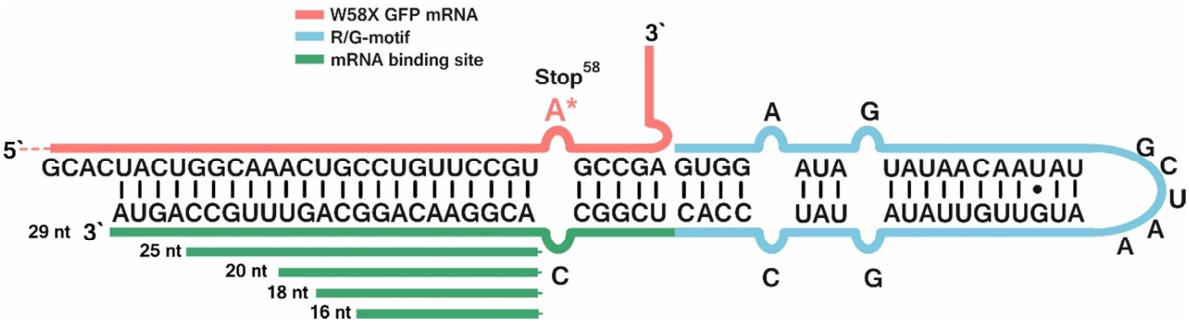

Figure 4-67: Variation of the lengths of the flexible part of the R/G-gRNA for mRNA hybridization. The starting point was a 16 nt long flexible part of the R/G-gRNA that hybridizes to the target region of the W58X eGFP mRNA (green). The length of the mRNA binding site was increased as follows: 18, 20, 25 and 29 nt.

All R/G-gRNAs with a length ranging from 18 to 29 nt contained a boxB-motif at the 3'-end of the flexible part. The background and details of attaching a boxB-motif to the 3'-end will be explained in chapter 4.2.2.6. The co-transfection experiment was performed with 300 ng W58X eGFP plasmid, 300 ng ADAR2 plasmid and 1300 ng of each tested R/G-gRNA in separate 24-well samples. The following R/G-gRNA plasmids were analyzed: 16 nt (pTS68), 18 nt (pTS87), 20 nt (pTS88), 25 nt (pTS93.1) and 29 nt (pTS94.1). The microscopy analysis and RNA isolation followed 48 hours after transfection. The sample with 16 nt and 18 nt long R/G-gRNA showed more and brighter fluorescence than the other R/G-gRNAs with longer mRNA binding site (Figure 4-68). Indeed, the editing yields of 41% for the 16 nt and 18 nt long R/G-gRNA were the highest yields of all tested guideRNA lengths in this experiment. The editing yields and the amount of fluorescent cells decreased slowly with the prolongation of the flexible part of the R/G-gRNA. A blockade of the translational process by longer guideRNAs does not seem likely. But a transfection experiment using eGFP plasmid together with longer guideRNAs would demonstrate if blocking of the translational process occurs, indicated by diverse fluorescence microscopy patterns.

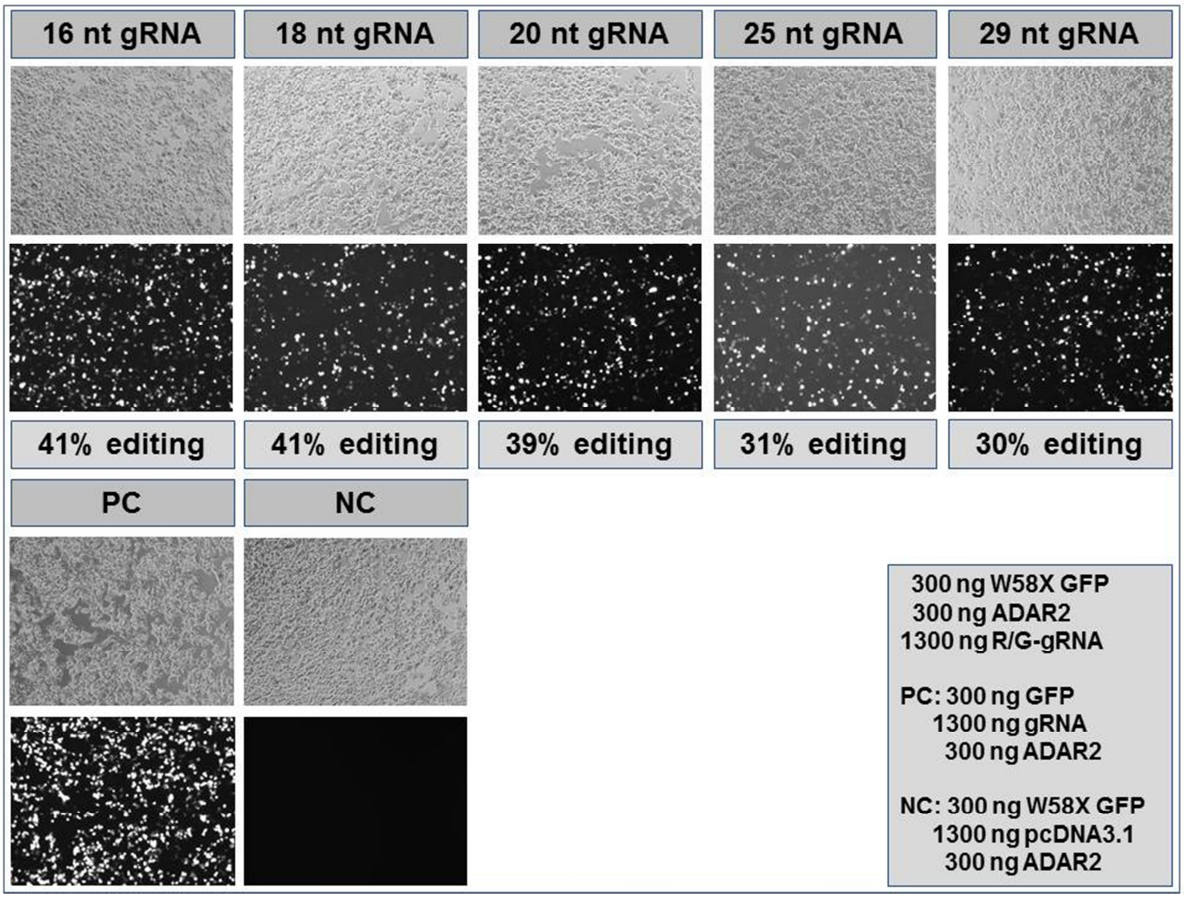

Figure 4-68: Microscopy analysis and editing yields of R/G-gRNAs with varying length of the flexible part. The strongest fluorescent signal and highest editing yield was obtained for a R/G-gRNA with 16 or 18 nt length of the mRNA binding site. With the prolongation of the mRNA binding site of the R/G-gRNA less fluorescent signal and lower editing yields were observed. Magnification: 100x, GFP exposure time: 50 ms.

Obviously, the prolongation of the guideRNAs does not result in a higher total editing yield of the target adenosine, but the possibility that longer guideRNAs might hybridize faster and stronger to the target mRNA is not assessable by these experiments. This effect could potentially be investigated by transfecting very low amounts of the respective guideRNA plasmid.

The prolongation of the flexile part of the R/G-gRNA did not affect the off-target editing of other adenosines in the whole W58X eGFP mRNA, except for specifically one adenosine in the bound mRNA region of the 25 nt and 29 nt long guideRNAs, which is edited up to 10% (Figure 4-69). The affected adenosine at position 53 and the 5'-neighbor base of this adenosine is only hybridized by a guideRNA of 25 nt and 29 nt length (Figure 4-70). These two guideRNAs put the most optimal anticodon (5'-GUU) for editing adenosine at position 53, according to the preferences of ADAR2

(49). In contrast, the 20 nt long guideRNA puts also the preferred counter base (uridine) into mismatch to the adenosine at position 53, but the adenosine is not edited. This observation is consistent with our findings from 2014 (59) that if an adenosine placed at the border of the RNA duplex this prevents the adenosine from being edited. That's why only the 25 nt and 29 nt long guideRNAs are affected from off-target editing within the RNA duplex, but not the 20 nt long guideRNA.

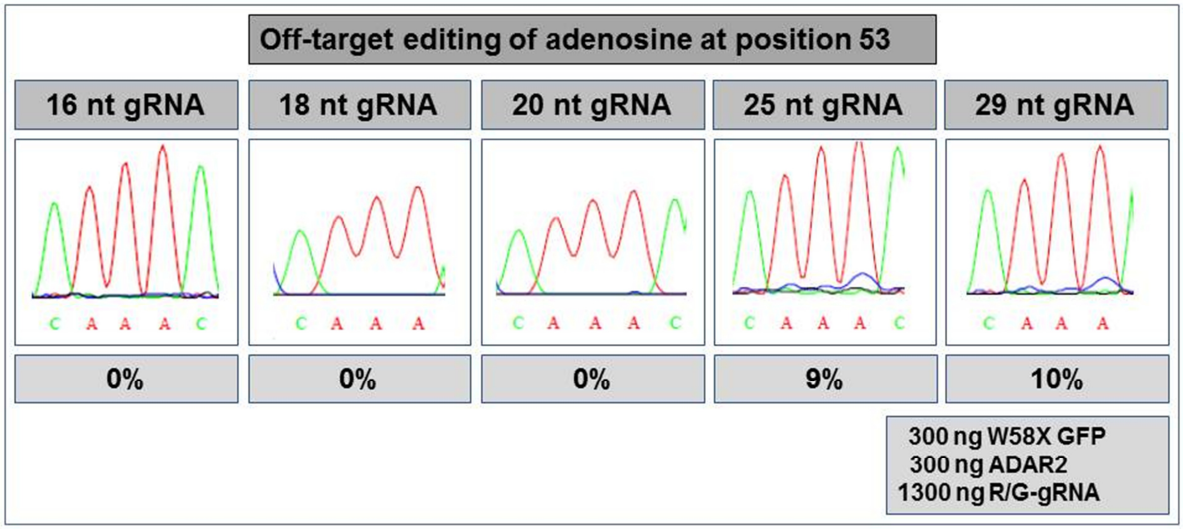

Figure 4-69: Off-target editing of adenosine at position 53. The off-target adenosine at position 53 is edited up to 10% if a 25 nt and 29 nt long R/G-gRNA is used for the editing reaction in cell culture.

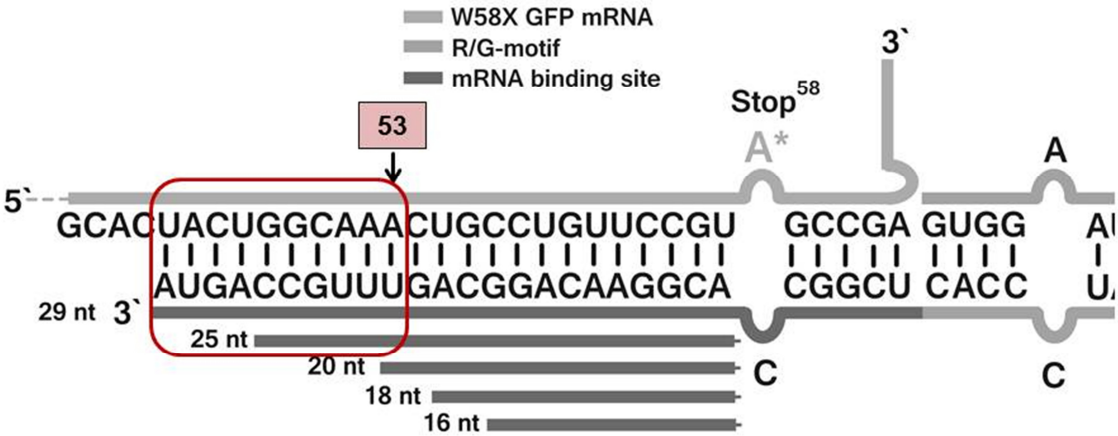

Figure 4-70: Off-target adenosine position 53 and the potentially hybridizing R/G-gRNAs. The off-target adenosine at position 53 and its 5'-neigbour are only bound by a 25 nt and 29 nt long R/G-gRNA.

The prolongation of the R/G-gRNA was not beneficial to increase the editing yield at the target adenosine of W58X eGFP mRNA. One assumption might be that a longer RNA duplex leads to higher levels of hyper-editing within the R/G-motif, resulting in a less competent guideRNA for steering ADAR2's activity. Therefore, the length of 16 nt for the flexible part seems to be optimal. This length was also optimal to protect other adenosines of the mRNA, that would have been bound by a guideRNA of 21 nt length or longer, so that off-target editing in the mRNA region could be completely prevented.

4.2.2.4 Positioning of the guideRNA with respect to the targeted adenosine

After determining the optimal length of the guideRNA to be 16 nt, the idea raise up to investigate the optimal distance of the target adenosine with respect to the 5'-end of the R/G-motif. Figure 4-71 demonstrates exemplarily the different possible positions of the adenosine with respect to the 5'-end R/G-gRNA W58X eGFP that were tested in the following cell culture experiments. The position 6, which means that the adenosine is positioned 6 nucleotides away from the 5'- end of the R/G-motif (blue), was the starting point for the in vitro, as well as for cell culture experiments. This position was chosen, since it corresponds to the target adenosine position of the natural GluR-B transcript. The NMR structure analysis from Stefl et al. (35) indicates that the position of the target adenosine might be relevant, since specific contacts between the dsRBM2 with the guanosine of the 5'-AA*G codon were observed. In contrast to the natural GluR-B transcript, our editing system with the engineered R/G-

gRNA works in *trans*, separating the target mRNA from R/G-motif, while creating an exit for the mRNA (chapter 4.1.1). In various experiments it was demonstrated that the position of the adenosine, either located more in the middle or at the border of the RNA duplex has a strong influence at the total editing yield. In the previous chapter it was demonstrated that placing an adenosine at the very border prevents this adenosine from being edited. These findings led to idea to systematically analyze the effect of various adenosine positions in respect to the 5'-end of the R/G-gRNA with the assumption that the total editing yield can be further increased.

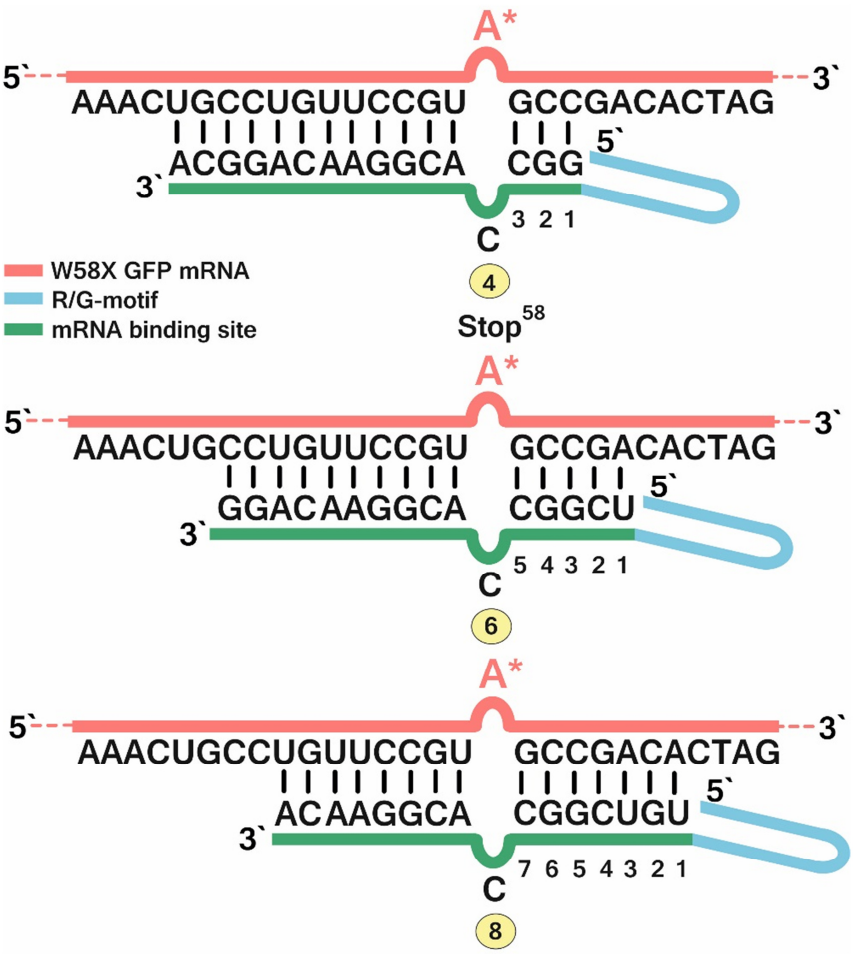

Figure 4-71: Positioning of the target adenosine with respect to the 5'-end of the R/G-gRNA. The target adenosine was varied in its position in respect to the 5'-end of the R/G-motif. Three examples are shown to illustrate the different mRNA binding regions and the variable distance of the target adenosine to the 5'-end of the R/G-gRNA. A R/G-gRNA position 6 places the target adenosine at position 6, which means that between the adenosine and the 5'-end of the R/G-gRNA is 5 a nucleotide distance.

The experiment was performed in a 24-well format and the result was analyzed 48 hours post transfection. Each well was co-transfected with 300 ng of W58X eGFP and ADAR2 plasmid and 1300 ng of tested position of R/G-gRNA plasmid. The R/G-gRNAs with the positions 3 up to 10 were tested for their editing efficiency: P3 – pTS155.2, P4 – pTS156.1, P5 – pTS157.1, P6 – pTS68, P7 – pTS185, P8 – pTS186, P9 – pTS187 and P10 – pTS188.

The microscopy analysis showed that the fluorescence signal and number of fluorescing cells was enhanced by increasing the position number of the R/G-gRNAs until position 8 (Figure 4-72). Then, the trend reversed, resulting in a decline of fluorescence of position 9 and 10. Thus, the microscopy analysis demonstrated that R/G-gRNA position 8 was the most effective guideRNA in restoring the functional tryptophan. GuideRNAs that placed the target adenosine further to the middle of the mRNA binding site, indicated by a higher position number, resulted in more and brighter fluorescent cells than guideRNAs that placed the adenosine closer to the 5'-end of the R/G-motif (lower position number). The editing yields confirmed the microscopy observation (Figure 4-73): the R/G-gRNA position 8 resulted with 52% in the highest editing yields. R/G-gRNA positions above or beneath position 8 resulted in lower editing yields. These results demonstrate that the assumption to improve the total editing yield further by varying the position of the target adenosine in respect to the 5'-end of the R/G-gRNA, was fulfilled. The target adenosine was more efficiently edited by placing it into the middle of the RNA duplex. Placing the target adenosine closer to the border of the RNA duplex, resulted in less efficient editing. Summarized, the most optimal R/G-gRNA W58X eGFP to place the adenosine was position 8.

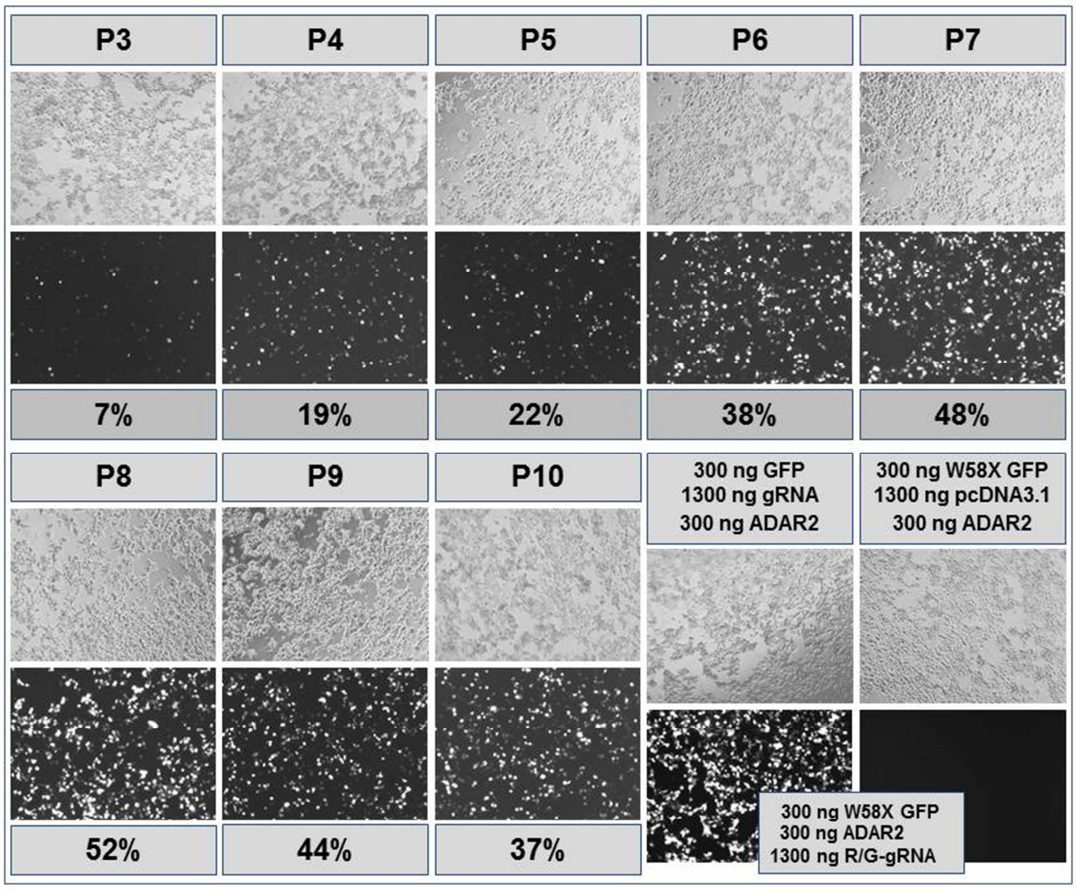

Figure 4-72: Varying positions of the W58X eGFP R/G-gRNA towards the target adenosine. To test the effect of the different R/G-gRNA positions towards the target adenosine a co-transfection with 300 ng of W58X eGFP and ADAR2 plasmid together with 1300 ng of the tested position of the R/G-gRNA was performed 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. 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 for achieving maximum editing yields. Magnification: 100x, GFP exposure time: 50 ms.

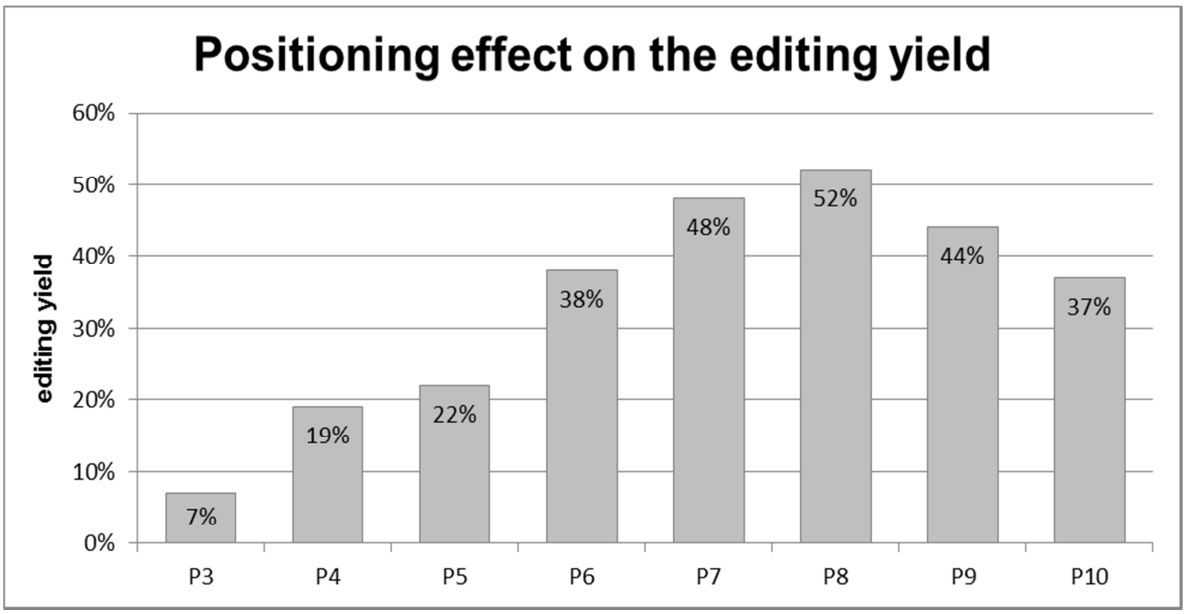

Figure 4-73: Positioning effect of the target adenosine in respect to the 5'-end of the R/G-gRNA on the editing yield. The R/G-gRNAs positions 3 until 10 are abbreviated by P3 – P10.

Our finding that position 8 is most effective in achieving high editing yields, may be specific for the particular sequence context of codon 58 in eGFP. To test if our finding is more general, the analysis of the editing efficiencies of variable positions was performed in the sequence context of the luciferase reporter gene, as well. These results confirm that the R/G-gRNA position 8 is the most suitable one for placing adenosine in ideal editing conditions for ADAR2. The details of the positioning experiment for the luciferase reporter gene will be discussed in chapter 4.2.3.1. Furthermore, a third positioning analysis of the R/G-gRNA was done with the PINK R407Q gene and was completely performed by Philipp Reautschnig. His results also confirm that position 8 is most suitable to place the target adenosine.

4.2.2.5 Varying the transfected R/G-gRNA position 8 plasmid amount

The previous experiments demonstrated that the most efficient R/G-gRNA to place the target adenosine was position 8. In the very beginning of the cell culture experiments was also shown that increasing amounts of transfected R/G-gRNA plasmid resulted in higher editing yields. The following transfection experiment was supposed to combine both – the optimal R/G-gRNA position 8 and the optimal amount of transfected guideRNA plasmid – to further increase the total editing yield. For the investigation of the optimal R/G-gRNA P8 plasmid amount, variable amounts

of it were co-transfected together with 300 ng of W58X eGFP plasmid and 300 ng of ADAR2 plasmid in a 24-well format. The microscopic analysis and RNA isolation with subsequent sequence analysis was performed 24 hours after transfection.

The differences in the fluorescence signal and amount were not distinguishable, when increasing the amount of R/G-gRNA P8 plasmid from 1000 ng up to 2000 ng (Figure 4-74). Lower amounts of transfected R/G-gRNA P8 plasmid (650 ng & 750 ng), however, showed less fluorescent and less bright cells. The corresponding editing yields confirmed the microscopy analysis that with increasing amounts of the R/G-gRNA P8 plasmid the editing level steadily increases until the amount of 1600 ng. The highest transfected amount (2000 ng) was not further improving the editing yield. Contrary, it decreased it by 9% compared to 1600 ng of R/G-gRNA P8 plasmid. The highest editing yield of 57% was obtained for an amount of 1600 ng R/G-gRNA P8 plasmid. Interestingly, when doubling the amount of R/G-gRNA (750 ng compared to 1600 ng) this does not result in a doubly editing yield. Either the editing machinery is already pushed to its maximum capacity with 1600 ng of R/G-gRNA and 300 ng of ADAR2 plasmid, or the transfection efficiency is limiting, too. The latter is definitely an issue. We tested a transfection reagent from another manufacturer (pecfect - not available anymore) that led to 63% editing yield at the same conditions, as for this experiment using 1600 ng R/G-gRNA P6 plasmid (data not shown). But this 63% editing yield was obtained for the R/G-gRNA position 6 and not for the optimal position 8. We assume that the transfection itself is a big limiting factor, as well. Another possibility to improve editing yield might potentially be to increase the amount of transfected ADAR2 plasmid. This was not pursued, since the final aim of this project is to steer endogenous ADAR2, which has lower concentrations in a living cell than in these transfection experiments. Further, a transfection of higher amounts of ADAR2 plasmid would run the risk for more and stronger off-target editing. Also the question arises, whether the usage of lower ADAR2 plasmid amounts reduces the amount of R/G-gRNA plasmid, which is necessary to achieve high editing yields. Since too little is known about the optimal stoichiometry of R/G-gRNA and ADAR2 in cell culture, we can only speculate about the effects of lowering or increasing one of the components.

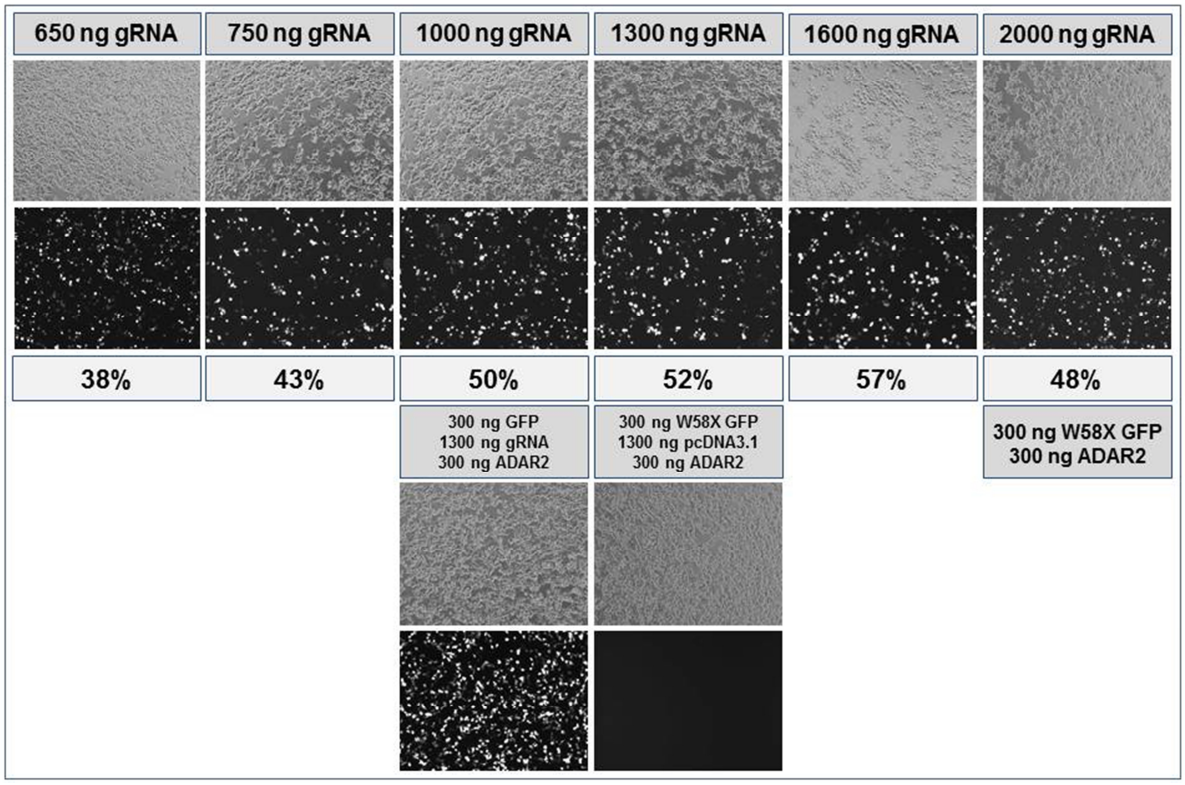

Figure 4-74: Varying amounts of ADAR2 and R/G-gRNA plasmids. Different amounts of R/G-gRNA P8 plasmid were transfected in 24-well plate format together with a constant amount of 300 ng W58X eGFP and ADAR2 plasmid. Higher amounts of R/G-gRNA plasmid resulted in more and stronger fluorescence, as well as in higher editing levels. Magnification: 100x, GFP exposure time: 50 ms.

Similar experiments of varying the R/G-gRNA plasmid amounts to edit luciferase W417X mRNA confirm that bigger amounts of transfected R/G-gRNA plasmid results in higher editing levels (details shown in chapter 4.2.3.3). The increasing editing yields for bigger amounts of R/G-gRNA plasmid demonstrate that the editing efficiency is mainly dependent on the presence of sufficient R/G-gRNA amounts.

4.2.2.6 Improving the stability of the R/G-gRNA by a boxB-motif

Considering the previous experiments, high amounts of guideRNA plasmid were required compared to the other two plasmids, implying that the availability of the R/G-gRNA might be a key to improve editing levels. One reason might be that the R/G-gRNA is unstable and being quickly degraded or the U6-promoter might express too little amounts of the guideRNA. Also the stoichiometry of the guideRNA and ADAR2 could determine improving the editing yields. We address all of these issues by stabilizing the guideRNA with a secondary structure at the 3'-end, and by the

creation of a plasmid containing guideRNA, as well as ADAR2 together in order to set a defined stoichiometry.

A ss-RNA is quickly degraded by 3'- and 5'- exonucleases, if the termini are not modified or protected by protein binding or secondary structure formation. In contrast to the 3'-end, the 5'-end is already a part of the secondary structure (R/G-motif). A modification of the 3'-end of the R/G-gRNA had to be chosen that enables the plasmid based editing system to be used as before, does not interfere or block the editing reaction and is auspicious in stabilizing the R/G-gRNA. For this purpose, a boxB-motif was attached to the 3'-end of the R/G-gRNA. The design and details of the attached boxB-motif to the 3'-end of the W58X eGFP R/G-gRNA P6 is demonstrated in Figure 4-75. Another issue is the presence of an undefined amount of poly(U)$_N$ at the 3'-end, which have the potential to represent a degradation signal. Since the poly(U)$_N$ -signal is needed as a termination signal for the RNA polymerase III this presence cannot be circumvented.

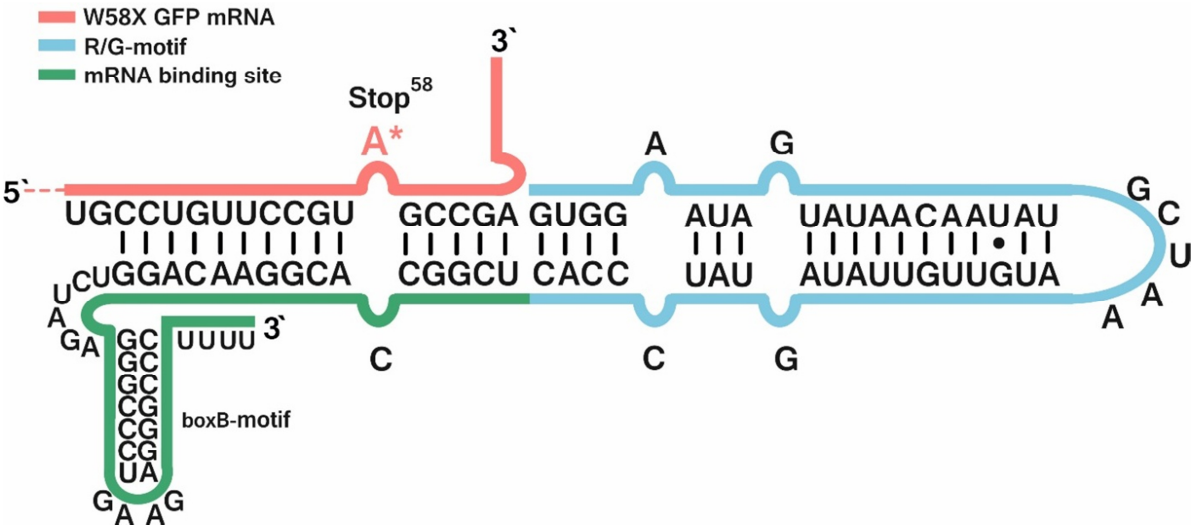

Figure 4-75: Design of the boxB-motif attached to the 3'-end of the W58X eGFP R/G-gRNA P6.
The flexible mRNA binding site of the R/G-gRNA is prolonged by a linker, which codes for the XbaI restriction site, followed by a boxB-motif.

The analysis of the effect of an attached boxB-motif was performed with the R/G-gRNAs position 6 and 8 in the context of W58X eGFP mRNA. The amount of 1300ng plasmid, coding for the R/G-gRNA, was co-transfected with 300 ng of W58X eGFP and ADAR2 plasmids in a 24-well format. The microscopic analysis and RNA isolation was performed 24 hours post transfection.

The addition of a boxB-motif to the R/G-gRNA position 6 seems to disturb the editing reaction, leading to a lower editing yield (28%) and less fluorescing cells compared to the R/G-gRNA P6 without the boxB-motif (41%) (Figure 4-76). However, the addition of the boxB-motif to the R/G-gRNA position 8 did not affect the editing level of this editing experiment significantly: 44% editing yield with boxB-motif and 42% without boxB-motif. The aim to drastically improve the editing yield by the addition of the boxB-motif to the R/G gRNA was not achieved, suggesting that the gRNA stability is not an issue. It remains elusive, why the attachment of the boxB-motif has a negative effect at the gRNA position 6 and not at the gRNA position 8 even though the distance of the boxB-element is 2 nt bigger to position 6 than position 8. The analysis of the boxB-motif attached to the guideRNA in the sequence context of W417X luciferase showed the same adverse effect on position 6 having a boxB-motif attached; whereas again no distinct effect is obtained for position 8 (details will be discussed in chapter 4.2.3.2).

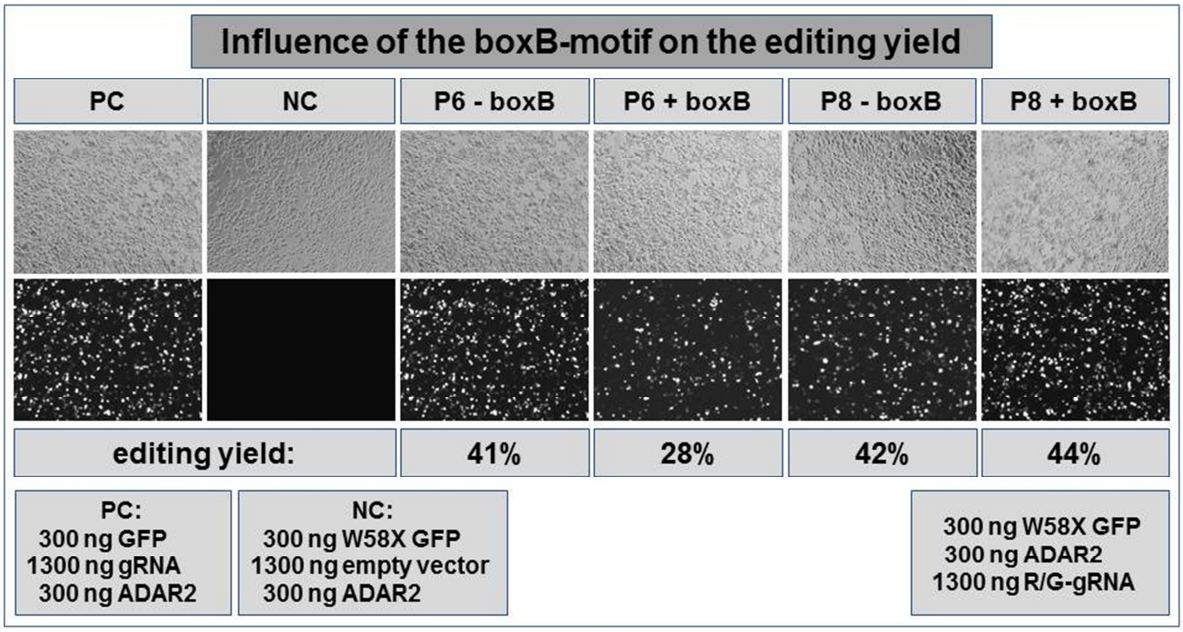

Figure 4-76: Influence of the boxB-motif on the editing efficiency in cell culture. The following amounts of plasmids were transfected in a 24-well plate format: 300 ng W58X eGFP, 300 ng ADAR2 and 1300 ng R/G-gRNA. The boxB-motif decreased the editing yield for the guideRNA targeting adenosine at position 6, whereas the editing levels of the guideRNA with position 8 stayed nearly. Magnification: 100x, GFP exposure time: 50 ms.

4.2.2.6.1 Gibson vector – the multiple copy vector

The second approach to positively influence the R/G-gRNA availability in the cell was to create a vector, which encodes for multiple copies of R/G-gRNA and U6 promoter and should overcome the problem that the R/G-gRNA could be limiting. This vector was cloned with the Gibson method and is called from now on Gibson vector (pTS137.2) (3.2.15). The ADAR2 gene was inserted on the same vector for two reasons: the first reason was the idea that a combination of multiple R/G-gRNA and U6 promoter copies together with the ADAR2 gene on the same vector lowers the total amount of transfected vectors, in order to exert less transfection stress on the cells. The other reason is that using one vector coding for both, the R/G-gRNA and the ADAR2 protein, assures that both components will always be present in the cell and in a defined stoichiometry.

By the time of this experiment, the effect of the different positions of the R/G-gRNA to the target adenosine at the editing yield has not been explored, therefore, this Gibson vector codes for four copies of the less effective R/G-gRNA position 6.

The transfection experiment included two different amounts of Gibson vector (300 ng and 1300 ng) together with 300 ng W58X eGFP plasmid in a 24-well format. The microscopic analysis and the RNA isolation were performed 48 hours post transfection. The sample containing 300 ng of Gibson vector demonstrates that editing of W58X eGFP mRNA in cell culture is possible (Figure 4-77). The amount and brightness of fluorescing cells was less intense than for the sample transfecting 1300 ng of Gibson vector. The sequence analysis showed the same result as the microscopic analysis; the editing yield of 1300 ng transfected Gibson vector was higher (54%) than for 300 ng (38%).

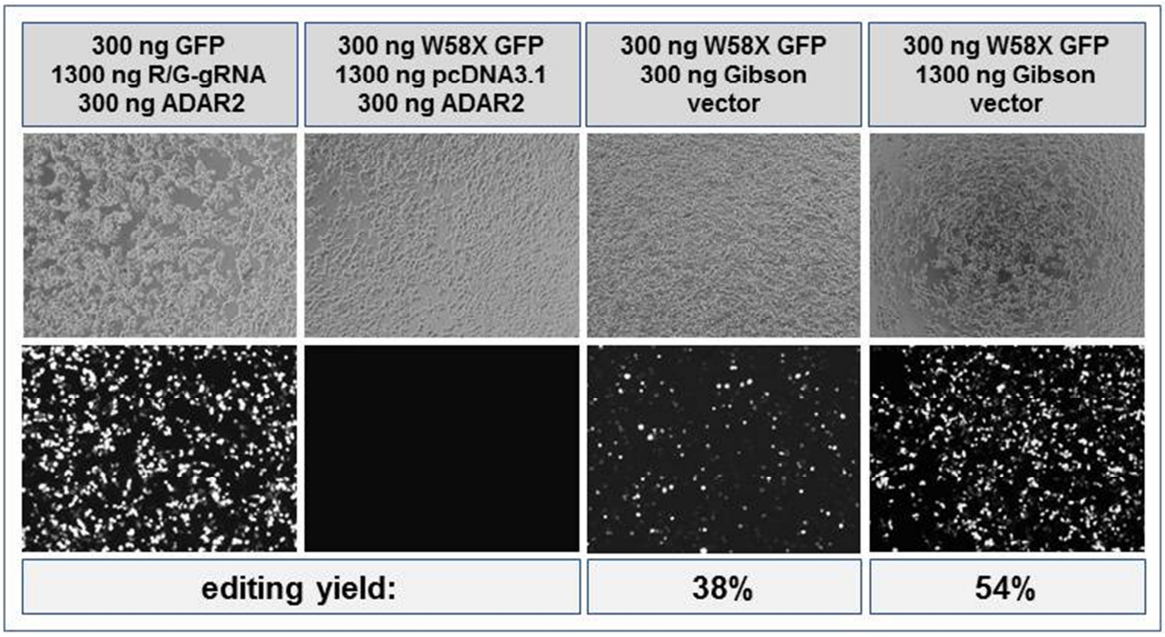

Figure 4-77: Editing efficiency of the Gibson vector in cell culture. The Gibson vector contains 4 copies of the R/G-gRNA P6 + U6 promoter construct and the ADAR2 gene. The Gibson vector was co-transfected with 300 ng of W58X eGFP plasmid in a 24-well format. The microscopic analysis and RNA isolation was performed 48 hours post transfection. Higher amounts of the Gibson vector (1300 ng) led to stronger fluorescent cells and amount. The sequence analysis confirmed the microscopic observation: 54% editing yield for 1300 ng of Gibson vector and 38% for 300 ng of Gibson vector. Magnification: 100x, GFP exposure time: 50 ms.

Unfortunately, the idea to reliably enhance the R/G-gRNA transcription by multiple copies on the same plasmid was not that successful: Various transfecting experiments were performed with the Gibson vector and the editing efficiency varied a lot. One possibility for this failure could be the size of the Gibson vector (9061 bp). First of all, the transfection efficiency with such a big plasmid is very low, and on the other hand the stability of such big plasmids is strongly reduced. These issues would result in lower editing yields and the evaluation of the potential of the multiple copy vector would mistakenly interpreted as ineffective. Therefore, always freshly produced Gibson vector without any freeze and thaw cycles were used for the latest experiments. The reduction of the plasmid size is a topic of a further PhD-project in the group and is already in the process.

It should be taken into account that the transfection of 1300 ng Gibson plasmid is not equimolar to the amount of the R/G-gRNA plasmid or even the ADAR2 plasmid, since the latter are much smaller in their plasmid size (4899 bp or 7547 bp). Also the usage

of the more efficient R/G-gRNA position 8 will definitely improve the editing yields further.

The idea to test another promoter for the guideRNA expression was tested by Pia Mach during her bachelor thesis (98). She compared the editing efficiencies of a R/G-gRNA expressed from either the U6- or H1-promoter cloned into a p*Silencer* vector. Unfortunately, no improvements for the editing yields were obtained. Contrary, the editing yields decreased using the H1-promoter for the R/G-gRNA expression in the cell culture experiment.

