

Establishing Site-Directed A-to-I RNA Editing in Cell Culture

Dissertation

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

vorgelegt von
Paul Vogel
aus Dresden

Tübingen
2018

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

Tag der mündlichen Qualifikation:	02.08.2018
Dekan:	Prof. Dr. Wolfgang Rosenstiel
1. Berichterstatter:	Prof. Dr. Thorsten Stafforst
2. Berichterstatter:	Prof. Dr. Ralf-Peter Jansen
3. Berichterstatter:	Prof. Dr. Michael F. Jantsch

Danksagung/Acknowledgments

Diese Arbeit wäre ohne die Unterstützung einer ganzen Reihe von Personen nicht möglich gewesen. Ich möchte hiermit die Gelegenheit nutzen, mich aufrichtig zu bedanken.

Mein großer Dank gilt meinem Doktorvater Prof. Dr. Thorsten Stafforst, der mich die ganze Zeit offenherzig unterstützt hat, zahlreiche Ideen gegeben hat, immer für wissenschaftliche Diskussionen verfügbar war und mir alle Freiheiten zur Verfügung gestellt hat, meine wissenschaftlichen Projekte zu verwirklichen. Ihm habe ich sehr viel zu verdanken und hoffe, mit meiner Arbeit etwas zurückgegeben zu haben. Ich möchte mich auch bei meinem Zweitbetreuer Prof. Dr. Ralf-Peter Jansen herzlichst für sein Interesse an meiner Arbeit und der Verfügbarkeit für den wissenschaftlichen Austausch bedanken.

Ein großes Dankschön möchte ich auch an meine ehemaligen und jetzigen Kollegen im Team Stafforst aussprechen. Danke für die gegenseitige Unterstützung und die super Atmosphäre innerhalb und außerhalb des Labors! Was ich besonders an dieser Arbeitsgruppe schätze ist, dass jeder so sein darf wie er ist. Dr. Jacqueline Wettengel danke ich für die Einführung in den ersten Tagen im Stafforst Labor. Mein Dank gilt auch Alfred Hanswillemenke, Martin Moschref, Tobias Merkle, Karthika Devi Selvasaravanan, Matthias Becker und Antonia Lott. Sie alle haben tatkräftig zum Gelingen meiner Projekte beigetragen.

I also want to thank Nupur Bhatt (NYU School of Medicine, USA) and Sinem Usluer (Sabanci University, Turkey) for their internship in our lab and their contributions to my projects. Special thanks go to Dr. Qin Li and Prof. Dr. Jin Billy Li at Stanford University, USA. Without their help and expertise in the NGS-based identification of global A-to-I sites, the publication in Nature Methods would not have been possible.

Ich möchte mich bei Lara Reichert für das Korrekturlesen meiner Arbeit bedanken. Außerdem gilt mein Dank meinen Freunden aus Tübingen und Dresden für die schöne gemeinsame Zeit und die dauerhafte Unterstützung.

Sophia, vielen Dank für deine Mitarbeit im Labor, das Korrekturlesen der Arbeit, deiner Geduld und deiner bedingungslosen Unterstützung. Danke, dass es dich gibt!

Schlussendlich möchte ich mich bei meinen Eltern und Geschwistern bedanken. Vielen Dank für eure Unterstützung und euren Rückhalt!

Table of contents

List of abbreviations	I
Summary	III
Zusammenfassung	V
List of publications and personal contributions	VII
1 Introduction	1
1.1 Inosine found in RNA.....	1
1.2 A-to-I RNA editing catalyzed by ADARs.....	1
1.2.1 ADAR enzymes.....	1
1.2.2 Physiological significance of ADARs.....	3
1.2.3 Regulation of A-to-I RNA editing	6
1.2.4 Structural and mechanistic insights into ADAR catalysis	8
1.3 Targeting RNA as therapeutic strategy	10
1.4 Novel concepts for manipulating genetic information	13
1.4.1 Genome engineering	13
1.4.2 Site-directed A-to-I RNA editing.....	17
2 Aims of this study	20
3 Results and discussion	21
3.1 SNAP-ADAR enzymes enable site-directed RNA editing in cell culture.....	21
3.2 SNAP-ADAR enzymes as a promising tool for medicine and life sciences	28
3.2.1 Correcting disease-causing mutations.....	28
3.2.2 Reversible manipulation of signaling networks.....	29
3.2.3 Switching protein localization by light	30
3.3 SNAP-ADAR enzymes have a promising safety profile.....	35
4 Conclusion	44
References	46
Curriculum vitae	56
Conference contributions	57
Manuscripts	58