4.2.2.7 Variation of incubation time for R/G-gRNA position 8

Many optimizations were performed with the aim to increase the total editing yield, starting from the optimal amount of the R/G-gRNA plasmid via analyzing the optimal length and position of the mRNA binding site to the target adenosine, the addition of a boxB motif to the 3'-end of the guideRNA and prolonging the editing time in cell culture. Now, the most optimal R/G-gRNA with 16 nt length, a boxB-motif attached to the 3'-end and placing the adenosine at position 8 relative to the beginning of the mRNA hybridizing part was used to repeat the time dependent editing experiment from the beginning. The co-transfection experiment was performed with 300 ng of W58X eGFP and ADAR2 plasmid, as well as two different amounts of R/G-gRNA P8 plasmid (1300 ng and 1600 ng). This time, the microscopic analysis and RNA isolation were performed 24, 48, 72 and 96 hours post transfection to figure out if a further prolongation of the editing time has the potential to increase the total editing yield or not.

The brightest fluorescence signal and the biggest number of eGFP expressing cells of the positive control were obtained after 48 hours of incubation (Figure 4-78). A shorter or longer incubation time decreased the eGFP signal. Both editing samples with varying R/G-gRNA P8 plasmid amounts showed a bigger number of eGFP expressing cells and a stronger fluorescent signal until 72 hours of incubation time. Afterwards, the eGFP signal was declining at 96 hours of incubation. The editing yields of the sequence analysis confirmed the microscopic observation: until 72 hours of editing the editing yields increased at both chosen R/G-gRNA P8 plasmid amounts up to 52% for each and decreased at 96 hours of incubation (44% for 1300 ng and 42% for 1600 ng of R/G-gRNA P8 plasmid).

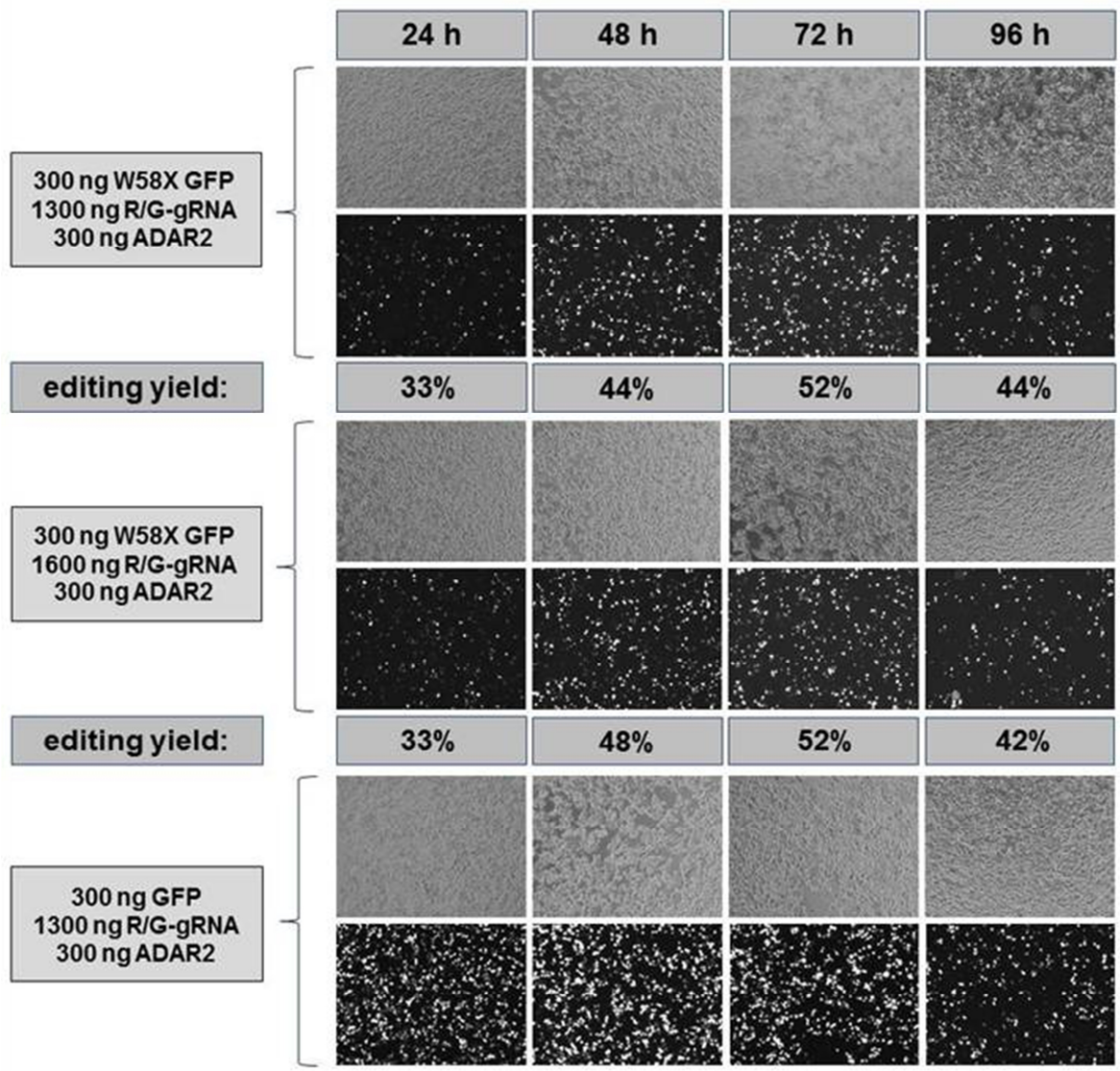

Figure 4-78: Prolongation of the editing time for the R/G-gRNA position 8. The co-transfection experiment of 300 ng W58X eGFP plasmid, 300 ng of ADAR2 plasmid and 1300 ng or 1600 ng R/G-gRNA P8 plasmid was performed in a 24-well format. The editing efficiency was analyzed 24, 48, 72 and 96 hours post 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 72 hours of incubation was visible. The eGFP intensity and amount of cells was declining after 96 hours of incubation. The sequence analysis confirmed the fluorescent microscopy: the editing yields increased until 72 hours of editing time and decreased after 96 hours again. Magnification: 100x, GFP exposure time: 50 ms.

In a next step, a FACS measurement was performed to analyze the effect of prolonging the editing time until 96 hours on the editing yield. The FACS analysis was supposed to complement the microscopic analysis by a quantification of the

fluorescent cells and their brightness distribution. First of all, the subpopulation of living cells (P2) was defined by plotting forward scatter (FSC) against side scatter (SSC) of untransfected cells and exclusion of cells with high granularity and size (Figure 4-79 a)). The same plot was performed for the positive control, which was transfected with 300 ng of eGFP plasmid, 300 ng of ADAR2 plasmid and 1300 ng R/G-gRNA P8 plasmid. In order to define the subpopulation of GFP positive cells (P1) the counts of the negative cells were plotted against the GFP-signal and the same was done for the positive control (b, e). The subpopulation P1 was adjusted in a way that as much as possible of the positive control belonged to it and as less as possible of the negative control. By this 0.3% of the negative control cells were defined as GFP fluorescent, which is tolerable as false positive (b, e). To determine the population of living and fluorescing cells (P3), the P1 subpopulation had to be separated from the cells defined as not living and the not fluorescing living cells. This was achieved by plotting the GFP-signal against FSC (c, f). This gating strategy, exemplary demonstrated in Figure 4-79, was applied to all the transfection samples of the experiment before. To cover all incubation time points of the range of 96 hours, each transfection sample (positive control and 2 editing samples) was started with four batches. The cells were harvested and fixed according to the protocol described in chapter 3.5.6 at each time point.

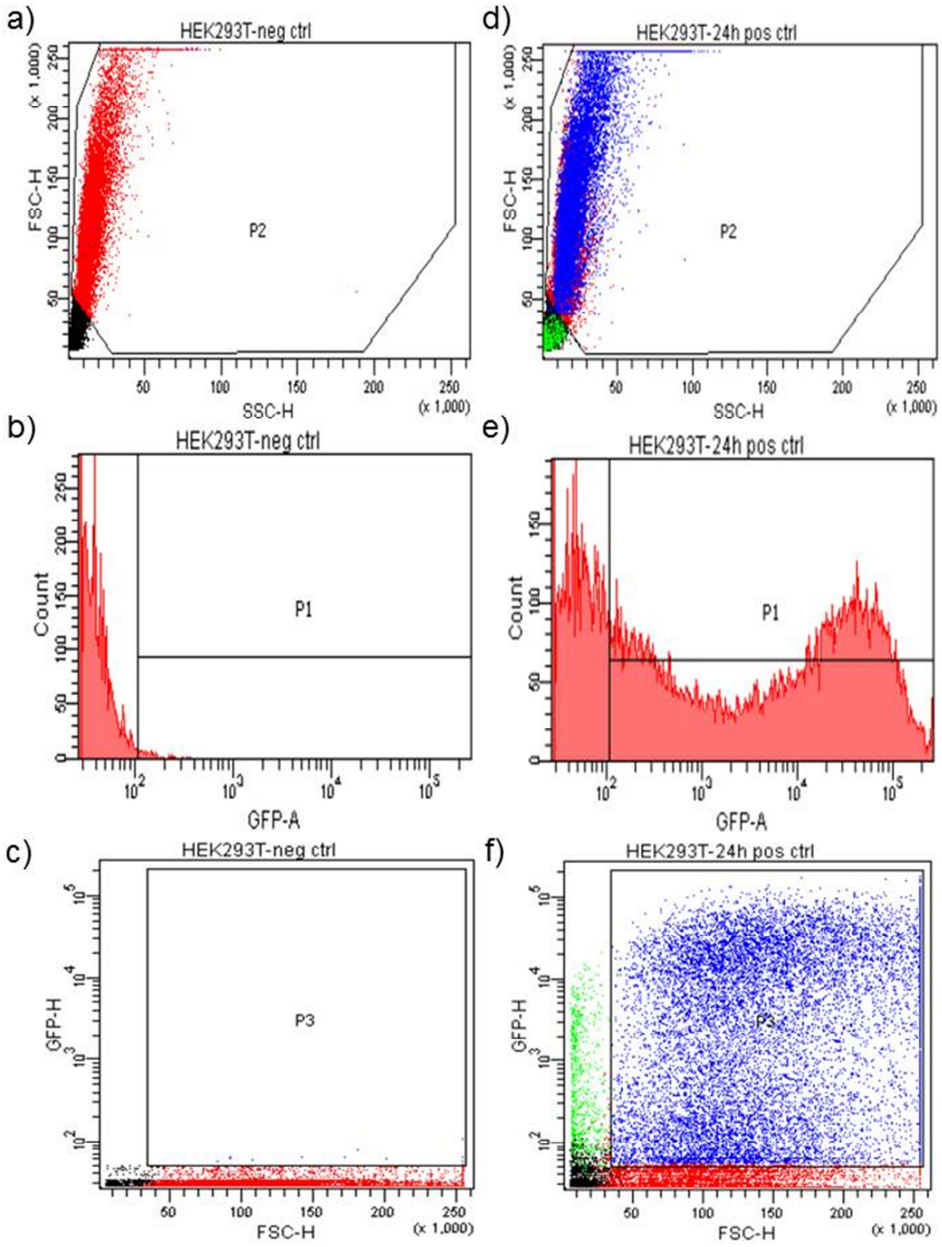

Figure 4-79: Gating strategy to define the subpopulations of living and GFP expressing cells.

Living cells were defined as a subpopulation (P2) by plotting forward scatter (FSC) against side scatter (SSC) for untransfected cells and excluding cells of a small size and high granularity (a) and d)). To define the subpopulation of GFP positive cells, all counts of the negative control were plotted against GFP-signal and served as a template to separate non fluorescing cells from fluorescing cells (P1) (b). The same gate was used for all counts of the positive control plotted against GFP-signal (e). To determine the amount of living and fluorescing cells, the GFP-signal was plotted against the forward scatter, to exclude either small size, high granular or not fluorescing cells (or all), resulting in the amount of living and fluorescing cells (P3) (c) and f)).

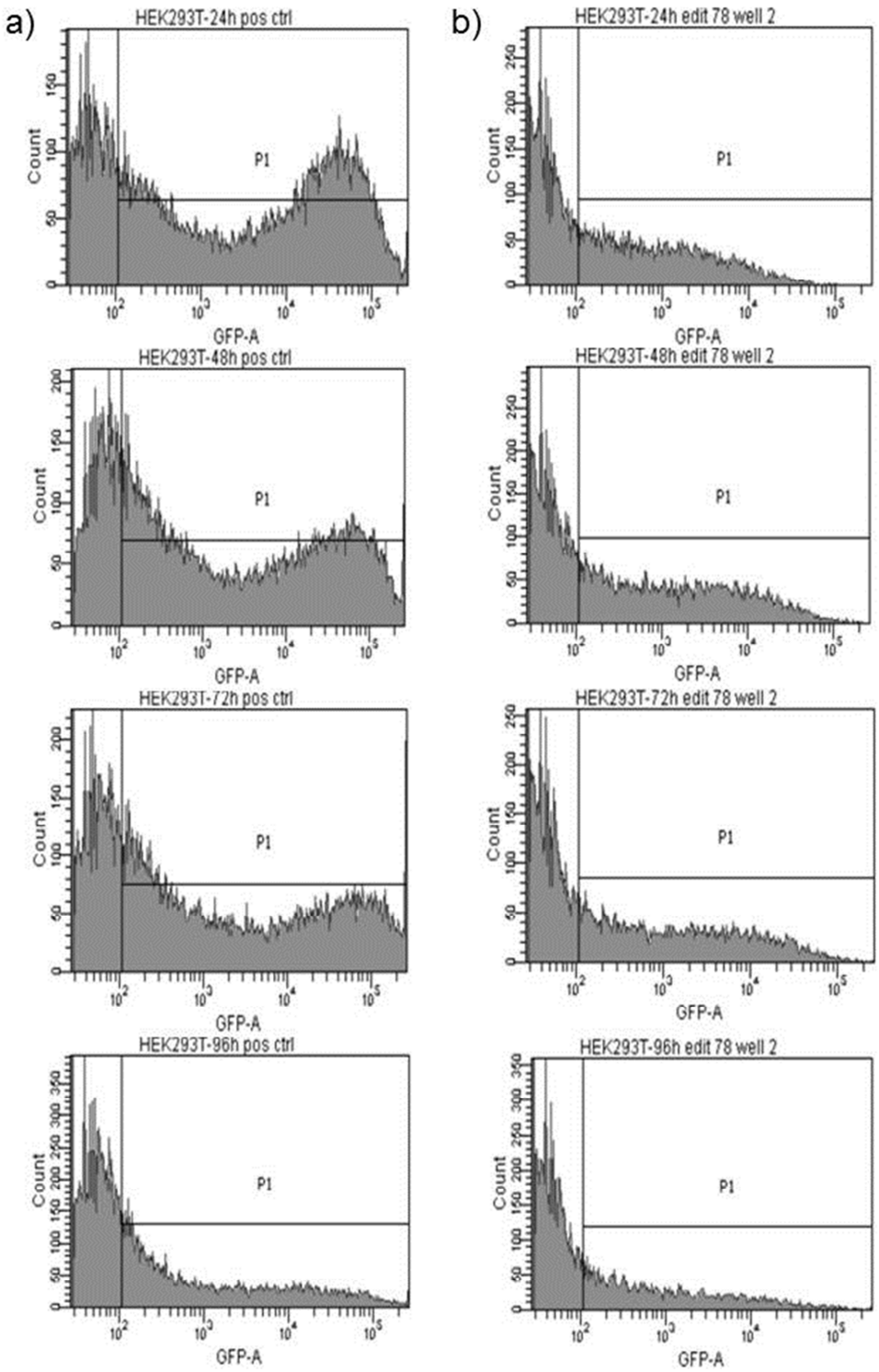

Figure 4-80: Comparison of the fluorescent intensities of the positive control with the editing sample. The intensity of the GFP signal of the positive control was decreasing with longer incubation times as well as the amount of cells that show a strong fluorescent signal (a). The fluorescent signal of the editing sample (1300 ng R/G-gRNA plasmid) was much lower compared to the positive control (b).

Comparing the fluorescing cells of the positive control for the different chosen editing end points, a clear declining trend of the GFP intensity was visible (Figure 4-80). The intensity and the number of GFP expressing cells of the editing sample (1300 ng R/G-gRNA plasmid) was distinctly lower for all chosen time points than the ones of the positive control.

Table 4-3: Overview of the subpopulations and editing yields. For each sample 20.000 events were measured. The FACS editing yield was determined by putting the amount of living and fluorescent cells (P3) of the editing sample in relation to the one obtained for the positive control.

Editing time		PC [%]	Edit sample 1300 ng gRNA [%]	Edit sample 1600 ng gRNA [%]
24h	P1 (fluorescent)	46.2	19.4	21
	P2 (living cells)	68.8	69.3	70.4
	P3 (living&fluorescent)	45.4	20.0	21.6
	FACS editing yield:		**44**	**48**
48h	P1 (fluorescent)	54.7	25.4	19.2
	P2 (living cells)	72.7	72.8	61.9
	P3 (living&fluorescent)	53.6	26.3	19.3
	FACS editing yield:		**49**	**36**
72h	P1 (fluorescent)	45.1	19.5	24.2
	P2 (living cells)	65.7	69.8	66.2
	P3 (living&fluorescent)	42.6	19.8	24.5
	FACS editing yield:		**47**	**58**
96h	P1 (fluorescent)	33.2	16.8	18.9
	P2 (living cells)	79.2	75.1	76
	P3 (living&fluorescent)	35.8	17.2	19.3
	FACS editing yield:		**48**	**54**

The intensity of the GFP signal of the positive control was decreasing over time, but the number of living and fluorescing cells increased from the time point 24 to 48 hours (P3), and was declining from 48 hours to 96 hours (Table 4-3). In the FACS analysis the same trend was visible for the editing sample transfecting 1300 ng of R/G-gRNA plasmid: the number of living and fluorescing cells (P3) increased from 24 to 48 hours increases, whereas the number decreased from 48 to 96 hours. But the editing sample transfecting 1600 ng of R/G-gRNA plasmid showed a different peak of living and fluorescing cells: the highest obtained number of living and fluorescing cells was

at 72 hours. Shorter and longer incubation times led to a lower number of living and fluorescing cells. To facilitate the comparison of the results obtained from the sequence analysis with the FACS measurement, the living and fluorescing cells (P3) of the editing samples were put into relation to the ones of the positive control, resulting in a percentage editing yield (Table 4-3). Table 4-4 contrasts the editing yields of the sequence and FACS analysis. The editing yields do not match in their absolute values, but both methods to determine the editing yields have in common that the editing yields increase from 24 to 48 hours. The amount of living cells of the editing sample transfected with 1600 ng guideRNA and 48 hours editing time (P2 - 61.9%) is markedly lower than the ones of the other editing sample (72.8%) or the positive control (72.7%). Obviously, the transfection led to lower amounts of healthy cells in this sample compared to others. Therefore, this sample of the FACS measurement is interpreted as an outlier (36% yield) and the trend would be the same as for the sequence analysis: the total editing yield increases up to 72 hours and declines again after this incubation time. At previous FACS analysis this trend was reliably obtained for the incubation times 24 hours, 48 to 72 hours (data not shown.) The reason for this decrease after 72 hours is based on the thinning out effect of the transfected plasmids. This effect is demonstrated by the FACS analysis of the positive control, expressing the GFP: A clear decrease of the fluorescent signal and amount of fluorescing cells is obtained after 48 hours (P1) (Table 4-3).

Table 4-4: Overview of the editing yields of the sequence and FACS analysis.

Editing time	Method	1300 ng gRNA editing yield [%]	1600 ng gRNA editing yield [%]
24h	Sequencing:	33	33
	FACS:	44	48
48h	Sequencing:	44	48
	FACS:	49	36
72h	Sequencing:	52	52
	FACS:	47	58
96h	Sequencing:	44	42
	FACS:	48	54

The FACS analysis is a good option to obtain comparable values for the editing yields than only the subject interpretation of microscopic pictures, but harbors a source of

errors through the process of cell fixing: loss of sample and destruction of GFP positive cells by cell fixing. On the other hand, the RNA isolation method harbors the same risks of loss of sample, but it has the advantage that the RNA of cells, which are not that healthy anymore, is isolated and analyzed, as well. One major aspect that has to be minded by comparing editing yields of FACS measurement and RNA isolation is the fact that by RNA isolation the transcript level is analyzed and by FACS the translated protein, which accumulates over time. Since the GFP expression is dependent on the editing at the transcript level, the FACS measurement will always deliver a slightly time delay of information about the editing yield. The distribution of the fluorescent cells of the positive control varies a lot between the editing times 24 hours and 96 hours: The emphasis of fluorescent cells is more located towards brighter fluorescent cells (24 hours), whereas the emphasis of fluorescent cells at 96 hours is more equal. Also the emphasis of the editing sample transfecting 1300 ng of guideRNA varies between the four editing durations: The emphasis shifts from 24 to 48 hours towards brighter fluorescent cells and is regressive after 96 hours, indicating that editing takes time.

A longer incubation time of 72 hours and 96 hours increased the editing yield of the off-target position 127 up to 13% (72h) and 11% (96h) (data not shown). It also created a second off-target site at position 102 (5'-AA<u>A</u>*). This position was edited up to 11% by 72 hours and 9% by 96 hours of incubation time (data not shown).

In summary, the total editing yields grow with increasing incubation time. But this effect is only obtained until 72 hours of incubation. Afterwards, the editing efficiencies decrease again, probably due to the thinning out of plasmids. The optimizations for the design of the R/G-gRNA for editing W58X eGFP mRNA came to an end with that set of experiments and in a next step the analysis of the feasibility to recruit the endogenously expressed ADAR2 was started.

4.2.2.8 Lowering ADAR2 plasmid

The engineered R/G-gRNA allows editing of W58X eGFP mRNA in cell culture by overexpression of the natural ADAR2 protein. The aim was not only to create an editing system which is encodable for all components of the editing system, but we rather aim to steer the endogenously expressed ADAR2 with the designed R/G-

gRNA. A next challenge for this editing system appears: Is the engineered R/G-gRNA able to steer endogenously expressed ADAR2?

In order to approach this answer, an experiment was performed with a constant amount of 300 ng W58X eGFP plasmid, a constant amount of 1300 ng R/G-gRNA-P8 plasmid and decreasing amounts of ADAR2 plasmid starting from 300 ng and lowering to 25 ng. The experiment was performed in the 24-well format and the microscopy analysis and RNA isolation were performed 48 hours post transfection.

The transfection of 300 ng over 200 ng to 100 ng of ADAR2 plasmid showed no marked difference in the editing yields ranging from 52% for 300 ng, 53% for 200 ng and 50% for 100 ng of ADAR2 plasmid (Figure 4-81). With a further decrease of ADAR2 plasmid down to 75 ng, 50 ng, and 25 ng the editing yields slowly diminished to 36%. The editing yield of 36% for the transfection of only 25 ng of ADAR2 plasmid raise the hope that even endogenous ADAR2 protein levels might be steerable for directed deamination activity. Of course, even the lowest amount of ADAR2 plasmid is not comparable with the basal ADAR2 expression in HEK293T cells, since the CMV-promoter enables a very strong expression and the HEK cells are obtained from human embryonic kidney cells, which are not supposed to have a high expression of ADAR2 compared to neuronal cells for instance. The off-target editing, which was mentioned in chapter 4.2.2.1 to be present for transfecting 300 ng ADAR2 plasmid and 48 hours of incubation time, can be decreased for transfecting less than 200 ng of ADAR2 plasmid and disappears entirely starting from 100 ng of ADAR2 plasmid (Figure 4-82).

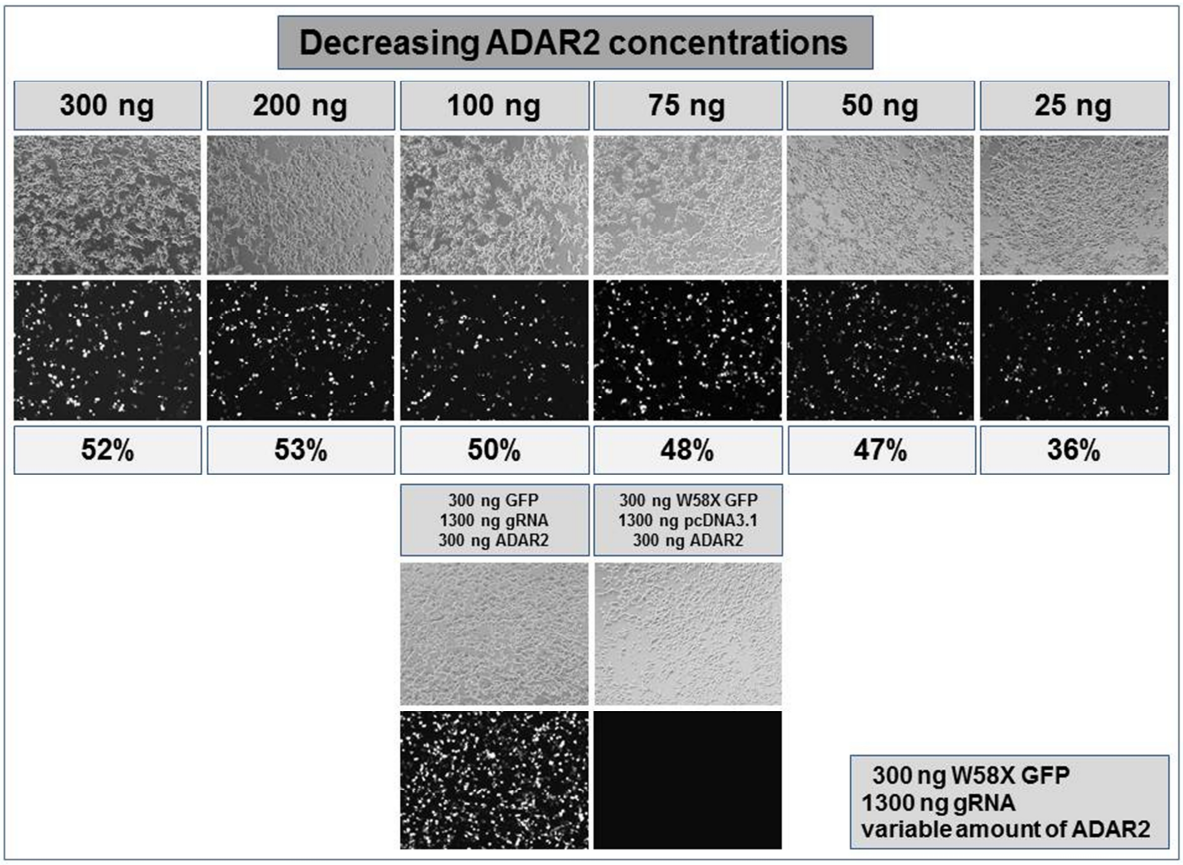

Figure 4-81: Decreasing amounts of transfected ADAR2 plasmid. In a 24-well format cells were co-transfected with 300 ng W58X eGFP and 1300 ng R/G-gRNA-P8 together with varying amounts of ADAR2 plasmid. The microscopic analysis and RNA isolation was carried out 48 hours post transfection. The decrease 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. Magnification: 100x, GFP exposure time: 50 ms.

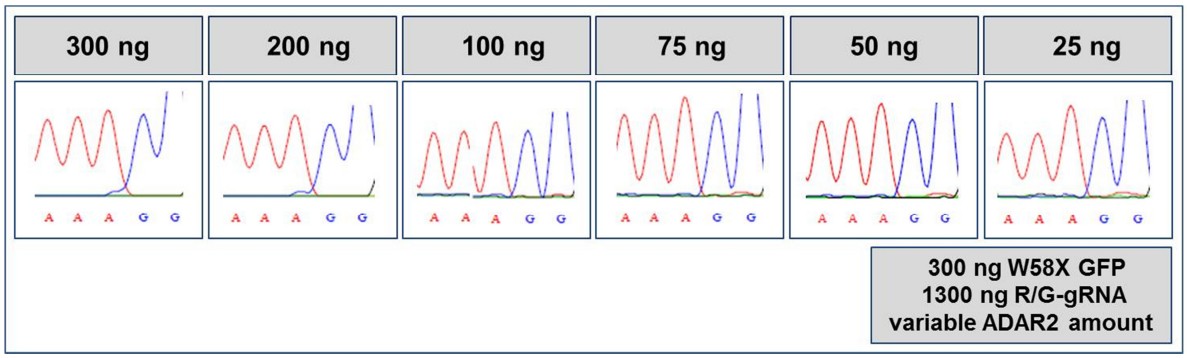

Figure 4-82: Decreasing off-target editing at position 127 with lower ADAR2 plasmid amounts. The reduction of the transfected ADAR2 plasmid amount leads after 48 hours of incubation to a decrease at the off-target site 127 eGFP. Transfection of 100 ng or lower ADAR2 plasmid amount completely prevents the off-target editing.

One of the fundamental experiments, which were performed in the beginning, was the testing if the two components, R/G-gRNA W58X GFP and ADAR2, are truly responsible for restoring the tryptophan codon and the fluorescent phenotype (4.2.2.2). The conclusion was that only the presence of these two components mediates editing of the W58X eGFP mRNA. But as already mentioned, some cells of the samples that were transfected only with W58X eGFP and R/G-gRNA plasmid but not with (functional) ADAR2 plasmid showed a low number of faintly fluorescing cells, suggesting that endogenous deamination activity may have been recruited in these cells. This experiment was repeated, transfecting 300 ng of W58X eGFP plasmid and 1600 ng of R/G-gRNA W58X eGFP P8 plasmid, to see if the result can be reliably repeated. Indeed, a few cells showed again slightly eGFP fluorescence signal (Figure 4-83). Obviously, taken the results from chapter 4.2.2.2 into account, the endogenous editing activity in HEK293T cells is steerable by the engineered R/G-gRNA. Whose editing activity is steered - ADAR2, ADAR1 or even ADAT - is not revealed by this experiment. Since both, ADAR1 and ADAR2 are expressed in kidney (8), only the fact that endogenous editing activity of one of them is steered could be demonstrated.

From a poster presentation by Chung Hachung at the Gordon RNA editing conference 2015 and the personal conversation with her, I was told that the basal expression of ADAR2 in HeLa cells is much stronger than for HEK293T cells. This was also confirmed by the findings from Galeano et al. 2010 (99). This would be an option to test this HeLa cell line with the same experimental setup. But more interestingly is to demonstrate that endogenous editing activity can be used for site-specific RNA editing in a neuronal cell line, neuroblastoma or for example PC-12 cells (100).

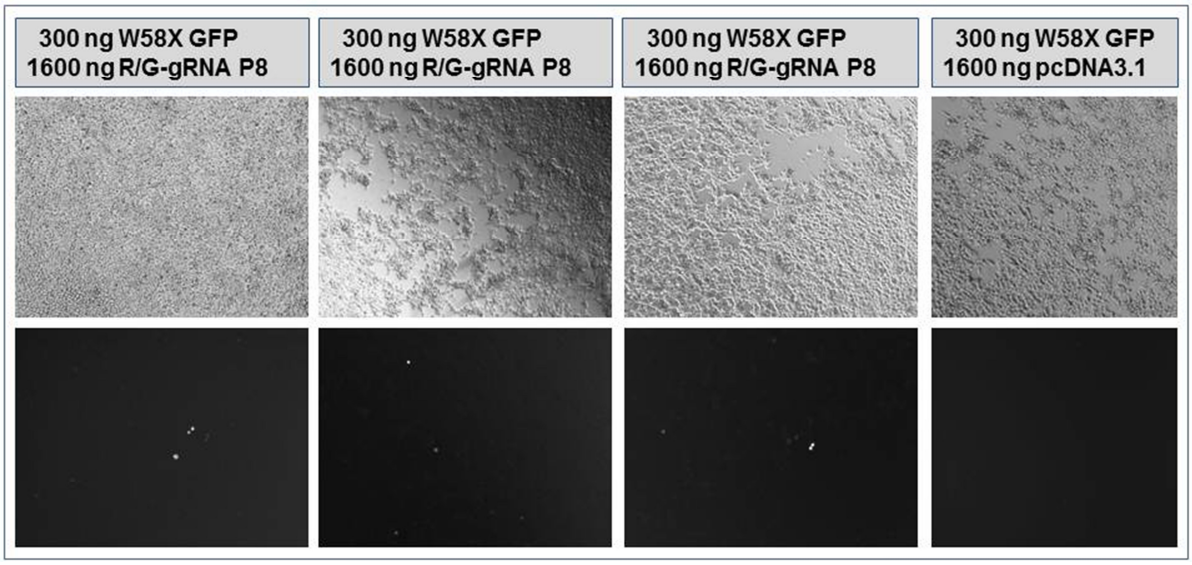

Figure 4-83: Steering the endogenously expressed ADAR2 of HEK293T cells. HEK293T cells were co-transfected with 300 ng W58X eGFP plasmid and 1600 ng of R/G-gRNA P8 plasmid in a 24-well format. The microscopic analysis was performed 48 hours post transfection. A few cells showed a slight fluorescence eGFP signal, demonstrating that the endogenous ADAR2 participates in the editing reaction of W58X eGFP mRNA. Magnification: 100x, GFP exposure time: 50 ms.

The attempt to create a stable cell line either by transient transfection and antibiotic resistance selection, or by establishment with the Flp-In T-REx system (one copy of the ADAR2 gene under Tet inducible promoter expression) or by viral transfection is still in process. Experiments with such a cell line will allow gaining further information of the ability to steer the ADAR2 enzyme, which is lower concentrated than by transient transfection, with the engineered R/G-gRNA.

4.2.3 Editing of a second reporter gene: luciferase

The feasibility of editing the W58X eGFP mRNA by transient transfection of all editing components, including the reporter and ADAR2 gene, as well as the R/G-gRNA plasmid, was shown. To demonstrate that other adenosines in a different sequence context can be edited similarly, a second reporter gene was analyzed. The luciferase gene was used for this purpose since it allows the usage of the luciferase reporter assay to analyze successful editing besides RNA isolation and sequencing. Again, but now with the luciferase mRNA, the analysis of the positioning effect of the R/G-gRNA to the target adenosine and the effect of the boxB-motif attached to the guideRNA was validated. Additionally, the amounts of transfected guideRNA and

ADAR2 plasmid were varied. Furthermore, we studied the simultaneous editing of the W58X eGFP mRNA and W417X luciferase mRNA in cell culture.

4.2.3.1 Revisiting the positioning effect

A stop codon at position W417 was inserted in the luciferase gene (Table 3-18, Table 3-19) to analyze the editing of this new target. An exemplary design of a R/G-gRNA P6 hybridized to the W417X luciferase mRNA is shown in Figure 4-84. The editing efficiency was assessed by performing a luciferase reporter assay (3.5.5) and RNA isolation followed by sequence analysis of the luciferase transcript. The luciferase assay was always performed with cells transfected in a 96-well format and transfected cells of a 24-well format were used for the RNA isolation and sequence analysis. The primer no. 144 was used as a bw primer for reverse transcription, primer no. 144 and no. 138 for Taq-PCR. The primer no. 335 was used for sequencing.

The bioluminescence results of the luciferase assay are given in relative light units (RLU), corresponding to the luciferase enzyme activity and were equated to the relative editing efficiency of each transfection experiment.

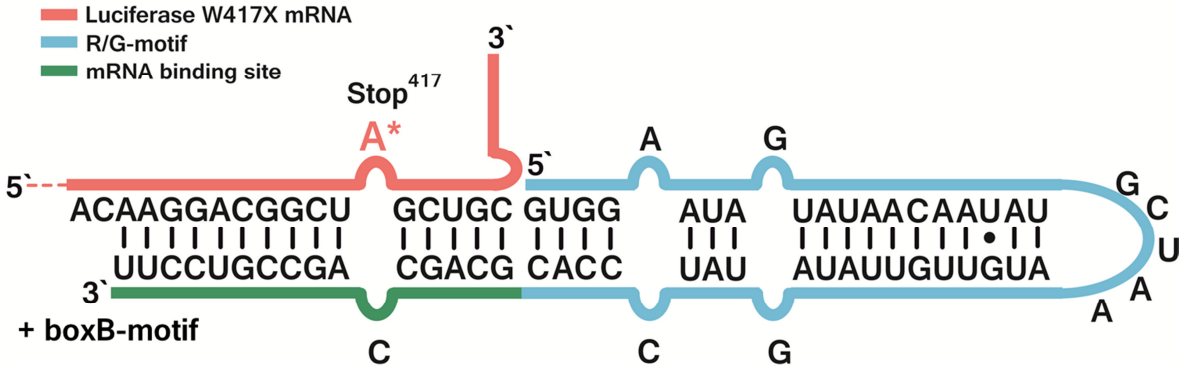

Figure 4-84: Illustration of R/G-gRNA position 6 hybridized to the W417X luciferase mRNA.
The W417X luciferase mRNA (red) is hybridized to the mRNA binding site (green) of the R/G-gRNA position 6 (blue).

In the following experiment we investigated, whether the editing system is addressable to another adenosine in a different mRNA sequence context and which position of the R/G-gRNA (with boxB-motif attached to the 3'-end) in respect to the target adenosine is most optimal. The cells of a 96-well format were transfected with 60 ng of W417X luciferase plasmid (pTS116.8), 60 ng of ADAR2 (pTS57) and 250 ng

of each tested position of R/G-gRNA plasmid in separate samples; P5 plasmid (pTS169), P6 plasmid (pTS121), P7 plasmid (pTS193), P8 plasmid (pTS170) and P9 plasmid (pTS171). The transfection experiment of the 24-well format included 300 ng of W417X luciferase, 300 ng of ADAR2 plasmid and 1300 ng of the tested position of R/G-gRNA plasmid in separate samples. The luciferase assay and the RNA isolation were performed 48 hours post transfection.

An overview of the obtained bioluminescence measurements (in RLU) of two luciferase assays, performed with two independent transfection experiments, is shown in Figure 4-85. The higher the RLU value, the higher was the luciferase enzyme activity of the sample, indicating a higher editing efficiency. Both experiments showed that R/G-gRNA position 8 is the most optimal position of the target adenosine followed by position 9. Taken the results of the sequence analysis into account (Figure 4-86), the second experiment of the luciferase assay matches completely with the ranking of the R/G-gRNA positions to the sequence results: position 8> position 9> position 5 > position 6> position 7 (most optimal>less optimal). The R/G-gRNA P8 achieved the highest editing level of the sequence analysis, whereas the position 7 resulted in the lowest editing yield. The same trend for the most optimal positions of R/G-gRNAs is obtained for the first luciferase assay experiment, too, with an outlier for position 7. This outlier might be reasoned for the error susceptibility of the luciferase assay. For example, variations in the cell amount that is disrupted for the assay, lead to different absolute RLU values. Both methods agree for the most optimal position of R/G-gRNA: position 8. The same result for the most optimal position (P8) was also obtained for the experiments of W58X eGFP mRNA editing (4.2.2.4). As mentioned in chapter 4.2.2.4, the positioning effect was analyzed for the PINK R407Q gene by Philipp Reautschnig, as well and demonstrates the same preference of the R/G-gRNA P8 as for W58X eGFP and W417X luciferase mRNA. Comparing the results of the editing yields obtained for different tested R/G-gRNA positions with W58X eGFP mRNA – a steady increase of the editing yield for placing the adenosine step by step into the middle of the RNA duplex – is not precisely visible for the tested positions of W417X luciferase mRNA. In order to define the real ranking of the R/G-gRNA positions to edit W417X luciferase mRNA, the experiments would have to be repeated. But the main message that was obtained from this set of experiments is clear: It is important to verify the optimal positions for high editing

yields, since position matters and R/G-gRNA position 8 is the most optimal guideRNA for editing W417X luciferase mRNA.

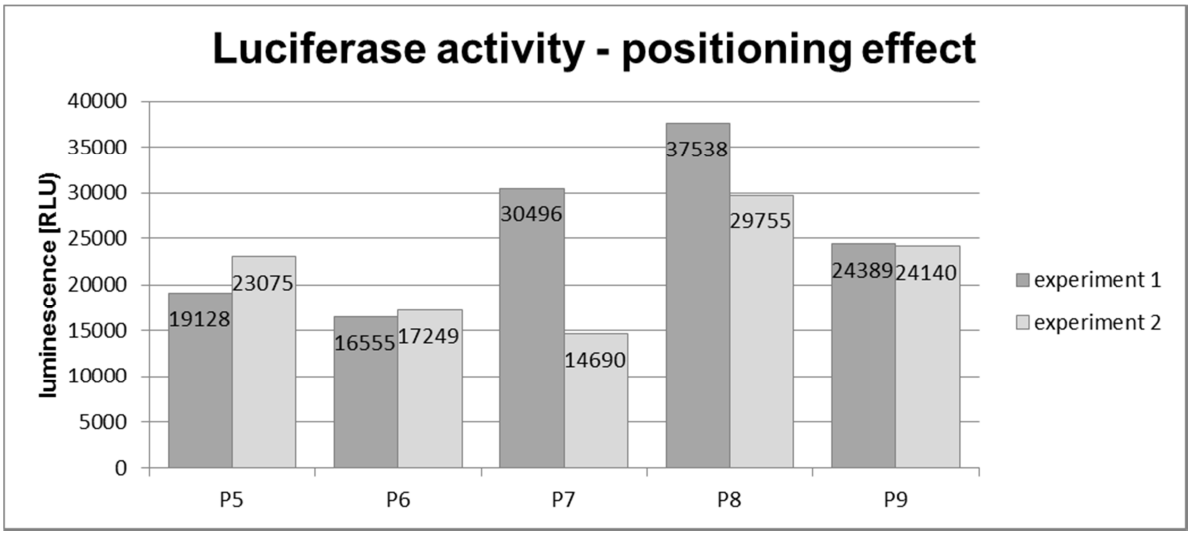

Figure 4-85: Luciferase activity of different positions of the target adenosine in respect to the 5'-end of the R/G-gRNA. The luciferase assay was performed with transfected cells of a 96-well format. The transfection included 60 ng of W417X luciferase and ADAR2 plasmid as well as 250 ng of each tested R/G-gRNA plasmid in separate samples. The R/G-gRNAs positions 5 until 9 are abbreviated by P5 – P9. The results for the ranking of the positions 5 to 7 are varying for the two independently performed experiments. Both experiments demonstrated that position 8 of the R/G-gRNA is the most optimal position. The bioluminescent values are given in relative light units (RLU).

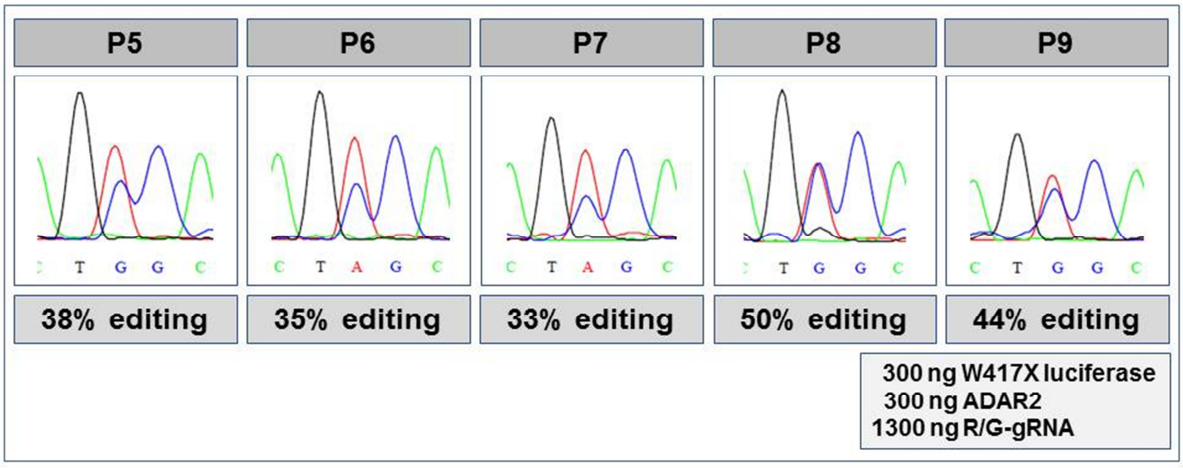

Figure 4-86: Sequence analysis of different positions of the luciferase R/G-gRNA towards the target adenosine with W417X luciferase mRNA. In a 24-well format cells were co-transfected with 300 ng W417X luciferase and ADAR2 plasmid, as well as 1300 ng of each tested R/G-gRNA plasmid. The positions 5 until 9 were tested, abbreviated by P5 – P9 and the RNA was isolated 48 hours post transfection. The highest obtained editing yield of 50% was achieved for the position 8 of R/G-gRNA and the editing yields from positions 5 to 7 are declining steadily.

4.2.3.2 Revisiting the effect of the 3'-terminal boxB-motif

The positioning experiment was performed with R/G-gRNAs containing a boxB-motif at the 3'-end. In order to analyze the effect of the 3'-terminal boxB-motif to the editing yield, the R/G-gRNA P6 and P8 were cloned into p*Silencer* vectors lacking the boxB-motif (3.2.14). As mentioned in chapter 4.2.2.6, the boxB-motif was added to the 3'-end of a R/G-gRNA with the aim to increase the guideRNA's stability and, therefore, enhancing the availability of the R/G-gRNA in the cell. The following experiment was supposed to confirm or to disprove the observations that a boxB-motif added to the R/G-gRNA P6 is rather disadvantageous, whereas the presence of the boxB-motif to the R/G-gRNA P8 has no drastical effect on the editing yield.

The experiment of the luciferase assay was performed with a co-transfection of 60 ng W417X luciferase and ADAR2 plasmid, as well as 250 ng of each tested R/G-gRNA plasmid in a 96-well format. The following R/G-gRNA plasmids were used for this purpose: P6+boxB-motif (pTS121), P6-boxB-motif (pTS194), P8+boxB-motif (pTS170) and P8-boxB-motif (pTS195). The experiment for the RNA isolation was performed with a transfection of 300 ng W417X luciferase and ADAR2 plasmid, as well as 1300 ng of each tested R/G-gRNA plasmid in a 24-well format. Two independent luciferase assays (separate transfection and assay) and the RNA isolation were performed 48 hours after transfection (Figure 4-87).

In both luciferase assay experiments, a higher editing efficiency (equated relatively from the RLU) was obtained without a boxB-motif added to the R/G-gRNA P6. The editing yields of the R/G-gRNA P6 are markedly lower than that of the R/G-gRNA P8 in both experiments, confirming again the results of the positioning experiment. The obtained luminescence values for the transfection of R/G-gRNA P8 with and without boxB-motif, are demonstrating that the effect of the boxB-motif attached to the R/G-gRNA P8 is less distinct than for position 6: The first experiment resulted in lower editing yields for the guideRNA with a boxB-motif, indicated by a lower luciferase activity and RLU values, whereas the second experiment obtained slightly higher editing levels for the addition of the boxB-motif.

The sequence results confirmed the trend that was obtained from the luciferase assay: The absence of the boxB-motif at the 3'-end of the guideRNA P6 results in higher editing yields, whereas this effect is negligible at position 8 (Figure 4-88).

The result that the W417X luciferase R/G-gRNA P6 benefits from the absence of the boxB-motif to achieve higher editing yields is equivalent to the results of the W58X eGFP mRNA investigation. We could also confirm that the boxB-motif at a R/G-gRNA P8 for editing W417X luciferase or W58X eGFP is less influencing on the editing yields (4.2.2.6). Obviously, there is a steric difference between the two positions of the R/G-gRNA owning a boxB-motif at their 3'-end.

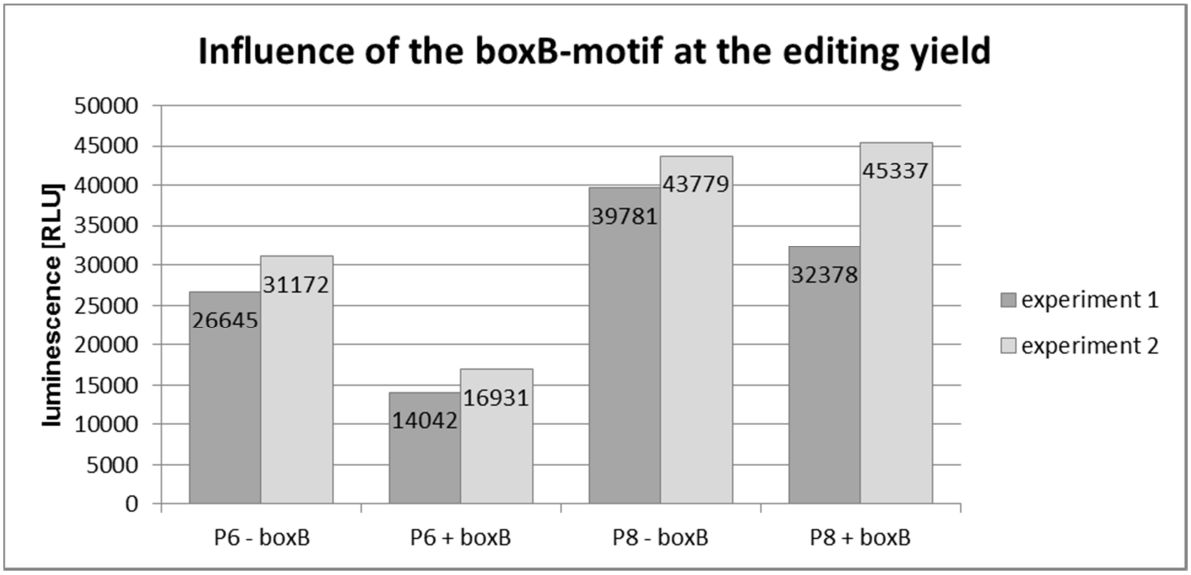

Figure 4-87: Influence of the boxB-motif at the editing yield of W417X luciferase mRNA. The following plasmids were co-transfected in a 96-well format: 60 ng W417X luciferase, 60 ng ADAR2 and 250 ng of the tested R/G-gRNA. The R/G-gRNAs P6 and P8 were selected for the analysis of the influence of the boxB-motif attached to the 3'-end of the guideRNA to the editing yield. In experiment 1 the R/G-gRNA lacking the boxB-motif achieved the highest luminescence values. In the second experiment the addition of the boxB-motif to the R/G-gRNA P6 decreased the luminescence value, whereas the addition of the boxB-motif to the R/G-gRNA P8 seemed to slightly increase it.

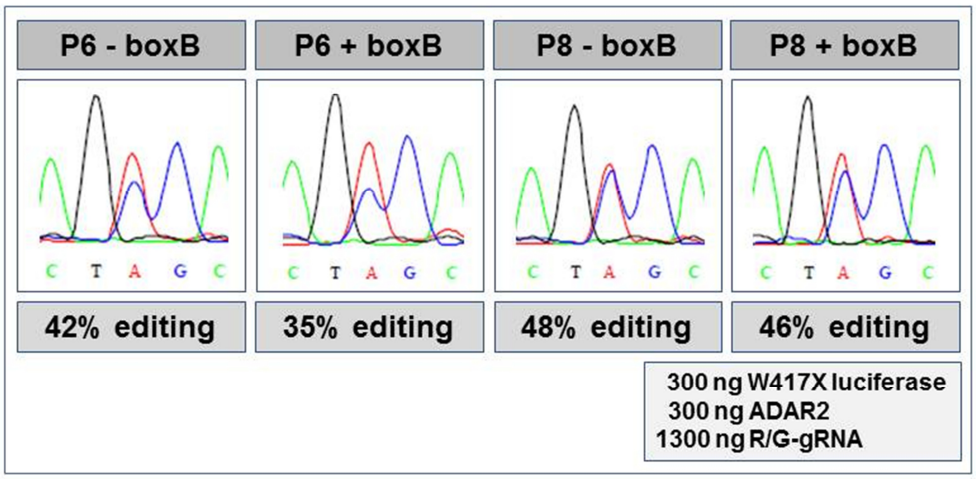

Figure 4-88: Influence of the boxB-motif attached to the R/G-gRNA P6 and P8 at the editing yield of W417X luciferase mRNA. RNA was isolated from cells of a 24-well format, which were transfected with 300 ng W417X luciferase, 300 ng of ADAR2 plasmid and 1300 ng of R/G-gRNA. The R/G-gRNAs P6 and P8 with and without the attached boxB-motif to the 3'-end of the guideRNA were tested. Both chosen positions of R/G-gRNA achieved without the boxB-motif higher editing yields. The highest editing yield of 48% was achieved for the R/G-gRNA P8 without boxB-motif.

4.2.3.3 Revisiting the effect of variable transfected amounts of R/G-gRNA and ADAR2 plasmid

Again, we wanted to analyze the optimal combination of the amount of R/G-gRNA plasmid and ADAR2 plasmid for editing a constant amount of W417X luciferase plasmid. Since the editing yields of W417X luciferase differ only slightly for the presence or absence of a boxB-motif at position 8, the following experiment was performed with a boxB-motif at the 3'-end of the R/G-gRNA P8. The luciferase assay was performed with cells, which were transfected with 60 ng W417X luciferase plasmid and variable amounts of the R/G-gRNA P8 (pTS170) and ADAR2 plasmid in a 96-well format. The RNA was isolated from cells (24-well format), which were transfected with 300 ng W417X luciferase plasmid and variable amounts of R/G-gRNA P8 and ADAR2 plasmid. Both methods were applied 48 hours post transfection.

The data of the luciferase assay demonstrated that at higher amounts of transfected ADAR2 plasmid with a fixed amount of 250 ng R/G-gRNA P8, the luminescence steadily decreased, indicating that the editing efficiency was decreasing with higher ADAR2 protein concentrations (Figure 4-89). The same trend was obtained for two different amounts of R/G-gRNA P8 plasmid: at a constant transfected amount of

200 ng or 125 ng R/G-gRNA plasmid, the double amount of ADAR2 plasmid (120 ng instead of 60 ng) decreased the editing yield. Interestingly, the editing efficiencies seem to be increased by transfecting lower amounts of R/G-gRNA plasmid than 250 ng as shown by higher luminescence. This effect is opposite to the microscopy analysis of the W58X eGFP editing, where it was shown that higher amounts of transfected R/G-gRNA are beneficial to increase the fluorescence signal and thus the editing yields (4.2.2 and 4.2.2.5). This trend that higher amounts of transfected R/G-gRNA increase the editing level was obtained in other luciferase assays as well, but not that reliably (data not shown).

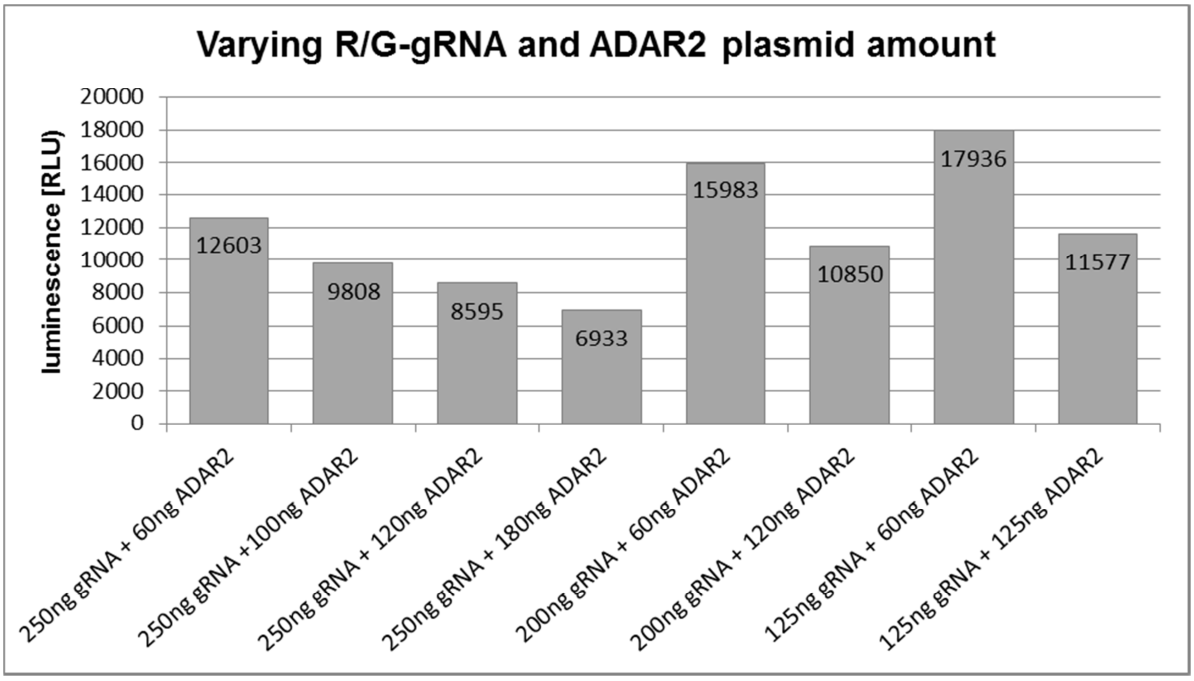

Figure 4-89: Varying R/G-gRNA and ADAR2 plasmid amount to edit W417X luciferase mRNA. The luciferase assay was performed with cells that were transfected in a 96-well format with the following plasmids: 60 ng W417X luciferase and variable amounts of R/G-gRNA P8 and ADAR2. The luciferase assay was performed 48 hours post transfection. The samples with a fixed amount of 250 ng of R/G-gRNA P8 plasmid and an increasing ADAR2 plasmid amount yielded in declining RLU values. The same effect was obtained for lower R/G-gRNA plasmid amounts of 200 ng and 125 ng. The highest RLU value was achieved for the sample transfecting 125 ng R/G-gRNA plasmid and 60 ng of ADAR2 plasmid.

The results of the RNA isolation of the editing of W417X luciferase mRNA in a 24-well format confirmed the findings of the W58X eGFP mRNA editing experiment: the editing yields increase from 38% to 48% by increasing amounts from 650 ng up to 1300 ng R/G-gRNA P8 plasmid at constantly 300 ng ADAR2 plasmid, demonstrating

that the transfection of more R/G-gRNA plasmid is beneficial to increase the total editing yield (Figure 4-90). But the increase of the transfected ADAR2 plasmid amount of the sample (600 ng – 900 ng) does not change the total editing yield of the target adenosine. The increase of the ADAR2 plasmid amount higher than 300 ng was not performed in the context of the W58X eGFP mRNA editing and can, therefore, not be compared. But this experiment leads to the impression that a minimum concentration of the ADAR2 protein is necessary to achieve high editing yields and a further increase of the ADAR2 plasmid amount is not improving the editing yield in the 24-well format. This assumption would fit to the observations that have been made in the editing experiment of the W58X eGFP mRNA with decreasing amounts of ADAR2 plasmid (4.2.2.8). Until 50 ng of ADAR2 plasmid, there is no distinct effect at the editing yield visible. Firstly, the decrease of the ADAR2 plasmid amount from 50 ng down to 25 ng in a 24-well format showed distinct changes in the editing yield (47% to 36%).

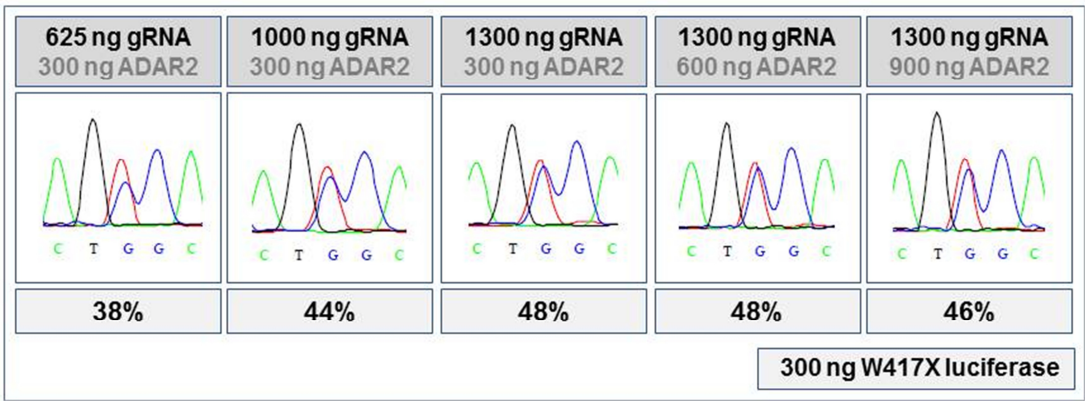

Figure 4-90: Varying amounts of ADAR2 and R/G-gRNA plasmids to edit W417 luciferase mRNA. Different combinations of the amounts of ADAR2 and R/G-gRNA P8 plasmid were transfected in 24-well plate format together with a constant amount of 300 ng W417X luciferase plasmid. The RNA isolation was performed 48 hours post transfection. Higher amounts of the R/G-gRNA plasmid together with 300 ng of ADAR2 plasmid resulted in higher editing levels. The transfection of higher ADAR2 plasmid amounts (double and three times) resulted in the same or slightly lowered editing yield.

But one point has not been taken into account for explaining discrepancy of the RNA sequencing versus luciferase assay in determining the editing efficiency at the target adenosine: the luciferase assay reflects the activity of the whole functional luciferase protein. By off-target editing, a change of the amino acid composition of the luciferase protein could possibly lead to an inactive protein. Therefore, the editing efficiency could potentially be misinterpreted when being judged only from the luminescence

values, disregarding the potential off-target editing in the luciferase mRNA. Off-target editing was already displayed for the editing of W58X eGFP mRNA, demonstrating a clear correlation to high ADAR2 concentrations. In order to prove this assumption, the luciferase transcript was analyzed for off-target editing. Indeed, two off-target sites, one at 291 (L) and second at the amino acid position 511 (K) were found. The amino acid leucine of the 5'-TAT codon at position 291 is unchanged by off-target editing. Whereas editing of the 5'-AAG codon at position 511 to 5'-AIG changes the amino acid lysine to arginine. Position 511 was edited at 7 % by 600 ng of ADAR2 plasmid and 9 % by 900 ng of ADAR2 (Figure 4-91). The transfection of 300 ng ADAR2 led in comparison to none editing at position 511. It is possible that this off-target editing influences the luciferase protein activity, resulting in decreased luminescence signals for higher ADAR2 plasmid amounts. We proved this assumption by the transfection of a plasmid coding for the mutant luciferase protein K511R and the comparison of the obtained luminescence signals with the wt luciferase. No markedly difference in the RLU values and protein activity was detectable.

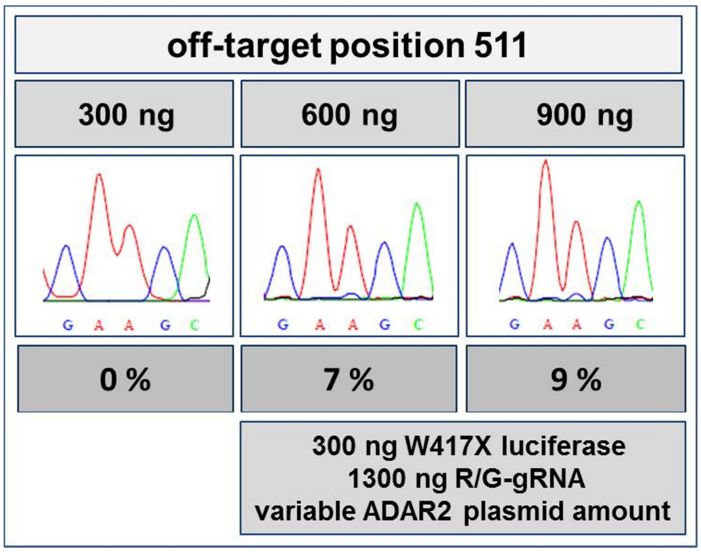

Figure 4-91: Off-target editing at position 511 of W417X luciferase mRNA. RNA was isolated from cells of a 24-well format, which were transfected with 300 ng W417X luciferase, 1300 ng of R/G-gRNA and a variable amount of ADAR2 plasmid. Higher amounts of transfected ADAR2 plasmid than 300 ng led to off-target editing at position 511 of W417X luciferase mRNA.

In our hands the luciferase assay showed a lot of variations in the same experimental setup. Thus, this method seems to be less reliable than the RNA isolation method and provides only trends with respect to editing efficiency. However, marked differences can be visualized using the luciferase assay but for reliable and distinct

information, the RNA isolation is the method of choice to quantify a single editing event.

In summary, the ability of the editing system to address W417X luciferase mRNA for site-specific RNA editing was proven. A maximum editing yield of 50% was achieved by transfecting 1300 ng amount of R/G-gRNA P8 plasmid. The analysis of the optimal position (P5 - P9) to edit W417X luciferase mRNA resulted in position 8 to be the most suitable position. This is in agreement with our research from the W58X eGFP editing. As already mentioned in chapter 4.2.2.6 for the analysis of the presence of the boxB-motif to edit W58X eGFP mRNA, the addition of a boxB-motif is not beneficial to increase the editing yield. The transfection of higher amounts of ADAR2 plasmid did not result in higher editing yields. Instead, it resulted in higher off-target editing, as demonstrated by the sequence analysis. The effect that a lower amount of transfected R/G-gRNA plasmid results in less editing yield was also confirmed for the sequence analysis.

4.2.3.4 Editing of W58X eGFP and W417X luciferase in parallel

In the previous chapter, it was shown that editing of the W417X luciferase mRNA in cell culture is possible at around 50%. In order to challenge the editing system further, an experiment tested if the parallel editing of W58X eGFP and W417X luciferase mRNA. If editing of two different target mRNAs is feasible in parallel at high editing yields, this offers the editing system an advantage over the CRISPR/Cas9 system for application: The CRISPR/Cas9 genome editing system has problems to address two or multiple targets in parallel, in a controllable way. This might be an advantage if multiple RNAs should to be targeted without potential germline intervention.

The transfection experiment of the luciferase assay was performed in a 96-well format always using 60 ng of W417X luciferase, W58X eGFP and ADAR2 plasmid together with the following combination of R/G-gRNA plasmids: 1) 250 ng of each R/G-gRNA (W58X eGFP gRNA - pTS186 and W417X luciferase gRNA - pTS170), 2) 200 ng of each R/G-gRNA, 3) only 250 ng of W58X eGFP R/G-gRNA and 4) only 250 ng of W417X luciferase R/G-gRNA. Since the luciferase assay will only provide information about the relative editing yield of the luciferase transcript, the editing yields of the eGFP transcript are determined by microscopy analysis. Both transcripts

are analyzed by RNA isolation followed by sequence analysis. The transfection experiment for microscopy analysis and RNA isolation was carried out in 24-well format using the same experimental design as for the luciferase assay, increasing only the amounts of the transfected plasmid components as follows: 300 ng of W417X luciferase, W58X eGFP and ADAR2 plasmid and 1300 ng or 1000 ng of each R/G-gRNA plasmid. The luciferase assay and the RNA isolation were applied 48 hours post transfection.

The results of the luciferase assay of four transfection samples performed by two separate transfection and luciferase assay experiments are given in Figure 4-92. If W417X luciferase gRNA is missing in the transfection samples, only background luminescence is obtained. This demonstrates the necessity for the matching guideRNA to address editing. The highest obtained luciferase activity was achieved for the sample transfecting W417X luciferase gRNA without W58X eGFP gRNA. This leads to the assumption that the single editing reaction results in higher editing yield than the parallel editing of two target mRNAs. Again, one has to note that the replicates of the luciferase assay do not agree in their ranking whether a higher or lower amount of both transfected R/G-gRNAs increase or lower the editing reaction of the luciferase transcript.

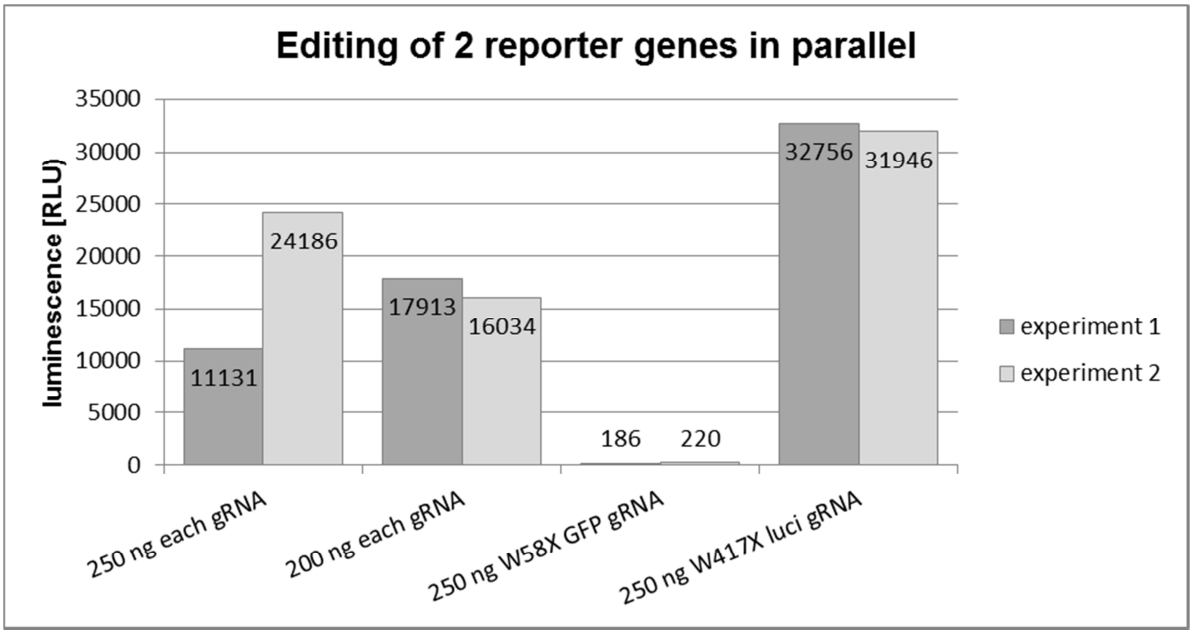

Figure 4-92: Luciferase assay of editing two reporter genes in parallel. The transfection experiment was performed in a 96-well format always using 60 ng of W417X luciferase, W58X eGFP and ADAR2 plasmid in combination with the following amounts of R/G-gRNA plasmids: 1) 250 ng of

each R/G-gRNA (W417X luciferase and W58X eGFP), 2) 200 ng of each R/G-gRNA, 3) only 250 ng of W58X eGFP R/G gRNA and 4) only 250 ng of W417X luciferase R/G-gRNA. The luciferase assay was performed 48 hours post transfection. Both independently performed experiments result for the sample lacking the W417X luciferase R/G-gRNA in a background luminescence, demonstrating that without the W417X luciferase gRNA no editing is performed. The sample transfecting W417X luciferase R/G-gRNA without the W58X eGFP R/G-gRNA results for both experiments in the highest obtained RLU, assuming that the highest editing yields for W417X luciferase are obtained with the absence of the second guideRNA. The combination of both transfected R/G-gRNAs decreases the RLU drastically.

The microscopy analyses only the eGFP expression and not the luciferase expression. Therefore, a further evaluation of the editing yield of the luciferase transcript can only be made on the basis of the sequence analysis. The microscopy pictures showed eGFP signal for each sample transfected with W58X eGFP gRNA, but no fluorescent signal was visible in absence of the W58X eGFP gRNA (Figure 4-93). The absence of the W58X eGFP gRNA resulted in no editing of W58X eGFP mRNA, as confirmed by the sequencing. The same effect was obtained for the sequence results of the luciferase transcript: Only the presence of the luciferase R/G-gRNA achieved editing of the luciferase transcript. The editing yields as determined by sequencing of the parallel editing reaction, were unchanged compared to the single editing reactions: The single editing sample of W58X eGFP mRNA with 1000 ng of R/G-gRNA plasmid resulted in 41% editing yield, whereas the parallel editing sample achieved 42% at the same amount of 1000 ng R/G-gRNA plasmid. The single editing experiment of the luciferase transcript achieved 39% (transfecting 1000 ng R/G-gRNA), whereas the parallel editing resulted in 35% editing yield. This stays in contrast to the luciferase assay data; there was a higher editing yield obtained by the drop out of the W58X eGFP gRNA leading to distinct higher luminescence. Taken the sequence results into account, that there are hardly differences in the editing yields for the parallel editing versus single editing setting, the following assumption for the parallel editing situation raise up: the parallel editing reaction creates a competition for the ADAR2 protein, so that a single editing situation like for the luciferase transcript is quicker, resulting in a faster protein translation and, therefore, higher luminescence. After 48 hours of incubation time this competing effect of the parallel editing seems to be negligible on the RNA level, but the protein level still represents the delayed editing reaction of the luciferase transcript.

The editing yields obtained for the W58X eGFP transcript were always slightly higher than for the luciferase transcript. This may be due to the specific sequence context or to the different secondary structure formation (accessibility) of both mRNA transcripts. Interestingly, the off-target editing that arises after 48 hours of incubation time for the W58X eGFP mRNA, which was observed and discussed in chapter 4.2.2.1, disappeared in the parallel editing reaction (Figure 4-94). This suggests that ADAR2 edits other adenosines, if its cellular concentration is very high and if no other substrates are available. Therefore, the amount of 300 ng ADAR2 plasmid for a single editing setting is too high, and provokes off-target editing. This is in agreement with the observation that the reduction of transfected ADAR2 plasmid quickly reduced and finally completely prevented off-target editing in the W58X eGFP transcript, whereas high editing yields at the target position were maintained (4.2.2.8).

It could be demonstrated that the parallel editing of two different mRNA transcripts is possible at high editing yields. The editing yields of both transcripts do not suffer from the parallel editing. However, the cells were strongly suffering by the immense amount of transfection agent, which was necessary to transfect all five plasmids at the desired amounts. Further investigations are planned to analyze a possible delaying effect by the parallel editing of two transcripts.

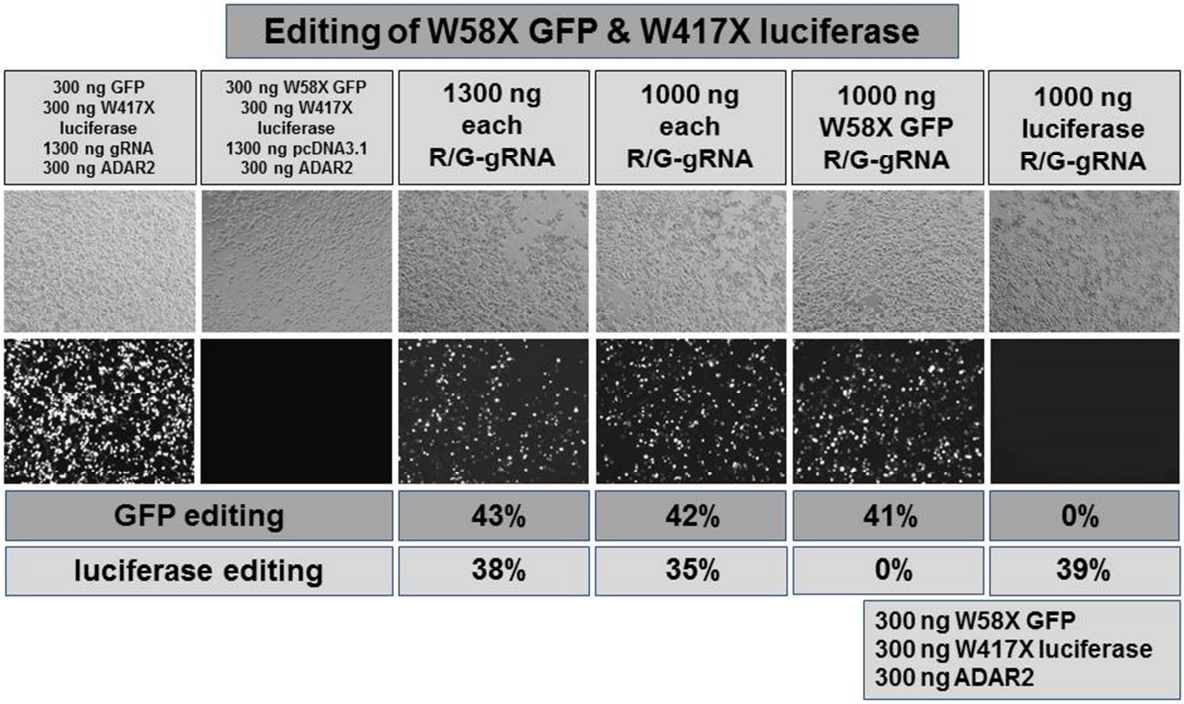

Figure 4-93: Microscopy and sequence analysis of the parallel editing of W58X eGFP and W417X luciferase mRNA. The transfection experiment was performed in a 24-well format always

using 300 ng of W417X luciferase, W58X eGFP and ADAR2 plasmid in combination with the following amounts of R/G-gRNA plasmids: 1) 1300 ng of each R/G-gRNA (W417X luciferase and W58X eGFP), 2) 1000 ng of each R/G-gRNA, 3) only 1000 ng of W58X eGFP R/G gRNA and 4) only 1000 ng of W417X luciferase R/G-gRNA. The microscopy analysis and RNA isolation were performed 48 hours post transfection. The microscopy analysis visualizes the presence of the eGFP expression for each sample transfecting the W58X eGFP R/G-gRNA. The absence of the W58X eGFP R/G-gRNA or luciferase R/G-gRNA results always in no editing of the corresponding mRNA transcript. Only the presence of the appropriate R/G-gRNA achieves editing of the respective mRNA transcripts. The editing yields of the parallel editing reaction are negligible lower than for the single editing reaction.

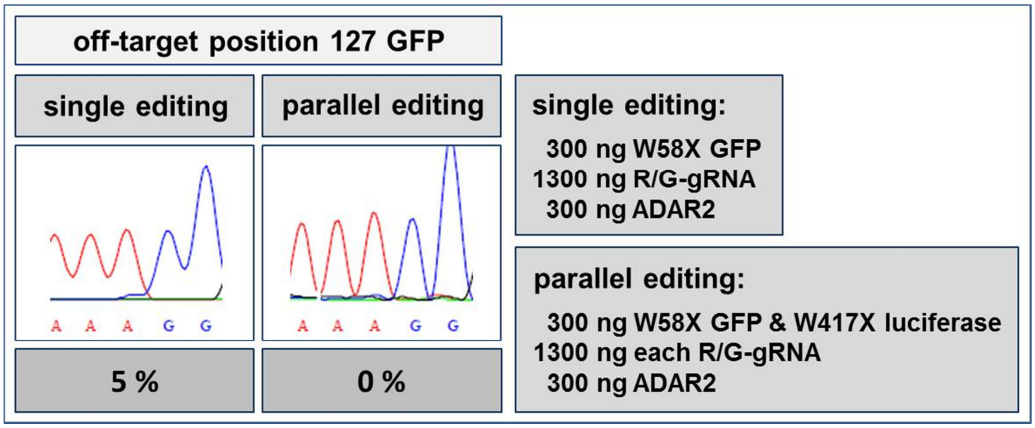

Figure 4-94: Off-target editing at position 127 of W58X eGFP mRNA in the single and parallel editing reaction. The RNA was isolated 48 hours post transfection from cells of a 24-well format. In the single editing reaction, the off-target position 127 of W58X eGFP mRNA is edited up to 5 %. Off-target editing at position 127 of W58X eGFP mRNA is not present at the parallel editing situation.

4.2.4 Off-target editing in cell culture

With the genetically encodable editing system off-target editing was demonstrated to happen in vitro, as well as in cell culture. There are two different modes of off-target editing to be found in cell culture, as we will discuss here for the off-target editing of eGFP. First off-target editing can happen at any positions of the mRNA transcript, independent from the presence of the R/G-gRNA, and was typically observed after 48 hours, but not before (4.2.2.2). However, the more active mutant E488Q edits the off-target position 127 of W58X eGFP mRNA already after 24 hours at 25% (4.2.2.2), and at 37% after 48 hours. For ADAR2, the prolongation of the incubation time increased the editing yield at the off-target position 127 up to 13% (72 hours / 96 hours), and also created a second off-target site at position 102 (11%) (4.2.2.7).

This mode of off-target editing can be influenced by the level of ADAR2 expression. The decrease of transfected ADAR2 plasmid strongly lowered the editing yield at the

off-target site 127 of eGFP and even prevented it, while maintaining high editing yields at the target position (4.2.2.8). This indicates that the off-target editing can be drastically reduced and even prevented by a tight control of the ADAR2 expression. This hypothesis is further strengthened by the observation that off-target editing is abolished during parallel editing in cell culture, when applying two editing substrates and their guideRNAs (4.2.3.4). This led to the assumption that transfection of 300 ng ADAR2 plasmid in a 24-well format provoked off-target editing by underemployment of ADAR2 capacity. Therefore, low ADAR2 plasmid amounts are recommended to prevent off-target editing.

The second mode of off-target editing is created by the mRNA/guideRNA duplex itself. As demonstrated in chapter 4.2.2.3, the prolongation of the mRNA binding site resulted in off-target editing of an adenosine that was now included in the mRNA/guideRNA duplex. Thinking of adenosine-rich sequence context, the following options can be pursued to prevent other adenosines in the RNA duplex from being edited:

Either, the adenosine, which should be protected from editing, should be positioned to the border of the RNA duplex. This strategy was demonstrated to prevent over editing in chapter 4.2.2.3 (57). Or the adenosine to be protected from over editing can be mispaired with guanosine (for details see Figure 2-11 and (49).

An important issue is the analysis of transcriptome wide off-target editing and how it could be controlled. Until now, we are overexpressing ADAR2 in cell culture and in literature this was demonstrated to change the natural editing of the endogenous targets (99). Also, our future aim is to steer endogenous ADAR for site-directed editing and therefore, it is too early to address the question for transcriptome wide off-target editing.

4.2.5 P53 gene

In the chapter of in vitro editing, it was demonstrated that the editing system has the potential and ability to edit less preferred codons like 5'-GAG and 5'-CAC (4.1.7). In order to challenge the editing system further, a new mRNA transcript – the tumor suppressor gene p53 (2.7.1) – was chosen to address less preferred 5'-CAC and 5'-CAG. The ability to restore p53 activity by editing mutated p53 transcripts could represent a further possibility in anti-tumor therapy. The R175H and R282Q

mutations of p53, which affect the DNA binding and sequence specificity of p53, were chosen for the cell culture experiment for representing two of six major hot spots of mutated p53 genes in cancer and are coding for a 5'-CAC and 5'-CAG triplet (66).
The experimental design to analyze the ability of the encodable editing system to edit R175H p53 / R282Q p53 in cell culture was planned as follows: Transient transfection of the R175H p53 plasmid (pTS217) or R282Q p53 plasmid (pTS269.4) together with the editing machinery (ADAR2 and R/G-gRNA) in a 24-well format. After 48 hours follows RNA isolation and sequence analysis of the p53 transcript. In consideration of the results of the previous editing studies in cell culture and in vitro, a R/G-gRNA of 16 nt length that places the target adenosine at position 8 with an 5'-GCG or 5'-GUA anticodon to target R175H p53 and a 5'-CCG anticodon to target R282Q p53 were chosen (Figure 4-95). All R/G-gRNAs were cloned into p*Silencer* together with an attached boxB-motif at the 3'-end resulting in pTS181 (5'-GCG), pTS182 (5'-GUA) and without a boxB-motif for R282Q gRNA pTS241 (5'-CCG). The details of the cloning strategy are given in 3.2.14.

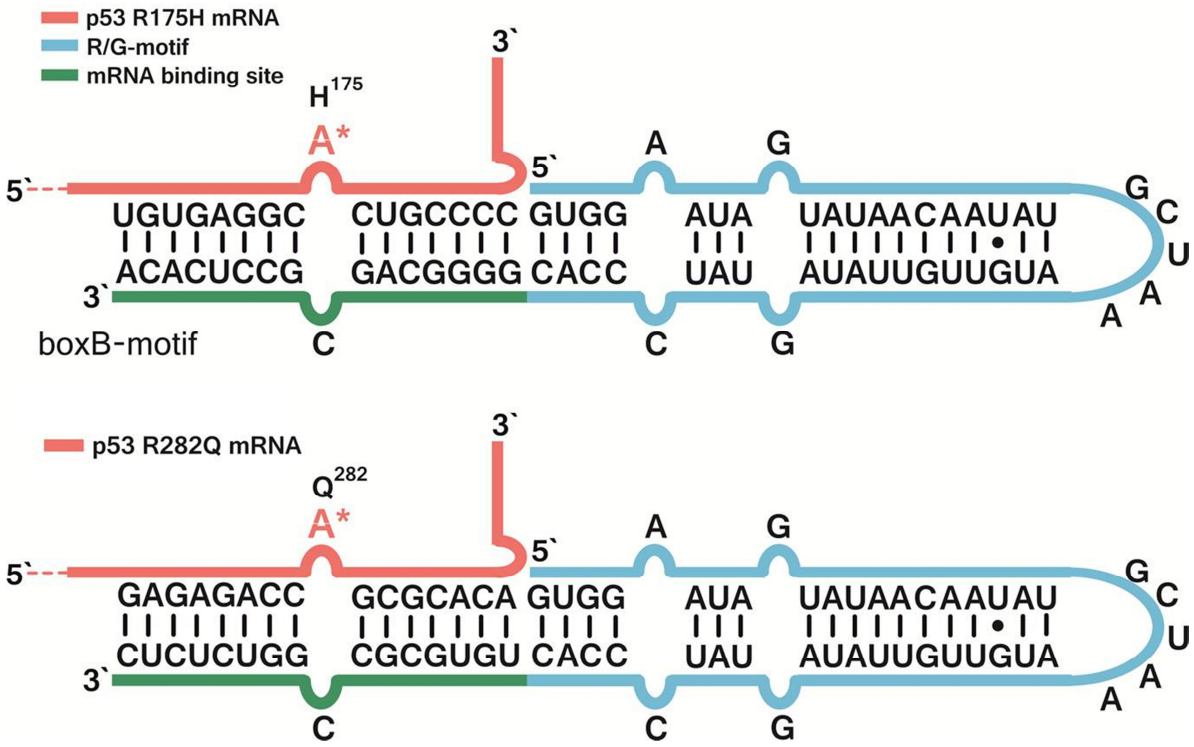

Figure 4-95: Design of the R/G-gRNA for editing R175H p53 and R282Q p53. The R175H p53 mRNA (red) is hybridized with the 16 nt long mRNA binding site (green) of the R/G-gRNA (blue). The target adenosine is placed at position 8 and hybridized by a 5'-GCG anticodon. A boxB-motif is attached to the R175H R/G-gRNA. The R282Q p53 5'-CAG codon is paired with a 5'-CCG anticodon.

The transfection experiment to address R175H p53 was performed with two different ADAR2 proteins, the wt ADAR2 (pTS57) and the mutant E488Q ADAR2 (pTS90), since the E488Q ADAR2 showed in general higher editing yields for all tested codons than wt ADAR2 in vitro (4.1.8) and was expected to show the same trend in cell culture. HEK293T cells were transfected with the following amounts of plasmids: 300 ng R175H p53, 300 ng ADAR2 and 1600 ng R/G-gRNA. The result of the sequence analysis (primer no. 586) is demonstrated in Figure 4-96 and shows that the editing yields of all four transfection samples vary only slightly. The highest editing yield obtained was 24% for the transfection sample with R/G-gRNA 5'-GCG and E488Q. As expected the p53 transcript of the E488Q transfection samples shows a few off-target sites, for example position 62 is edited up to 40% (Figure 4-97), whereas wt ADAR2 showed none. The expectation that this mutant achieves distinctly higher editing yields for the 5'-CAC codon than the wt ADAR2 is not the case. The same effect in cell culture that the wt ADAR2 achieves higher editing yields than the E488Q mutant was observed for the 5'-UAG codon of the W58X eGFP transcript, too (4.2.2.2).

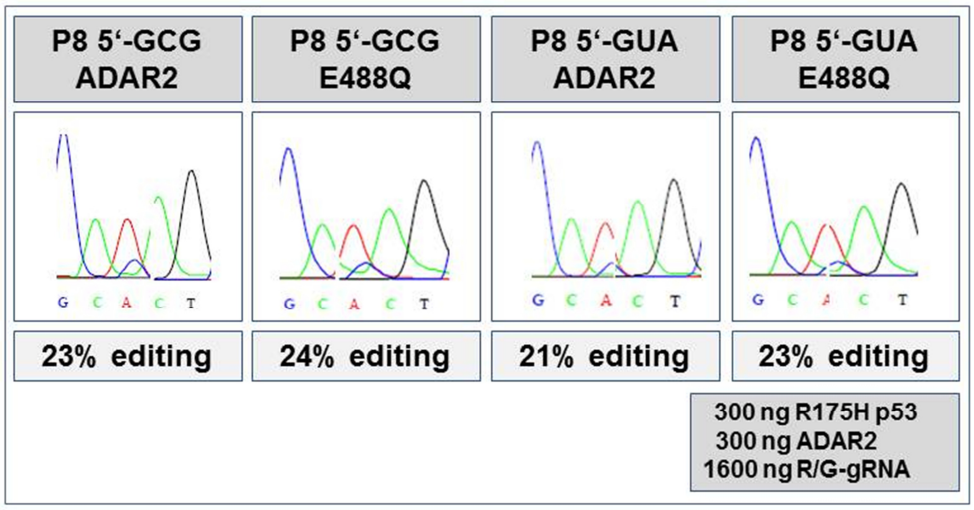

Figure 4-96: Editing efficiency of R175H p53 mRNA by transiently transfection in cell culture. The transfection experiment was performed with 300 ng R175H p53 plasmid (pTS217), 300 ng of ADAR2 (pTS57) or E488Q (pTS90) and 1600 ng of R/G-gRNA 5'-GCG anticodon (pTS181) or 5'GUA anticodon (pTS182) in a 24-well format. The RNA was isolated 48 hours post transfection followed by sequence analysis. The editing yields vary only slightly between the different chosen R/G-gRNAs and ADAR2 proteins. The maximum editing yield of 24% was achieved with a R/G-gRNA 5'-GCG and E488Q.

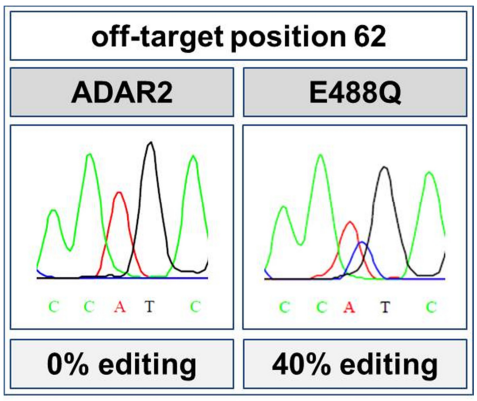

Figure 4-97: Off-target editing of wt ADAR2 and E488Q at position 62 of R175H p53 mRNA. The transfection experiment was performed with 300 ng R175H p53 plasmid (pTS217), 300 ng of ADAR2 (pTS57) or E488Q (pTS90) and 1600 ng of R/G-gRNA 5'-GCG anticodon (pTS181) in a 24-well format. The RNA was isolated 48 hours post transfection followed by sequence analysis. E488Q edited the off-target position 62 up to 40%, whereas the editing reaction with ADAR2 showed no off-target editing at p53 mRNA.

Editing of R282Q p53 in cell culture was also performed in a 24-well format, transfecting 300 ng of R282Q p53 plasmid (pTS269.4), 1300 ng of R/G-gRNA 5'-CCG (pTS241) and 300 ng of ADAR2. The editing resulted in 37% after 48 hours of incubation (Figure 4-98). As observed in the editing experiment of R175H p53 no over editing for editing the R282Q p53 transcript with ADAR2 was visible.

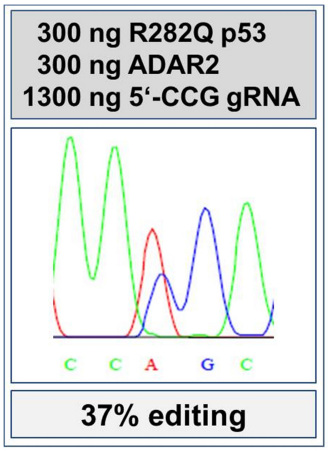

Figure 4-98: Editing efficiency of R282Q p53 mRNA by transiently transfection in cell culture. The transfection experiment was performed with 300 ng R282Q p53 plasmid (pTS269.4), 300 ng of ADAR2 (pTS57) and 1300 ng of R/G-gRNA 5'-CCG anticodon (pTS241) in a 24-well format. The RNA was isolated 48 hours post transfection followed by sequence analysis. An editing yield of 37% was achieved with a R/G-gRNA 5'-CCG and ADAR2.