List of abbreviations

ACTB	β -actin
ADAR	adenosine deaminase acting on RNA
ADAR-D	ADAR deaminase domain
AMPA	α -amino-3-hydroxy-5-methyl-4-isoxazolepropionic acid
A-to-I	adenosine-to-inosine
BG	benzylguanine
cEt	constrained ethyl
CLOCK	circadian locomotor output cycles kaput
CREB	cAMP response element-binding protein
CRISPR	clustered regularly interspaced short palindromic repeats
dsRBD	dsRNA-binding domain
dsRNA	double-stranded RNA
eCFP	enhanced cyan fluorescent protein
eGFP	enhanced green fluorescent protein
ENA	ethylene-bridged nucleic acid
2'-F	2'-fluoro
FBS	fetal bovine serum
GAPDH	glyceraldehyde 3-phosphate dehydrogenase
GRIA2	glutamate receptor subunit B
gRNA	guide RNA
GUSB	β -glucuronidase
IgK	immunoglobulin K
LNA	locked nucleic acid
MAVS	mitochondrial antiviral signaling protein
MDA5	melanoma differentiation-associated gene 5
NGS	next-generation sequencing
NLS	nuclear localization signal
2'-OMe	2'- <i>O</i> -methyl
ORF	open reading frame
PDGF	platelet-derived growth factor
Pin1	peptidylprolyl cis/trans isomerase, NIMA-interacting 1
PRPS1	phosphoribosyl pyrophosphate synthetase 1

II *LIST OF ABBREVIATIONS*

PS	phosphorothioate
SA	SNAP-ADAR
SNP	single nucleotide polymorphism
STAT1	signal transducer and activator of transcription 1
UTR	untranslated region
WWP2	WW domain-containing E3 ubiquitin protein ligase 2

Summary

Tools to manipulate genetic information without interfering at the DNA level are highly desirable in medicine and the life sciences. Recently, our group introduced the first engineered, RNA-guided deaminase. The approach relies on the *in situ* covalent bond formation between a benzylguanine-modified guide RNA (BG-gRNA) and a SNAP-tagged deaminase (SNAP-ADAR). Once the gRNA-deaminase conjugate is formed, it enables specific adenosine-to-inosine (A-to-I) substitutions in target RNAs. Since inosine is biochemically interpreted as guanosine by the cellular machinery, site-directed A-to-I editing provides the possibility to manipulate RNA and protein function.

In this PhD project, it was aimed at elucidating the potential of the SNAP-ADAR approach for future applications. Therefore, the performance of the editing system in mammalian cells was comprehensively characterized.

It could be shown that efficient site-directed RNA editing with SNAP-ADAR enzymes in cell culture requires the chemical modification of the BG-gRNA. A strong performance in the editing of endogenous transcripts was demonstrated in engineered cell lines stably expressing SNAP-ADAR enzymes. Editing yields up to 90% were achieved and remained stable even when several transcripts or multiple sites on a single transcript were concurrently targeted. Maximum editing was reached after 3 hours of BG-gRNA transfection and stayed unchanged for several days. Additionally, low concentrations (≥ 1.25 pmol/96-well) of the BG-gRNA were sufficient to obtain highest editing levels. The SNAP-ADAR approach holds great promise for the recoding of many functionally important amino acid residues as 11 out of the 16 adenosine-containing 5'-NAN triplets were editable between 50% and 90%.

First evidence was provided that the editing system might be a valuable tool for the correction of disease-causing mutations. Moreover, the possibility of manipulating entire signaling networks was highlighted by the efficient and concurrent editing of two disease-relevant transcripts, *KRAS* and *STAT1*. Photo-controlled A-to-I editing was applied to direct protein localization within the cell by introducing alternative start or stop codons which allowed the expression of signals for nuclear and membrane translocation.

NGS-based analysis revealed that wild-type SNAP-ADAR enzymes are highly precise editing machines. Their more active versions (SNAP-ADARQ enzymes) produced some off-target edits into the transcriptome, but the observed off-target activity appeared to be reducible by lowering the SNAP-ADAR protein amounts without great

IV *SUMMARY*

inhibition of the on-target editing. Nevertheless, these enzymes were one order more precise than editing machines applied by competing approaches. The chemical modification of the BG-gRNA was shown to suppress the off-target editing within a duplex formed by the BG-gRNA and the target RNA.

The SNAP-ADAR approach outcompetes all well-characterized approaches for site-directed RNA editing due the best balance between efficiency and specificity.