Obviously, the challenge to address the p53 transcript with a less preferred 5'-CAC or 5'-CAG codon was successful, but the aim to achieve comparably high editing yields as obtained for the 5'-UAG codon of W58X eGFP or W417X luciferase is not reached yet with the first experimental trials. This set of data only highlights the starting point for further experiments to improve the ability of less preferred codons in cell culture.

4.2.6 The phenomenon of enhanced eGFP signals

During the various experiments when optimizing the editing efficiency of W58X eGFP mRNA, the phenomenon was repeatedly observed that the fluorescence signal of positive controls was enhanced, depending on the transfection of the editing components. Whereas a repression of eGFP expression from an antisense effect of the editing machinery would have made sense, the observed activation or derepression of eGFP expression was entirely unexpected. An instructive transfection experiment was performed in a 24-well format containing the different compositions of plasmids listed in Table 4-5 and was analyzed 48 hours post transfection by fluorescence microscopy.

Table 4-5: Transfection samples and composition of different positive controls.

	Sample and plasmid composition
1)	300 ng GFP (pTS57) + 1900 ng empty pcDNA3.1 vector (pTS28)
2)	300 ng GFP (pTS57) + 300 ng ADAR2 (pTS57) + 1600 ng empty pcDNA3.1 vector (pTS28)
3)	300 ng GFP (pTS57) + 1300 ng R/G-gRNA P6 (pTS68) + 300 ng empty pcDNA3.1 vector (pTS28)
4)	300 ng GFP (pTS57) + 1300 ng R/G-gRNA P6 (pTS68) + 300 ng ADAR2 (pTS57)

To disentangle the effect of the individual components of the editing machinery, we performed various transfections. The co-transfection with ADAR2 plasmid alone increased the number of eGFP expressing cells and the strength of the eGFP signal slightly (Figure 4-99). An even higher fluorescence signal was observed if the ADAR2 plasmid was replaced by the R/G-gRNA plasmid. The strongest fluorescence signal was demonstrated by the sample transfecting eGFP plasmid together with R/G-gRNA plasmid and ADAR2. This leads to the following ranking of the individual components with respect to their ability to affect eGFP expression:
ADAR2< R/G-gRNA< ADAR2 & R/G-gRNA.

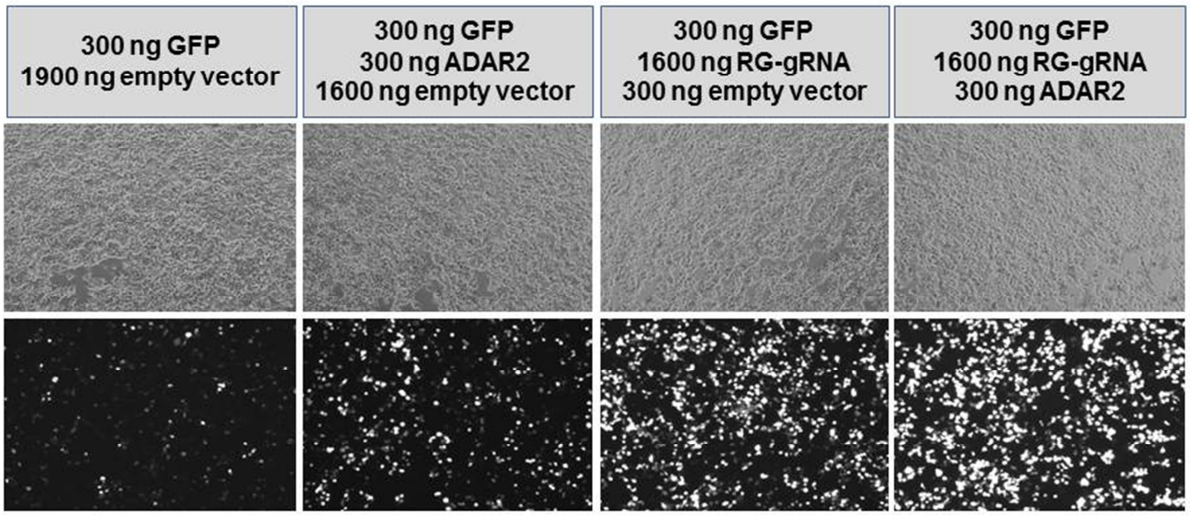

Figure 4-99: Enhancing fluorescent signal of different transfected positive controls. Different compositions of plasmids were transfected in a 24-well format. The microscopy analysis was performed 48 hours after transfection. The transfection of eGFP plasmid together with ADAR2 plasmid increases slightly the amount and strength of the eGFP signal. A distinct stronger fluorescence signal was obtained for the replacement of ADAR2 by the R/G-gRNA plasmid. The strongest and biggest amount of fluorescing cells was obtained for the combination of ADAR2 and R/G-gRNA plasmid. Magnification: 100x, GFP exposure time: 50 ms.

This enhancing effect was so unordinary and unexpected that we reproduced and further analyzed the effect, to find out which component is the reason of this enhanced eGFP signal. Table 4-6 lists the transfection experiments that were performed in a 96-well format together with Pia Mach during her bachelor thesis:

Table 4-6: Transfection samples and composition to analyze the enhancing eGFP signal effect.

	Sample and plasmid composition
1)	60 ng GFP (pTS57) + 60 ng empty pcDNA3.1 vector (pTS28)
2)	60 ng GFP (pTS57) + 60 ng ADAR2 (pTS57)
3)	60 ng GFP (pTS57) + 250 ng W58X eGFP R/G-gRNA (pTS68)
4)	60 ng GFP (pTS57) + 250 ng W417X luciferase R/G-gRNA (pTS121)
5)	60 ng GFP (pTS57) + 250 ng W437X PINK R/G-gRNA (pTS67)
6)	60 ng GFP (pTS57) + 250 ng W58X eGFP R/G-gRNA (pTS68) + 60 ng ADAR2 (pTS57)
7)	60 ng GFP (pTS57) + 250 ng W417X luciferase R/G-gRNA (pTS121) + 60 ng ADAR2 (pTS57)
8)	60 ng GFP (pTS57) + 250 ng W437X PINK R/G-gRNA (pTS67) + 60 ng ADAR2 (pTS57)
9)	60 ng GFP (pTS57) + 250 ng W58X eGFP R/G-gRNA (pTS68) + 60 ng E396A ADAR2 (pTS63.1)

The overview of the microscopy analysis demonstrates that the enhancement of the eGFP signal was only visible for specific R/G-gRNAs (Figure 4-100). This includes

the gRNAs of W58X eGFP and W417X luciferase but not of W437X PINK, demonstrating that the choice of the R/G-gRNA has a distinct contribution to the increasing fluorescence signal. This enhanced fluorescence signal is further increased for the combination of R/G-gRNA and ADAR2 protein, as previously seen and is again restricted only to the W58X eGFP and W417X luciferase gRNA. Even for the PINK gRNA is an increment visible for its combined transfection with the ADAR2 plasmid, indicating that both components of the editing system (R/G-gRNA and ADAR2) contribute to this enhancing effect. A further proof that ADAR2 contributes to the enhancement demonstrates the transfection of only eGFP and ADAR2 plasmid, resulting in an increased eGFP signal compared to the transfection of eGFP plasmid alone. But this increment is less strong than for the sample transfecting eGFP and W58X eGFP R/G-gRNA plasmid. The combination of both editing components – ADAR2 and the R/G-gRNA – leads again to the strongest obtained eGFP signal. Interestingly, the replacement of the ADAR2 protein by the not enzymatically active mutant E396A ADAR2 does not reduce this enhancing effect of the eGFP signal, indicating that this effect is not based on the editing activity of the enzyme.

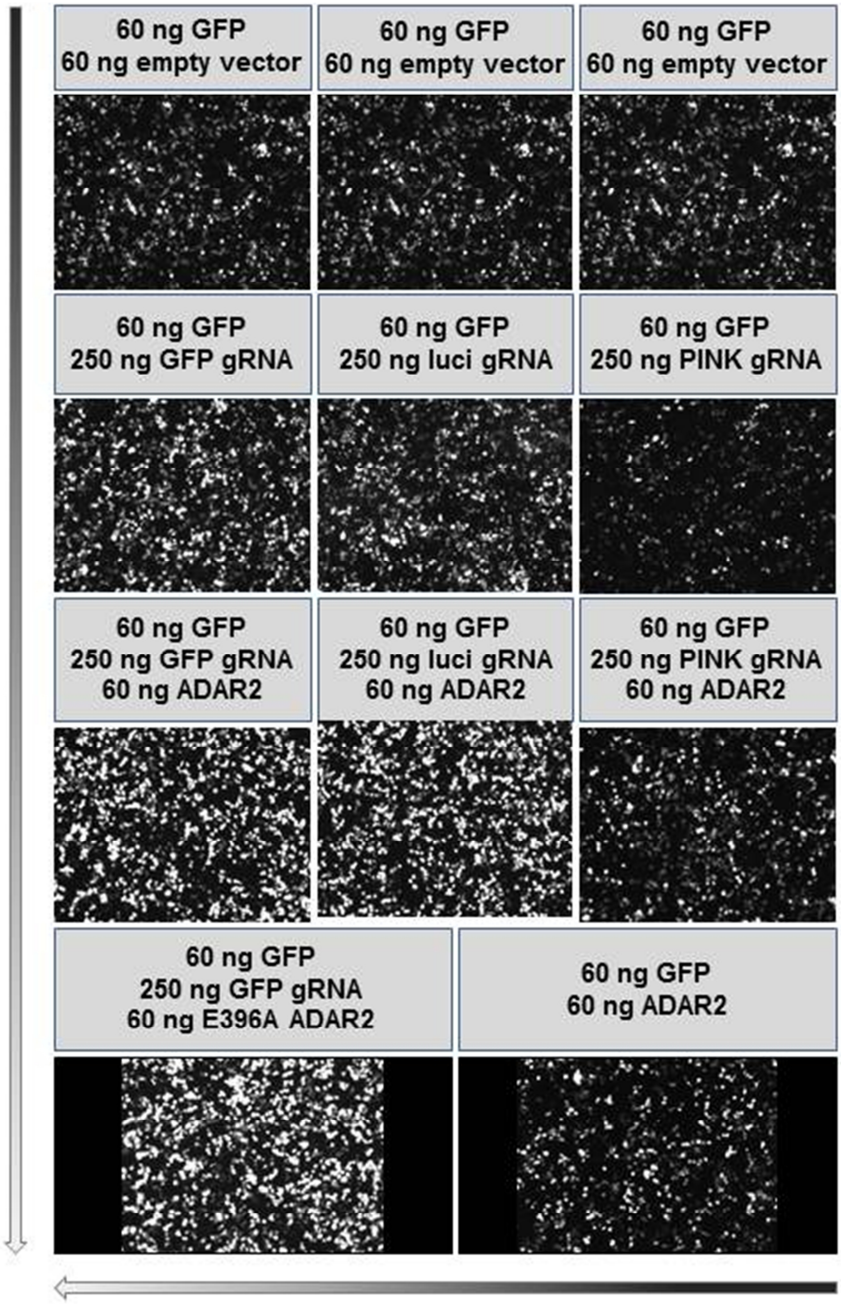

Figure 4-100: Enhancing eGFP signal by the presence of the editing system. Transfections of variable compositions of plasmids together with the eGFP plasmid were performed in a 96-well format. The microscopy analysis demonstrates that only certain R/G-gRNAs are increasing the eGFP signal, including W58X GFP gRNA and W417X luciferase. The PINK R/G-gRNA has no effect on the eGFP signal. As soon as ADAR2 plasmid is combined with the transfection an increment of the fluorescence signal is obtained, even for the sample with PINK R/G-gRNA. The sample of transfected eGFP and ADAR2 plasmid demonstrates that each editing component – ADAR2 protein and the R/G-gRNA – are contributing to this increment. The replacement of ADAR2 by the not enzymatically active mutant E396A ADAR2 does not lower the fluorescence signal. The arrows are indicating the direction of the increment of the eGFP signal. Magnification: 100x, GFP exposure time: 50 ms.

According to the microscopy analysis and the reproducibility it can be reasoned that the enhancing effect of the eGFP signal is caused by the presence of the R/G-gRNA and ADAR2 protein. One theory assumes that by the presence of the W58X eGFP R/G-gRNA and ADAR2 protein the increment of the eGFP signal is caused by an increased transcript level of the eGFP mRNA, which results in an enhanced protein biosynthesis. In order to prove this theory of enhanced transcript levels, a quantitative PCR (RTqPCR) was performed by Pia Mach with the same experimental setup as for the microscopy analysis in a 24-well format. Indeed, she found that the eGFP transcript level of the sample transfecting eGFP plasmid (300 ng) and W58X eGFP gRNA (1300 ng) was 2.29-fold increased compared to the sample transfecting eGFP plasmid alone. Furthermore, she obtained a 1.57-fold increased eGFP transcript level in the RTqPCR experiment when transfecting eGFP plasmid together with W417X luciferase gRNA. Co-transfection of eGFP plasmid with W437X PINK gRNA did not change the eGFP transcript level. These results correspond to the microscopy analysis and confirm the finding that only with distinct R/G-gRNAs the eGFP transcript levels are increased, resulting in higher eGFP protein levels. The mechanism that leads to the increased eGFP transcript level remains unclear. Most probably, the transcript is stabilized by the binding of the R/G-gRNA and is thus prevented from degradation. This binding and stabilization theory is strengthened by the fact that W58X eGFP gRNA can hybridize to the transcript, whereas PINK gRNA is not supposed to be able to bind the eGFP mRNA at all according to binding analysis with the DNAMAN. But the binding analysis for luciferase gRNA did not show any mRNA binding region, too, but obviously there exists an interaction with the transcript, as well. Many questions arise from these results: Are other transcripts affected by the presence of a complementary gRNA, as well? Does the SNAP-tag BG-system, lacking the two dsRBMs, influence the transcript level, too, or is this phenomenon only restricted to the R/G-gRNA ADAR2 editing system? What is exactly the mechanism behind this?

4.3 Platynereis Dumerilii

The encodable editing system was proven to work successfully in vitro, as well as in cell culture for numerous mRNA transcripts. The next milestone in establishing this new editing method is the transfer from cell culture implementation to a living organism. The organism of choice was the worm *Platynereis Dumerilii*. It offers many advantages, like the possibility to microinject fertilized eggs, the characteristics that the eggs and larvae are transparent and, therefore, an analysis by light microscopy is possible (101,102) and that about 2000 eggs can be obtained from one batch. A collaboration with the group of Dr. Gáspár Jékely (Max-Planck-Institut für Entwicklungsbiologie Tübingen) was started, who are experts in the research of this model organism. The implementation of the SNAP-tag BG-system was started during the supervision of the bachelor thesis of Tahsin Kuzdere pursuing the same aim, as set for the encodable R/G-gRNA ADAR2 editing system to edit W58X eGFP mRNA in the larvae of *Platynereis Dumerilii*.

In order to obtain a uniform expression pattern, the (W58X) GFP and ADAR2 gene were microinjected as RNA components into fertilized eggs. Therefore, the genes had to be cloned into pUC57 vector (3.6.1). This vector contains a 169 bp 5'-UTR from the Platynereis 60 S acidic ribosomal protein P2 downstream of the T7 promoter. This 5'-UTR enhances protein expression of the transcripts. The obtained plasmids (pTS82 for GFP, pTS83 for W58X eGFP and pTS101.1 for ADAR2) had to be linearized with Eco53KI in order to achieve high in vitro RNA transcription yields with the mMESSAGE mMACHINE T7 Ultra Kit (3.6.2). The polyadenylation state and size of the transcripts were tested by a 1.0% TBE agarose gel electrophoresis (3.2.6). An exemplary agarose gel picture is shown in Figure 4-101. A distinct mass shift is visible for the ADAR2 sample treated with *E. coli* Poly(A) Polymerase (EPAP) compared to the untreated sample, indicating an effective polyadenylation reaction.

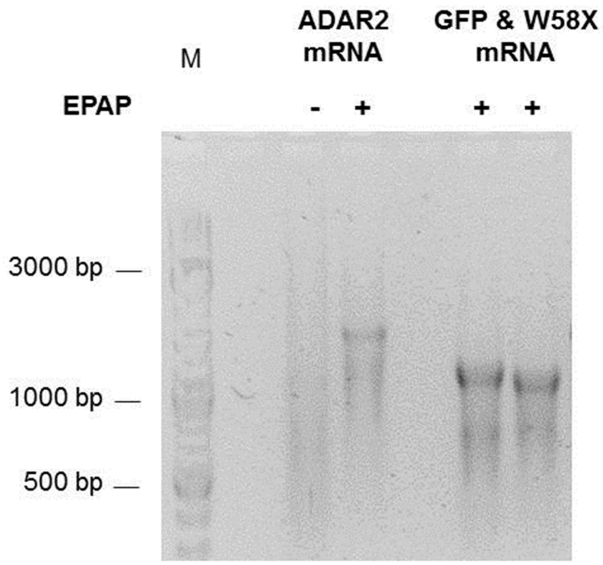

Figure 4-101: TBE agarose gel for visualization of the polyadenylation state of synthesized mRNA.

A 1.0% TBE agarose gel electrophoresis of denatured mRNAs was performed to visualize the polyadenylation state of ADAR2 mRNA after in vitro transcription and to analyze the quality of eGFP/W58X transcripts. The size of the products could be approximately estimated by comparison with the 2-Log DNA Ladder. The mRNA is polyadenylated, indicated by a bigger product band and capped at its 5'-end. All mRNAs harbor a 169 bp 5'-UTR from the Platynereis 60 S acidic ribosomal protein P2 downstream of the T7 promoter to enhance protein expression in the worm.

The first sample of the microinjection consisted of 150 ng/µl W58X eGFP mRNA, 500 ng/µl ADAR2 mRNA, 40 µM R/G-gRNA P6 and 2.25 µg/µl of Rhodamine-labeled dextran as an injection control. In contrast to the SNAP-tag BG-system (59), restoring of the fluorescence signal with the R/G-gRNA ADAR2 system was initially not successful. Possible reasons are 1) inactivation of ADAR2 in the eggs environment, 2) localization of ADAR2 in the nucleus or 3) the instability of the guideRNA. The first reason seems unlikely, since the catalytic activity of SNAP-ADAR1 was already proven in the eggs environment (60). The reduced stability of the R/G-gRNA is much more probable compared to the BG-gRNA, since the transfection of unmodified ss-gRNA was already found to be instable in cell culture (59). In order to exclude the inactivation of ADAR2 in the worm's egg and the solely presence of ADAR2 in the nucleus, it was tested if a chemically stabilized ss-gRNA can induce editing of ADAR2. The ss-gRNA was stabilized by 2'O-methylation and the insertion of phosphorthioates (NH2-W58X-23nt-PTO), but lacks the R/G-motif. Injected eggs were separated from uninjected ones by the parallel analysis of Rhodamine-dextran signal by 6 ms exposure time. The ADAR2's activity could be confirmed by the

microscopy analysis together with the stabilized ss-gRNA by restoring the 5'-UGG codon of W58X eGFP mRNA (Figure 4-102). Therefore, the assumption that the unstabilized R/G-gRNA might be the reason for the ineffective editing was strengthened. Differences in the brightness of the Rhodamine-dextran signal of the positive control transfected with 150 ng/µl of eGFP mRNA compared to the editing sample are due to different injected amounts of sample into the egg. Therefore, the Rhodamine-dextran brightness at constant exposure time (6 ms) enables the comparison of injected amounts of the samples. The eGFP exposure time was increased for the editing sample up to 383 ms (PC 100 ms), since the eGFP signal was much lower than for the positive control, although the amount of injected sample was markedly higher for the editing sample. This is in accordance with the data from chapter (4.2.2.2), which demonstrated that ADAR2's activity is addressable by ss-gRNAs lacking the R/G-motif, however, at a very lousy editing yield.

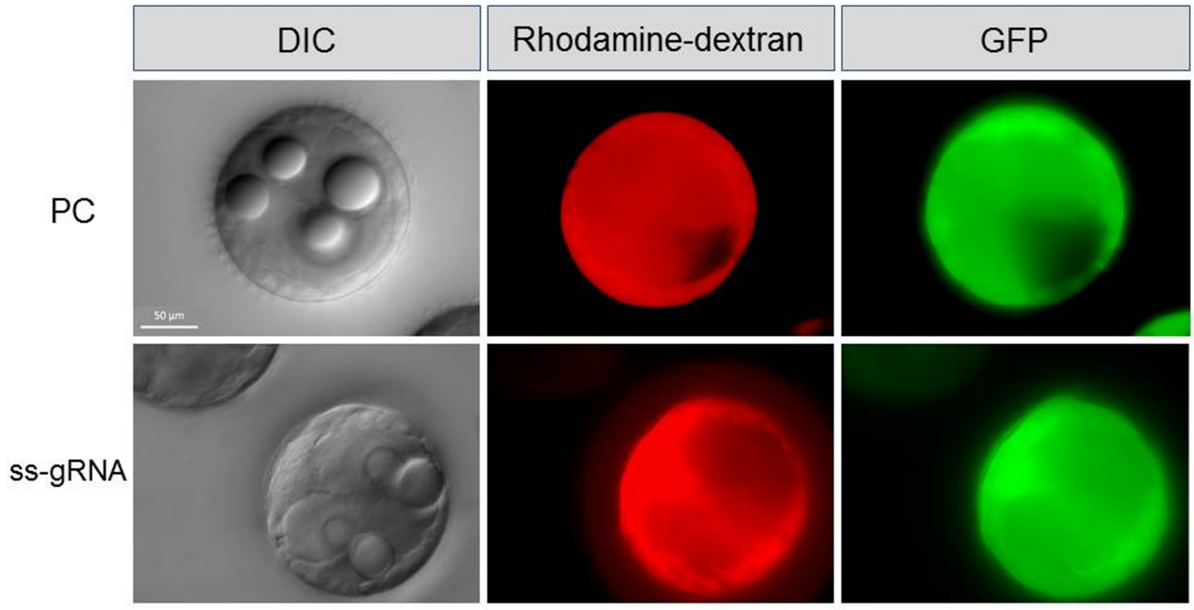

Figure 4-102: Differential interference contrast (DIC) and fluorescence pictures to analyze editing of W58X eGFP mRNA with a ss-gRNA in Platynereis zygotes. The zygotes were injected with 150 ng/µl of eGFP mRNA for the positive control (PC) and with 150 ng/µl of W58X eGFP mRNA, 500 ng/µl of ADAR2 mRNA and 40 µM NH2-W58X-21nt-PTO gRNA (ss-gRNA). Each injection sample contained 2.25 µg/µl of Rhodamine-labeled dextran as an injection control. Magnification: 400x; exposure time: DIC 5 ms, Rhodamine-dextran 5,6 ms, GFP 100 ms for PC and 383 ms for editing.

Then the stability of the R/G-gRNA should be improved by two options: a) the addition of a boxB-motif to the 3'-end of the gRNA, and b) the incorporation of modified nucleoside-tri-phosphates (NTPs). The first option a) was not successful. As

already shown in cell culture, the boxB-motif did not seem to have any beneficial effect on the editing reaction. In some cases the editing efficiency was even decreased (4.2.2.6).

The incorporation of modified nucleotide substrates (option b)) was restricted to UTP and GTP since CTP and ATP are part of the 5'-CCA anticodon and it should be excluded that the modifications hamper the editing reaction. The 2'-O-methyl-NTPs and 5'-O-(1-thiotriphosphate)-NTPs (rNTPαS) were selected for both nucleotides (Figure 4-103), tested for their incorporation into the R/G-gRNA and monitored by Urea-PAGE gel (3.2.18).

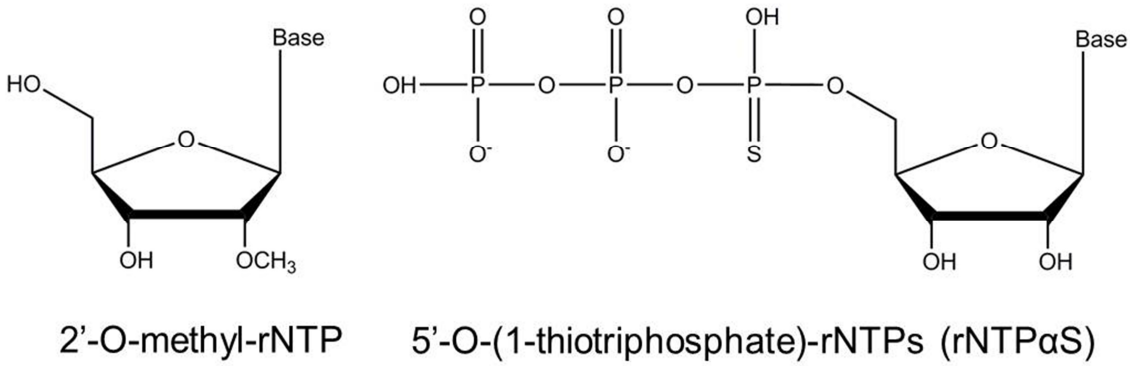

Figure 4-103: 2'-O-methyl-rNTP and 5'-O-(1-thiotriphosphate)-rNTPs (rNTPαS).

The assembly of the 2'-O-methyl modified UTP or GTP was not successful, demonstrated by the absence of transcript at all (Figure 4-104). A low incorporation efficiency for 2'-O-methyl-NTPs could be expected (103), as 2'-O-methyl-NTPs are poor substrates for T7 RNA polymerase. According to the findings from Conrad et al. (104), the incorporation of 2'-O-methyl-NTPs could have been enhanced by addition of low manganese concentrations (2.5 mM) to the transcription mix or evolved polymerases (105). However, the full replacement of the GTP nucleotide by GTPαS and UTP by UTPαS was completely tolerated by the T7 RNA polymerase but the hammerhead (HH) enzyme was not cleaving itself off from the GTPαS containing transcript, demonstrated by the presence of only one strong transcript band. This makes sense, since the cleavage site of the HH-enzyme is constituted of a CTP and GTP nucleotide, and the GTP is replaced by the GTPαS assembly, which hampers the cleavage activity. The presence of the phosphorthioate hampers the hydrolysis of the phosphate group between the two nucleotides, resulting in an absent cleavage reaction.

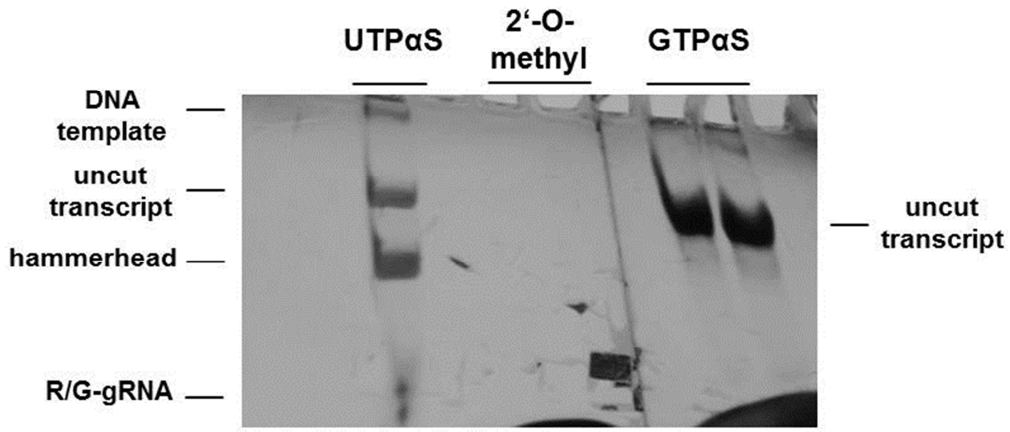

Figure 4-104: Urea PAGE-gel of R/G-gRNA transcription with 2'-O-methyl-NTPs and NTPαS. Either UTP or GTP nucleotides were completely replaced by a 2'-O-methyl UTP/GTP modified substrates or 5'-O-(1-thiotriphosphate)-UTP/GTP in the transcription mix. The usage of only 2'-O-methyl UTP/GTP modified substrates resulted in the absence of any generated transcript, whereas the assembly of UTPαS/GTPαS was successful. The hammerhead ribozyme was not able to cut itself out from the GTPαS transcript, resulting in a single uncut transcript band.

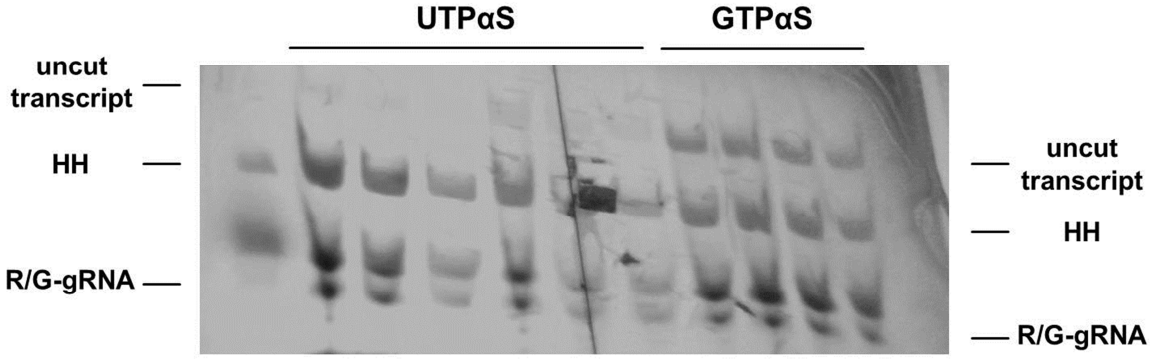

Figure 4-105: Urea PAGE-gel of R/G-gRNA transcription with UTPαS and GTPαS. UTP or GTP nucleotides were partly replaced by 5'-O-(1-thiotriphosphate)-UTP/GTP (UTPαS/GTPαS) in the transcription mix. A ratio of 1:1 was chosen for UTP and 1:2 for GTP. The usage of mixed nucleotides with modified substrates led to an increased amount of R/G-gRNA product for both batches. The hammerhead ribozyme was able to cut itself out from the GTPαS transcript when choosing a 1:2 ratio.

A partial replacement of UTP and GTP nucleotides by phosphorthioate containing NTPs in the transcription mix – 1:1 for UTP/UTPαS and 1:2 for GTP/GTPαS – allowed the HH-enzyme to cut itself off from the transcript again and increased the R/G-gRNA product amount drastically for both samples (Figure 4-105). Still, the HH-enzyme was not able to fully cleave itself off from the GTP/GTPαS containing transcript, indicated by a stronger uncut transcript band compared to the UTP/UTPαS containing transcript. In contrast to former R/G-gRNA productions, without the use of

modified nucleotide substrates, a second band was obtained. In order to distinguish, which band represents the desired R/G-gRNA construct, both products were tested in the standard in vitro editing reaction mix (Table 3-23) without spermidine addition, with the result that the lower band is the correct functional R/G-gRNA (Figure 4-106): The upper gRNA-band achieved only 25% editing yield, whereas the lower band of the gRNA resulted in 93% editing. This test editing was also performed to prove that the assembly of phosphorthioate containing NTPs into the R/G-gRNA does not hamper the editing reaction.

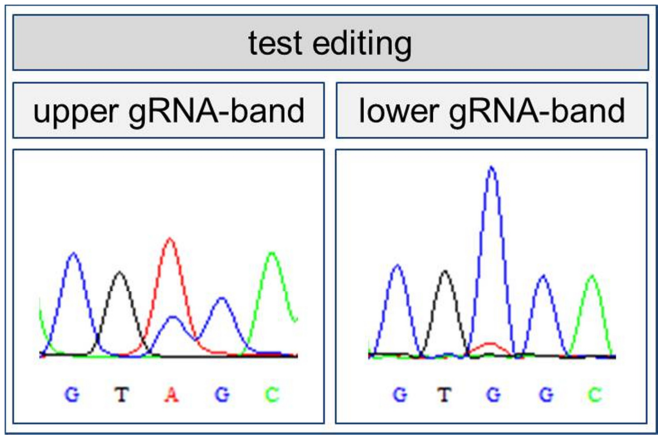

Figure 4-106: Test editing of the two obtained gRNA-bands at the in vitro transcription. The test editing was performed at standard editing reaction mixture without spermidine addition. The upper gRNA-band resulted in less editing yield (25%) than the lower gRNA-band (93%), proving that the lower band is the correct R/G-gRNA band.

After confirming the feasibility to perform editing with the PTO-stabilized R/G-gRNA three different microinjection experiments were performed, using as before 150 ng/µl W58X eGFP mRNA, 500 ng/µl ADAR2 mRNA and 40 µM of the following guideRNAs: 1) ss-gRNA (unmodified NH2-W58X-23nt-gRNA) 2) unmodified R/G-gRNA and 3) stabilized UTPαS containing R/G-gRNA. Unfortunately, the microinjection experiment with the stabilized GTPαS containing R/G-gRNA could not be performed due to lacking of enough material for it. An overview of the microscopy analysis is shown in Figure 4-107, demonstrating that the injection sample with the stabilized R/G-gRNA restored the eGFP signal, whereas the microinjection sample of unstabilized R/G-gRNA and unstabilized ss-gRNA did not edit the W58X eGFP mRNA, indicated by the absence of fluorescence signal. This confirms the assumption that the stability of the gRNA is the reason for the missing editing reaction.

The injected amount of the samples for the stabilized ss-gRNA and the stabilized R/G-gRNA were very different, indicated by a diverging rhodamine intensity. Therefore, it is very difficult to compare the intensity of the fluorescence signal of both injection samples, in order to estimate the contribution of the R/G-motif. But taken into account that the amount of transfected sample of the stabilized ss-gRNA was higher, as well as the exposure time for eGFP than of the stabilized R/G-gRNA, the presence of the R/G-motif seems beneficial for the editing efficiency. But further experiments have to confirm these observations.

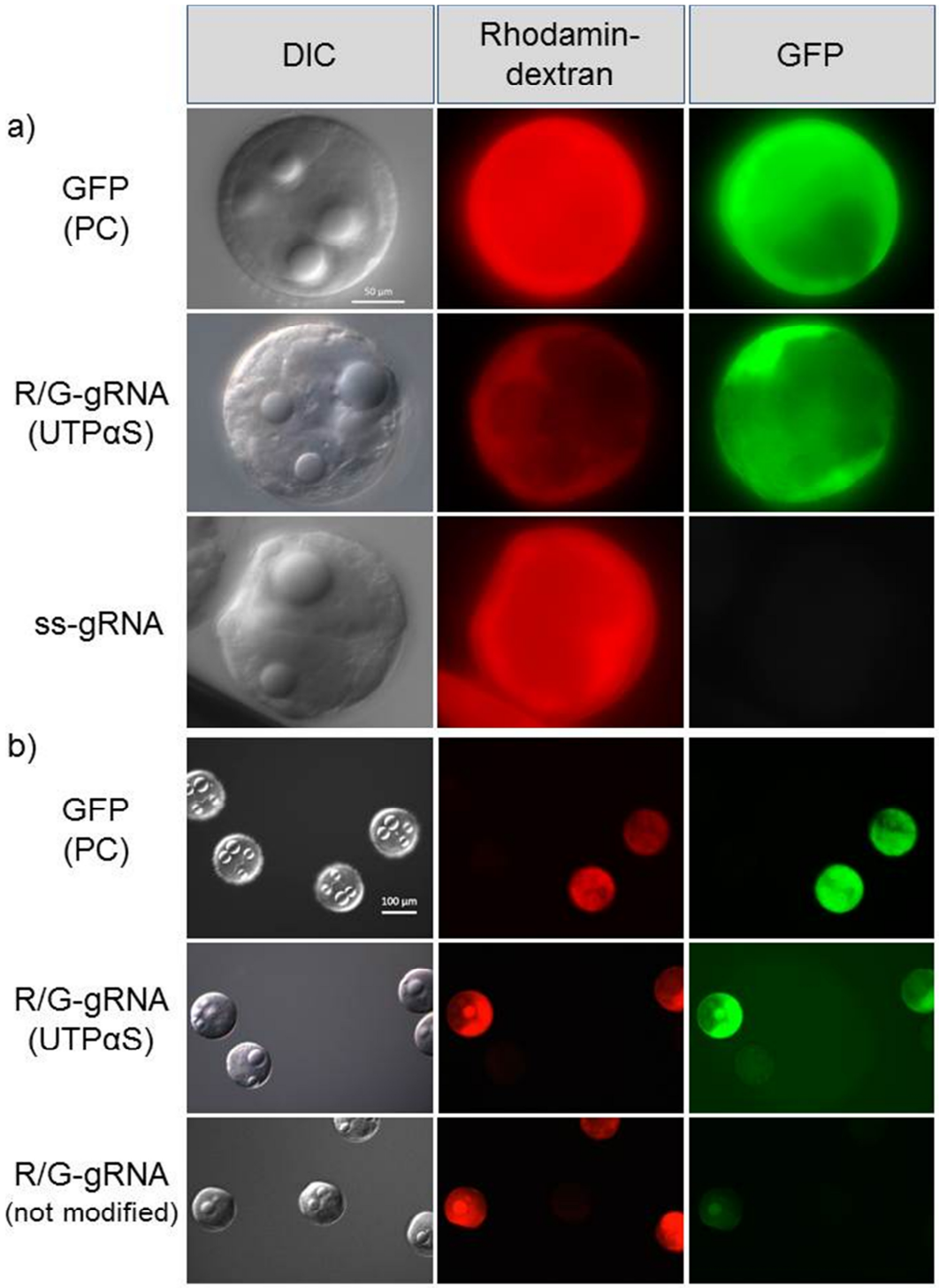

Figure 4-107: Differential interference contrast (DIC) and fluorescence pictures to analyze editing of W58X eGFP mRNA with different gRNAs in Platynereis zygotes. The zygotes were injected with 150 ng/µl of eGFP mRNA for the positive control (PC) and with 150 ng/µl of W58X eGFP mRNA, 500 ng/µl of ADAR2 mRNA and 40 µM of the following gRNAs: 1) NH2-W58X-23nt gRNA (ss-gRNA) 2) unmodified R/G-gRNA and 3) stabilized R/G-gRNA UTPαS. Each injection sample contained 2.25 µg/µl of Rhodamine-labeled dextran as an injection control.

a) Magnification: 400x; exposure time: DIC 5 ms, Rhodamine-dextran 6 ms, GFP 70 ms for PC and 152 ms for editing.
b) Magnification: 100x; exposure time: DIC 2,8 ms, Rhodamine-dextran 60 ms, GFP 2 s for PC and 1 s for editing.

This experiment demonstrates that editing of W58X eGFP mRNA with the encodable editing is possible in *Platynereis* zygotes, representing the proof-of-principle in a living organism, but only with a stabilized R/G-gRNA. Previously, we demonstrated the ability of the SNAP-tag BG-system, to achieve high editing efficiency in *Platynereis* zygotes (60). In comparison the editing efficiency of the R/G-gRNA ADAR2 system is until now markedly lower. The benefit of the SNAP-ADAR system is the possibility to use chemically modified guideRNAs, which are much more stable in general. With this kind of setting the advantage of a fully genetically editing system is not exploited. The stable integration of the R/G-gRNA ADAR2 system into the genome might be one option to improve the efficiency.

5 Conclusion and Outlook

The site-specific manipulation of mRNA processing and its encoded content could dramatically help to elucidate fundamental biological processes and to develop novel therapies for the treatment of various diseases, including genetic disorders that are currently difficult to treat. The idea is to harness A-to-I editing activity for this purpose; in contrast with two other competitive approaches (2.6), we focus on the development of a novel guideRNA architecture that allows recruiting endogenous ADAR for site-specific RNA editing, which makes expression of engineered proteins unnecessary, and that is fully genetically encodable. We have engineered the guideRNA from two components, a single-stranded mRNA binding template, which is reverse complementary to the target mRNA sequence and a complex RNA secondary fold, based on the naturally occurring R/G-loop motif of the glutamate receptor transcript GluR-B (4.1.1). This guideRNA design fulfills two tasks: specific binding of the target mRNA by RNA duplex formation and the recruitment of ADARs via their dsRBDs (2.3.1) to elicit the deamination reaction (2.1). The system was studied for optimization in the PCR tube with purified compounds and then transferred to cell culture.

Firstly, we proved the necessity of our engineered guideRNA working in *trans* to specifically harness ADAR2's editing activity to defined target mRNAs in the PCR reaction tube (4.1.6 & 4.1.10). The advantage provided by the R/G-motif could be demonstrated at stringent editing conditions (4.1.3). Thereby, we found ADAR2 to be much more tolerant towards high magnesia concentrations (3 mM / 5 mM) than SNAP-ADAR2 (0-1.0 mM) (4.1.2).

In contrast to the SNAP-ADAR enzymes, ADAR2 showed markedly more off-target editing at the same editing conditions. We reason the two dsRBDs of ADAR2, which are replaced by the SNAP-tag in SNAP-ADAR2 (2.6), for the increased tendency to edit the mRNA independently of the guideRNA. Since increased magnesia concentrations alone could not prevent off-target editing (4.1.2), we tested two additives in the reaction mix to control off-target editing; heparin and spermidine (4.1.3). Heparin was not found to be suitable to prevent off-target editing, since ADAR2 turned out to be much more sensitive towards heparin than SNAP-ADAR2, resulting in distinctly lower editing yields at the target adenosine. ADAR2 is bound by heparin at the deaminase domain and the two dsRBMs. Since ADAR2 binds

reversibly to the R/G-gRNA, ADAR2 will diffuse away from the RNA substrate duplex upon heparin complexation. The editing efficiency is less affected by SNAP-ADAR2, since SNAP-ADAR2 is covalently linked to the guideRNA and thus doesn't diffuse away upon heparin binding (Figure 4-13 in 4.1.3). The second additive – spermidine – turned out to prevent off-target editing and in parallel to maintain high editing yields at the target adenosine position for the R/G-gRNA ADAR2 system.

The duplex recognition and binding of the dsRBM2 of ADAR2 is influenced by the exit of the mRNA, which is created by the application of the guideRNA in *trans* (4.1.9 & 4.1.10). But nevertheless, high editing yields - up to 100% in the PCR tube and 57% in cell culture were obtained for a 5'-UAG codon (4.1.10 & 4.2.2.5).

Pursuing the aim to achieve high editing yields (≥ 50% conversion) in the PCR tube, the following points should be taken into account. 1) The ratio of enzyme to gRNA and mRNA is determining the editing efficiency, since lowering the ADAR2 protein concentration steadily in particular below stoichiometric amounts decreased the editing yield at the target adenosine (4.1.4). 2) The editing reaction mix should be incubated at least for 45 min without spermidine and 90 min with spermidine addition (4.1.5). 3) The choice of the optimal anticodon preferences of ADAR2 determines the editing efficiency of the targeted adenosine (Figure 2-11 in 2.6) (49): Considering the preferred anticodons of ADAR2 of our engineered guideRNA, even less preferred codons, such as 5'-CAG, 5'-CAC and 5'-GAG, were editable to at least 50% in the PCR reaction tube.

In order to increase the editing yield of less preferred codons even further, a more active ADAR2 mutant, namely E488Q (2.5.1), was tested. This mutant is also steerable with the engineered guideRNA, but shows a reverse preference for the tested anticodons than wt ADAR2 to edit a 5'-GAG codon: wt ADAR2 edits the 5'-GAG codon best with a 5'-CCG anticodon, followed by 5'-CUA and no editing is obtained for a 5'-CCC anticodon. E488Q has the following ranking of anticodon preferences: 5'-CCC>>5'-CUA>5'-CCG (4.1.8). Not only the anticodon preferences are differently to wt ADAR2, also higher editing yields could be confirmed for E488Q. This is reasoned by the interference of the glutamate to glutamine substitution at position 488 with the base flipping (2.5.1), which is facilitated by the point mutation of glutamic acid to glutamine at position 488 of ADAR2. According to the studies of

Wang et al. (54), this residue promotes not only base flipping, it also contacts and stabilizes the orphaned counter base. It is very likely that this residue is also interacting or influenced by the 3'-neighbor of the counter base, leading to different anticodon preferences of the mutant and wt ADAR2.
The downside of the enhanced editing activity of E488Q is the presence of more off-target editing (Figure 4-37 in 4.1.8) compared to the wt ADAR2. Spermidine addition to the reaction mix was not found to be sufficient to control off-target editing by E488Q (Figure 4-40 in 4.1.8).

Since mRNAs form secondary structures, the accessibility of complementary oligonucleotides is dependent on the targeted sequence. This diverse mRNA accessibility might explain why identical codons of different mRNAs are edited to a variable extend (4.1.7.1). But also the corresponding R/G-gRNA can tend to self-hybridization or hybridization with each other. We observed that with lower magnesia concentrations the editing yield of W437X PINK mRNA (5'-UAG) could be improved (Figure 4-31). A lower magnesia concentration might not only lead to less strong secondary structure formation of the PINK mRNA, but also might reduce the hybridization of R/G-gRNA with each other or itself. This issue could turn out as a major limitation of an editing system applied in *trans*.

To further analyze the characteristics of the editing system in the PCR reaction tube, we reduced the components of the editing system down to a sub-nanomolar concentration (4.1.10). We found that ADAR2 cannot be diluted below 100 nM in a PCR tube without significant loss of enzyme activity. Therefore, a concentration of 187 nM was used for further investigations. The reduction of the components was not affecting the peculiarities and characteristics of the editing system: The system showed to be specific, efficient and robust towards stringent conditions and to be efficient on a broad range of component concentrations & stoichiometries. This was an important prerequisite for starting experiments with this editing system in cell culture.

In cell culture we transfected all components of the editing system coded on plasmids into HEK293T cells. Again, we demonstrated the necessity and specificity of the designed guideRNA for steering ADAR2's activity to the target mRNA (4.2.2.2).

When applying the incubation time (24h) that was established for the SNAP-ADAR BG-system, we found that the R/G-gRNA ADAR2 system achieved 30%, whereas the SNAP-ADAR BG-system achieved 63% (4.2.2). One assumption for this different response of both editing system might be the production and provision of sufficiently high guideRNA amount. However, a main difference is the covalent and stable conjugate formation of SNAP-ADAR and the BG-gRNA, as soon as they encounter each other, whereas the contact and binding of ADAR2 with the R/G-gRNA is reversible.

With the aim to increase the total editing yield in cell culture further, we optimized details of the guideRNA and the experimental editing conditions:
1) Prolonging the length of the mRNA binding site was not successful to increase the total editing yield (4.2.2.3). Instead, the prolongation up to 25 nt or longer created off-target editing in the RNA-duplex. Therefore, the length of 16 nt for the mRNA binding site was found to be most suitable. 2) The analysis of the optimal distance of the target adenosine with respect to the 5'-end of the R/G-motif revealed that the position 8 was the most suitable one to achieve highest editing yields (4.2.2.4). This finding is more general, since the investigation of the positioning effect at two other target mRNAs confirmed position 8 to be the optimal position (4.2.3.1). 3) The transfection of bigger amounts of R/G-gRNA resulted in higher editing yields (4.2.2.5 & 4.2.3.1). This suggests that the editing yield might be limited by the presence of sufficient amounts of guideRNA. 4) Therefore, we conclude that the stability of the guideRNA is a major factor to optimize the provision of sufficient guideRNA. The idea that a boxB-motif attached to the 3'-end could improve the editing yield by protecting the guideRNA from being degraded could not be verified (4.2.2.6 & 4.2.3.2). By these experiments it could not be determined, if the degradation of the guideRNA is in fact a limiting point or if the boxB-motif stabilizes the guideRNA but hampers the editing reaction. 5) The exchange of the U6-promoter against a H1-promoter for the guideRNA transcription did not improve the editing yield either (4.2.2.6). 6) But the transfection of one single vector, encoding for multiple copies of the R/G-gRNA and the ADAR2 protein, indicates to be promising to improve the stoichiometry of the editing components and thus the editing yield (4.2.2.6.1). 7) Contrary to our expectations, the transfection of E488Q instead of wt ADAR2 resulted in markedly lower editing yields (4.2.2.2). This result was unexpected, since in vitro it was clearly

shown that E488Q is not only advantegous in editing of less preferred codons, but also in general being faster in editing the target position compared to wt ADAR2 (4.1.8). We speculate that the failure of E488Q might come from hyper-editing and thus destruction of the R/G-motif itself. This phenomenon was already observed by editing in the PCR tube (4.1.9). 8) We found that the editing yields in the PCR reaction tube and in cell culture depend on the incubation time: The total editing yield accumulates with increasing incubation times from 24h to 72h (4.2.2 & 4.2.2.7). After longer incubation (>96 h) the trend reverses again (4.2.2.7).
By improving the experimental conditions and the guideRNA design, the highest obtained editing yield in cell culture was 57%, whereas 50% was the common standard to edit a 5'-UAG codon. These editing yields are markedly higher than the editing yields of the genetically encodable editing system of Montiel-Gonzales et al. (50) with 20% for the same W58X GFP codon. With our optimized experimental conditions the disease-promoting mutation R175H p53 (5'-CAC codon) was edited up to 23% and the R282Q mutation on p53 (5'-CAG triplet) at 37% yield (0).

As previously mentioned the accessibility of the mRNA plays an important role to determine the editing yields and varies a lot between genes. Whereas secondary structure formation might be the main factor in the PCR reaction tube, the binding of various RNA binding proteins to the mRNA transcript further intensifies this problem in cell culture and might explain why lower editing yields are reproducibly obtained inside the cell. Furthermore, it is more difficult to adjust the optimal stoichiometry of the components inside the cell and the time for editing to occur is also limited. Finally, as long as the promoter steadily transcribes mRNA one will never achieve 100% editing.

The issue of off-target editing is present in cell culture as well (4.2.4). There are two different modes of off-target editing. The first off-target editing can appear at any position of the mRNA transcript, is independent of the presence of the guideRNA, and was observed starting from an incubation time of 48 hours in the GFP mRNA transcript (4.2.2.2). Further prolongation of the incubation time increased the off-target editing (4.2.2.7). This mode of off-target editing can be controlled by shorter incubation times and a reduction of the ADAR protein concentration (4.2.2.8). The second mode of off-target editing is created by the mRNA/guideRNA duplex itself (4.2.2.3). Adenosines that are included in the mRNA/guideRNA duplex can potentially

be edited as well. By a careful design of the complementary mRNA binding part, other adenosines in the RNA duplex can be prevented from being edited (4.2.4) (59): Placing the adenosine, which should be protected from editing, to the border of the RNA duplex or mispairing it with guanosine (Figure 2-11 in 2.6) (49).

We could also observe that the R/G-motif itself is highly edited by the ADAR proteins (4.1.9), possibly leading to a decreased ability to harness ADAR activity, due to the loss of the structural integrity of the guideRNA. A prevention of hyper-editing at the R/G-motif by single nucleotide changes, while maintaining the secondary structure offers the potential to further increase the editing yield, especially for the E488Q mutant, which is more active than the wild type ADAR2, but hasn't been done yet.

Our future aim of this project is to steer the endogenously expressed ADAR2 with the designed R/G-gRNA. First experimental results in HEK293T cells indicate that endogenous deamination activity could be harnessed for site-specific RNA editing and opens the perspective to recruit ADAR activity of neuronal cells, too (4.2.2.8). To prove this observation, a cell line with a single stable integrated copy of ADAR2 is under development and the viral production to transfect neuronal cells with the engineered guideRNA.

We challenge the editing system by eliciting editing at two different reporter genes in parallel. This was possible while maintaining high editing yields comparable to the editing yields obtained with single reporter genes only, but now for both transcripts in parallel (4.2.3.4). This demonstrates that redirecting this editing system is not limited to only one single point mutation or one single gene and may offer the application of RNA editing for the treatment of multigenetic diseases. However, one limiting point for this parallel editing was the total amount of transfected plasmid and the corresponding amount of transfection agent, which was necessary to achieve high editing yields. The transfected cells were strongly suffering from this amount of transfection agent. But the idea to encode multiple copies of guideRNAs and enzyme on the same vector (4.2.2.6.1) offers an option to circumvent the co-transfection of many plasmids and to reduce the amount of the transfection reagent.

In cell culture, we could observe that the presence of the editing system – ADAR protein and the complementary guideRNA – increased the respective transcript level and the protein expression (4.2.6). Still, the origin remains unclear how the presence

of the editing components leads to an increased GFP fluorescence signal and has to be further analyzed.

With regard to our last objective we could successfully demonstrate the functioning of the R/G-gRNA ADAR2 editing system in the living organism *Platynereis dumerilii.* Compared to the SNAP-tag BG-system (60), the editing efficiency of the R/G-gRNA ADAR2 system in *Platynereis* zygotes is markedly lower, probably due to the missing chemical modifications for its stabilization. Therefore, these experiments only highlight the proof-of-principle of the R/G-gRNA ADAR2 system and a stable integration of the R/G-gRNA ADAR2 system into worm's genome would exploit the advantage of this system. But the chemical modification of the R/G-gRNA, as we have already initiated in this work, represents an alternative strategy to harness ADAR2 in a living organism.

Our results demonstrate that with the genetically encodable editing system we are able to correct different codons on the RNA-level, resulting in a restored protein function and we are able to address disease-promoting genetic mutations. For further development of this editing system, the creation of a cell line with a genomic copy of endogenous ADAR2 is necessary to investigate the efficiency of our editing system with very low ADAR2 amounts and to come closer to natural conditions. An obvious expansion of our editing strategy is the recruitment of hADAR1 instead of hADAR2. ADAR1 is expressed from two different promoters, resulting in two different isoforms: constitutive and inducible by inflammatory factors (22). This would be beneficial, since the inducible isoform of ADAR1 is also localized in the cytoplasm, giving editing more time than being timely limited to the nucleus, and recruiting this isoform would not misbalance natural substrate editing.
If the stability and efficiency of the guideRNA should be improved by chemical modification, the same delivery strategies as for other antisense-oligonucleotides are available. With regard to the delivery problems of oligonucleotide-based drugs (80), we could circumvent this issue by using viral vectors that genetically encode our editing system. Therefore, the production of viral vectors encoding the engineered guideRNA and/or ADAR protein is in the process.

With the advent of CRISPR/Cas9 and TALEN the therapeutic correction of disease causing point mutations via genome editing might come into reach. However, RNA

editing might be an attractive alternative and complementation due to its advantages: Firstly, RNA editing can restore the wild type function of a protein on the RNA level, thus avoiding ethical issues related to the manipulation of the somatic or even germline genome. Secondly, off-target editing is expected to be less severe and reversible compared to off-target genome slicing. Thirdly, current gene therapy is limited to non-toxic alterations, since only a complete and irreversible switch between two isoforms can be performed. Both, the completeness or the irreversibility of the genomic intervention run the risk of severe adverse effects or lethality. Fourthly, intermediate positions can be achieved by RNA editing, in contrast to the genomic level, where either a complete or none change can be carried out. This allows us to generate many protein isoforms or short term changes by RNA editing. And finally, RNA editing changes genetic function without interfering with their endogenous regulation via transcription, splicing and transport, as this is necessary for many proteins to maintain function.

Therefore, the results presented here pave the way for the development of novel therapeutic strategies for treating diseases that complement current strategies, which rely on genome editing, RNAi, splice alteration, antibodies, small molecules, and others.

6 Appendix

6.1 Vector sequences

Table 6-1: Vector sequences

pRJ114

```
   1 AAAAGCTGGA GCTCGTGACA GCCCTCCGAA GGAAGACTCT CCTCCGTGCG TCCTCGTCTT
  61 CACCGGTCGC GTTCCTGAAA CGCAGATGTG CCTCGCGCCG CACTGCTCCG AACAATAAAG
 121 ATTCTACAAT ACTAGCTTTT ATGGTTATGA AGAGGAAAAA TTGGCAGTAA CCTGGCCCCA
 181 CAAACCTTCA AATGAACGAA TCAAATTAAC AACCATAGGA TGATAATGCG ATTAGTTTTT
 241 TAGCCTTATT TCTGGGGTAA TTAATCAGCG AAGCGATGAT TTTTGATCTA TTAACGGATA
 301 TATAAATGCA AAAACTGCAT AACCACTTTA ACTAATACTT TCAACATTTT CGGTTTGTAT
 361 TACTTCTTAT TCAAATGTAA TAAAAGTATC AACAAAAAAT TGTTAATATA CCTCTATACT
 421 TTAACGTCAA GGAGAAAAAA CCCCGGATTC TAGAACTAGT GGATCCCCCG GGCTGCAGGA
 481 ATTCGATATC AAGCTTATCG ATACCGTCGA CCTCGAGTCA TGTAATTAGT TATGTCACGC
 541 TTACATTCAC GCCCTCCCCC CACATCCGCT CTAACCGAAA AGGAAGGAGT TAGACAACCT
 601 GAAGTCTAGG TCCCTATTTA TTTTTTTATA GTTATGTTAG TATTAAGAAC GTTATTTATA
 661 TTTCAAATTT TTCTTTTTTT TCTGTACAGA CGCGTGTACG CATGTAACAT TATACTGAAA
 721 ACCTTGCTTG AGAAGGTTTT GGGACGCTCG AAGGCTTTAA TTTGCGGCCG GTACCCAATT
 781 CGCCCTATAG TGAGTCGTAT TACGCGCGCT CACTGGCCGT CGTTTTACAA CGTCGTGACT
 841 GGGAAAACCC TGGCGTTACC CAACTTAATC GCCTTGCAGC ACATCCCCCT TTCGCCAGCT
 901 GGCGTAATAG CGAAGAGGCC CGCACCGATC GCCCTTCCCA ACAGTTGCGC AGCCTGAATG
 961 GCGAATGGCG CGACGCGCCC TGTAGCGGCG CATTAAGCGC GGCGGGTGTG GTGGTTACGC
1021 GCAGCGTGAC CGCTACACTT GCCAGCGCCC TAGCGCCCGC TCCTTTCGCT TTCTTCCCTT
1081 CCTTTCTCGC CACGTTCGCC GGCTTTCCCC GTCAAGCTCT AAATCGGGGG CTCCCTTTAG
1141 GGTTCCGATT TAGTGCTTTA CGGCACCTCG ACCCCAAAAA ACTTGATTAG GGTGATGGTT
1201 CACGTAGTGG GCCATCGCCC TGATAGACGG TTTTTCGCCC TTTGACGTTG GAGTCCACGT
1261 TCTTTAATAG TGGACTCTTG TTCCAAACTG GAACAACACT CAACCCTATC TCGGTCTATT
1321 CTTTTGATTT ATAAGGGATT TTGCCGATTT CGGCCTATTG GTTAAAAAAT GAGCTGATTT
1381 AACAAAAATT TAACGCGAAT TTTAACAAAA TATTAACGTT TACAATTTCC TGATGCGGTA
1441 TTTTCTCCTT ACGCATCTGT GCGGTATTTC ACACCGCATA GGGTAATAAC TGATATAATT
1501 AAATTGAAGC TCTAATTTGT GAGTTTAGTA TACATGCATT TACTTATAAT ACAGTTTTTT
1561 AGTTTTGCTG GCCGCATCTT CTCAAATATG CTTCCCAGCC TGCTTTTCTG TAACGTTCAC
1621 CCTCTACCTT AGCATCCCTT CCCTTTGCAA ATAGTCCTCT TCCAACAATA ATAATGTCAG
1681 ATCCTGTAGA GACCACATCA TCCACGGTTC TATACTGTTG ACCCAATGCG TCTCCCTTGT
1741 CATCTAAACC CACACCGGGT GTCATAATCA ACCAATCGTA ACCTTCATCT CTTCCACCCA
1801 TGTCTCTTTG AGCAATAAAG CCGATAACAA AATCTTTGTC GCTCTTCGCA ATGTCAACAG
1861 TACCCTTAGT ATATTCTCCA GTAGATAGGG AGCCCTTGCA TGACAATTCT GCTAACATCA
1921 AAAGGCCTCT AGGTTCCTTT GTTACTTCTT CTGCCGCCTG CTTCAAACCG CTAACAATAC
1981 CTGGGCCCAC CACACCGTGT GCATTCGTAA TGTCTGCCCA TTCTGCTATT CTGTATACAC
2041 CCGCAGAGTA CTGCAATTTG ACTGTATTAC CAATGTCAGC AAATTTTCTG TCTTCGAAGA
2101 GTAAAAAATT GTACTTGGCG GATAATGCCT TTAGCGGCTT AACTGTGCCC TCCATGGAAA
2161 AATCAGTCAA GATATCCACA TGTGTTTTTA GTAAACAAAT TTTGGGACCT AATGCTTCAA
2221 CTAACTCCAG TAATTCCTTG GTGGTACGAA CATCCAATGA AGCACACAAG TTTGTTTGCT
2281 TTTCGTGCAT GATATTAAAT AGCTTGGCAG CAACAGGACT AGGATGAGTA GCAGCACGTT
2341 CCTTATATGT AGCTTTCGAC ATGATTTATC TTCGTTTCCT GCAGGTTTTT GTTCTGTGCA
2401 GTTGGGTTAA GAATACTGGG CAATTTCATG TTTCTTCAAC ACTACATATG CGTATATATA
2461 CCAATCTAAG TCTGTGCTCC TTCCTTCGTT CTTCCTTCTG TTCGGAGATT ACCGAATCAA
2521 AAAAATTTCA AAGAAACCGA AATCAAAAAA AAGAATAAAA AAAAAATGAT GAATTGAATT
2581 GAAAAGCTGT GGTATGGTGC ACTCTCAGTA CAATCTGCTC TGATGCCGCA TAGTTAAGCC
2641 AGCCCCGACA CCCGCCAACA CCCGCTGACG CGCCCTGACG GGCTTGTCTG CTCCCGGCAT
2701 CCGCTTACAG ACAAGCTGTG ACCGTCTCCG GGAGCTGCAT GTGTCAGAGG TTTTCACCGT
2761 CATCACCGAA ACGCGCGAGA CGAAAGGGCC TCGTGATACG CCTATTTTTA TAGGTTAATG
2821 TCATGATAAT AATGGTTTCT TAGGACGGAT CGCTTGCCTG TAACTTACAC GCGCCTCGTA
2881 TCTTTTAATG ATGGAATAAT TTGGGAATTT ACTCTGTGTT TATTTATTTT TATGTTTTGT
2941 ATTTGGATTT TAGAAAGTAA ATAAAGAAGG TAGAAGAGTT ACGGAATGAA GAAAAAAAAA
3001 TAAACAAAGG TTTAAAAAAT TTCAACAAAA AGCGTACTTT ACATATATAT TTATTAGACA
3061 AGAAAAGCAG ATTAAATAGA TATACATTCG ATTAACGATA AGTAAAATGT AAAATCACAG
3121 GATTTTCGTG TGTGGTCTTC TACACAGACA AGATGAAACA ATTCGGCATT AATACCTGAG
3181 AGCAGGAAGA GCAAGATAAA AGGTAGTATT TGTTGGCGAT CCCCCTAGAG TCTTTTACAT
3241 CTTCGGAAAA CAAAAACTAT TTTTTCTTTA ATTTCTTTTT TTACTTTCTA TTTTTAATTT
3301 ATATATTTAT ATTAAAAAAT TTAAATTATA ATTATTTTTA TAGCACGTGA TGAAAAGGAC
3361 CCAGGTGGCA CTTTTCGGGG AAATGTGCGC GGAACCCCTA TTTGTTTATT TTTCTAAATA
```

```
3421 CATTCAAATA TGTATCCGCT CATGAGACAA TAACCCTGAT AAATGCTTCA ATAATATTGA
3481 AAAAGGAAGA GTATGAGTAT TCAACATTTC CGTGTCGCCC TTATTCCCTT TTTTGCGGCA
3541 TTTTGCCTTC CTGTTTTTGC TCACCCAGAA ACGCTGGTGA AAGTAAAAGA TGCTGAAGAT
3601 CAGTTGGGTG CACGAGTGGG TTACATCGAA CTGGATCTCA ACAGCGGTAA GATCCTTGAG
3661 AGTTTTCGCC CCGAAGAACG TTTTCCAATG ATGAGCACTT TTAAAGTTCT GCTATGTGGC
3721 GCGGTATTAT CCCGTATTGA CGCCGGGCAA GAGCAACTCG GTCGCCGCAT ACACTATTCT
3781 CAGAATGACT TGGTTGAGTA CTCACCAGTC ACAGAAAAGC ATCTTACGGA TGGCATGACA
3841 GTAAGAGAAT TATGCAGTGC TGCCATAACC ATGAGTGATA ACACTGCGGC CAACTTACTT
3901 CTGACAACGA TCGGAGGACC GAAGGAGCTA ACCGCTTTTT TGCACAACAT GGGGGATCAT
3961 GTAACTCGCC TTGATCGTTG GGAACCGGAG CTGAATGAAG CCATACCAAA CGACGAGCGT
4021 GACACCACGA TGCCTGTAGC AATGGCAACA ACGTTGCGCA AACTATTAAC TGGCGAACTA
4081 CTTACTCTAG CTTCCCGGCA ACAATTAATA GACTGGATGG AGGCGGATAA AGTTGCAGGA
4141 CCACTTCTGC GCTCGGCCCT TCCGGCTGGC TGGTTTATTG CTGATAAATC TGGAGCCGGT
4201 GAGCGTGGGT CTCGCGGTAT CATTGCAGCA CTGGGGCCAG ATGGTAAGCC CTCCCGTATC
4261 GTAGTTATCT ACACGACGGG GAGTCAGGCA ACTATGGATG AACGAAATAG ACAGATCGCT
4321 GAGATAGGTG CCTCACTGAT TAAGCATTGG TAACTGTCAG ACCAAGTTTA CTCATATATA
4381 CTTTAGATTG ATTTAAAACT TCATTTTTAA TTTAAAAGGA TCTAGGTGAA GATCCTTTTT
4441 GATAATCTCA TGACCAAAAT CCCTTAACGT GAGTTTTCGT TCCACTGAGC GTCAGACCCC
4501 GTAGAAAAGA TCAAAGGATC TTCTTGAGAT CCTTTTTTTC TGCGCGTAAT CTGCTGCTTG
4561 CAAACAAAAA AACCACCGCT ACCAGCGGTG GTTTGTTTGC CGGATCAAGA GCTACCAACT
4621 CTTTTTCCGA AGGTAACTGG CTTCAGCAGA GCGCAGATAC CAAATACTGT CCTTCTAGTG
4681 TAGCCGTAGT TAGGCCACCA CTTCAAGAAC TCTGTAGCAC CGCCTACATA CCTCGCTCTG
4741 CTAATCCTGT TACCAGTGGC TGCTGCCAGT GGCGATAAGT CGTGTCTTAC CGGGTTGGAC
4801 TCAAGACGAT AGTTACCGGA TAAGGCGCAG CGGTCGGGCT GAACGGGGGG TTCGTGCACA
4861 CAGCCCAGCT TGGAGCGAAC GACCTACACC GAACTGAGAT ACCTACAGCG TGAGCTATGA
4921 GAAAGCGCCA CGCTTCCCGA AGGGAGAAAG GCGGACAGGT ATCCGGTAAG CGGCAGGGTC
4981 GGAACAGGAG AGCGCACGAG GGAGCTTCCA GGGGGAAACG CCTGGTATCT TTATAGTCCT
5041 GTCGGGTTTC GCCACCTCTG ACTTGAGCGT CGATTTTTGT GATGCTCGTC AGGGGGGCGG
5101 AGCCTATGGA AAAACGCCAG CAACGCGGCC TTTTTACGGT TCCTGGCCTT TTGCTGGCCT
5161 TTTGCTCACA TGTTCTTTCC TGCGTTATCC CCTGATTCTG TGGATAACCG TATTACCGCC
5221 TTTGAGTGAG CTGATACCGC TCGCCGCAGC CGAACGACCG AGCGCAGCGA GTCAGTGAGC
5281 GAGGAAGCGG AAGAGCGCCC AATACGCAAA CCGCCTCTCC CCGCGCGTTG GCCGATTCAT
5341 TAATGCAGCT GGCACGACAG GTTTCCCGAC TGGAAAGCGG GCAGTGAGCG CAACGCAATT
5401 AATGTGAGTT ACCTCACTCA TTAGGCACCC CAGGCTTTAC ACTTTATGCT TCCGGCTCCT
5461 ATGTTGTGTG GAATTGTGAG CGGATAACAA TTTCACACAG GAAACAGCTA TGACCATGAT
5521 TACGCCAAGC GCGCAATTAA CCCTCACTAA AGGGAAC
```

pMG211

```
1        TACAATTTCA GGTGGCACTT TTCGGGGAAA TGTGCGCGGA ACCCCTATTT GTTTATTTTT
61       CTAAATACAT TCAAATATGT ATCCGCTCAT GAGACAATAA CCCTGATAAA TGCTTCAATA
121      ATATTGAAAA AGGAAGAGTA TGAGTATTCA ACATTTCCGT GTCGCCCTTA TTCCCTTTTT
181      TGCGGCATTT TGCCTTCCTG TTTTTGCTCA CCCAGAAACG CTGGTGAAAG TAAAAGATGC
241      TGAAGATCAG TTGGGTGCAC GAGTGGGTTA CATCGAACTG GATCTCAACA GCGGTAAGAT
301      CCTTGAGAGT TTTCGCCCCG AAGAACGTTT TCCAATGATG AGCACTTTTA AAGTTCTGCT
361      ATGTGGCGCG GTATTATCCC GTATTGACGC CGGGCAAGAG CAACTCGGTC GCCGCATACA
421      CTATTCTCAG AATGACTTGG TTGAGTACTC ACCAGTCACA GAAAAGCATC TTACGGATGG
481      CATGACAGTA AGAGAATTAT GCAGTGCTGC CATAACCATG AGTGATAACA CTGCGGCCAA
541      CTTACTTCTG ACAACGATCG GAGGACCGAA GGAGCTAACC GCTTTTTTGC ACAACATGGG
601      GGATCATGTA ACTCGCCTTG ATCGTTGGGA ACCGGAGCTG AATGAAGCCA TACCAAACGA
661      CGAGCGTGAC ACCACGATGC CTGCAGCAAT GGCAACAACG TTGCGCAAAC TATTAACTGG
721      CGAACTACTT ACTCTAGCTT CCCGGCAACA ATTAATAGAC TGGATGGAGG CGGATAAAGT
781      TGCAGGACCA CTTCTGCGCT CGGCCCTTCC GGCTGGCTGG TTTATTGCTG ATAAATCTGG
841      AGCCGGTGAG CGTGGGTCTC GCGGTATCAT TGCAGCACTG GGGCCAGATG GTAAGCCCTC
901      CCGTATCGTA GTTATCTACA CGACGGGGAG TCAGGCAACT ATGGATGAAC GAAATAGACA
961      GATCGCTGAG ATAGGTGCCT CACTGATTAA GCATTGGTAA CTGTCAGACC AAGTTTACTC
1021     ATATATACTT TAGATTGATT TAAAACTTCA TTTTTAATTT AAAAGGATCT AGGTGAAGAT
1081     CCTTTTTGAT AATCTCATGA CCAAAATCCC TTAACGTGAG TTTTCGTTCC ACTGAGCGTC
1141     AGACCCCGTA GAAAAGATCA AAGGATCTTC TTGAGATCCT TTTTTTCTGC GCGTAATCTG
1201     CTGCTTGCAA ACAAAAAAAC CACCGCTACC AGCGGTGGTT TGTTTGCCGG ATCAAGAGCT
1261     ACCAACTCTT TTTCCGAAGG TAACTGGCTT CAGCAGAGCG CAGATACCAA ATACTGTTCT
1321     TCTAGTGTAG CCGTAGTTAG GCCACCACTT CAAGAACTCT GTAGCACCGC CTACATACCT
1381     CGCTCTGCTA ATCCTGTTAC CAGTGGCTGC TGCCAGTGGC GATAAGTCGT GTCTTACCGG
1441     GTTGGACTCA AGACGATAGT TACCGGATAA GGCGCAGCGG TCGGGCTGAA CGGGGGGTTC
1501     GTGCACACAG CCCAGCTTGG AGCGAACGAC CTACACCGAA CTGAGATACC TACAGCGTGA
1561     GCTATGAGAA AGCGCCACGC TTCCCGAAGG GAGAAAGGCG GACAGGTATC CGGTAAGCGG
1621     CAGGGTCGGA ACAGGAGAGC GCACGAGGGA GCTTCCAGGG GGAAACGCCT GGTATCTTTA
```

```
1681    TAGTCCTGTC GGGTTTCGCC ACCTCTGACT TGAGCGTCGA TTTTTGTGAT GCTCGTCAGG
1741    GGGGCGGAGC CTATGGAAAA ACGCCAGCAA CGCGGCCTTT TTACGGTTCC TGGCCTTTTG
1801    CTGGCCTTTT GCTCATAGTC ATGCCCCGCG CCCACCGGAA GGAGCTGACT GGGTTGAAGG
1861    CTCTCAAGGG CATCGGTCGA GATCCCGGTG CCTAATGAGT GAGCTAACTT ACATTAATTG
1921    CGTTGCGCTC ACTGCCCGCT TTCCAGTCGG GAAACCTGTC GTGCCAGCTG CATTAATGAA
1981    TCGGCCAACG CGCGGGGAGA GGCGGTTTGC GTATTGGGCG CCAGGGTGGT TTTTCTTTTC
2041    ACCAGTGAGA CGGGCAACAG CTGATTGCCC TTCACCGCCT GGCCCTGAGA GAGTTGCAGC
2101    AAGCGGTCCA CGCTGGTTTG CCCCAGCAGG CGAAAATCCT GTTTGATGGT GGTTAACGGC
2161    GGGATATAAC ATGAGCTGTC TTCGGTATCG TCGTATCCCA CTACCGAGAT ATCCGCACCA
2221    ACGCGCAGCC CGGACTCGGT AATGGCGCGC ATTGCGCCCA GCGCCATCTG ATCGTTGGCA
2281    ACCAGCATCG CAGTGGGAAC GATGCCCTCA TTCAGCATTT GCATGGTTTG TTGAAAACCG
2341    GACATGGCAC TCCAGTCGCC TTCCCGTTCC GCTATCGGCT GAATTTGATT GCGAGTGAGA
2401    TATTTATGCC AGCCAGCCAG ACGCAGACGC GCCGAGACAG AACTTAATGG GCCCGCTAAC
2461    AGCGCGATTT GCTGGTGACC CAATGCGACC AGATGCTCCA CGCCCAGTCG CGTACCGTCT
2521    TCATGGGAGA AAATAATACT GTTGATGGGT GTCTGGTCAG AGACATCAAG AAATAACGCC
2581    GGAACATTAG TGCAGGCAGC TTCCACAGCA ATGGCATCCT GGTCATCCAG CGGATAGTTA
2641    ATGATCAGCC CACTGACGCG TTGCGCGAGA AGATTGTGCA CCGCCGCTTT ACAGGCTTCG
2701    ACGCCGCTTC GTTCTACCAT CGACACCACC ACGCTGGCAC CCAGTTGATC GGCGCGAGAT
2761    TTAATCGCCG CGACAATTTG CGACGGCGCG TGCAGGGCCA GACTGGAGGT GGCAACGCCA
2821    ATCAGCAACG ACTGTTTGCC CGCCAGTTGT TGTGCCACGC GGTTGGGAAT GTAATTCAGC
2881    TCCGCCATCG CCGCTTCCAC TTTTTCCCGC GTTTTCGCAG AAACGTGGCT GGCCTGGTTC
2941    ACCACGCGGG AAACGGTCTG ATAAGAGACA CCGGCATACT CTGCGACATC GTATAACGTT
3001    ACTGGTTTCA CATTCACCAC CCTGAATTGA CTCTCTTCCG GGCGCTATCA TGCCATACCG
3061    CGAAAGGTTT TGCGCCATTC GATGGTGTCC GGGATCTCGA CGCTCTCCCT TATGCGACTC
3121    CTGCATTAGG AAGCAGCCCA GTAGTAGGTT GAGGCCGTTG AGCACCGCCG CCGCAAGGAA
3181    TGGTGCATGC TAGATCAATC CGTAAACAGG TCAAACATCA GTTGCCGCAA CCAAATATTG
3241    GCTAGGTCCT TGTGGTACTT CGCATGCCAG AACATGTTGA TGGCTATTTC AGGCAAGACG
3301    ACTGGGTGCG GCAAGGCGCT TAGGCCGAAG GGCTCCACGC AGCAGTCGGC TAAACGTATC
3361    GGCACAGTGG CGAGCAGATC GGTGCGCTGG AGGATGTGGC CAACGGCGGC GAAGTGCGGC
3421    ACTTCCAGAC GGATGTCGCG CCGGATGCCG ACCCGTGTCA TGTACGTGTC CACCTCGCCG
3481    TGGCCGGTGC CAGCGGCGAT GACACGCACG TGGCCGTAGG AACAAAAGCG CTCCAGAGTC
3541    AGGGGTTCGC GGGTGACTGG ATGGTCCTTG CGACATAGGC ACACGTAGTG ATTCTGGAGC
3601    AGCCGGCGCT GAAAGAAGCC AGTTTGCAGA TTGGGAAGCA GGCCCACGGC CAAGTCCACG
3661    GTTCCGTTCT GCAAGGCCTG CATCAGGCTC ATCGAACTGT CGCGCACCGT ACTGATCACG
3721    CAATTGGGGG CCTGGTGAGC CAGCACATCC ATCAGCCGCG GCATGAAGTA GATCTCGCCA
3781    ATGTCGGTCA TGGCCAGGGT GAAGGTACGC TCGCTGGTCA GCGGATCGAA GCTTTCATGG
3841    TGCTGTAGGG CGTTGCGCAG TGCGTGCATG GCCGAAGTGA CGGGCTCGGC CAGATGCGCG
3901    GCATAGGGTG TGGGTTCCAT TCCCTGATGT GTGCGCACGA AGAGTGGGTC CTGTAGCGAG
3961    GTGCGCAGGC GTTTCAGCGC ATTGCTCACG GCAGGCTGGG TCAGGCCCAG GTTCTCCGCA
4021    GTGATAGAGA CGCGTCTGTC GACCAGCAAC TGGTTGAACA CCACCAGCAG GTTTAAATCC
4081    AGGTCACGCA GTTCCATGTG GCCTCGCTTG GGTTATTGCT GGTGCCCGGC CCTCGAGCAC
4141    CACCACCACC ACCACTAATA ATGACTAGTC AGCTGATCCG GCTGCTAACA AAGCCCGAAA
4201    GGAAGCTGAG TTGGCTGCTG CCACCGCTGA GCAATAACTA GCATAACCCC TTGGGGCCTC
4261    TAAACGGGTC TTGAGGGGTT TTTTGCTGAA AGGAGGAACT ATATCCGGAT
```

pcDNA3.1

```
1       GACGGATCGG GAGATCTCCC GATCCCCTAT GGTGCACTCT CAGTACAATC TGCTCTGATG
61      CCGCATAGTT AAGCCAGTAT CTGCTCCCTG CTTGTGTGTT GGAGGTCGCT GAGTAGTGCG
121     CGAGCAAAAT TTAAGCTACA ACAAGGCAAG GCTTGACCGA CAATTGCATG AAGAATCTGC
181     TTAGGGTTAG GCGTTTTGCG CTGCTTCGCG ATGTACGGGC CAGATATACG CGTTGACATT
241     GATTATTGAC TAGTTATTAA TAGTAATCAA TTACGGGGTC ATTAGTTCAT AGCCCATATA
301     TGGAGTTCCG CGTTACATAA CTTACGGTAA ATGGCCCGCC TGGCTGACCG CCCAACGACC
361     CCCGCCCATT GACGTCAATA ATGACGTATG TTCCCATAGT AACGCCAATA GGGACTTTCC
421     ATTGACGTCA ATGGGTGGAG TATTTACGGT AAACTGCCCA CTTGGCAGTA CATCAAGTGT
481     ATCATATGCC AAGTACGCCC CCTATTGACG TCAATGACGG TAAATGGCCC GCCTGGCATT
541     ATGCCCAGTA CATGACCTTA TGGGACTTTC CTACTTGGCA GTACATCTAC GTATTAGTCA
601     TCGCTATTAC CATGGTGATG CGGTTTTGGC AGTACATCAA TGGGCGTGGA TAGCGGTTTG
661     ACTCACGGGG ATTTCCAAGT CTCCACCCCA TTGACGTCAA TGGGAGTTTG TTTTGGCACC
721     AAAATCAACG GGACTTTCCA AAATGTCGTA ACAACTCCGC CCCATTGACG CAAATGGGCG
781     GTAGGCGTGT ACGGTGGGAG GTCTATATAA GCAGAGCTCT CTGGCTAACT AGAGAACCCA
841     CTGCTTACTG GCTTATCGAA ATTAATACGA CTCACTATAG GGAGACCCAA GCTTGGTACC
901     GAGCTCGGAT CCACTAGTAA CGGCCGCCAG TGTGCTGGAA TTCTGCAGAT ATCCATCACA
961     CTGGCGGCCG CTCGAGCATG CATCTAGAGG GCCCTATTCT ATAGTGTCAC CTAAATGCTA
1021    GAGCTCGCTG ATCAGCCTCG ACTGTGCCTT CTAGTTGCCA GCCATCTGTT GTTTGCCCCT
1081    CCCCCGTGCC TTCCTTGACC CTGGAAGGTG CCACTCCCAC TGTCCTTTCC TAATAAAATG
1141    AGGAAATTGC ATCGCATTGT CTGAGTAGGT GTCATTCTAT TCTGGGGGGT GGGGTGGGGC
```

```
1201    AGGACAGCAA GGGGGAGGAT TGGGAAGACA ATAGCAGGCA TGCTGGGGAT GCGGTGGGCT
1261    CTATGGCTTC TGAGGCGGAA AGAACCAGCT GGGGCTCTAG GGGGTATCCC CACGCGCCCT
1321    GTAGCGGCGC ATTAAGCGCG GCGGGTGTGG TGGTTACGCG CAGCGTGACC GCTACACTTG
1381    CCAGCGCCCT AGCGCCCGCT CCTTTCGCTT TCTTCCCTTC CTTTCTCGCC ACGTTCGCCG
1441    GCTTTCCCCG TCAAGCTCTA AATCGGGGGC TCCCTTTAGG GTTCCGATTT AGTGCTTTAC
1501    GGCACCTCGA CCCCAAAAAA CTTGATTAGG GTGATGGTTC ACGTAGTGGG CCATCGCCCT
1561    GATAGACGGT TTTTCGCCCT TTGACGTTGG AGTCCACGTT CTTTAATAGT GGACTCTTGT
1621    TCCAAACTGG AACAACACTC AACCCTATCT CGGTCTATTC TTTTGATTTA TAAGGGATTT
1681    TGCCGATTTC GGCCTATTGG TTAAAAAATG AGCTGATTTA ACAAAAATTT AACGCGAATT
1741    AATTCTGTGG AATGTGTGTC AGTTAGGGTG TGGAAAGTCC CCAGGCTCCC CAGCAGGCAG
1801    AAGTATGCAA AGCATGCATC TCAATTAGTC AGCAACCAGG TGTGGAAAGT CCCCAGGCTC
1861    CCCAGCAGGC AGAAGTATGC AAAGCATGCA TCTCAATTAG TCAGCAACCA TAGTCCCGCC
1921    CCTAACTCCG CCCATCCCGC CCCTAACTCC GCCCAGTTCC GCCCATTCTC CGCCCCATGG
1981    CTGACTAATT TTTTTTATTT ATGCAGAGGC CGAGGCCGCC TCTGCCTCTG AGCTATTCCA
2041    GAAGTAGTGA GGAGGCTTTT TTGGAGGCCT AGGCTTTTGC AAAAAGCTCC CGGGAGCTTG
2101    TATATCCATT TTCGGATCTG ATCAAGAGAC AGGATGAGGA TCGTTTCGCA TGATTGAACA
2161    AGATGGATTG CACGCAGGTT CTCCGGCCGC TTGGGTGGAG AGGCTATTCG GCTATGACTG
2221    GGCACAACAG ACAATCGGCT GCTCTGATGC CGCCGTGTTC CGGCTGTCAG CGCAGGGGCG
2281    CCCGGTTCTT TTTGTCAAGA CCGACCTGTC CGGTGCCCTG AATGAACTGC AGGACGAGGC
2341    AGCGCGGCTA TCGTGGCTGG CCACGACGGG CGTTCCTTGC GCAGCTGTGC TCGACGTTGT
2401    CACTGAAGCG GGAAGGGACT GGCTGCTATT GGGCGAAGTG CCGGGGCAGG ATCTCCTGTC
2461    ATCTCACCTT GCTCCTGCCG AGAAAGTATC CATCATGGCT GATGCAATGC GGCGGCTGCA
2521    TACGCTTGAT CCGGCTACCT GCCCATTCGA CCACCAAGCG AAACATCGCA TCGAGCGAGC
2581    ACGTACTCGG ATGGAAGCCG GTCTTGTCGA TCAGGATGAT CTGGACGAAG AGCATCAGGG
2641    GCTCGCGCCA GCCGAACTGT TCGCCAGGCT CAAGGCGCGC ATGCCCGACG GCGAGGATCT
2701    CGTCGTGACC CATGGCGATG CCTGCTTGCC GAATATCATG GTGGAAAATG GCCGCTTTTC
2761    TGGATTCATC GACTGTGGCC GGCTGGGTGT GGCGGACCGC TATCAGGACA TAGCGTTGGC
2821    TACCCGTGAT ATTGCTGAAG AGCTTGGCGG CGAATGGGCT GACCGCTTCC TCGTGCTTTA
2881    CGGTATCGCC GCTCCCGATT CGCAGCGCAT CGCCTTCTAT CGCCTTCTTG ACGAGTTCTT
2941    CTGAGCGGGA CTCTGGGGTT CGAAATGACC GACCAAGCGA CGCCCAACCT GCCATCACGA
3001    GATTTCGATT CCACCGCCGC CTTCTATGAA AGGTTGGGCT TCGGAATCGT TTTCCGGGAC
3061    GCCGGCTGGA TGATCCTCCA GCGCGGGGAT CTCATGCTGG AGTTCTTCGC CCACCCCAAC
3121    TTGTTTATTG CAGCTTATAA TGGTTACAAA TAAAGCAATA GCATCACAAA TTTCACAAAT
3181    AAAGCATTTT TTTCACTGCA TTCTAGTTGT GGTTTGTCCA AACTCATCAA TGTATCTTAT
3241    CATGTCTGTA TACCGTCGAC CTCTAGCTAG AGCTTGGCGT AATCATGGTC ATAGCTGTTT
3301    CCTGTGTGAA ATTGTTATCC GCTCACAATT CCACACAACA TACGAGCCGG AAGCATAAAG
3361    TGTAAAGCCT GGGGTGCCTA ATGAGTGAGC TAACTCACAT TAATTGCGTT GCGCTCACTG
3421    CCCGCTTTCC AGTCGGGAAA CCTGTCGTGC CAGCTGCATT AATGAATCGG CCAACGCGCG
3481    GGGAGAGGCG GTTTGCGTAT TGGGCGCTCT TCCGCTTCCT CGCTCACTGA CTCGCTGCGC
3541    TCGGTCGTTC GGCTGCGGCG AGCGGTATCA GCTCACTCAA AGGCGGTAAT ACGGTTATCC
3601    ACAGAATCAG GGGATAACGC AGGAAAGAAC ATGTGAGCAA AAGGCCAGCA AAAGGCCAGG
3661    AACCGTAAAA AGGCCGCGTT GCTGGCGTTT TTCCATAGGC TCCGCCCCCC TGACGAGCAT
3721    CACAAAAATC GACGCTCAAG TCAGAGGTGG CGAAACCCGA CAGGACTATA AAGATACCAG
3781    GCGTTTCCCC CTGGAAGCTC CCTCGTGCGC TCTCCTGTTC CGACCCTGCC GCTTACCGGA
3841    TACCTGTCCG CCTTTCTCCC TTCGGGAAGC GTGGCGCTTT CTCATAGCTC ACGCTGTAGG
3901    TATCTCAGTT CGGTGTAGGT CGTTCGCTCC AAGCTGGGCT GTGTGCACGA ACCCCCCGTT
3961    CAGCCCGACC GCTGCGCCTT ATCCGGTAAC TATCGTCTTG AGTCCAACCC GGTAAGACAC
4021    GACTTATCGC CACTGGCAGC AGCCACTGGT AACAGGATTA GCAGAGCGAG GTATGTAGGC
4081    GGTGCTACAG AGTTCTTGAA GTGGTGGCCT AACTACGGCT ACACTAGAAG AACAGTATTT
4141    GGTATCTGCG CTCTGCTGAA GCCAGTTACC TTCGGAAAAA GAGTTGGTAG CTCTTGATCC
4201    GGCAAACAAA CCACCGCTGG TAGCGGTTTT TTTGTTTGCA AGCAGCAGAT TACGCGCAGA
4261    AAAAAAGGAT CTCAAGAAGA TCCTTTGATC TTTTCTACGG GGTCTGACGC TCAGTGGAAC
4321    GAAAACTCAC GTTAAGGGAT TTTGGTCATG AGATTATCAA AAAGGATCTT CACCTAGATC
4381    CTTTTAAATT AAAAATGAAG TTTTAAATCA ATCTAAAGTA TATATGAGTA AACTTGGTCT
4441    GACAGTTACC AATGCTTAAT CAGTGAGGCA CCTATCTCAG CGATCTGTCT ATTTCGTTCA
4501    TCCATAGTTG CCTGACTCCC CGTCGTGTAG ATAACTACGA TACGGGAGGG CTTACCATCT
4561    GGCCCCAGTG CTGCAATGAT ACCGCGAGAC CCACGCTCAC CGGCTCCAGA TTTATCAGCA
4621    ATAAACCAGC CAGCCGGAAG GGCCGAGCGC AGAAGTGGTC CTGCAACTTT ATCCGCCTCC
4681    ATCCAGTCTA TTAATTGTTG CCGGGAAGCT AGAGTAAGTA GTTCGCCAGT TAATAGTTTG
4741    CGCAACGTTG TTGCCATTGC TACAGGCATC GTGGTGTCAC GCTCGTCGTT TGGTATGGCT
4801    TCATTCAGCT CCGGTTCCCA ACGATCAAGG CGAGTTACAT GATCCCCCAT GTTGTGCAAA
4861    AAAGCGGTTA GCTCCTTCGG TCCTCCGATC GTTGTCAGAA GTAAGTTGGC CGCAGTGTTA
4921    TCACTCATGG TTATGGCAGC ACTGCATAAT TCTCTTACTG TCATGCCATC CGTAAGATGC
4981    TTTTCTGTGA CTGGTGAGTA CTCAACCAAG TCATTCTGAG AATAGTGTAT GCGGCGACCG
5041    AGTTGCTCTT GCCCGGCGTC AATACGGGAT AATACCGCGC CACATAGCAG AACTTTAAAA
5101    GTGCTCATCA TTGGAAAACG TTCTTCGGGG CGAAAACTCT CAAGGATCTT ACCGCTGTTG
5161    AGATCCAGTT CGATGTAACC CACTCGTGCA CCCAACTGAT CTTCAGCATC TTTTACTTTC
```

```
5221    ACCAGCGTTT CTGGGTGAGC AAAAACAGGA AGGCAAAATG CCGCAAAAAA GGGAATAAGG
5281    GCGACACGGA AATGTTGAAT ACTCATACTC TTCCTTTTTC AATATTATTG AAGCATTTAT
5341    CAGGGTTATT GTCTCATGAG CGGATACATA TTTGAATGTA TTTAGAAAAA TAAACAAATA
5401    GGGGTTCCGC GCACATTTCC CCGAAAAGTG CCACCTGACG TC
```