Zusammenfassung

Werkzeuge, die die Veränderung genetischer Informationen ohne den Eingriff auf DNA-Ebene gewährleisten, sind sehr attraktiv für die Medizin und den Lebenswissenschaften. Kürzlich präsentierte unsere Arbeitsgruppe die erste künstliche RNA-gesteuerte Deaminase. Der Ansatz baut auf die kovalente Verknüpfung einer Benzylguanin-modifizierten guide-RNA (BG-gRNA) und einer SNAP-Tag-fusionierten Deaminase (SNAP-ADAR). Sobald sich das gRNA-Deaminasekonjugat bildet, ermöglicht es die spezifische Umwandlung von Adenosin zu Inosin (A-zu-I) in Ziel-RNAs. Da Inosin von der zellulären Maschinerie biochemisch als Guanosin interpretiert wird, bietet die gerichtete A-zu-I-Editierung die Möglichkeit, Funktionen von RNAs und Proteinen zu verändern.

Ziel dieser Arbeit war es, das Potenzial des SNAP-ADAR-Ansatzes für zukünftige Anwendungen zu beleuchten. Dazu wurde die Leistungsfähigkeit des Editierungssystems in Säugetierzellen umfangreich charakterisiert.

Es konnte gezeigt werden, dass die gerichtete RNA-Editierung mit SNAP-ADAR-Enzymen die chemische Modifizierung der BG-gRNA benötigt, um in der Zellkultur effizient zu sein. Eine starke Leistung in der Editierung endogener Transkripte wurde in Zelllinien demonstriert, die stabil SNAP-ADAR-Enzyme exprimierten. Editierungsausbeuten bis zu 90% wurden erreicht und blieben stabil, selbst als mehrere Transkripte oder mehrere Stellen auf einem Transkript gleichzeitig editiert wurden. Das Maximum an Editierung wurde 3 Stunden nach der BG-gRNA-Transfektion erreicht und blieb mehrere Tage lang unverändert. Zusätzlich waren geringe Konzentrationen ($\geq 1,25$ pmol/96-Well) an BG-gRNA ausreichend, um die höchsten Editierungsausbeuten zu erhalten. Der SNAP-ADAR-Ansatz ist sehr vielversprechend für die Umcodierung funktionell wichtiger Aminosäurereste, da 11 von den 16 Adenosin-enthaltenden 5'-NAN Triplets zwischen 50% und 90% editierbar waren.

Erste Ergebnisse zeigten, dass das Editierungssystem ein wertvolles Werkzeug für Korrektur von krankheitsverursachenden Mutationen sein könnte. Des Weiteren wurde die Möglichkeit zur Manipulation ganzer Signalnetzwerke durch das effiziente und gleichzeitige Editieren von zwei krankheitsrelevanten Transkripten (*KRAS* und *STAT1*) hervorgehoben. Photokontrollierte A-zu-I-Editierung wurde angewendet, um die Proteinlokalisierung in der Zelle durch das Einbringen alternativer Start- oder Stopcodonen

VI ZUSAMMENFASSUNG

zu steuern. Dies erlaubte die Expression von Signalen für die Zellkern- und Membranlokalisation.

Die NGS-basierte Analyse zeigte, dass Wildtyp-SNAP-ADAR-Enzyme sehr präzise Editiermaschinen sind. Ihre aktiveren Versionen (SNAP-ADARQ-Enzyme) produzierten einige Off-Target-Editierungen im Transkriptom, jedoch erschien die Off-Target-Aktivität durch die Verringerung der SNAP-ADAR-Proteinmenge reduzierbar, ohne dabei großartig die Editierung an der Zielstelle zu inhibieren. Nichtsdestotrotz waren diese Enzyme um eine Größenordnung präziser als Editiermaschinen, die von konkurrierenden Ansätzen genutzt werden. Es wurde gezeigt, dass die chemische Modifikation der BG-gRNA die Off-Target-Editierung innerhalb eines Duplexes unterdrückt, der von der BG-gRNA und Ziel-RNA gebildet wird.

Der SNAP-ADAR-Ansatz übertrifft alle gut charakterisierten Ansätze zur gerichteten RNA-Editierung aufgrund des besten Verhältnisses zwischen Effizienz und Spezifität.

List of publications and personal contributions

Original research articles (peer-reviewed)

Manuscript (Man.) 1: accepted

Vogel, P., Schneider, M. F., Wettengel, J. & Stafforst, T. Improving site-directed RNA editing in vitro and in cell culture by chemical modification of the guideRNA. *Angew. Chem. Int. Ed.* **53**, 6267-6271 (2014).