pcDNA3.1 puro

```
1       GACGGATCGG GAGATCTCCC GATCCCCTAT GGTGCACTCT CAGTACAATC TGCTCTGATG
61      CCGCATAGTT AAGCCAGTAT CTGCTCCCTG CTTGTGTGTT GGAGGTCGCT GAGTAGTGCG
121     CGAGCAAAAT TTAAGCTACA ACAAGGCAAG GCTTGACCGA CAATTGCATG AAGAATCTGC
181     TTAGGGTTAG GCGTTTTGCG CTGCTTCGCG ATGTACGGGC CAGATATACG CGTTGACATT
241     GATTATTGAC TAGTTATTAA TAGTAATCAA TTACGGGGTC ATTAGTTCAT AGCCCATATA
301     TGGAGTTCCG CGTTACATAA CTTACGGTAA ATGGCCCGCC TGGCTGACCG CCCAACGACC
361     CCCGCCCATT GACGTCAATA ATGACGTATG TTCCCATAGT AACGCCAATA GGGACTTTCC
421     ATTGACGTCA ATGGGTGGAG TATTTACGGT AAACTGCCCA CTTGGCAGTA CATCAAGTGT
481     ATCATATGCC AAGTACGCCC CCTATTGACG TCAATGACGG TAAATGGCCC GCCTGGCATT
541     ATGCCCAGTA CATGACCTTA TGGGACTTTC CTACTTGGCA GTACATCTAC GTATTAGTCA
601     TCGCTATTAC CATGGTGATG CGGTTTTGGC AGTACATCAA TGGGCGTGGA TAGCGGTTTG
661     ACTCACGGGG ATTTCCAAGT CTCCACCCCA TTGACGTCAA TGGGAGTTTG TTTTGGCACC
721     AAAATCAACG GGACTTTCCA AAATGTCGTA ACAACTCCGC CCCATTGACG CAAATGGGCG
781     GTAGGCGTGT ACGGTGGGAG GTCTATATAA GCAGAGCTCT CTGGCTAACT AGAGAACCCA
841     CTGCTTACTG GCTTATCGAA ATTAATACGA CTCACTATAG GGAGACCCAA GCTTGGTACC
901     GAGCTCGGAT CCACTAGTAA CGGCCGCCAG TGTGCTGGAA TTCTGCAGAT ATCCATCACA
961     CTGGCGGCCG CTCGAGCATG CATCTAGAGG GCCCTATTCT ATAGTGTCAC CTAAATGCTA
1021    GAGCTCGCTG ATCAGCCTCG ACTGTGCCTT CTAGTTGCCA GCCATCTGTT GTTTGCCCCT
1081    CCCCCGTGCC TTCCTTGACC CTGGAAGGTG CCACTCCCAC TGTCCTTTCC TAATAAAATG
1141    AGGAAATTGC ATCGCATTGT CTGAGTAGGT GTCATTCTAT TCTGGGGGGT GGGGTGGGGC
1201    AGGACAGCAA GGGGGAGGAT TGGGAAGACA ATAGCAGGCA TGCTGGGGAT GCGGTGGGCT
1261    CTATGGCTTC TGAGGCGGAA AGAACCAGCT GGGGCTCTAG GGGGTATCCC CACGCGCCCT
1321    GTAGCGGCGC ATTAAGCGCG GCGGGTGTGG TGGTTACGCG CAGCGTGACC GCTACACTTG
1381    CCAGCGCCCT AGCGCCCGCT CCTTTCGCTT TCTTCCCTTC CTTTCTCGCC ACGTTCGCCG
1441    GCTTTCCCCG TCAAGCTCTA AATCGGGGGC TCCCTTTAGG GTTCCGATTT AGTGCTTTAC
1501    GGCACCTCGA CCCCAAAAAA CTTGATTAGG GTGATGGTTC ACGTAGTGGG CCATCGCCCT
1561    GATAGACGGT TTTTCGCCCT TTGACGTTGG AGTCCACGTT CTTTAATAGT GGACTCTTGT
1621    TCCAAACTGG AACAACACTC AACCCTATCT CGGTCTATTC TTTTGATTTA TAAGGGATTT
1681    TGCCGATTTC GGCCTATTGG TTAAAAAATG AGCTGATTTA ACAAAAATTT AACGCGAATT
1741    AATTCTGTGG AATGTGTGTC AGTTAGGGTG TGGAAAGTCC CCAGGCTCCC CAGCAGGCAG
1801    AAGTATGCAA AGCATGCATC TCAATTAGTC AGCAACCAGG TGTGGAAAGT CCCCAGGCTC
1861    CCCAGCAGGC AGAAGTATGC AAAGCATGCA TCTCAATTAG TCAGCAACCA TAGTCCCGCC
1921    CCTAACTCCG CCCATCCCGC CCCTAACTCC GCCCAGTTCC GCCCATTCTC CGCCCCATGG
1981    CTGACTAATT TTTTTTATTT ATGCAGAGGC CGAGGCCGCC TCTGCCTCTG AGCTATTCCA
2041    GAAGTAGTGA GGAGGCTTTT TTGGAGGCCT AGGCTTTTGC AAAAAGCTCC CATGACCGAG
2101    TACAAGCCCA CGGTGCGCCT CGCCACCCGC GACGACGTCC CCAGGGCCGT ACGCACCCTC
2161    GCCGCCGCGT TCGCCGACTA CCCCGCCACG CGCCACACCG TCGATCCGGA CCGCCACATC
2221    GAGCGGGTCA CCGAGCTGCA AGAACTCTTC CTCACGCGCG TCGGGCTCGA CATCGGCAAG
2281    GTGTGGGTCG CGGACGACGG CGCCGCGGTG GCGGTCTGGA CCACGCCGGA GAGCGTCGAA
2341    GCGGGGGCGG TGTTCGCCGA GATCGGCCCG CGCATGGCCG AGTTGAGCGG TTCCCGGCTG
2401    GCCGCGCAGC AACAGATGGA AGGCCTCCTG GCGCCGCACC GGCCCAAGGA GCCCGCGTGG
2461    TTCCTGGCCA CCGTCGGCGT CTCGCCCGAC CACCAGGGCA AGGGTCTGGG CAGCGCCGTC
2521    GTGCTCCCCG GAGTGGAGGC GGCCGAGCGC GCCGGGGTGC CCGCCTTCCT GGAGACCTCC
2581    GCGCCCCGCA ACCTCCCCTT CTACGAGCGG CTCGGCTTCA CCGTCACCGC CGACGTCGAG
2641    GTGCCCGAAG GACCGCGCAC CTGGTGCATG ACCCGCAAGC CCGGTGCCTG AGCGGGACTC
2701    TGGGGTTCGA AATGACCGAC CAAGCGACGC CCGAAATGAC CGACCAAGCG ACGCCCAACC
2761    TGCCATCACG AGATTTCGAT TCCACCGCCG CCTTCTATGA AAGGTTGGGC TTCGGAATCG
2821    TTTTCCGGGA CGCCGGCTGG ATGATCCTCC AGCGCGGGGA TCTCATGCTG GAGTTCTTCG
2881    CCCACCCCAA CTTGTTTATT GCAGCTTATA ATGGTTACAA ATAAAGCAAT AGCATCACAA
2941    ATTTCACAAA TAAAGCATTT TTTTCACTGC ATTCTAGTTG TGGTTTGTCC AAACTCATCA
3001    ATGTATCTTA TCATGTCTGT ATACCGTCGA CCTCTAGCTA GAGCTTGGCG TAATCATGGT
3061    CATAGCTGTT TCCTGTGTGA AATTGTTATC CGCTCACAAT TCCACACAAC ATACGAGCCG
3121    GAAGCATAAA GTGTAAAGCC TGGGGTGCCT AATGAGTGAG CTAACTCACA TTAATTGCGT
3181    TGCGCTCACT GCCCGCTTTC CAGTCGGGAA ACCTGTCGTG CCAGCTGCAT TAATGAATCG
3241    GCCAACGCGC GGGGAGAGGC GGTTTGCGTA TTGGGCGCTC TTCCGCTTCC TCGCTCACTG
3301    ACTCGCTGCG CTCGGTCGTT CGGCTGCGGC GAGCGGTATC AGCTCACTCA AAGGCGGTAA
3361    TACGGTTATC CACAGAATCA GGGGATAACG CAGGAAAGAA CATGTGAGCA AAAGGCCAGC
3421    AAAAGGCCAG GAACCGTAAA AAGGCCGCGT TGCTGGCGTT TTTCCATAGG CTCCGCCCCC
3481    CTGACGAGCA TCACAAAAAT CGACGCTCAA GTCAGAGGTG GCGAAACCCG ACAGGACTAT
3541    AAAGATACCA GGCGTTTCCC CCTGGAAGCT CCCTCGTGCG CTCTCCTGTT CCGACCCTGC
3601    CGCTTACCGG ATACCTGTCC GCCTTTCTCC CTTCGGGAAG CGTGGCGCTT TCTCATAGCT
```

```
3661    CACGCTGTAG GTATCTCAGT TCGGTGTAGG TCGTTCGCTC CAAGCTGGGC TGTGTGCACG
3721    AACCCCCCGT TCAGCCCGAC CGCTGCGCCT TATCCGGTAA CTATCGTCTT GAGTCCAACC
3781    CGGTAAGACA CGACTTATCG CCACTGGCAG CAGCCACTGG TAACAGGATT AGCAGAGCGA
3841    GGTATGTAGG CGGTGCTACA GAGTTCTTGA AGTGGTGGCC TAACTACGGC TACACTAGAA
3901    GAACAGTATT TGGTATCTGC GCTCTGCTGA AGCCAGTTAC CTTCGGAAAA AGAGTTGGTA
3961    GCTCTTGATC CGGCAAACAA ACCACCGCTG GTAGCGGTTT TTTTGTTTGC AAGCAGCAGA
4021    TTACGCGCAG AAAAAAAGGA TCTCAAGAAG ATCCTTTGAT CTTTTCTACG GGGTCTGACG
4081    CTCAGTGGAA CGAAAACTCA CGTTAAGGGA TTTTGGTCAT GAGATTATCA AAAAGGATCT
4141    TCACCTAGAT CCTTTTAAAT TAAAAATGAA GTTTTAAATC AATCTAAAGT ATATATGAGT
4201    AAACTTGGTC TGACAGTTAC CAATGCTTAA TCAGTGAGGC ACCTATCTCA GCGATCTGTC
4261    TATTTCGTTC ATCCATAGTT GCCTGACTCC CCGTCGTGTA GATAACTACG ATACGGGAGG
4321    GCTTACCATC TGGCCCCAGT GCTGCAATGA TACCGCGAGA CCCACGCTCA CCGGCTCCAG
4381    ATTTATCAGC AATAAACCAG CCAGCCGGAA GGGCCGAGCG CAGAAGTGGT CCTGCAACTT
4441    TATCCGCCTC CATCCAGTCT ATTAATTGTT GCCGGGAAGC TAGAGTAAGT AGTTCGCCAG
4501    TTAATAGTTT GCGCAACGTT GTTGCCATTG CTACAGGCAT CGTGGTGTCA CGCTCGTCGT
4561    TTGGTATGGC TTCATTCAGC TCCGGTTCCC AACGATCAAG GCGAGTTACA TGATCCCCCA
4621    TGTTGTGCAA AAAAGCGGTT AGCTCCTTCG GTCCTCCGAT CGTTGTCAGA AGTAAGTTGG
4681    CCGCAGTGTT ATCACTCATG GTTATGGCAG CACTGCATAA TTCTCTTACT GTCATGCCAT
4741    CCGTAAGATG CTTTTCTGTG ACTGGTGAGT ACTCAACCAA GTCATTCTGA GAATAGTGTA
4801    TGCGGCGACC GAGTTGCTCT TGCCCGGCGT CAATACGGGA TAATACCGCG CCACATAGCA
4861    GAACTTTAAA AGTGCTCATC ATTGGAAAAC GTTCTTCGGG GCGAAAACTC TCAAGGATCT
4921    TACCGCTGTT GAGATCCAGT TCGATGTAAC CCACTCGTGC ACCCAACTGA TCTTCAGCAT
4981    CTTTTACTTT CACCAGCGTT TCTGGGTGAG CAAAAACAGG AAGGCAAAAT GCCGCAAAAA
5041    AGGGAATAAG GGCGACACGG AAATGTTGAA TACTCATACT CTTCCTTTTT CAATATTATT
5101    GAAGCATTTA TCAGGGTTAT TGTCTCATGA GCGGATACAT ATTTGAATGT ATTTAGAAAA
5161    ATAAACAAAT AGGGGTTCCG CGCACATTTC CCCGAAAAGT GCCACCTGAC GTC
```

pcDNA6v5Hismyc

```
1       GACGGATCGG GAGATCTCCC GATCCCCTAT GGTGCACTCT CAGTACAATC TGCTCTGATG
61      CCGCATAGTT AAGCCAGTAT CTGCTCCCTG CTTGTGTGTT GGAGGTCGCT GAGTAGTGCG
121     CGAGCAAAAT TTAAGCTACA ACAAGGCAAG GCTTGACCGA CAATTGCATG AAGAATCTGC
181     TTAGGGTTAG CGTTTTGCG CTGCTTCGCG ATGTACGGGC CAGATATACG CGTTGACATT
241     GATTATTGAC TAGTTATTAA TAGTAATCAA TTACGGGGTC ATTAGTTCAT AGCCCATATA
301     TGGAGTTCCG CGTTACATAA CTTACGGTAA ATGGCCCGCC TGGCTGACCG CCCAACGACC
361     CCCGCCCATT GACGTCAATA ATGACGTATG TTCCCATAGT AACGCCAATA GGGACTTTCC
421     ATTGACGTCA ATGGGTGGAG TATTTACGGT AAACTGCCCA CTTGGCAGTA CATCAAGTGT
481     ATCATATGCC AAGTACGCCC CCTATTGACG TCAATGACGG TAAATGGCCC GCCTGGCATT
541     ATGCCCAGTA CATGACCTTA TGGGACTTTC CTACTTGGCA GTACATCTAC GTATTAGTCA
601     TCGCTATTAC CATGGTGATG CGGTTTTGGC AGTACATCAA TGGGCGTGGA TAGCGGTTTG
661     ACTCACGGGG ATTTCCAAGT CTCCACCCCA TTGACGTCAA TGGGAGTTTG TTTTGGCACC
721     AAAATCAACG GGACTTTCCA AAATGTCGTA ACAACTCCGC CCCATTGACG CAAATGGGCG
781     GTAGGCGTGT ACGGTGGGAG GTCTATATAA GCAGAGCTCT CTGGCTAACT AGAGAACCCA
841     CTGCTTACTG GCTTATCGAA ATTAATACGA CTCACTATAG GGAGACCCAA GCTGGCTAGC
901     GTTTAAACTT AAGCTTGGTA CCGAGCTCGG ATCCACTAGT CCAGTGTGGT GGAATTCTGC
961     AGATATCCAG CACAGTGGCG GCCGCTCGAG TCTAGAGGGC CCTTCGAAGG TAAGCCTATC
1021    CCTAACCCTC TCCTCGGTCT CGATTCTACG CGTACCGGTC ATCATCACCA TCACCATTGA
1081    GTTTAAACCC GCTGATCAGC CTCGACTGTG CCTTCTAGTT GCCAGCCATC TGTTGTTTGC
1141    CCCTCCCCCG TGCCTTCCTT GACCCTGGAA GGTGCCACTC CCACTGTCCT TTCCTAATAA
1201    AATGAGGAAA TTGCATCGCA TTGTCTGAGT AGGTGTCATT CTATTCTGGG GGGTGGGGTG
1261    GGGCAGGACA GCAAGGGGGA GGATTGGGAA GACAATAGCA GGCATGCTGG GGATGCGGTG
1321    GGCTCTATGG CTTCTGAGGC GGAAAGAACC AGCTGGGGCT CTAGGGGGTA TCCCCACGCG
1381    CCCTGTAGCG GCGCATTAAG CGCGGCGGGT GTGGTGGTTA CGCGCAGCGT GACCGCTACA
1441    CTTGCCAGCG CCCTAGCGCC CGCTCCTTTC GCTTTCTTCC CTTCCTTTCT CGCCACGTTC
1501    GCCGGCTTTC CCCGTCAAGC TCTAAATCGG GGGCTCCCTT TAGGGTTCCG ATTTAGTGCT
1561    TTACGGCACC TCGACCCCAA AAAACTTGAT TAGGGTGATG GTTCACGTAG TGGGCCATCG
1621    CCCTGATAGA CGGTTTTTCG CCCTTTGACG TTGGAGTCCA CGTTCTTTAA TAGTGGACTC
1681    TTGTTCCAAA CTGGAACAAC ACTCAACCCT ATCTCGGTCT ATTCTTTTGA TTTATAAGGG
1741    ATTTTGCCGA TTTCGGCCTA TTGGTTAAAA AATGAGCTGA TTTAACAAAA ATTTAACGCG
1801    AATTAATTCT GTGGAATGTG TGTCAGTTAG GGTGTGGAAA GTCCCCAGGC TCCCCAGCAG
1861    GCAGAAGTAT GCAAAGCATG CATCTCAATT AGTCAGCAAC CAGGTGTGGA AAGTCCCCAG
1921    GCTCCCCAGC AGGCAGAAGT ATGCAAAGCA TGCATCTCAA TTAGTCAGCA ACCATAGTCC
1981    CGCCCCTAAC TCCGCCCATC CCGCCCCTAA CTCCGCCCAG TTCCGCCCAT TCTCCGCCCC
2041    ATGGCTGACT AATTTTTTTT ATTTATGCAG AGGCCGAGGC CGCCTCTGCC TCTGAGCTAT
2101    TCCAGAAGTA GTGAGGAGGC TTTTTTGGAG GCCTAGGCTT TTGCAAAAAG CTCCCGGGAG
2161    CTTGTATATC CATTTTCGGA TCTGATCAGC ACGTGTTGAC AATTAATCAT CGGCATAGTA
2221    TATCGGCATA GTATAATACG ACAAGGTGAG GAACTAAACC ATGGCCAAGC CTTTGTCTCA
2281    AGAAGAATCC ACCCTCATTG AAAGAGCAAC GGCTACAATC AACAGCATCC CCATCTCTGA
```

```
2341    AGACTACAGC GTCGCCAGCG CAGCTCTCTC TAGCGACGGC CGCATCTTCA CTGGTGTCAA
2401    TGTATATCAT TTTACTGGGG GACCTTGTGC AGAACTCGTG GTGCTGGGCA CTGCTGCTGC
2461    TGCGGCAGCT GGCAACCTGA CTTGTATCGT CGCGATCGGA AATGAGAACA GGGGCATCTT
2521    GAGCCCCTGC GGACGGTGCC GACAGGTGCT TCTCGATCTG CATCCTGGGA TCAAAGCCAT
2581    AGTGAAGGAC AGTGATGGAC AGCCGACGGC AGTTGGGATT CGTGAATTGC TGCCCTCTGG
2641    TTATGTGTGG GAGGGCTAAG CACTTCGTGG CCGAGGAGCA GGACTGACAC GTGCTACGAG
2701    ATTTCGATTC CACCGCCGCC TTCTATGAAA GGTTGGGCTT CGGAATCGTT TTCCGGGACG
2761    CCGGCTGGAT GATCCTCCAG CGCGGGGATC TCATGCTGGA GTTCTTCGCC CACCCCAACT
2821    TGTTTATTGC AGCTTATAAT GGTTACAAAT AAAGCAATAG CATCACAAAT TTCACAAATA
2881    AAGCATTTTT TTCACTGCAT TCTAGTTGTG GTTTGTCCAA ACTCATCAAT GTATCTTATC
2941    ATGTCTGTAT ACCGTCGACC TCTAGCTAGA GCTTGGCGTA ATCATGGTCA TAGCTGTTTC
3001    CTGTGTGAAA TTGTTATCCG CTCACAATTC CACACAACAT ACGAGCCGGA AGCATAAAGT
3061    GTAAAGCCTG GGGTGCCTAA TGAGTGAGCT AACTCACATT AATTGCGTTG CGCTCACTGC
3121    CCGCTTTCCA GTCGGGAAAC CTGTCGTGCC AGCTGCATTA ATGAATCGGC CAACGCGCGG
3181    GGAGAGGCGG TTTGCGTATT GGGCGCTCTT CCGCTTCCTC GCTCACTGAC TCGCTGCGCT
3241    CGGTCGTTCG GCTGCGGCGA GCGGTATCAG CTCACTCAAA GGCGGTAATA CGGTTATCCA
3301    CAGAATCAGG GGATAACGCA GGAAAGAACA TGTGAGCAAA AGGCCAGCAA AAGGCCAGGA
3361    ACCGTAAAAA GGCCGCGTTG CTGGCGTTTT TCCATAGGCT CCGCCCCCCT GACGAGCATC
3421    ACAAAAATCG ACGCTCAAGT CAGAGGTGGC GAAACCCGAC AGGACTATAA AGATACCAGG
3481    CGTTTCCCCC TGGAAGCTCC CTCGTGCGCT CTCCTGTTCC GACCCTGCCG CTTACCGGAT
3541    ACCTGTCCGC CTTTCTCCCT TCGGGAAGCG TGGCGCTTTC TCATAGCTCA CGCTGTAGGT
3601    ATCTCAGTTC GGTGTAGGTC GTTCGCTCCA AGCTGGGCTG TGTGCACGAA CCCCCCGTTC
3661    AGCCCGACCG CTGCGCCTTA TCCGGTAACT ATCGTCTTGA GTCCAACCCG GTAAGACACG
3721    ACTTATCGCC ACTGGCAGCA GCCACTGGTA ACAGGATTAG CAGAGCGAGG TATGTAGGCG
3781    GTGCTACAGA GTTCTTGAAG TGGTGGCCTA ACTACGGCTA CACTAGAAGA ACAGTATTTG
3841    GTATCTGCGC TCTGCTGAAG CCAGTTACCT TCGGAAAAAG AGTTGGTAGC TCTTGATCCG
3901    GCAAACAAAC CACCGCTGGT AGCGGTGGTT TTTTTGTTTG CAAGCAGCAG ATTACGCGCA
3961    GAAAAAAAGG ATCTCAAGAA GATCCTTTGA TCTTTTCTAC GGGGTCTGAC GCTCAGTGGA
4021    ACGAAAACTC ACGTTAAGGG ATTTTGGTCA TGAGATTATC AAAAAGGATC TTCACCTAGA
4081    TCCTTTTAAA TTAAAAATGA AGTTTTAAAT CAATCTAAAG TATATATGAG TAAACTTGGT
4141    CTGACAGTTA CCAATGCTTA ATCAGTGAGG CACCTATCTC AGCGATCTGT CTATTTCGTT
4201    CATCCATAGT TGCCTGACTC CCCGTCGTGT AGATAACTAC GATACGGGAG GGCTTACCAT
4261    CTGGCCCCAG TGCTGCAATG ATACCGCGAG ACCCACGCTC ACCGGCTCCA GATTTATCAG
4321    CAATAAACCA GCCAGCCGGA AGGGCCGAGC GCAGAAGTGG TCCTGCAACT TTATCCGCCT
4381    CCATCCAGTC TATTAATTGT TGCCGGGAAG CTAGAGTAAG TAGTTCGCCA GTTAATAGTT
4441    TGCGCAACGT TGTTGCCATT GCTACAGGCA TCGTGGTGTC ACGCTCGTCG TTTGGTATGG
4501    CTTCATTCAG CTCCGGTTCC CAACGATCAA GGCGAGTTAC ATGATCCCCC ATGTTGTGCA
4561    AAAAAGCGGT TAGCTCCTTC GGTCCTCCGA TCGTTGTCAG AAGTAAGTTG GCCGCAGTGT
4621    TATCACTCAT GGTTATGGCA GCACTGCATA ATTCTCTTAC TGTCATGCCA TCCGTAAGAT
4681    GCTTTTCTGT GACTGGTGAG TACTCAACCA AGTCATTCTG AGAATAGTGT ATGCGGCGAC
4741    CGAGTTGCTC TTGCCCGGCG TCAATACGGG ATAATACCGC GCCACATAGC AGAACTTTAA
4801    AAGTGCTCAT CATTGGAAAA CGTTCTTCGG GGCGAAAACT CTCAAGGATC TTACCGCTGT
4861    TGAGATCCAG TTCGATGTAA CCCACTCGTG CACCCAACTG ATCTTCAGCA TCTTTTACTT
4921    TCACCAGCGT TTCTGGGTGA GCAAAAACAG GAAGGCAAAA TGCCGCAAAA AAGGGAATAA
4981    GGGCGACACG GAAATGTTGA ATACTCATAC TCTTCCTTTT TCAATATTAT TGAAGCATTT
5041    ATCAGGGTTA TTGTCTCATG AGCGGATACA TATTTGAATG TATTTAGAAA AATAAACAAA
5101    TAGGGGTTCC GCGCACATTT CCCCGAAAAG TGCCACCTGA CGTC
```

puc57

```
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 CCAGTGAATT CGAGCTCGGT ACCTCGCGAA
421     TGCATCTAGA TCCAATTTTC TCTTAAACAA CTCCAGACCA GTAAGTTACA TGTTTTTTTT
481     TTATTGGAAA GATGTTACAG ACTGATGTTT GTAAAAGTGT TTCTTTTAAC CGGTTTTGTA
541     CAATTCGTCC ATTCCCAATG TGATTCCAGC AGCTGTCACG AATTCCAACA ACACCATGTG
601     GTCTCTTTTT TCGTTTGGGT CTTTAGACAA AGCAGACTGT GTAGACAAGT AGTGGTTGTC
661     TGGCAACAAC ACTGGTCCGT CTCCGATTGG TGTGTTCTGC TGGTAGTGGT CAGCCAACTG
721     CACAGATCCG TCTTCGATGT TGTGTCTGAT TTTGAAGTTC ACTTTGATTC CGTTTTTCTG
781     TTTGTCAGCC ATGATGTACA CGTTGTGAGA GTTGTAGTTG TATTCCAATT TGTGTCCCAA
841     GATGTTTCCG TCTTCTTTGA AGTCGATTCC TTTCAATTCG ATTCTGTTCA CCAATGTGTC
901     TCCTTCGAAT TTCACTTCAG CTCTTGTTTT GTAGTTTCCG TCGTCTTTGA AGAAGATTGT
961     TCTTTCCTGC ACGTATCCTT CTGGCATAGC AGATTTGAAG AAGTCGTGCT GTTTCATGTG
1021    GTCTGGGTAT CTAGAGAAAC ACTGCACTCC GTATGTCAAT GTTGTCACCA ATGTTGGCCA
```

```
1081    TGGCACTGGC AATTTTCCTG TTGTACAGAT GAATTTCAAT GTCAATTTTC CGTATGTAGC
1141    GTCTCCTTCT CCTTCTCCAG ACACAGAGAA TTTGTGTCCG TTCACGTCTC CGTCCAATTC
1201    CACCAAGATT GGCACCACTC CTGTGAACAA TTCTTCTCCT TTAGACACCA GGCGCGCCAT
1261    TTTTTAAGTC TTCCAAGAGA CACGAAAACG TGTGCACACC AGAAAATTTC TCAAACAGAA
1321    AGAAAGATTT GGGGGCCACA GCTTCCTGAA CATGTGGGTT TGAGACACCA AAATAATTTC
1381    CACCTAAATT ATTAACAATT TGTTTATAAA TAGCCCTGTA AACATCAAAT CTCCCTATAG
1441    TGAGTCGTAT TAGGATCCCG GGCCCGTCGA CTGCAGAGGC CTGCATGCAA GCTTGGCGTA
1501    ATCATGGTCA TAGCTGTTTC CTGTGTGAAA TTGTTATCCG CTCACAATTC CACACAACAT
1561    ACGAGCCGGA AGCATAAAGT GTAAAGCCTG GGGTGCCTAA TGAGTGAGCT AACTCACATT
1621    AATTGCGTTG CGCTCACTGC CCGCTTTCCA GTCGGGAAAC CTGTCGTGCC AGCTGCATTA
1681    ATGAATCGGC CAACGCGCGG GGAGAGGCGG TTTGCGTATT GGGCGCTCTT CCGCTTCCTC
1741    GCTCACTGAC TCGCTGCGCT CGGTCGTTCG GCTGCGGCGA GCGGTATCAG CTCACTCAAA
1801    GGCGGTAATA CGGTTATCCA CAGAATCAGG GGATAACGCA GGAAAGAACA TGTGAGCAAA
1861    AGGCCAGCAA AAGGCCAGGA ACCGTAAAAA GGCCGCGTTG CTGGCGTTTT TCCATAGGCT
1921    CCGCCCCCCT GACGAGCATC ACAAAAATCG ACGCTCAAGT CAGAGGTGGC GAAACCCGAC
1981    AGGACTATAA AGATACCAGG CGTTTCCCCC TGGAAGCTCC CTCGTGCGCT CTCCTGTTCC
2041    GACCCTGCCG CTTACCGGAT ACCTGTCCGC CTTTCTCCCT TCGGGAAGCG TGGCGCTTTC
2101    TCATAGCTCA CGCTGTAGGT ATCTCAGTTC GGTGTAGGTC GTTCGCTCCA AGCTGGGCTG
2161    TGTGCACGAA CCCCCCGTTC AGCCCGACCG CTGCGCCTTA TCCGGTAACT ATCGTCTTGA
2221    GTCCAACCCG GTAAGACACG ACTTATCGCC ACTGGCAGCA GCCACTGGTA ACAGGATTAG
2281    CAGAGCGAGG TATGTAGGCG GTGCTACAGA GTTCTTGAAG TGGTGGCCTA ACTACGGCTA
2341    CACTAGAAGA ACAGTATTTG GTATCTGCGC TCTGCTGAAG CCAGTTACCT TCGGAAAAAG
2401    AGTTGGTAGC TCTTGATCCG GCAAACAAAC CACCGCTGGT AGCGGTGGTT TTTTTGTTTG
2461    CAAGCAGCAG ATTACGCGCA GAAAAAAAGG ATCTCAAGAA GATCCTTTGA TCTTTTCTAC
2521    GGGGTCTGAC GCTCAGTGGA ACGAAAACTC ACGTTAAGGG ATTTTGGTCA TGAGATTATC
2581    AAAAAGGATC TTCACCTAGA TCCTTTTAAA TTAAAAATGA AGTTTTAAAT CAATCTAAAG
2641    TATATATGAG TAAACTTGGT CTGACAGTTA CCAATGCTTA ATCAGTGAGG CACCTATCTC
2701    AGCGATCTGT CTATTTCGTT CATCCATAGT TGCCTGACTC CCCGTCGTGT AGATAACTAC
2761    GATACGGGAG GGCTTACCAT CTGGCCCCAG TGCTGCAATG ATACCGCGAG ACCCACGCTC
2821    ACCGGCTCCA GATTTATCAG CAATAAACCA GCCAGCCGGA AGGGCCGAGC GCAGAAGTGG
2881    TCCTGCAACT TTATCCGCCT CCATCCAGTC TATTAATTGT TGCCGGGAAG CTAGAGTAAG
2941    TAGTTCGCCA GTTAATAGTT TGCGCAACGT TGTTGCCATT GCTACAGGCA TCGTGGTGTC
3001    ACGCTCGTCG TTTGGTATGG CTTCATTCAG CTCCGGTTCC CAACGATCAA GGCGAGTTAC
3061    ATGATCCCCC ATGTTGTGCA AAAAAGCGGT TAGCTCCTTC GGTCCTCCGA TCGTTGTCAG
3121    AAGTAAGTTG GCCGCAGTGT TATCACTCAT GGTTATGGCA GCACTGCATA ATTCTCTTAC
3181    TGTCATGCCA TCCGTAAGAT GCTTTTCTGT GACTGGTGAG TACTCAACCA AGTCATTCTG
3241    AGAATAGTGT ATGCGGCGAC CGAGTTGCTC TTGCCCGGCG TCAATACGGG ATAATACCGC
3301    GCCACATAGC AGAACTTTAA AAGTGCTCAT CATTGGAAAA CGTTCTTCGG GGCGAAAACT
3361    CTCAAGGATC TTACCGCTGT TGAGATCCAG TTCGATGTAA CCCACTCGTG CACCCAACTG
3421    ATCTTCAGCA TCTTTTACTT TCACCAGCGT TTCTGGGTGA GCAAAAACAG GAAGGCAAAA
3481    TGCCGCAAAA AAGGGAATAA GGGCGACACG GAAATGTTGA ATACTCATAC TCTTCCTTTT
3541    TCAATATTAT TGAAGCATTT ATCAGGGTTA TTGTCTCATG AGCGGATACA TATTTGAATG
3601    TATTTAGAAA AATAAACAAA TAGGGGTTCC GCGCACATTT CCCCGAAAAG TGCCACCTGA
3661    CGTCTAAGAA ACCATTATTA TCATGACATT AACCTATAAA AATAGGCGTA TCACGAGGCC
3721    CTTTCGTC
```

p*Silencer*

```
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 GCTTTCCAAA AAAGTGGTGG
421     GATACTATAA CAACATTTAG CATATTGTTA TACTATTCCA CGGATCCCGC GTCCTTTCCA
481     CAAGATATAT AAACCCAAGA AATCGAAATA CTTTCAAGTT ACGGTAAGCA TATGATAGTC
541     CATTTTAAAA CATAATTTTA AAACTGCAAA CTACCCAAGA AATTATTACT TTCTACGTCA
601     CGTATTTTGT ACTAATATCT TTGTGTTTAC AGTCAAATTA ATTCTAATTA TCTCTCTAAC
661     AGCCTTGTAT CGTATATGCA AATATGAAGG AATCATGGGA AATAGGCCCT CTTCCTGCCC
721     GACCTTGGCG CGCGCTCGGC GCGCGGTCAC GCTCCGTCAC GTGGTGCGTT TTGCCTGCGC
781     GTCTTTCCAC TGGGGAATTC ATGCTTCTCC TCCCTTTAGT GAGGGTAATT CTCTCTCTCT
841     CCCTATAGTG AGTCGTATTA ATTCCTTCTC TTCTATAGTG TCACCTAAAT CGTTGCAATT
901     CGTAATCATG TCATAGCTGT TTCCTGTGTG AAATTGTTAT CCGCTCACAA TTCCACACAA
961     CATACGAGCC GGAAGCATAA AGTGTAAAGC CTGGGGTGCC TAATGAGTGA GCTAACTCAC
1021    ATTAATTGCG TTGCGCTCAC TGCCCGCTTT CCAGTCGGGA AACCTGTCGT GCCAGCTGCA
1081    TTAATGAATC GGCCAACGCG CGGGGAGAGG CGGTTTGCGT ATTGGGCGCT CTTCCGCTTC
1141    CTCGCTCACT GACTCGCTGC GCTCGGTCGT TCGGCTGCGG CGAGCGGTAT CAGCTCACTC
```

```
1201    AAAGGCGGTA ATACGGTTAT CCACAGAATC AGGGGATAAC GCAGGAAAGA ACATGTGAGC
1261    AAAAGGCCAG CAAAAGGCCA GGAACCGTAA AAAGGCCGCG TTGCTGGCGT TTTTCCATAG
1321    GCTCCGCCCC CCTGACGAGC ATCACAAAAA TCGACGCTCA AGTCAGAGGT GGCGAAACCC
1381    GACAGGACTA TAAAGATACC AGGCGTTTCC CCCTGGAAGC TCCCTCGTGC GCTCTCCTGT
1441    TCCGACCCTG CCGCTTACCG GATACCTGTC CGCCTTTCTC CCTTCGGGAA GCGTGGCGCT
1501    TTCTCATAGC TCACGCTGTA GGTATCTCAG TTCGGTGTAG GTCGTTCGCT CCAAGCTGGG
1561    CTGTGTGCAC GAACCCCCCG TTCAGCCCGA CCGCTGCGCC TTATCCGGTA ACTATCGTCT
1621    TGAGTCCAAC CCGGTAAGAC ACGACTTATC GCCACTGGCA GCAGCCACTG GTAACAGGAT
1681    TAGCAGAGCG AGGTATGTAG GCGGTGCTAC AGAGTTCTTG AAGTGGTGGC CTAACTACGG
1741    CTACACTAGA AGAACAGTAT TTGGTATCTG CGCTCTGCTG AAGCCAGTTA CCTTCGGAAA
1801    AAGAGTTGGT AGCTCTTGAT CCGGCAAAAA AACCACCGCT GGTAGCGGTG GTTTTTTTGT
1861    TTGCAAGCAG CAGATTACGC GCAGAAAAAA AGGATCTCAA GAAGATCCTT TGATCTTTTC
1921    TACGGGGTCT GACGCTCAGT GGAACGAAAA CTCACGTTAA GGGATTTTGG TCATGAGATT
1981    ATCAAAAAGG ATCTTCACCT AGATCCTTTT AAATTAAAAA TGAAGTTTTA AATCAATCTA
2041    AAGTATATAT GAGTAAACTT GGTCTGACAG TTACCAATGC TTAATCAGTG AGGCACCTAT
2101    CTCAGCGATC TGTCTATTTC GTTCATCCAT AGTTGCCTGA CTCCCCGTCG TGTAGATAAC
2161    TACGATACGG GAGGGCTTAC CATCTGGCCC CAGTGCTGCA ATGATACCGC GAGACCCACG
2221    CTCACCGGCT CCAGATTTAT CAGCAATAAA CCAGCCAGCC GGAAGGGCCG AGCGCAGAAG
2281    TGGTCCTGCA ACTTTATCCG CCTCCATCCA GTCTATTAAT TGTTGCCGGG AAGCTAGAGT
2341    AAGTAGTTCG CCAGTTAATA GTTTGCGCAA CGTTGTTGCC ATTGCTACAG GCATCGTGGT
2401    GTCACGCTCG TCGTTTGGTA TGGCTTCATT CAGCTCCGGT TCCCAACGAT CAAGGCGAGT
2461    TACATGATCC CCCATGTTGT GCAAAAAAGC GGTTAGCTCC TTCGGTCCTC CGATCGTTGT
2521    CAGAAGTAAG TTGGCCGCAG TGTTATCACT CATGGTTATG GCAGCACTGC ATAATTCTCT
2581    TACTGTCATG CCATCCGTAA GATGCTTTTC TGTGACTGGT GAGTACTCAA CCAAGTCATT
2641    CTGAGAATAG TGTATGCGGC GACCGAGTTG CTCTTGCCCG GCGTCAATAC GGGATAATAC
2701    CGCGCCACAT AGCAGAACTT TAAAAGTGCT CATCATTGGA AAACGTTCTT CGGGGCGAAA
2761    ACTCTCAAGG ATCTTACCGC TGTTGAGATC CAGTTCGATG TAACCCACTC GTGCACCCAA
2821    CTGATCTTCA GCATCTTTTA CTTTCACCAG CGTTTCTGGG TGAGCAAAAA CAGGAAGGCA
2881    AAATGCCGCA AAAAAGGGAA TAAGGGCGAC ACGGAAATGT TGAATACTCA TACTCTTCCT
2941    TTTTCAATAT TCAGGCTTAC CTTTATGCTT CCGGCTCGTA TGTTGTGTGG AATTGTGAGC
3001    GGATAACAAT TTCACACAGG AAACAGCTAT GACCATGATT ACGCCAAGCT CTAGCTAGAG
3061    GTCGACGGTA TACAGACATG ATAAGATACA TTGATGAGTT TGGACAAACC ACAACTAGAA
3121    TGCAGTGAAA AAAATGCTTT ATTTGTGAAA TTTGTGATGC TATTGCTTTA TTTGTAACCA
3181    TTATAAGCTG CAATAAACAA GTTGGGGTGG GCGAAGAACT CCAGCATGAG ATCCCCGCGC
3241    TGGAGGATCA TCCAGCCGGC GTCCCGGAAA ACGATTCCGA AGCCCAACCT TTCATAGAAG
3301    GCGGCGGTGG AATCGAAATC TCGTAGCACG TGCTATTCCT TTGCCCTCGG ACGAGTGCTG
3361    GGGCGTCGGT TTCCACTATC GGCGAGTACT TCTACACAGC CATCGGTCCA GACGGCCGCG
3421    CTTCTGCGGG CGATTTGTGT ACGCCCGACA GTCCCGGCTC CGGATCGGAC GATTGCGTCG
3481    CATCGACCCT GCGCCCAAGC TGCATCATCG AAATTGCCGT CAACCAAGCT CTGATAGAGT
3541    TGGTCAAGAC CAATGCGGAG CATATACGCC CGGAGCCGCG GCGATCCTGC AAGCTCCGGA
3601    TGCCTCCGCT CGAAGTAGCG CGTCTGCTGC TCCATACAAG CCAACCACGG CCTCCAGAAG
3661    AAGATGTTGG CGACCTCGTA TTGGGAATCC CCGAACATCG CCTCGCTCCA GTCAATGACC
3721    GCTGTTATGC GGCCATTGTC CGTCAGGACA TTGTTGGAGC CGAAATCCGC GTGCACGAGG
3781    TGCCGGACTT CGGGGCAGTC CTCGGCCCAA AGCATCAGCT CATCGAGAGC CTGCGCGACG
3841    GACGCACTGA CGGTGTCGTC CATCACAGTT TGCCAGTGAT ACACATGGGG ATCAGCAATC
3901    GCGCATATGA AATCACGCCA TGTAGTGTAT TGACCGATTC CTTGCGGTCC GAATGGGCCG
3961    AACCCGCTCG TCTGGCTAAG ATCGGCCGCA GCGATCGCAT CCATGGCCTC CGCGACCGGC
4021    TGCAGAACAG CGGGCAGTTC GGTTTCAGGC AGGTCTTGCA ACGTGACACC CTGTGCACGG
4081    CGGGAGATGC AATAGGTCAG GCTCTCGCTG AATTCCCCAA TGTCAAGCAC TTCCGGAATC
4141    GGGAGCGCGG CCGATGCAAA GTGCCGATAA ACATAACGAT CTTTGTAGAA ACCATCGGCG
4201    CAGCTATTTA CCCGCAGGAC ATATCCACGC CCTCCTACAT CGAAGCTGAA AGCACGAGAT
4261    TCTTCGCCCT CCGAGAGCTG CATCAGGTCG GAGACGCTGT CGAACTTTTC GATCAGAAAC
4321    TTCTCGACAG ACGTCGCGGT GAGTTCAGGC TTTTTCATCA CGTGCTGATC AGATCCGAAA
4381    ATGGATATAC AAGCTCCCGG GAGCTTTTTG CAAAAGCCTA GGCCTCCAAA AAAGCCTCCC
4441    CACTACTTCT GGAATAGCTC AGAGGCAGAG GCGGCCTCGG CCTCTGCATA AATAAAAAAA
4501    ATTAGTCAGC CATGGGGCGG AGAATGGGCG GAACTGGGCG GAGTTAGGGG CGGGATGGGC
4561    GGAGTTAGGG GCGGGACTAT GGTTGCTGAC TAATTGAGAT GCATGCTTTG CATACTTCTG
4621    CCTGCTGGGG AGCCTGGGGA CTTTCCACAC CTGGTTGCTA ACTAATTGAG ATGCATGCTT
4681    TGCATACTTC TGCCTGCTGG AATATTATTG AAGCATTTAT CAGGGTTATT GTCTCATGAG
4741    CGGATACATA TTTGAATGTA TTTAGAAAAA TAAACAAATA GGGGTTCCGC GCACATTTCC
4801    CCGAAAAGTG CCACCTGACG TCTAAGAAAC CATTATTATC ATGACATTAA CCTATAAAAA
4861    TAGGCGTATC ACGAGGCCCT TTCGTC
```

6.2 mRNAs sequences

The following table lists the mRNA sequences that were produced for PCR reaction tube experiments and/or expressed in cell culture (Table 6-2).

Table 6-2: mRNA sequences.

E66 eCFP mRNA

```
1       AUGGCUAGCA AAGGAGAAGA ACUCUUCACU GGAGUUGUCC CAAUUCUUGU UGAAUUAGAU
61      GGUGAUGUUA ACGGCCACAA GUUCUCUGUC AGUGGAGAGG GUGAAGGUGA UGCAACAUAC
121     GGAAAACUUA CCCUGAAGUU CAUCUGCACU ACUGGCAAAC UGCCUGUUCC GUGGCCAACA
181     CUAGUCACUA CUCUGUGCGA GGGUGUUCAA UGCUUUUCAA GAUACCCGGA UCACAUGAAA
241     CGGCAUGACU UUUUCAAGAG UGCCAUGCCC GAAGGUUAUG UACAGGAAAG GACCAUCUUC
301     UUCAAAGAUG ACGGCAACUA CAAGACACGU GCUGAAGUCA AGUUUGAAGG UGAUACCCUU
361     GUUAAUAGAA UCGAGUUAAA AGGUAUUGAC UUCAAGGAAG AUGGCAACAU UCUGGGACAC
421     AAAUUGGAAU ACAACUAUAU CUCACACAAU GUAUACAUCA CCGCAGACAA ACAAAAGAAU
481     GGAAUCAAAG CCCACUUCAA GACCCGCCAC AACAUUGAAG AUGGAAGCGU UCAACUAGCA
541     GACCAUUAUC AACAAAAUAC UCCAAUUGGC GAUGGCCCUG UCCUUUUACC AGACAACCAU
601     UACCUGUCCA CACAAUCUGC CCUUUCGAAA GAUCCCAACG AAAAGAGAGA CCACAUGGUC
661     CUUCUUGAGU UUGUAACAGC UGCUGGGAUU ACACAUGGCA UGGAUGAACU AUACAAAUCC
721     GGCGGCUCCA UGGCGCUCGA G
```

Q66 eCFP mRNA

```
1       AUGGCUAGCA AAGGAGAAGA ACUCUUCACU GGAGUUGUCC CAAUUCUUGU UGAAUUAGAU
61      GGUGAUGUUA ACGGCCACAA GUUCUCUGUC AGUGGAGAGG GUGAAGGUGA UGCAACAUAC
121     GGAAAACUUA CCCUGAAGUU CAUCUGCACU ACUGGCAAAC UGCCUGUUCC GUGGCCAACA
181     CUAGUCACUA CUCUGUGCCA GGGUGUUCAA UGCUUUUCAA GAUACCCGGA UCACAUGAAA
241     CGGCAUGACU UUUUCAAGAG UGCCAUGCCC GAAGGUUAUG UACAGGAAAG GACCAUCUUC
301     UUCAAAGAUG ACGGCAACUA CAAGACACGU GCUGAAGUCA AGUUUGAAGG UGAUACCCUU
361     GUUAAUAGAA UCGAGUUAAA AGGUAUUGAC UUCAAGGAAG AUGGCAACAU UCUGGGACAC
421     AAAUUGGAAU ACAACUAUAU CUCACACAAU GUAUACAUCA CCGCAGACAA ACAAAAGAAU
481     GGAAUCAAAG CCCACUUCAA GACCCGCCAC AACAUUGAAG AUGGAAGCGU UCAACUAGCA
541     GACCAUUAUC AACAAAAUAC UCCAAUUGGC GAUGGCCCUG UCCUUUUACC AGACAACCAU
601     UACCUGUCCA CACAAUCUGC CCUUUCGAAA GAUCCCAACG AAAAGAGAGA CCACAUGGUC
661     CUUCUUGAGU UUGUAACAGC UGCUGGGAUU ACACAUGGCA UGGAUGAACU AUACAAAUCC
721     GGCGGCUCCA UGGCGCUCGA G
```

Stop66 eCFP mRNA

```
1       AUGGCUAGCA AAGGAGAAGA ACUCUUCACU GGAGUUGUCC CAAUUCUUGU UGAAUUAGAU
61      GGUGAUGUUA ACGGCCACAA GUUCUCUGUC AGUGGAGAGG GUGAAGGUGA UGCAACAUAC
121     GGAAAACUUA CCCUGAAGUU CAUCUGCACU ACUGGCAAAC UGCCUGUUCC GUGGCCAACA
181     CUAGUCACUA CUCUGUGCUA GGGUGUUCAA UGCUUUUCAA GAUACCCGGA UCACAUGAAA
241     CGGCAUGACU UUUUCAAGAG UGCCAUGCCC GAAGGUUAUG UACAGGAAAG GACCAUCUUC
301     UUCAAAGAUG ACGGCAACUA CAAGACACGU GCUGAAGUCA AGUUUGAAGG UGAUACCCUU
361     GUUAAUAGAA UCGAGUUAAA AGGUAUUGAC UUCAAGGAAG AUGGCAACAU UCUGGGACAC
421     AAAUUGGAAU ACAACUAUAU CUCACACAAU GUAUACAUCA CCGCAGACAA ACAAAAGAAU
481     GGAAUCAAAG CCCACUUCAA GACCCGCCAC AACAUUGAAG AUGGAAGCGU UCAACUAGCA
541     GACCAUUAUC AACAAAAUAC UCCAAUUGGC GAUGGCCCUG UCCUUUUACC AGACAACCAU
601     UACCUGUCCA CACAAUCUGC CCUUUCGAAA GAUCCCAACG AAAAGAGAGA CCACAUGGUC
661     CUUCUUGAGU UUGUAACAGC UGCUGGGAUU ACACAUGGCA UGGAUGAACU AUACAAAUCC
721     GGCGGCUCCA UGGCGCUCGA G
```

mRNA+RG construct

```
1       AUGGCUAGCA AAGGAGAAGA ACUCUUCACU GGAGUUGUCC CAAUUCUUGU UGAAUUAGAU
61      GGUGAUGUUA ACGGCCACAA GUUCUCUGUC AGUGGAGAGG GUGAAGGUGA UGCAACAUAC
121     GGAAAACUUA CCCUGAAGUU CAUCUGCACU ACUGGCAAAC UGCCUGUUCC GUGGCCAACA
181     CUAGUCACUA CUCUGUGCUA GGGUGGUGGA AUAGUAUAAC AAUAUGCUAA AUGUUGUUAU
241     AGUAUCCCAC CACCCCAGCA CAGAGUUUCA AUGCUUUUCA AGAUACCCGG AUCACAUGAA
301     ACGGCAUGAC UUUUUCAAGA GUGCCAUGCC CGAAGGUUAU GUACAGGAAA GGACCAUCUU
361     CUUCAAAGAU GACGGCAACU ACAAGACACG UGCUGAAGUC AAGUUUGAAG GUGAUACCCU
421     UGUUAAUAGA AUCGAGUUAA AAGGUAUUGA CUUCAAGGAA GAUGGCAACA UUCUGGGACA
481     CAAAUUGGAA UACAACUAUA UCUCACACAA UGUAUACAUC ACCGCAGACA AACAAAAGAA
```

```
541     UGGAAUCAAA GCCCACUUCA AGACCCGCCA CAACAUUGAA GAUGGAAGCG UUCAACUAGC
601     AGACCAUUAU CAACAAAAUA CUCCAAUUGG CGAUGGCCCU GUCCUUUUAC CAGACAACCA
661     UUACCUGUCC ACACAAUCUG CCCUUUCGAA AGAUCCCAAC GAAAAGAGAG ACCACAUGGU
721     CCUUCUUGAG UUUGUAACAG CUGCUGGGAU UACACAUGGC AUGGAUGAAC UAUACAAAUC
781     CGGCGGCUCC AUGGCGCUCG AGCACCACCA CCACCACCAC UAAUAAUGAC UAGU
```

W58X eGFP mRNA

```
1       AUGGCUAGCA AAGGAGAAGA ACUCUUCACU GGAGUUGUCC CAAUUCUUGU UGAAUUAGAU
61      GGUGAUGUUA ACGGCCACAA GUUCUCUGUC AGUGGAGAGG GUGAAGGUGA UGCAACAUAC
121     GGAAAACUUA CCCUGAAGUU CAUCUGCACU ACUGGCAAAC UGCCUGUUCC GUAGCCGACA
181     CUAGUGACGA CGCUCUGCUA UGGCGUCCAG UGCUUUUCAA GAUACCCGGA UCACAUGAAA
241     CGGCAUGACU UUUUCAAGAG UGCCAUGCCC GAAGGUUAUG UACAGGAAAG GACCAUCUUC
301     UUCAAAGAUG ACGGCAACUA CAAGACACGU GCUGAAGUCA AGUUUGAAGG UGAUACCCUU
361     GUUAAUAGAA UCGAGUUAAA AGGUAUUGAC UUCAAGGAAG AUGGCAACAU UCUGGGACAC
421     AAAUUGGAAU ACAACUAUAA CUCACACAAU GUAUACAUCA UGGCAGACAA ACAAAAGAAU
481     GGAAUCAAAG UGAACUUCAA GACCCGCCAC AACAUUGAAG AUGGAAGCGU UCAACUAGCA
541     GACCAUUAUC AACAAAAUAC UCCAAUUGGC GAUGGCCCUG UCCUUUUACC AGACAACCAU
601     UACCUGUCCA CACAAUCUGC CCUUUCGAAA GAUCCCAACG AAAAGAGAGA CCACAUGGUC
661     CUUCUUGAGU UUGUAACAGC UGCUGGGAUU ACACAUGGCA UGGAUGAACU AUACAAAUCC
721     GGCGGCUCCA UGGCGCUCGA G
```

eGFP mRNA

```
1       AUGGCUAGCA AAGGAGAAGA ACUCUUCACU GGAGUUGUCC CAAUUCUUGU UGAAUUAGAU
61      GGUGAUGUUA ACGGCCACAA GUUCUCUGUC AGUGGAGAGG GUGAAGGUGA UGCAACAUAC
121     GGAAAACUUA CCCUGAAGUU CAUCUGCACU ACUGGCAAAC UGCCUGUUCC GTGGCCGACA
181     CUAGUGACGA CGCUCUGCUA UGGCGUCCAG UGCUUUUCAA GAUACCCGGA UCACAUGAAA
241     CGGCAUGACU UUUUCAAGAG UGCCAUGCCC GAAGGUUAUG UACAGGAAAG GACCAUCUUC
301     UUCAAAGAUG ACGGCAACUA CAAGACACGU GCUGAAGUCA AGUUUGAAGG UGAUACCCUU
361     GUUAAUAGAA UCGAGUUAAA AGGUAUUGAC UUCAAGGAAG AUGGCAACAU UCUGGGACAC
421     AAAUUGGAAU ACAACUAUAA CUCACACAAU GUAUACAUCA UGGCAGACAA ACAAAAGAAU
481     GGAAUCAAAG UGAACUUCAA GACCCGCCAC AACAUUGAAG AUGGAAGCGU UCAACUAGCA
541     GACCAUUAUC AACAAAAUAC UCCAAUUGGC GAUGGCCCUG UCCUUUUACC AGACAACCAU
601     UACCUGUCCA CACAAUCUGC CCUUUCGAAA GAUCCCAACG AAAAGAGAGA CCACAUGGUC
661     CUUCUUGAGU UUGUAACAGC UGCUGGGAUU ACACAUGGCA UGGAUGAACU AUACAAAUCC
721     GGCGGCUCCA UGGCGCUCGA G
```

W417X amber luciferase mRNA

```
1       AUGGAAGAUG CCAAAAACAU UAAGAAGGGC CCAGCGCCAU UCUACCCACU CGAAGACGGG
61      ACCGCCGGCG AGCAGCUGCA CAAAGCCAUG AAGCGCUACG CCCUGGUGCC CGGCACCAUC
121     GCCUUUACCG ACGCACAUAU CGAGGUGGAC AUUACCUACG CCGAGUACUU CGAGAUGAGC
181     GUUCGGCUGG CAGAAGCUAU GAAGCGCUAU GGGCUGAAUA CAAACCAUCG GAUCGUGGUG
241     UGCAGCGAGA AUAGCUUGCA GUUCUUCAUG CCCGUGUUGG GUGCCCUGUU CAUCGGUGUG
301     GCUGUGGCCC CAGCUAACGA CAUCUACAAC GAGCGCGAGC UGCUGAACAG CAUGGGCAUC
361     AGCCAGCCCA CCGUCGUAUU CGUGAGCAAG AAAGGGCUGC AAAAGAUCCU CAACGUGCAA
421     AAGAAGCUAC CGAUCAUACA AAAGAUCAUC AUCAUGGAUA GCAAGACCGA CUACCAGGGC
481     UUCCAAAGCA UGUACACCUU CGUGACUUCC CAUUUGCCAC CCGGCUUCAA CGAGUACGAC
541     UUCGUGCCCG AGAGCUUCGA CCGGGACAAA ACCAUCGCCC UGAUCAUGAA CAGUAGUGGC
601     AGUACCGGAU UGCCCAAGGG CGUAGCCCUA CCGCACCGCA CCGCUUGUGU CCGAUUCAGU
661     CAUGCCCGCG ACCCCAUCUU CGGCAACCAG AUCAUCCCCG ACACCGCUAU CCUCAGCGUG
721     GUGCCAUUUC ACCACGGCUU CGGCAUGUUC ACCACGCUGG GCUACUUGAU CUGCGGCUUU
781     CGGGUCGUGC UCAUGUACCG CUUCGAGGAG GAGCUAUUCU UGCGCAGCUU GCAAGACUAU
841     AAGAUUCAAU CUGCCCUGCU GGUGCCCACA CUAUUUAGCU UCUUCGCUAA GAGCACUCUC
901     AUCGACAAGU ACGACCUAAG CAACUUGCAC GAGAUCGCCA GCGGCGGGGC GCCGCUCAGC
961     AAGGAGGUAG GUGAGGCCGU GGCCAAACGC UUCCACCUAC CAGGCAUCCG CCAGGGCUAC
1021    GGCCUGACAG AAACAACCAG CGCCAUUCUG AUCACCCCCG AAGGGGACGA CAAGCCUGGC
1081    GCAGUAGGCA AGGUGGUGCC CUUCUUCGAG GCUAAGGUGG UGGACUUGGA CACCGGUAAG
1141    ACACUGGGUG UGAACCAGCG CGGCGAGCUG UGCGUCCGUG GCCCCAUGAU CAUGAGCGGC
1201    UACGUUAACA ACCCCGAGGC UACAAACGCU CUCAUCGACA AGGACGGCUA GGCUGCACAGC
1261    GGCGACAUCG CCUACUGGGA CGAGGACGAG CACUUCUUCA UCGUGGACCG GCUGAAGAGC
1321    CUGAUCAAAU ACAAGGGCUA CCAGGUAGCC CCAGCCGAAC UGGAGAGCAU CCUGCUGCAA
1381    CACCCCAACA UCUUCGACGC CGGGGUCGCC GGCCUGCCCG ACGACGAUGC CGGCGAGCUG
1441    CCCGCCGCAG UCGUCGUGCU GGAACACGGU AAAACCAUGA CCGAGAAGGA GAUCGUGGAC
1501    UAUGUGGCCA GCCAGGUCAC AACCGCCAAG AAGCUGCGCG GUGGUGUUGU GUUCGUGGAC
1561    GAGGUGCCUA AAGGACUGAC CGGCAAGUUG GACGCCCGCA AGAUCCGCGA GAUUCUCAUU
1621    AAGGCCAAGA AGGGCGGCAA GAUCGCCGUG
```

PINK UAG mRNA

```
1       AUGGCGGUGC GACAGGCGCU GGGCCGCGGC CUGCAGCUGG GUCGAGCGCU GCUGCUGCGC
61      UUCACGGGCA AGCCCGGCCG GGCCUACGGC UUGGGGCGGC CGGGCCCGGC GGCGGGCUGU
121     GUCCGCGGGG AGCGUCCAGG CUGGGCCGCA GGACCGGGCG CGGAGCCUCG CAGGGUCGGG
181     CUCGGGCUCC CUAACCGUCU CCGCUUCUUC CGCCAGUCGG UGGCCGGGCU GGCGGCGCGG
241     UUGCAGCGGC AGUUCGUGGU GCGGGCCUGG GGCUGCGCGG GCCCUUGCGG CCGGGCAGUC
301     UUUCUGGCCU UCGGGCUAGG GCUGGGCCUC AUCGAGGAAA AACAGGCGGA GAGCCGGCGG
361     GCGGUCUCGG CCUGUCAGGA GAUCCAGGCA AUUUUUACCC AGAAAAGCAA GCCGGGGCCU
421     GACCCGUUGG ACACGAGACG CUUGCAGGGC UUUCGGCUGG AGGAGUAUCU GAUAGGGCAG
481     UCCAUUGGUA AGGGCUGCAG UGCUGCUGUG UAUGAAGCCA CCAUGCCUAC AUUGCCCCAG
541     AACCUGGAGG UGACAAAGAG CACCGGGUUG CUUCCAGGGA GAGGCCCAGG UACCAGUGCA
601     CCAGGAGAAG GGCAGGAGCG AGCUCCGGGG GCCCCUGCCU UCCCCUUGGC CAUCAAGAUG
661     AUGUGGAACA UCUCGGCAGG UUCCUCCAGC GAAGCCAUCU UGAACACAAU GAGCCAGGAG
721     CUGGUCCCAG CGAGCCGAGU GGCCUUGGCU GGGGAGUAUG GAGCAGUCAC UUACAGAAAA
781     UCCAAGAGAG GUCCCAAGCA ACUAGCCCCU CACCCCAACA UCAUCCGGGU UCUCCGCGCC
841     UUCACCUCUU CCGUGCCGCU GCUGCCAGGG GCCCUGGUCG ACUACCCUGA UGUGCUGCCC
901     UCACGCCUCC ACCCUGAAGG CCUGGGCCAU GGCCGGACGC UGUUCCUCGU UAUGAAGAAC
961     UAUCCCUGUA CCCUGCGCCA GUACCUUUGU GUGAACACAC CCAGCCCCCG CCUCGCCGCC
1021    AUGAUGCUGC UGCAGCUGCU GGAAGGCGUG GACCAUCUGG UUCAACAGGG CAUCGCGCAC
1081    AGAGACCUGA AAUCCGACAA CAUCCUUGUG GAGCUGGACC CAGACGGCUG CCCCUGGCUG
1141    GUGAUCGCAG AUUUUGGCUG CUGCCUGGCU GAUGAGAGCA UCGGCCUGCA GUUGCCCUUC
1201    AGCAGCUGGU ACGUGGAUCG GGGCGGAAAC GGCUGUCUGA UGGCCCCAGA GGUGUCCACG
1261    GCCCGUCCUG GCCCCAGGGC AGUGAUUGAC UACAGCAAGG CUGAUGCCUA GGCAGUGGGA
1321    GCCAUCGCCU AUGAAAUCUU CGGGCUUGUC AAUCCCUUCU ACGGCCAGGG CAAGGCCCAC
1381    CUUGAAAGCC GCAGCUACCA AGAGGCUCAG CUACCUGCAC UGCCCGAGUC AGUGCCUCCA
1441    GACGUGAGAC AGUUGGUGAG GGCACUGCUC CAGCGAGAGG CCAGCAAGAG ACCAUCUGCC
1501    CGAGUAGCCG CAAAUGUGCU UCAUCUAAGC CUCUGGGGUG AACAUAUUCU AGCCCUGAAG
1561    AAUCUGAAGU UAGACAAGAU GGUUGGCUGG CUCCUCCAAC AAUCGGCCGC CACUUUGUUG
1621    GCCAACAGGC UCACAGAGAA GUGUUGUGUG GAAACAAAAA UGAAGAUGCU CUUUCUGGCU
1681    AACCUGGAGU GUGAAACGCU CUGCCAGGCA GCCCUCCUCC UCUGCUCAUG GAGGGCAGCC
1741    CUGUGAUGUC CCUGCAUGGA GCUGGUGAAU UACUAAAAGA ACAUGGCAUC CUCUGUGUCG
1801    UGAUGGUCUG UGAAUGGUGA GGGUGGGAGU CAGGAGACAA GACAGCGCAG AGAGGGCUGG
1861    UUAGCCGGAA AAGGCCUCGG GCUUGGCAAA UGGAAGAACU UGAGUGAGAG UUCAGUCUGC
1921    AGUCCUCUGC UCACAGACAU CUGAAAAGUG AAUGGCCAAG CUGGUCUAGU AGAUGAGGCU
1981    GGACUGAGGA GGGGUAGGCC UGCAUCCACA GAGAGGAUCC AGGCCAAGGC ACUGGCUGUC
2041    AGUGGCAGAG UUUGGCUGUG ACCUUUGCCC CUAACACGAG GAACUCGUUU GAAGGGGGCA
2101    GCGUAGCAUG UCUGAUUUGC CACCUGGAUG AAGGCAGACA UCAACAUGGG UCAGCACGUU
2161    CAGUUACGGG AGUGGGAAAU UACAUGAGGC CUGGGCCUCU GCGUUCCCAA GCUGUGCGUU
2221    CUGGACCAGC UACUGAAUUA UUAAUCUCAC UUAGCGAAAG UGACGGAUGA GCAGUAAGUA
2281    AGUAAGUGUG GGGAUUUAAA CUUGAGGGUU UCCCUCCUGA CUAGCCUCUC UUACAGGAAU
2341    UGUGAAAUAU UAAAUGCAAA UUUACAACUG CAGAUGACGU AUGUGCCUUG AACUGAAUAU
2401    UUGGCUUUAA GAAUGAUUCU UAUACUCUGA AGGUGAGAAU AUUUUGUGGG CAGGUAUCAA
2461    CAUUGGGGAA GAGAUUUCAU GUCUAACUAA CUAACUUUAU ACAUGAUUUU UAGGAAGCUA
2521    UUGCCUAAAU CAGCGUCAAC AUGCAGUAAA GGUUGUCUUC AACUGAAAAA AAAAAAAAAA
2581    AAAAAA
```

R175H p53 mRNA

```
1       AUGGAGGAGC CGCAGUCAGA UCCUAGCGUC GAGCCCCCUC UGAGUCAGGA AACAUUUUCA
61      GACCUAUGGA AACUACUUCC UGAAAACAAC GUUCUGUCCC CCUUGCCGUC CCAAGCAAUG
121     GAUGAUUUGA UGCUGUCCCC GGACGAUAUU GAACAAUGGU UCACUGAAGA CCCAGGUCCA
181     GAUGAAGCUC CCAGAAUGCC AGAGGCUGCU CCCCCCGUGG CCCCUGCACC AGCAGCUCCU
241     ACACCGGCGG CCCCUGCACC AGCCCCCUCC UGGCCCCUGU CAUCUUCUGU CCCUUCCCAG
301     AAAACCUACC AGGGCAGCUA CGGUUUCCGU CUGGGCUUCU UGCAUUCUGG GACAGCCAAG
361     UCUGUGACUU GCACGUACUC CCCUGCCCUC AACAAGAUGU UUUGCCAACU GGCCAAGACC
421     UGCCCUGUGC AGCUGUGGGU UGAUUCCACA CCCCCGCCCG GCACCCGCGU CCGCGCCAUG
481     GCCAUCUACA AGCAGUCACA GCACAUGACG GAGGUUGUGA GGCACUGCCC CCACCAUGAG
541     CGCUGCUCAG AUAGCGAUGG UCUGGCCCCU CCUCAGCAUC UUAUCCGAGU GGAAGGAAAU
601     UUGCGUGUGG AGUAUUUGGA UGACAGAAAC ACUUUUCGAC AUAGUGUGGU GGUGCCCUAU
661     GAGCCGCCUG AGGUUGGCUC UGACUGUACC ACCAUCCACU ACAACUACAU GUGUAACAGU
721     UCCUGCAUGG GCGGCAUGAA CCGGAGGCCC AUCCUCACCA UCAUCACACU GGAAGACUCC
781     AGUGGUAAUC UACUGGGACG GAACAGCUUU GAGGUGCGUG UUUGUGCCUG UCCUGGGAGA
841     GACCGGCGCA CAGAGGAAGA GAAUCUCCGC AAGAAAGGGG AGCCUCACCA CGAGCUGCCC
901     CCAGGGAGCA CUAAGCGAGC ACUGCCCAAC AACACCAGCU CCUCUCCCCA GCCAAAGAAG
961     AAACCACUGG AUGGAGAAUA UUUCACCCUU CAGAUCCGUG GGCGUGAGCG CUUCGAGAUG
1021    UUCCGAGAGC UGAAUGAGGC CUUGGAACUC AAGGAUGCCC AGGCUGGGAA GGAGCCAGGG
1081    GGGAGCAGGG CUCACUCCAG CCACCUGAAG UCCAAAAAGG GUCAGUCUAC CUCCCGCCAU
```

1141	AAAAAACUCA UGUUCAAGAC AGAAGGGCCU GACUCAGACU GA

R282Q p53 mRNA

1	AUGGAGGAGC CGCAGUCAGA UCCUAGCGUC GAGCCCCCUC UGAGUCAGGA AACAUUUUCA
61	GACCUAUGGA AACUACUUCC UGAAAACAAC GUUCUGUCCC CCUUGCCGUC CCAAGCAAUG
121	GAUGAUUUGA UGCUGUCCCC GGACGAUAUU GAACAAUGGU UCACUGAAGA CCCAGGUCCA
181	GAUGAAGCUC CCAGAAUGCC AGAGGCUGCU CCCCCCGUGG CCCCUGCACC AGCAGCUCCU
241	ACACCGGCGG CCCCUGCACC AGCCCCCUCC UGGCCCCUGU CAUCUUCUGU CCCUUCCCAG
301	AAAACCUACC AGGGCAGCUA CGGUUUCCGU CUGGGCUUCU UGCAUUCUGG GACAGCCAAG
361	UCUGUGACUU GCACGUACUC CCCUGCCCUC AACAAGAUGU UUUGCCAACU GGCCAAGACC
421	UGCCCUGUGC AGCUGUGGGU UGAUUCCACA CCCCCGCCCG GCACCCGCGU CCGCGCCAUG
481	GCCAUCUACA AGCAGUCACA GCACAUGACG GAGGUUGUGA GGCGCUGCCC CCACCAUGAG
541	CGCUGCUCAG AUAGCGAUGG UCUGGCCCCU CCUCAGCAUC UUAUCCGAGU GGAAGGAAAU
601	UUGCGUGUGG AGUAUUUGGA UGACAGAAAC ACUUUUCGAC AUAGUGUGGU GGUGCCCUAU
661	GAGCCGCCUG AGGUUGGCUC UGACUGUACC ACCAUCCACU ACAACUACAU GUGUAACAGU
721	UCCUGCAUGG GCGGCAUGAA CCGGAGGCCC AUCCUCACCA UCAUCACACU GGAAGACUCC
781	AGUGGUAAUC UACUGGGACG GAACAGCUUU GAGGUGCGUG UUUGUGCCUG UCCUGGGAGA
841	GAC**CAG**CGCA CAGAGGAAGA GAAUCUCCGC AAGAAAGGGG AGCCUCACCA CGAGCUGCCC
901	CCAGGGAGCA CUAAGCGAGC ACUGCCCAAC AACACCAGCU CCUCUCCCCA GCCAAAGAAG
961	AAACCACUGG AUGGAGAAUA UUUCACCCUU CAGAUCCGUG GGCGUGAGCG CUUCGAGAUG
1021	UUCCGAGAGC UGAAUGAGGC CUUGGAACUC AAGGAUGCCC AGGCUGGGAA GGAGCCAGGG
1081	GGGAGCAGGG CUCACUCCAG CCACCUGAAG UCCAAAAAGG GUCAGUCUAC CUCCCGCCAU
1141	AAAAAACUCA UGUUCAAGAC AGAAGGGCCU GACUCAGACU GA

R521H FUS mRNA

1	AUGGCCUCAA ACGAUUAUAC CCAACAAGCA ACCCAAAGCU AUGGGGCCUA CCCCACCCAG
61	CCCGGGCAGG GCUAUUCCCA GCAGAGCAGU CAGCCCUACG GACAGCAGAG UUACAGUGGU
121	UAUAGCCAGU CCACGGACAC UUCAGGCUAU GGCCAGAGCA GCUAUUCUUC UUAUGGCCAG
181	AGCCAGAACA CAGGCUAUGG AACUCAGUCA ACUCCCCAGG GAUAUGGCUC GACUGGCGGC
241	UAUGGCAGUA GCCAGAGCUC CCAAUCGUCU UACGGGCAGC AGUCCUCCUA CCCUGGCUAU
301	GGCCAGCAGC CAGCUCCCAG CAGCACCUCG GGAAGUUACG GUAGCAGUUC UCAGAGCAGC
361	AGCUAUGGGC AGCCCCAGAG UGGGAGCUAC AGCCAGCAGC CUAGCUAUGG UGGACAGCAG
421	CAAAGCUAUG GACAGCAGCA AAGCUAUAAU CCCCCUCAGG GCUAUGGACA GCAGAACCAG
481	UACAACAGCA GCAGUGGUGG UGGAGGUGGA GGUGGAGGUG GAGGUAACUA UGGCCAAGAU
541	CAAUCCUCCA UGAGUAGUGG UGGUGGCAGU GGUGGCGGUU AUGGCAAUCA AGACCAGAGU
601	GGUGGAGGUG GCAGCGGUGG CUAUGGACAG CAGGACCGUG GAGGCCGCGG CAGGGGUGGC
661	AGUGGUGGCG GCGGCGGCGG CGGCGGUGGU GGUUACAACC GCAGCAGUGG UGGCUAUGAA
721	CCCAGAGGUC GUGGAGGUGG CCGUGGAGGC AGAGGUGGCA UGGGCGGAAG UGACCGUGGU
781	GGCUUCAAUA AAUUUGGUGG CCCUCGGGAC CAAGGAUCAC GUCAUGACUC CGAACAGGAU
841	AAUUCAGACA ACAACACCAU CUUUGUGCAA GGCCUGGGUG AGAAUGUUAC AAUUGAGUCU
901	GUGGCUGAUU ACUUCAAGCA GAUUGGUAUU AUUAAGACAA ACAAGAAAAC GGGACAGCCC
961	AUGAUUAAUU UGUACACAGA CAGGGAAACU GGCAAGCUGA AGGGAGAGGC AACGGUCUCU
1021	UUUGAUGACC CACCUUCAGC UAAAGCAGCU AUUGACUGGU UUGAUGGUAA AGAAUUCUCC
1081	GGAAAUCCUA UCAAGGUCUC AUUUGCUACU CGCCGGGCAG ACUUUAAUCG GGGUGGUGGC
1141	AAUGGUCGUG GAGGCCGAGG GCGAGGAGGA CCCAUGGGCC GUGGAGGCUA UGGAGGUGGU
1201	GGCAGUGGUG GUGGUGGCCG AGGAGGAUUU CCCAGUGGAG GUGGUGGCGG UGGAGGACAG
1261	CAGCGAGCUG GUGACUGGAA GUGUCCUAAU CCCACCUGUG AGAAUAUGAA CUUCUCUUGG
1321	AGGAAUGAAU GCAACCAGUG UAAGGCCCCU AAACCAGAUG GCCCAGGAGG GGGACCAGGU
1381	GGCUCUCACA UGGGGGGUAA CUACGGGGAU GAUCGUCGUG GUGGCAGAGG AGGCUAUGAU
1441	CGAGGCGGCU ACCGGGGCCG CGGCGGGGAC CGUGGAGGCU UCCGAGGGGG CCGGGGUGGU
1501	GGGGACAGAG GUGGCUUUGG CCCUGGCAAG AUGGAUUCCA GGGGUGAGCA CAGACAGGAU
1561	**CAC**AGGGAGA GGCCGUAUUA ACUCGAGCAC CACCACCACC ACCAC

6.3 R/G-gRNA sequences

All synthesized R/G-gRNAs for in vitro editing experiments are listed in Table 6-3, R/G-gRNAs for cell culture application in Table 6-4 and the in vitro transcription PCR templates for R/G-gRNA transcription in Table 6-5.

Table 6-3: Synthesized R/G-gRNAs and their sequences of the mRNA binding site

Name of R/G-gRNA	Sequence of the mRNA binding site 5'→3'	Oligonucleotide for template generation
stop66 eCFP CCA	CACC**CCA**GCACAGAGT	92
E66 eCFP CCG	CACC**CCG**GCACAGAGT	169
E66 eCFP CUA	CAC**CCU**AGCACAGAGT	170
E66 eCFP CCC	CAC**CCC**AGCACAGAGT	171
W58X eGFP CCA	UCGG**CCA**CGGAACAGG	264
W417X luciferase CCA	GCAG**CCA**GCCGUCCUU	157
W437X PINK CCA	CUGC**CCA**GGCAUCAGC	201
H521 FUS GCG 15nt	CCCU**GCG**AUCCUGUC	139
H521 FUS GCG 16nt	CCCU**GCG**AUCCUGUCU	141
H521 FUS GUG 15nt	CCCU**GUG**AUCCUGUC	140
H521 FUS GUG 16nt	CCCU**GUG**AUCCUGUCU	142
H521 FUS GUA 16nt	CCCU**GUA**AUCCUGUCU	152

Table 6-4: R/G-gRNAs used in cell culture and the design of the mRNA binding site

Name of R/G-gRNA	Sequence of the mRNA binding site 5'→3'	Internal plasmid name
W58X GFP P6 16nt	UCGG**CCA**CGGAACAGG	pTS68
W58X GFP P6 18nt	UCGG**CCA**CGGAACAGGCA	pTS87
W58X GFP P6 20nt	UCGG**CCA**CGGAACAGGCAGU	pTS88
W58X GFP P6 25nt	UCGG**CCA**CGGAACAGGCAGUUUGCC	pTS93.1
W58X GFP P6 29nt	UCGG**CCA**CGGAACAGGCAGUUUGCCAGUA	pTS94.1
W58X GFP P3 16 nt	GG**CCA**CGGAACAGGCA	pTS155.2
W58X GFP P4 16 nt	CGG**CCA**CGGAACAGGC	pTS156.1
W58X GFP P5 16 nt	UCGG**CCA**CGGAACAGG	pTS157.1
W58X GFP P7 16 nt	GUCGG**CCA**CGGAACAG	pTS185
W58X GFP P8 16 nt	UGUCGG**CCA**CGGAACA	pTS186
W58X GFP P9 16 nt	GUGUCGG**CCA**CGGAAC	pTS187
W58X GFP P10 16 nt	AGUGUCGG**CCA**CGGAA	pTS188
W417X luciferase P5	CAG**CCA**GCCGUCCUUG	pTS169
W417X luciferase P6	GCAG**CCA**GCCGUCCUU	pTS121
W417X luciferase P7	UGCAG**CCA**GCCGUCCU	pTS193
W417X luciferase P8	GUGCAG**CCA**GCCGUCC	pTS170
W417X luciferase P9	UGUGCAG**CCA**GCCGUC	pTS171
R175H p53 5'-GCG	GGGGCA**GCG**CCTCACA	pTS181

R175H p53 5‘-GUA	GGGGCA**GUA**CCTCACA	pTS182
R282Q p53 5‘-CCG	UGUGCG**CCG**GUCUCUC	pTS241
R407Q PINK P6	CGCC**CCG**AUCCACGUA	pTS164.1

Table 6-5: R/G-gRNA in vitro transcription template

Standard transcription template for R/G-gRNAs without gRNA sequence	
1	GGTCAGGCCC AGGTTCTCCG CAGTGATAGA GACGCGTCTG TCGACCAGCA ACTGGTTGAA
61	CACCACCAGC AGGTTTAAAT CCAGGTCACG CAGTTCCATG TGGCCTCGCT TGGGTTATTG
121	CTGGTGCCCG GCCGGGCGCA ATATTCATGT TGATGATTTA TTATATATCG AGTGGTGTAT
181	TTATCAATAT TGTTTGCTCC GTTATCGTTA TTAACAAGTC ATCAATAAAG CCATCACGAG
241	TACAGCATGA AAAACAATAA ACTTGGCTTG CGCATCTAAG GATCTCGATC CCGCGAAATT
301	AATACGACTC ACTATAGGGG AATTGTGAGC GGATAACAAT TCCCCTCTAG AAATAATTTT
361	GTTTAACTTT AAGAAGGAGA TATACATATG GCTAGCTATT CCACCTGATG AGTTTTTACG
421	AAACGTTCCC GTGAGGGAAC GTCGTGGAAT AGTATAACAA TATGCTAAAT GTTGTTATAG
481	TATC
3‘ UTR shortened R/G-gRNA template	
1	TAATACGACT CACTATAGGG GAAGCTAGCTA TTCCACCTGA TGAGTTTTTA CGAAACGTTC
61	CCGTGAGGGA ACGTCGTGGA ATAGTATAAC AATATGCTAA ATGTTGTTAT AGTATCCCAC
121	CACCCCAGCA CAGAGT

6.4 ADAR protein sequences

ADAR1	
1	ATGGATATAG AAGATGAAGA AAACATGAGT TCCAGCAGCA CTGATGTGAA GGAAAACCGC
61	AATCTGGACA ACGTGTCCCC CAAGGATGGC AGCACACCTG GGCCTGGCGA GGGCTCTCAG
121	CTCTCCAATG GGGGTGGTGG TGGCCCCGGC AGAAAGCGGC CCCTGGAGGA GGGCAGCAAT
181	GGCCACTCCA AGTACCGCCT GAAGAAAAGG AGGAAAACAC CAGGGCCCGT CCTCCCCAAG
241	AACGCCCTGA TGCAGCTGAA TGAGATCAAG CCTGGTTTGC AGTACACACT CCTGTCCCAG
301	ACTGGGCCCG TGCACGCGCC TTTGTTTGTC ATGTCTGTGG AGGTGAATGG CCAGGTTTTT
361	GAGGGCTCTG GTCCCACAAA GAAAAAGGCA AAACTCCATG CTGCTGAGAA GGCCTTGAGG
421	TCTTTCGTTC AGTTTCCTAA TGCCTCTGAG GCCCACCTGG CCATGGGGAG GACCCTGTCT
481	GTCAACACGG ACTTCACATC TGACCAGGCC GACTTCCCTG ACACGCTCTT CAATGGTTTT
541	GAAACTCCTG ACAAGGCGGA GCCTCCCTTT TACGTGGGCT CCAATGGGGA TGACTCCTTC
601	AGTTCCAGCG GGGACCTCAG CTTGTCTGCT TCCCCGGTGC CTGCCAGCCT AGCCCAGCCT
661	CCTCTCCCTG CCTTACCACC ATTCCCACCC CCGAGTGGGA AGAATCCCGT GATGATCTTG
721	AACGAACTGC GCCCAGGACT CAAGTATGAC TTCCTCTCCG AGAGCGGGGA GAGCCATGCC
781	AAGAGCTTCG TCATGTCTGT GGTCGTGGAT GGTCAGTTCT TTGAAGGCTC AAGGCAGAAC
841	GCATGGGTTT CACAGAGGTA ACCCCAGTGA CAGGGGCCAG TCTCAGAAGA ACTATGCTCC
901	TCCTCTCAAG GTCCCCAGAA GCACAGCCAA AGACACTCCC TCTCACTGGC AGCACCTTCC
961	ATGACCAGAT AGCCATGCTG AGCCACCGGT GCTTCAACAC TCTGACTAAC AGCTTCCAGC
1021	CCTCCTTGCT CGGCCGCAAG ATTCTGGCCG CCATCATTAT GAAAAAAGAC TCTGAGGACA
1081	TGGGTGTCGT CGTCAGCTTG GGAACAGGGA ATCGCTGTGT AAAAGGAGAT TCTCTCAGCC
1141	TAAAAGGAGA AACTGTCAAT GACTGCCATG CAGAAATAAT CTCCCGGAGA GGCTTCATCA
1201	GGTTTCTCTA CAGTGAGTTA ATGAAATACA ACTCCCAGAC TGCGAAGGAT AGTATATTTG
1261	AACCTGCTAA GGGAGGAGAA AAGCTCCAAA TAAAAAAGAC TGTGTCATTC CATCTGTATA
1321	TCAGCACTGC TCCGTGTGGA GATGGCGCCC TCTTTGACAA GTCCTGCAGC GACCGTGCTA
1381	TGGAAAGCAC AGAATCCCGC CACTACCCTG TCTTCGAGAA TCCCAAACAA GGAAAGCTCC
1441	GCACCAAGGT GGAGAACGGA GAAGGCACAA TCCCTGTGGA ATCCAGTGAC ATTGTGCCTA
1501	CGTGGGATGG CATTCGGCTC GGGGAGAGAC TCCGTACCAT GTCCTGTAGT GACAAAATCC
1561	TACGCTGGAA CGTGCTGGGC CTGCAAGGGG CACTGTTGAC CCACTTCCTG CAGCCCATTT
1621	ATCTCAAATC TGTCACATTG GGTTACCTTT TCAGCCAAGG GCATCTGACC CGTGCTATTT
1681	GCTGTCGTGT GACAAGAGAT GGGAGTGCAT TTGAGGATGG ACTACGACAT CCCTTTATTG
1741	TCAACCACCC CAAGGTTGGC AGAGTCAGCA TATATGATTC CAAAAGGCAA TCCGGGAAGA
1801	CTAAGGAGAC AAGCGTCAAC TGGTGTCTGG CTGATGGCTA TGACCTGGAG ATCCTGGACG
1861	GTACCAGAGG CACTGTGGAT GGGCCACGGA ATGAATTGTC CCGGGTCTCC AAAAAGAACA
1921	TTTTTCTTCT ATTTAAGAAG CTCTGCTCCT TCCGTTACCG CAGGGATCTA C

ADAR2

```
1       ATGGATATAG AAGATGAAGA AAACATGAGT TCCAGCAGCA CTGATGTGAA GGAAAACCGC
61      AATCTGGACA ACGTGTCCCC CAAGGATGGC AGCACACCTG GGCCTGGCGA GGGCTCTCAG
121     CTCTCCAATG GGGGTGGTGG TGGCCCCGGC AGAAAGCGGC CCCTGGAGGA GGGCAGCAAT
181     GGCCACTCCA AGTACCGCCT GAAGAAAAGG AGGAAAACAC CAGGGCCCGT CCTCCCCAAG
241     AACGCCCTGA TGCAGCTGAA TGAGATCAAG CCTGGTTTGC AGTACACACT CCTGTCCCAG
301     ACTGGGCCCG TGCACGCGCC TTTGTTTGTC ATGTCTGTGG AGGTGAATGG CCAGGTTTTT
361     GAGGGCTCTG GTCCCACAAA GAAAAAGGCA AAACTCCATG CTGCTGAGAA GGCCTTGAGG
421     TCTTTCGTTC AGTTTCCTAA TGCCTCTGAG GCCCACCTGG CCATGGGGAG GACCCTGTCT
481     GTCAACACGG ACTTCACATC TGACCAGGCC GACTTCCCTG ACACGCTCTT CAATGGTTTT
541     GAAACTCCTG ACAAGGCGGA GCCTCCCTTT TACGTGGGCT CCAATGGGGA TGACTCCTTC
601     AGTTCCAGCG GGGACCTCAG CTTGTCTGCT TCCCCGGTGC CTGCCAGCCT AGCCCAGCCT
661     CCTCTCCCTG CCTTACCACC ATTCCCACCC CCGAGTGGGA AGAATCCCGT GATGATCTTG
721     AACGAACTGC GCCCAGGACT CAAGTATGAC TTCCTCTCCG AGAGCGGGGA GAGCCATGCC
781     AAGAGCTTCG TCATGTCTGT GGTCGTGGAT GGTCAGTTCT TTGAAGGCTC GGGGAGAAAC
841     AAGAAGCTTG CCAAGGCCCG GGCTGCGCAG TCTGCCCTGG CCGCCATTTT TAACTTGCAC
901     TTGGATCAGA CGCCATCTCG CCAGCCTATT CCCAGTGAGG GTCTTCAGCT GCATTTACCG
961     CAGGTTTTAG CTGACGCTGT CTCACGCCTG GTCCTGGGTA AGTTTGGTGA CCTGACCGAC
1021    AACTTCTCCT CCCCTCACGC TCGCAGAAAA GTGCTGGCTG GAGTCGTCAT GACAACAGGC
1081    ACAGATGTTA AAGATGCCAA GGTGATAAGT GTTTCTACAG GAACAAAATG TATTAATGGT
1141    GAATACATGA GTGATCGTGG CCTTGCATTA AATGACTGCC ATGCAGAAAT AATATCTCGG
1201    AGATCCTTGC TCAGATTTCT TTATACACAA CTTGAGCTTT ACTTAAATAA CAAAGATGAT
1261    CAAAAAAGAT CCATCTTTCA GAAATCAGAG CGAGGGGGGT TTAGGCTGAA GGAGAATGTC
1321    CAGTTTCATC TGTACATCAG CACCTCTCCC TGTGGAGATG CCAGAATCTT CTCACCACAT
1381    GAGCCAATCC TGGAAGAACC AGCAGATAGA CACCCAAATC GTAAAGCAAG AGGACAGCTA
1441    CGGACCAAAA TAGAGTCTGG TGAGGGGACG ATTCCAGTGC GCTCCAATGC GAGCATCCAA
1501    ACGTGGGACG GGGTGCTGCA AGGGGAGCGG CTGCTCACCA TGTCCTGCAG TGACAAGATT
1561    GCACGCTGGA ACGTGGTGGG CATCCAGGGT TCCCTGCTCA GCATTTTCGT GGAGCCCATT
1621    TACTTCTCGA GCATCATCCT GGGCAGCCTT TACCACGGGG ACCACCTTTC CAGGGCCATG
1681    TACCAGCGGA TCTCCAACAT AGAGGACCTG CCACCTCTCT ACACCCTCAA CAAGCCTTTG
1741    CTCAGTGGCA TCAGCAATGC AGAAGCACGG CAGCCAGGGA AGGCCCCCAA CTTCAGTGTC
1801    AACTGGACGG TAGGCGACTC CGCTATTGAG GTCATCAACG CCACGACTGG GAAGGATGAG
1861    CTGGGCCGCG CGTCCCGCCT GTGTAAGCAC GCGTTGTACT GTCGCTGGAT GCGTGTGCAC
1921    GGCAAGGTTC CCTCCCACTT ACTACGCTCC AAGATTACCA AACCCAACGT GTACCATGAG
1981    TCCAAGCTGG CGGCAAAGGA GTACCAGGCC GCCAAGGCGC GTCTGTTCAC AGCCTTCATC
2041    AAGGCGGGGC TGGGGGCCTG GGTGGAGAAG CCCACCGAGC AGGACCAGTT CTCACTCACG
2101    CCC
```

Short ADAR2 ΔdsRBM1

```
1       ATGCCGAGTG GGAAGAATCC CGTGATGATC TTGAACGAAC TGCGCCCAGG ACTCAAGTAT
61      GACTTCCTCT CCGAGAGCGG GGAGAGCCAT GCCAAGAGCT TCGTCATGTC TGTGGTCGTG
121     GATGGTCAGT TCTTTGAAGG CTCGGGGAGA AACAAGAAGC TTGCCAAGGC CCGGGCTGCG
181     CAGTCTGCCC TGGCCGCCAT TTTTAACTTG CACTTGGATC AGACGCCATC TCGCCAGCCT
241     ATTCCCAGTG AGGGTCTTCA GCTGCATTTA CCGCAGGTTT TAGCTGACGC TGTCTCACGC
301     CTGGTCCTGG GTAAGTTTGG TGACCTGACC GACAACTTCT CCTCCCCTCA CGCTCGCAGA
361     AAAGTGCTGG CTGGAGTCGT CATGACAACA GGCACAGATG TTAAAGATGC CAAGGTGATA
421     AGTGTTTCTA CAGGAACAAA ATGTATTAAT GGTGAATACA TGAGTGATCG TGGCCTTGCA
481     TTAAATGACT GCCATGCAGA AATAATATCT CGGAGATCCT TGCTCAGATT TCTTTATACA
541     CAACTTGAGC TTTACTTAAA TAACAAAGAT GATCAAAAAA GATCCATCTT TCAGAAATCA
601     GAGCGAGGGG GGTTTAGGCT GAAGGAGAAT GTCCAGTTTC ATCTGTACAT CAGCACCTCT
661     CCCTGTGGAG ATGCCAGAAT CTTCTCACCA CATGAGCCAA TCCTGGAAGA ACCAGCAGAT
721     AGACACCCAA ATCGTAAAGC AAGAGGACAG CTACGGACCA AAATAGAGTC TGGTGAGGGG
781     ACGATTCCAG TGCGCTCCAA TGCGAGCATC CAAACGTGGG ACGGGGTGCT GCAAGGGGAG
841     CGGCTGCTCA CCATGTCCTG CAGTGACAAG ATTGCACGCT GGAACGTGGT GGGCATCCAG
901     GGATCCCTGC TCAGCATTTT CGTGGAGCCC ATTTACTTCT CGAGCATCAT CCTGGGCAGC
961     CTTTACCACG GGGACCACCT TTCCAGGGCC ATGTACCAGC GGATCTCCAA CATAGAGGAC
1021    CTGCCACCTC TCTACACCCT CAACAAGCCT TTGCTCAGTG GCATCAGCAA TGCAGAAGCA
1081    CGGCAGCCAG GGAAGGCCCC CAACTTCAGT GTCAACTGGA CGGTAGGCGA CTCCGCTATT
1141    GAGGTCATCA ACGCCACGAC TGGGAAGGAT GAGCTGGGCC GCGCGTCCCG CCTGTGTAAG
1201    CACGCGTTGT ACTGTCGCTG GATGCGTGTG CACGGCAAGG TTCCCTCCCA CTTACTACGC
1261    TCCAAGATTA CCAAACCCAA CGTGTACCAT GAGTCCAAGC TGGCGGCAAA GGAGTACCAG
1321    GCCGCCAAGG CGCGTCTGTT CACAGCCTTC ATCAAGGCGG GGCTGGGGGC CTGGGTGGAG
1381    AAGCCCACCG AGCAGGACCA GTTCTCACTC ACGCCC
```

E488Q ADAR2

```
1       ATGGATATAG AAGATGAAGA AAACATGAGT TCCAGCAGCA CTGATGTGAA GGAAAACCGC
```

```
61      AATCTGGACA ACGTGTCCCC CAAGGATGGC AGCACACCTG GGCCTGGCGA GGGCTCTCAG
121     CTCTCCAATG GGGGTGGTGG TGGCCCCGGC AGAAAGCGGC CCCTGGAGGA GGGCAGCAAT
181     GGCCACTCCA AGTACCGCCT GAAGAAAAGG AGGAAAACAC CAGGGCCCGT CCTCCCCAAG
241     AACGCCCTGA TGCAGCTGAA TGAGATCAAG CCTGGTTTGC AGTACACACT CCTGTCCCAG
301     ACTGGGCCCG TGCACGCGCC TTTGTTTGTC ATGTCTGTGG AGGTGAATGG CCAGGTTTTT
361     GAGGGCTCTG GTCCCACAAA GAAAAAGGCA AAACTCCATG CTGCTGAGAA GGCCTTGAGG
421     TCTTTCGTTC AGTTTCCTAA TGCCTCTGAG GCCCACCTGG CCATGGGGAG GACCCTGTCT
481     GTCAACACGG ACTTCACATC TGACCAGGCC GACTTCCCTG ACACGCTCTT CAATGGTTTT
541     GAAACTCCTG ACAAGGCGGA GCCTCCCTTT TACGTGGGCT CCAATGGGGA TGACTCCTTC
601     AGTTCCAGCG GGGACCTCAG CTTGTCTGCT TCCCCGGTGC CTGCCAGCCT AGCCCAGCCT
661     CCTCTCCCTG CCTTACCACC ATTCCCACCC CCGAGTGGGA AGAATCCCGT GATGATCTTG
721     AACGAACTGC GCCCAGGACT CAAGTATGAC TTCCTCTCCG AGAGCGGGGA GAGCCATGCC
781     AAGAGCTTCG TCATGTCTGT GGTCGTGGAT GGTCAGTTCT TTGAAGGCTC GGGGAGAAAC
841     AAGAAGCTTG CCAAGGCCCG GGCTGCGCAG TCTGCCCTGG CCGCCATTTT TAACTTGCAC
901     TTGGATCAGA CGCCATCTCG CCAGCCTATT CCCAGTGAGG GTCTTCAGCT GCATTTACCG
961     CAGGTTTTAG CTGACGCTGT CTCACGCCTG GTCCTGGGTA AGTTTGGTGA CCTGACCGAC
1021    AACTTCTCCT CCCCTCACGC TCGCAGAAAA GTGCTGGCTG GAGTCGTCAT GACAACAGGC
1081    ACAGATGTTA AAGATGCCAA GGTGATAAGT GTTTCTACAG GAACAAAATG TATTAATGGT
1141    GAATACATGA GTGATCGTGG CCTTGCATTA AATGACTGCC ATGCAGAAAT AATATCTCGG
1201    AGATCCTTGC TCAGATTTCT TTATACACAA CTTGAGCTTT ACTTAAATAA CAAAGATGAT
1261    CAAAAAAGAT CCATCTTTCA GAAATCAGAG CGAGGGGGGT TTAGGCTGAA GGAGAATGTC
1321    CAGTTTCATC TGTACATCAG CACCTCTCCC TGTGGAGATG CCAGAATCTT CTCACCACAT
1381    GAGCCAATCC TGGAAGAACC AGCAGATAGA CACCCAAATC GTAAAGCAAG AGGACAGCTA
1441    CGGACCAAAA TA**CAG**TCTGG TGAGGGGACG ATTCCAGTGC GCTCCAATGC GAGCATCCAA
1501    ACGTGGGACG GGGTGCTGCA AGGGGAGCGG CTGCTCACCA TGTCCTGCAG TGACAAGATT
1561    GCACGCTGGA ACGTGGTGGG CATCCAGGGT TCCCTGCTCA GCATTTTCGT GGAGCCCATT
1621    TACTTCTCGA GCATCATCCT GGGCAGCCTT TACCACGGGG ACCACCTTTC CAGGGCCATG
1681    TACCAGCGGA TCTCCAACAT AGAGGACCTG CCACCTCTCT ACACCCTCAA CAAGCCTTTG
1741    CTCAGTGGCA TCAGCAATGC AGAAGCACGG CAGCCAGGGA AGGCCCCCAA CTTCAGTGTC
1801    AACTGGACGG TAGGCGACTC CGCTATTGAG GTCATCAACG CCACGACTGG GAAGGATGAG
1861    CTGGGCCGCG CGTCCCGCCT GTGTAAGCAC GCGTTGTACT GTCGCTGGAT GCGTGTGCAC
1921    GGCAAGGTTC CCTCCCACTT ACTACGCTCC AAGATTACCA AACCCAACGT GTACCATGAG
1981    TCCAAGCTGG CGGCAAAGGA GTACCAGGCC GCCAAGGCGC GTCTGTTCAC AGCCTTCATC
2041    AAGGCGGGGC TGGGGGCCTG GGTGGAGAAG CCCACCGAGC AGGACCAGTT CTCACTCACG
2101    CCC
```

E396A ADAR2

```
1       ATGGATATAG AAGATGAAGA AAACATGAGT TCCAGCAGCA CTGATGTGAA GGAAAACCGC
61      AATCTGGACA ACGTGTCCCC CAAGGATGGC AGCACACCTG GGCCTGGCGA GGGCTCTCAG
121     CTCTCCAATG GGGGTGGTGG TGGCCCCGGC AGAAAGCGGC CCCTGGAGGA GGGCAGCAAT
181     GGCCACTCCA AGTACCGCCT GAAGAAAAGG AGGAAAACAC CAGGGCCCGT CCTCCCCAAG
241     AACGCCCTGA TGCAGCTGAA TGAGATCAAG CCTGGTTTGC AGTACACACT CCTGTCCCAG
301     ACTGGGCCCG TGCACGCGCC TTTGTTTGTC ATGTCTGTGG AGGTGAATGG CCAGGTTTTT
361     GAGGGCTCTG GTCCCACAAA GAAAAAGGCA AAACTCCATG CTGCTGAGAA GGCCTTGAGG
421     TCTTTCGTTC AGTTTCCTAA TGCCTCTGAG GCCCACCTGG CCATGGGGAG GACCCTGTCT
481     GTCAACACGG ACTTCACATC TGACCAGGCC GACTTCCCTG ACACGCTCTT CAATGGTTTT
541     GAAACTCCTG ACAAGGCGGA GCCTCCCTTT TACGTGGGCT CCAATGGGGA TGACTCCTTC
601     AGTTCCAGCG GGGACCTCAG CTTGTCTGCT TCCCCGGTGC CTGCCAGCCT AGCCCAGCCT
661     CCTCTCCCTG CCTTACCACC ATTCCCACCC CCGAGTGGGA AGAATCCCGT GATGATCTTG
721     AACGAACTGC GCCCAGGACT CAAGTATGAC TTCCTCTCCG AGAGCGGGGA GAGCCATGCC
781     AAGAGCTTCG TCATGTCTGT GGTCGTGGAT GGTCAGTTCT TTGAAGGCTC GGGGAGAAAC
841     AAGAAGCTTG CCAAGGCCCG GGCTGCGCAG TCTGCCCTGG CCGCCATTTT TAACTTGCAC
901     TTGGATCAGA CGCCATCTCG CCAGCCTATT CCCAGTGAGG GTCTTCAGCT GCATTTACCG
961     CAGGTTTTAG CTGACGCTGT CTCACGCCTG GTCCTGGGTA AGTTTGGTGA CCTGACCGAC
1021    AACTTCTCCT CCCCTCACGC TCGCAGAAAA GTGCTGGCTG GAGTCGTCAT GACAACAGGC
1081    ACAGATGTTA AAGATGCCAA GGTGATAAGT GTTTCTACAG GAACAAAATG TATTAATGGT
1141    GAATACATGA GTGATCGTGG CCTTGCATTA AATGACTGCC ATGCA**GCA**AT AATATCTCGG
1201    AGATCCTTGC TCAGATTTCT TTATACACAA CTTGAGCTTT ACTTAAATAA CAAAGATGAT
1261    CAAAAAAGAT CCATCTTTCA GAAATCAGAG CGAGGGGGGT TTAGGCTGAA GGAGAATGTC
1321    CAGTTTCATC TGTACATCAG CACCTCTCCC TGTGGAGATG CCAGAATCTT CTCACCACAT
1381    GAGCCAATCC TGGAAGAACC AGCAGATAGA CACCCAAATC GTAAAGCAAG AGGACAGCTA
1441    CGGACCAAAA TAGAGTCTGG TGAGGGGACG ATTCCAGTGC GCTCCAATGC GAGCATCCAA
1501    ACGTGGGACG GGGTGCTGCA AGGGGAGCGG CTGCTCACCA TGTCCTGCAG TGACAAGATT
1561    GCACGCTGGA ACGTGGTGGG CATCCAGGGT TCCCTGCTCA GCATTTTCGT GGAGCCCATT
1621    TACTTCTCGA GCATCATCCT GGGCAGCCTT TACCACGGGG ACCACCTTTC CAGGGCCATG
1681    TACCAGCGGA TCTCCAACAT AGAGGACCTG CCACCTCTCT ACACCCTCAA CAAGCCTTTG
1741    CTCAGTGGCA TCAGCAATGC AGAAGCACGG CAGCCAGGGA AGGCCCCCAA CTTCAGTGTC
```

1801	AACTGGACGG TAGGCGACTC CGCTATTGAG GTCATCAACG CCACGACTGG GAAGGATGAG
1861	CTGGGCCGCG CGTCCCGCCT GTGTAAGCAC GCGTTGTACT GTCGCTGGAT GCGTGTGCAC
1921	GGCAAGGTTC CCTCCCACTT ACTACGCTCC AAGATTACCA AACCCAACGT GTACCATGAG
1981	TCCAAGCTGG CGGCAAAGGA GTACCAGGCC GCCAAGGCGC GTCTGTTCAC AGCCTTCATC
2041	AAGGCGGGGC TGGGGGCCTG GGTGGAGAAG CCCACCGAGC AGGACCAGTT CTCACTCACG
2101	CCC

SNAP-ADAR2

1	ATGGACAAAG ATTGCGAAAT GAAACGTACC ACCCTGGATA GCCCGCTGGG CAAACTGGAA
61	CTGAGCGGCT GCGAACAGGG CCTGCATGAA ATTAAACTGC TGGGTAAAGG CACCAGCGCG
121	GCCGATGCGG TTGAAGTTCC GGCCCCGGCC GCCGTGCTGG GTGGTCCGGA ACCGCTGATG
181	CAGGCGACCG CGTGGCTGAA CGCGTATTTT CATCAGCCGG AAGCGATTGA AGAATTTCCG
241	GTTCCGGCGC TGCATCATCC GGTGTTTCAG CAGGAGAGCT TTACCCGTCA GGTGCTGTGG
301	AAACTGCTGA AAGTGGTTAA ATTTGGCGAA GTGATTAGCT ATCAGCAGCT GGCGGCCCTG
361	GCGGGTAATC CGGCGGCCAC CGCCGCCGTT AAAACCGCGC TGAGCGGTAA CCCGGTGCCG
421	ATTCTGATTC CGTGCCATCG TGTGGTTAGC TCTAGCGGTG CGGTTGGCGG TTATGAAGGT
481	GGTCTGGCGG TGAAAGAGTG GCTGCTGGCC CATGAAGGTC ATCGTCTGGG TAAACCGGGT
541	CTGGGACCTG CAGGTATAGG CGCGCCAGGG TCTGGCGGCG GCAGTAAGAA GCTTGCCAAG
601	GCCCGGGCTG CGCAGTCTGC CCTGGCCGCC ATTTTTAACT TGCACTTGGA TCAGACGCCA
661	TCTCGCCAGC CTATTCCCAG TGAGGGTCTT CAGCTGCATT TACCGCAGGT TTTAGCTGAC
721	GCTGTCTCAC GCCTGGTCCT GGGTAAGTTT GGTGACCTGA CCGACAACTT CTCCTCCCCT
781	CACGCTCGCA GAAAAGTGCT GGCTGGAGTC GTCATGACAA CAGGCACAGA TGTTAAAGAT
841	GCCAAGGTGA TAAGTGTTTC TACAGGAACA AAATGTATTA ATGGTGAATA CATGAGTGAT
901	CGTGGCCTTG CATTAAATGA CTGCCATGCA GAAATAATAT CTCGGAGATC CTTGCTCAGA
961	TTTCTTTATA CACAACTTGA GCTTTACTTA AATAACAAAG ATGATCAAAA AAGATCCATC
1021	TTTCAGAAAT CAGAGCGAGG GGGGTTTAGG CTGAAGGAGA ATGTCCAGTT TCATCTGTAC
1081	ATCAGCACCT CTCCCTGTGG AGATGCCAGA ATCTTCTCAC CACATGAGCC AATCCTGGAA
1141	GAACCAGCAG ATAGACACCC AAATCGTAAA GCAAGAGGAC AGCTACGGAC CAAAATAGAG
1201	TCTGGTGAGG GGACGATTCC AGTGCGCTCC AATGCGAGCA TCCAAACGTG GGACGGGGTG
1261	CTGCAAGGGG AGCGGCTGCT CACCATGTCC TGCAGTGACA AGATTGCACG CTGGAACGTG
1321	GTGGGCATCC AGGGATCCCT GCTCAGCATT TTCGTGGAGC CCATTTACTT CTCGAGCATC
1381	ATCCTGGGCA GCCTTTACCA CGGGGACCAC CTTTCCAGGG CCATGTACCA GCGGATCTCC
1441	AACATAGAGG ACCTGCCACC TCTCTACACC CTCAACAAGC CTTTGCTCAG TGGCATCAGC
1501	AATGCAGAAG CACGGCAGCC AGGGAAGGCC CCCAACTTCA GTGTCAACTG GACGGTAGGC
1561	GACTCCGCTA TTGAGGTCAT CAACGCCACG ACTGGGAAGG ATGAGCTGGG CCGCGCGTCC
1621	CGCCTGTGTA AGCACGCGTT GTACTGTCGC TGGATGCGTG TGCACGGCAA GGTTCCCTCC
1681	CACTTACTAC GCTCCAAGAT TACCAAGCCC AACGTGTACC ATGAGTCCAA GCTGGCGGCA
1741	AAGGAGTACC AGGCCGCCAA GGCGCGTCTG TTCACAGCCT TCATCAAGGC GGGGCTGGGG
1801	GCCTGGGTGG AGAAGCCCAC CGAGCAGGAC CAGTTCTCAC TCACGCCCCA CCATCACCAT
1861	CACCAT

6.5 ADAR protein productions

The following proteins were produced in this work: SNAP-ADAR2, wt ADAR2, E488Q and short ADAR2. Of each protein a SDS-PAGE gel of their production is exemplary shown.

The main fractions of the heparin eluates of SNAP-ADAR2 were obtained by a concentration of 250 mM and 300 mM NaCl (Figure 6-1). These fractions were pooled and resulted in a protein concentration of 23 µM.

After heparin column purification the main fractions of ADAR2 (400 mM – 500 mM) (Figure 6-2) and E488Q (450mM – 600 mM) (Figure 6-3) were pooled and resulted in a protein concentration of 36.5 µM for ADAR2 and 29.81 µM for E488Q.

The main fractions (300 mM and 350 mM NaCl) of short ADAR2 were pooled and resulted in a protein concentration of 13 µM (Figure 6-4).

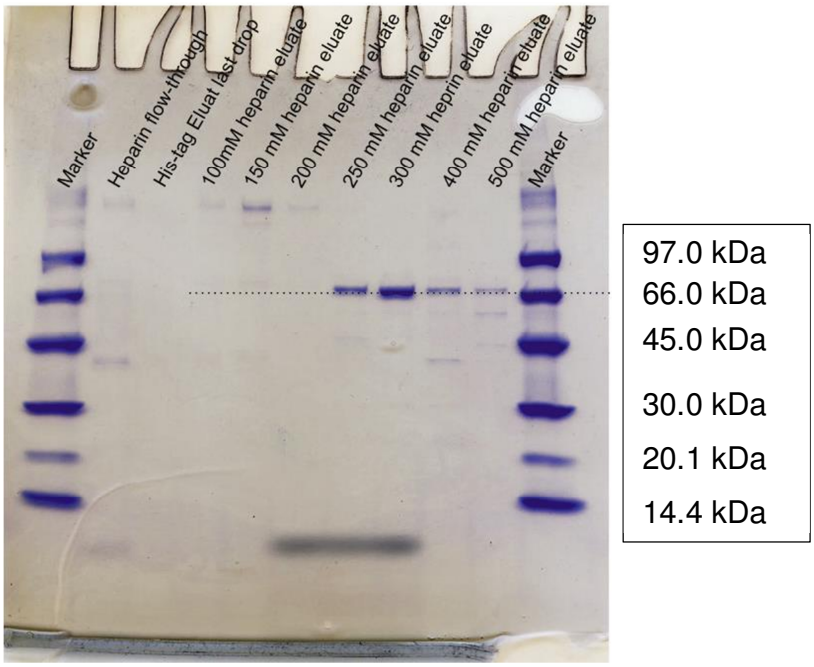

Figure 6-1: SDS-PAGE of purified SNAP-ADAR2.

The protein was purified from cell disruption using a Ni-NTA column. Subsequently, the eluate was further purified with a Heparin column and eluted by increasing concentration of NaCl.

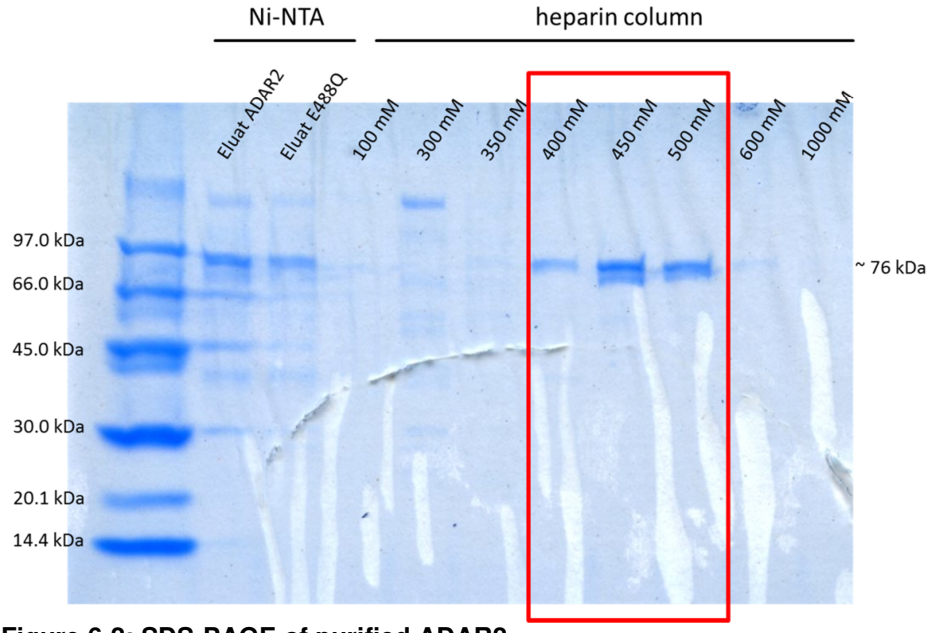

Figure 6-2: SDS-PAGE of purified ADAR2.

The protein was purified from cell disruption using a Ni-NTA column. Subsequently, the eluate was further purified with a Heparin column and eluted by increasing concentration of NaCl. The fractions 400 mM – 500 mM were pooled and resulted in a protein concentration of 36.5 µM.

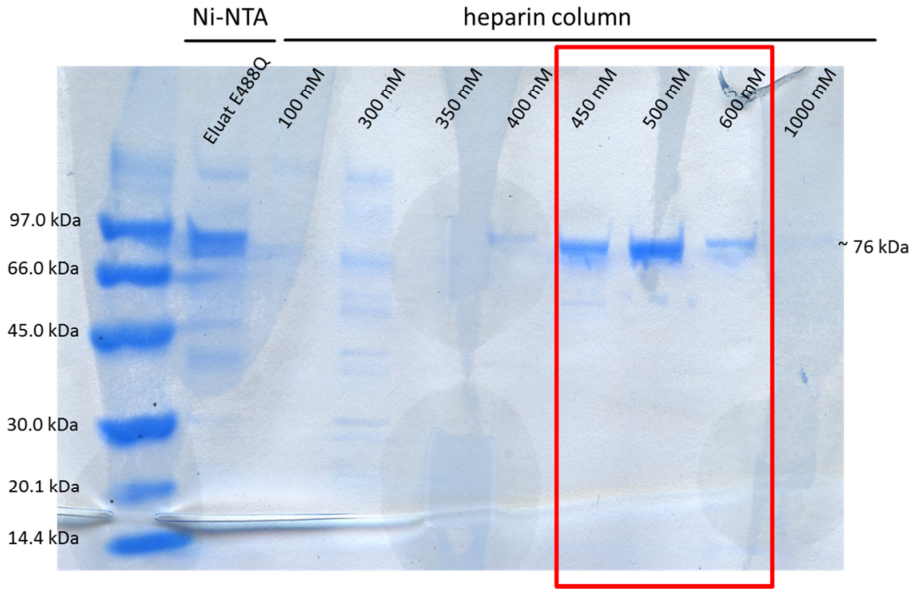

Figure 6-3: SDS-PAGE of purified E488Q ADAR2.

The protein was purified from cell disruption using a Ni-NTA column. Subsequently, the eluate was further purified with a Heparin column and eluted by increasing concentration of NaCl. The fractions 450 mM – 600 mM were pooled and resulted in a protein concentration of 29.81 µM.

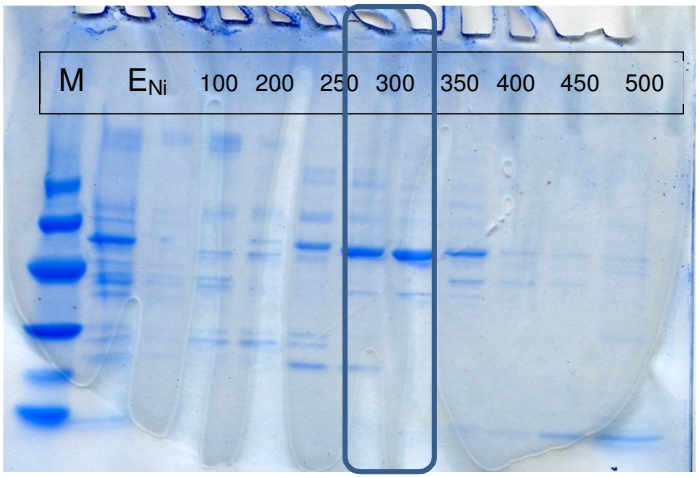

Figure 6-4: SDS-PAGE of purified short ADAR2.

The protein was purified from cell disruption using a Ni-NTA column. Subsequently, the eluate was further purified with a Heparin column and eluted by increasing concentration of NaCl. The fractions 300 mM and 350 mM were pooled and resulted in a protein concentration of 13 µM.

7 References

1. Bass, B.L. (2001) *RNA editing*. Oxford University Press.
2. Gott, J.M. and Emeson, R.B. (2000) Functions and mechanisms of RNA editing. *Annual review of genetics*, **34**, 499-531.
3. Wahba, A.J., Basilio, C., Speyer, J.F., Lengyel, P., Miller, R.S. and Ochoa, S. (1962) Synthetic polynucleotides and the amino acid code. VI. *Proceedings of the National Academy of Sciences of the United States of America*, **48**, 1683.
4. Bass, B.L. (2002) RNA editing by adenosine deaminases that act on RNA. *Annu. Rev. Biochem.*, **71**, 817.
5. Hundley, H.A. and Bass, B.L. (2010) ADAR editing in double-stranded UTRs and other noncoding RNA sequences. *Trends in biochemical sciences*, **35**, 377-383.
6. Levanon, E.Y., Eisenberg, E., Yelin, R., Nemzer, S., Hallegger, M., Shemesh, R., Fligelman, Z.Y., Shoshan, A., Pollock, S.R. and Sztybel, D. (2004) Systematic identification of abundant A-to-I editing sites in the human transcriptome. *Nature biotechnology*, **22**, 1001-1005.
7. Smith, H.C. (2008) *RNA and DNA editing: molecular mechanisms and their integration into biological systems*. John Wiley & Sons.
8. Jantsch, M.F. and Öhman, M. (2008) In Göringer, H. U. (ed.), *RNA Editing*. Springer Berlin Heidelberg, Vol. 20, pp. 51-84.
9. Scadden, A. and Smith, C.W. (2001) RNAi is antagonized by A→ I hyper-editing. *EMBO reports*, **2**, 1107-1111.
10. 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.
11. Li, J.B. and Church, G.M. (2013) Deciphering the functions and regulation of brain-enriched A-to-I RNA editing. *Nat. Neurosci.*, **16**, 1518-1522.
12. 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. *Nature structural & molecular biology*, **13**, 13-21.
13. Nishikura, K. (2006) Editor meets silencer: crosstalk between RNA editing and RNA interference. *Nature Reviews Molecular Cell Biology*, **7**, 919-931.
14. Bass, B. (2006), *Cold Spring Harbor symposia on quantitative biology*. Cold Spring Harbor Laboratory Press, Vol. 71, pp. 285-292.
15. Heale, B.S., Keegan, L.P., McGurk, L., Michlewski, G., Brindle, J., Stanton, C.M., Caceres, J.F. and O'Connell, M.A. (2009) Editing independent effects of ADARs on the miRNA/siRNA pathways. *EMBO J.*, **28**, 3145-3156.
16. Ota, H., Sakurai, M., Gupta, R., Valente, L., Wulff, B.-E., Ariyoshi, K., Iizasa, H., Davuluri, R.V. and Nishikura, K. (2013) ADAR1 forms a complex with Dicer to promote microRNA processing and RNA-induced gene silencing. *Cell*, **153**, 575-589.
17. Luciano, D.J., Mirsky, H., Vendetti, N.J. and Maas, S. (2004) RNA editing of a miRNA precursor. *RNA*, **10**, 1174-1177.

18. Kawahara, Y., Megraw, M., Kreider, E., Iizasa, H., Valente, L., Hatzigeorgiou, A.G. and Nishikura, K. (2008) Frequency and fate of microRNA editing in human brain. *Nucleic acids research*, **36**, 5270-5280.
19. Nishikura, K. (2010) Functions and regulation of RNA editing by ADAR deaminases. *Annu. Rev. Biochem.*, **79**, 321.
20. Chen, C.-X., Cho, D., Wang, Q., Lai, F., Carter, K.C. and Nishikura, K. (2000) A third member of the RNA-specific adenosine deaminase gene family, ADAR3, contains both single-and double-stranded RNA binding domains. *Rna*, **6**, 755-767.
21. Patterson, J.B. and Samuel, C.E. (1995) Expression and regulation by interferon of a double-stranded-RNA-specific adenosine deaminase from human cells: evidence for two forms of the deaminase. *Mol. Cell. Biol.*, **15**, 5376-5388.
22. Keegan, L.P., Leroy, A., Sproul, D. and O'Connell, M.A. (2004) Adenosine deaminases acting on RNA (ADARs): RNA-editing enzymes. *Genome Biol*, **5**, 209.
23. Kim, U., Wang, Y., Sanford, T., Zeng, Y. and Nishikura, K. (1994) Molecular cloning of cDNA for double-stranded RNA adenosine deaminase, a candidate enzyme for nuclear RNA editing. *Proceedings of the National Academy of Sciences*, **91**, 11457-11461.
24. Melcher, T., Maas, S., Herb, A., Sprengel, R., Seeburg, P.H. and Higuchi, M. (1996) A mammalian RNA editing enzyme.
25. Palladino, M.J., Keegan, L.P., O'Connell, M.A. and Reenan, R.A. (2000) dADAR, a Drosophila double-stranded RNA-specific adenosine deaminase is highly developmentally regulated and is itself a target for RNA editing. *Rna*, **6**, 1004-1018.
26. Tonkin, L.A., Saccomanno, L., Morse, D.P., Brodigan, T., Krause, M. and Bass, B.L. (2002) RNA editing by ADARs is important for normal behavior in Caenorhabditis elegans. *The EMBO journal*, **21**, 6025-6035.
27. 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.
28. 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. *Journal of Biological Chemistry*, **279**, 4894-4902.
29. 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. *Journal of Biological Chemistry*, **279**, 4952-4961.
30. Köhler, M., Burnashev, N., Sakmann, B. and Seeburg, P.H. (1993) Determinants of Ca 2+ permeability in both TM1 and TM2 of high affinity kainate receptor channels: diversity by RNA editing. *Neuron*, **10**, 491-500.
31. Brusa, R., Zimmermann, F., Koh, D.-S., Feldmeyer, D., Gass, P., Seeburg, P.H. and Sprengel, R. (1995) Early-onset epilepsy and postnatal lethality associated with an editing-deficient GluR-B allele in mice. *Science*, **270**, 1677-1680.

32. Seeburg, P.H., Higuchi, M. and Sprengel, R. (1998) RNA editing of brain glutamate receptor channels: mechanism and physiology. *Brain research reviews*, **26**, 217-229.
33. Lomeli, H., Mosbacher, J., Melcher, T., Hoger, T., Kuner, T., Monyer, H., Higuchi, M., Bach, A. and Seeburg, P. (1994) Control of kinetic properties of AMPA receptor channels by nuclear RNA editing. *Science*, **266**, 1709-1713.
34. Stefl, R., Xu, M., Skrisovska, L., Emeson, R.B. and Allain, F.H.-T. (2006) Structure and specific RNA binding of ADAR2 double-stranded RNA binding motifs. *Structure*, **14**, 345-355.
35. Stefl, R., Oberstrass, F.C., Hood, J.L., Jourdan, M., Zimmermann, M., Skrisovska, L., Maris, C., Peng, L., Hofr, C. and Emeson, R.B. (2010) The solution structure of the ADAR2 dsRBM-RNA complex reveals a sequence-specific readout of the minor groove. *Cell*, **143**, 225-237.
36. Tian, B., Bevilacqua, P.C., Diegelman-Parente, A. and Mathews, M.B. (2004) The double-stranded-RNA-binding motif: interference and much more. *Nature reviews Molecular cell biology*, **5**, 1013-1023.
37. Masliah, G., Barraud, P. and Allain, F.H.-T. (2013) RNA recognition by double-stranded RNA binding domains: a matter of shape and sequence. *Cell. Mol. Life Sci.*, **70**, 1875-1895.
38. Ramos, A., Grünert, S., Adams, J., Micklem, D.R., Proctor, M.R., Freund, S., Bycroft, M., St Johnston, D. and Varani, G. (2000) RNA recognition by a Staufen double-stranded RNA-binding domain. *The EMBO Journal*, **19**, 997-1009.
39. Wilson, D.K., Rudolph, F.B. and Quiocho, F.A. (1991) Atomic structure of adenosine deaminase complexed with a transition-state analog: understanding catalysis and immunodeficiency mutations. *Science*, **252**, 1278-1284.
40. Betts, L., Xiang, S., Short, S.A., Wolfenden, R. and Carter, C.W. (1994) Cytidine deaminase. The 2· 3 Å crystal structure of an enzyme: transition-state analog complex. *Journal of molecular biology*, **235**, 635-656.
41. Macbeth, M.R., Schubert, H.L., VanDemark, A.P., Lingam, A.T., Hill, C.P. and Bass, B.L. (2005) Inositol hexakisphosphate is bound in the ADAR2 core and required for RNA editing. *Science*, **309**, 1534-1539.
42. Eifler, T., Chan, D. and Beal, P.A. (2012) A Screening Protocol for Identification of Functional Mutants of RNA Editing Adenosine Deaminases. *Current protocols in chemical biology*, 357-369.
43. Polson, A.G. and Bass, B.L. (1994) Preferential selection of adenosines for modification by double-stranded RNA adenosine deaminase. *The EMBO journal*, **13**, 5701.
44. Hough, R.F. and Bass, B.L. (1997) Analysis of Xenopus dsRNA adenosine deaminase cDNAs reveals similarities to DNA methyltransferases. *Rna*, **3**, 356-370.
45. Stephens, O.M., Yi-Brunozzi, H.Y. and Beal, P.A. (2000) Analysis of the RNA-editing reaction of ADAR2 with structural and fluorescent analogues of the GluR-B R/G editing site. *Biochemistry*, **39**, 12243-12251.
46. Cho, D.-S.C., Yang, W., Lee, J.T., Shiekhattar, R., Murray, J.M. and Nishikura, K. (2003) Requirement of dimerization for RNA editing activity of adenosine deaminases acting on RNA. *Journal of Biological Chemistry*, **278**, 17093-17102.

47. Nishikura, K., Yoo, C., Kim, U., Murray, J., Estes, P., Cash, F. and Liebhaber, S. (1991) Substrate specificity of the dsRNA unwinding/modifying activity. *The EMBO journal*, **10**, 3523.
48. Eggington, J.M., Greene, T. and Bass, B.L. (2011) Predicting sites of ADAR editing in double-stranded RNA. *Nature communications*, **2**, 319.
49. 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.*, DOI: 10.1093/nar/gku1272.
50. Montiel-Gonzalez, M.F., Vallecillo-Viejo, I., Yudowski, G.A. and Rosenthal, J.J. (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.
51. Stephens, O.M., Haudenschild, B.L. and Beal, P.A. (2004) The binding selectivity of ADAR2's dsRBMs contributes to RNA-editing selectivity. *Chemistry & biology*, **11**, 1239-1250.
52. Carlson, C.B., Stephens, O.M. and Beal, P.A. (2003) Recognition of double-stranded RNA by proteins and small molecules. *Biopolymers*, **70**, 86-102.
53. Kuttan, A. and Bass, B.L. (2012) Mechanistic insights into editing-site specificity of ADARs. *Proc. Natl. Acad. Sci. U.S.A.*, **109**, E3295-E3304.
54. Wang, Y., Havel, J. and Beal, P.A. (2015) A Phenotypic Screen for Functional Mutants of Human Adenosine Deaminase Acting on RNA 1. *ACS Chemical Biology*.
55. Daujotyt, D., Serva, S., Vilkaitis, G., Venclovas, Č. and Klimašauskas, S. (2004) HhaI DNA methyltransferase uses the protruding Gln237 for active flipping of its target cytosine. *Structure*, **12**, 1047-1055.
56. Klimasauskas, S., Kumar, S., Roberts, R.J. and Cheng, X. (1994) Hhal methyltransferase flips its target base out of the DNA helix. *Cell*, **76**, 357-369.
57. Stafforst, T. and Schneider, M.F. (2012) An RNA–deaminase conjugate selectively repairs point mutations. *Angew. Chem., Int. Ed.*, **51**, 11166-11169.
58. Keppler, A., Gendreizig, S., Gronemeyer, T., Pick, H., Vogel, H. and Johnsson, K. (2003) A general method for the covalent labeling of fusion proteins with small molecules in vivo. *Nature biotechnology*, **21**, 86-89.
59. 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.*, **53**, 6267-6271.
60. Hanswillemenke, A., Kuzdere, T., Vogel, P., Jekely, G. and Stafforst, T. (2015) Site-directed RNA editing in vivo can be triggered by the light-driven assembly of an artificial riboprotein. *Journal of the American Chemical Society*.
61. Chattopadhyay, S., Garcia-Mena, J., DeVito, J., Wolska, K. and Das, A. (1995) Bipartite function of a small RNA hairpin in transcription antitermination in bacteriophage lambda. *Proceedings of the National Academy of Sciences*, **92**, 4061-4065.
62. Austin, R.J., Xia, T., Ren, J., Takahashi, T.T. and Roberts, R.W. (2002) Designed arginine-rich RNA-binding peptides with picomolar affinity. *Journal of the American Chemical Society*, **124**, 10966-10967.
63. LaFountaine, J.S., Fathe, K. and Smyth, H.D. (2015) Delivery and therapeutic applications of gene editing technologies ZFNs, TALENs, and CRISPR/Cas9. *International journal of pharmaceutics*, **494**, 180-194.

64. Ginn, S.L., Alexander, I.E., Edelstein, M.L., Abedi, M.R. and Wixon, J. (2013) Gene therapy clinical trials worldwide to 2012–an update. *The journal of gene medicine*, **15**, 65-77.
65. Richards, S. (2012) Gene therapy arrives in Europe. *The Scientist*, **1**, 2.
66. Essmann, F. and Schulze-Osthoff, K. (2012) Translational approaches targeting the p53 pathway for anti-cancer therapy. *British journal of pharmacology*, **165**, 328-344.
67. Fischer, U. and Schulze-Osthoff, K. (2005) New approaches and therapeutics targeting apoptosis in disease. *Pharmacological reviews*, **57**, 187-215.
68. Zhou, Y., Liu, S., Liu, G., Öztürk, A. and Hicks, G.G. (2013) ALS-associated FUS mutations result in compromised FUS alternative splicing and autoregulation.
69. 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. *Nature cell biology*, **12**, 119-131.
70. Capecchi, M.R. (1989) Altering the genome by homologous recombination. *Science*, **244**, 1288-1292.
71. Bibikova, M., Beumer, K., Trautman, J.K. and Carroll, D. (2003) Enhancing gene targeting with designed zinc finger nucleases. *Science*, **300**, 764-764.
72. Bibikova, M., Golic, M., Golic, K.G. and Carroll, D. (2002) Targeted chromosomal cleavage and mutagenesis in Drosophila using zinc-finger nucleases. *Genetics*, **161**, 1169-1175.
73. Bitinaite, J., Wah, D.A., Aggarwal, A.K. and Schildkraut, I. (1998) FokI dimerization is required for DNA cleavage. *Proceedings of the national academy of sciences*, **95**, 10570-10575.
74. Wijshake, T., Baker, D.J. and van de Sluis, B. (2014) Endonucleases: new tools to edit the mouse genome. *Biochimica et Biophysica Acta (BBA)-Molecular Basis of Disease*, **1842**, 1942-1950.
75. Boch, J., Scholze, H., Schornack, S., Landgraf, A., Hahn, S., Kay, S., Lahaye, T., Nickstadt, A. and Bonas, U. (2009) Breaking the code of DNA binding specificity of TAL-type III effectors. *Science*, **326**, 1509-1512.
76. Moscou, M.J. and Bogdanove, A.J. (2009) A simple cipher governs DNA recognition by TAL effectors. *Science*, **326**, 1501-1501.
77. Mussolino, C., Morbitzer, R., Lütge, F., Dannemann, N., Lahaye, T. and Cathomen, T. (2011) A novel TALE nuclease scaffold enables high genome editing activity in combination with low toxicity. *Nucleic acids research*, **39**, 9283-9293.
78. Mali, P., Yang, L., Esvelt, K.M., Aach, J., Guell, M., DiCarlo, J.E., Norville, J.E. and Church, G.M. (2013) RNA-guided human genome engineering via Cas9. *Science*, **339**, 823-826.
79. Cho, S.W., Kim, S., Kim, Y., Kweon, J., Kim, H.S., Bae, S. and Kim, J.-S. (2014) Analysis of off-target effects of CRISPR/Cas-derived RNA-guided endonucleases and nickases. *Genome research*, **24**, 132-141.
80. Kole, R., Krainer, A.R. and Altman, S. (2012) RNA therapeutics: beyond RNA interference and antisense oligonucleotides. *Nature reviews Drug discovery*, **11**, 125-140.
81. Lee, R.C., Feinbaum, R.L. and Ambros, V. (1993) The C. elegans heterochronic gene lin-4 encodes small RNAs with antisense complementarity to lin-14. *cell*, **75**, 843-854.

82. Bora, R.S., Gupta, D., Mukkur, T.K.S. and Saini, K.S. (2012) RNA interference therapeutics for cancer: challenges and opportunities (review). *Molecular medicine reports*, **6**, 9-15.
83. Hutvágner, G. and Zamore, P.D. (2002) A microRNA in a multiple-turnover RNAi enzyme complex. *Science*, **297**, 2056-2060.
84. (!!! INVALID CITATION !!! {}).
85. Kappel, S., Matthess, Y., Kaufmann, M. and Strebhardt, K. (2007) Silencing of mammalian genes by tetracycline-inducible shRNA expression. *Nature protocols*, **2**, 3257-3269.
86. Kibbe, W.A. (2007) OligoCalc: an online oligonucleotide properties calculator. *Nucleic acids research*, **35**, W43-W46.
87. Ho, S.N., Hunt, H.D., Horton, R.M., Pullen, J.K. and Pease, L.R. (1989) Site-directed mutagenesis by overlap extension using the polymerase chain reaction. *Gene*, **77**, 51-59.
88. Gibson, D.G., Young, L., Chuang, R.-Y., Venter, J.C., Hutchison, C.A. and Smith, H.O. (2009) Enzymatic assembly of DNA molecules up to several hundred kilobases. *Nature methods*, **6**, 343-345.
89. Gould, S.J. and Subramani, S. (1988) Firefly luciferase as a tool in molecular and cell biology. *Analytical biochemistry*, **175**, 5-13.
90. Masliah, G., Barraud, P. and Allain, F.H.-T. (2013) RNA recognition by double-stranded RNA binding domains: a matter of shape and sequence. *Cellular and Molecular Life Sciences*, **70**, 1875-1895.
91. Milligan, J.F. and Uhlenbeck, O.C. (1989) Synthesis of small RNAs using T7 RNA polymerase. *Methods in enzymology*, **180**, 51.
92. Pleiss, J.A., Derrick, M.L. and UHLENBECK, O.C. (1998) T7 RNA polymerase produces 5′ end heterogeneity during in vitro transcription from certain templates. *Rna*, **4**, 1313-1317.
93. Helm, M., Brule, H., Giege, R. and Florentz, C. (1999) More mistakes by T7 RNA polymerase at the 5'ends of in vitro-transcribed RNAs. *RNA*, **5**, 618-621.
94. 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.
95. Stefl, R., Skrisovska, L. and Allain, F.H.T. (2005) RNA sequence-and shape-dependent recognition by proteins in the ribonucleoprotein particle. *EMBO reports*, **6**, 33-38.
96. Öhman, M., Källman, A.M. and Bass, B.L. (2000) In vitro analysis of the binding of ADAR2 to the pre-mRNA encoding the GluR-B R/G site. *RNA*, **6**, 687-697.
97. Kerstan, D. and Quamme, G. (2002), *Calcium in internal medicine*. Springer, pp. 171-183.
98. Mach, P. (2015) Optimaziation of the encodable RNA Editing System.
99. Galeano, F., Leroy, A., Rossetti, C., Gromova, I., Gautier, P., Keegan, L.P., Massimi, L., Di Rocco, C., O'Connell, M.A. and Gallo, A. (2010) Human BLCAP transcript: new editing events in normal and cancerous tissues. *International Journal of Cancer*, **127**, 127-137.
100. Isom, G.E. and Borowitz, J. (2013) " PC-12 Cells. *In Vitro Biological Systems: Methods in Toxicology*, **1**.
101. Fischer, A.H., Henrich, T. and Arendt, D. (2010) The normal development of Platynereis dumerilii (Nereididae, Annelida). *Frontiers in zoology*, **7**, 31.

102. Zantke, J., Bannister, S., Rajan, V.B.V., Raible, F. and Tessmar-Raible, K. (2014) Genetic and genomic tools for the marine annelid Platynereis dumerilii. *Genetics*, **197**, 19-31.
103. Hartmann, R.K., Bindereif, A., Sch_n, A. and Westhof, E. (2015) *Handbook of RNA biochemistry*. John Wiley & Sons.
104. Conrad, F., Hanne, A., Gaur, R.K. and Krupp, G. (1995) Enzymatic synthesis of 2′-modified nucleic acids: Identification of important phosphate and ribose moieties in RNase P substrates. *Nucleic acids research*, **23**, 1845-1853.
105. Sousa, R. and Padilla, R. (1995) A mutant T7 RNA polymerase as a DNA polymerase. *The EMBO journal*, **14**, 4609.