Personal contribution: I designed and conducted all cell culture experiments. I analyzed and interpreted the data which were generated from these experiments (Fig. 2, Suppl. Fig. 6,7). I contributed to the preparation of the supporting information. This included the preparation of Suppl. Fig. 6,7 and the description of the methods used for the cell culture experiments. Additionally, I was involved in the discussion of the manuscript.

Man. 2: accepted

Hanswillemenke, A., Kuzdere, T., Vogel, P., Jékely, G. & Stafforst, T. Site-directed RNA editing in vivo can be triggered by the light-driven assembly of an artificial riboprotein. *J. Am. Chem. Soc.* **137**, 15875-15881 (2015).

Personal contribution: Together with A.H., I optimized the assay for performing site-directed A-to-I RNA editing in mammalian cells. I gave scientific advice for conducting the final cell culture experiments and created the *eGFP W58X* reporter plasmid used in these experiments. I was involved in the discussion of the manuscript. In the revision process of the manuscript, I performed additional experiments to examine the influence of the SNAP-ADAR editing system on known editing sites of ADAR substrates. The data were only shown to the reviewers and were not included into the manuscript.

Man. 3: accepted

Vogel, P., Hanswillemenke, A. & Stafforst, T. Switching protein localization by site-directed RNA editing under control of light. *ACS Synth. Biol.* **6**, 1642-1649 (2017).

Personal contribution: I generated the stably transfected Flp-In T-REx cell lines, created the used plasmids and synthesized the BG-modified gRNAs. I designed all experiments under the assistance of all authors and conducted them. Together with all authors, I analyzed and interpreted the generated data. I wrote the methods section of the manuscript and prepared the supporting information. Additionally, I was involved in the discussion of the manuscript.

Man. 4: accepted

Vogel, P., Moschref, M., Li, Q., Merkle, T., Selvasarayanan, K. D., Li, J. B. & Stafforst, T. Efficient and precise editing of endogenous transcripts with SNAP-tagged ADARs. *Nat. Methods* **15**, 535-538 (2018).

Personal contribution: I generated the stably transfected Flp-In T-REx cell lines and synthesized the BG-modified gRNAs, excepting the BG-gRNA targeting STAT1 Y701. I designed and conducted the experiments shown in Fig. 1a,1b,1d,2a,2e (2e under the assistance of K.D.S.), Suppl. Fig. 2-4,7,9,12,13,15 and Suppl. Note 1,3,4. I analyzed and interpreted the data which were generated from these experiments. I prepared the Fig. 1,2 (both figures under the assistance of T.S.) together with Table 1 in the manuscript. I wrote parts (BG-gRNA synthesis, SNAP-ADAR-expressing cell lines and RNA editing experiments) of the methods section and prepared the supporting information, excepting Suppl. Note. 1,2 and Suppl. Table 2,3. I was involved in the discussion of the manuscript. In the revision process of the manuscript, I performed additional experiments to examine the effect of the nuclear localization of SA enzymes on the editing yield. The data were only shown to the reviewers and were not included into the manuscript. However, the results (termed as non-included data) can be found attached to the manuscript.

Review articles (peer-reviewed)

Man. 5: accepted

Vogel, P. & Stafforst, T. Site-directed RNA editing with antagomir deaminases - A tool to study protein and RNA function. *ChemMedChem* **9**, 2021-2025 (2014).

Personal contribution: I wrote a first draft and was involved in the discussion of the final manuscript version.

Man. 6: accepted

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

Personal contribution: I contributed to the concept of the review article and wrote the parts: “RNaseH-dependent antisense oligonucleotides”, “Splice-switching oligonucleotides” and “Aptamers”. I prepared Fig. 1,2 and was involved in the discussion of the manuscript.

1 Introduction

1.1 Inosine found in RNA

According to the MODOMICS database, more than 150 different modifications have been found in RNA molecules.¹ Over the last 60 years, biochemical studies have well documented the occurrence and effects of modifications in tRNA and rRNA. High-throughput sequencing technologies have recently provided a global view on RNA modifications within transcripts and have led to a revival of interest in the epitranscriptome. RNA modifications have been shown to be abundant, conserved and dynamically regulated. Besides other nucleosides with base modifications, inosine is currently being studied intensively to learn more about its functions in RNA.

The conversion of adenosine to inosine (A-to-I editing) frequently happens within double-stranded RNA (dsRNA) regions in the transcriptome of metazoans (**Fig. 1a**).^{2,3} In contrast to adenosine, inosine preferentially interacts with cytidine and is biochemically interpreted as guanosine by the cellular machinery. Editing sites within a known RNA substrate can be validated by Sanger sequencing to compare the sequences of genomic DNA (gDNA) and complementary DNA (cDNA). Since A-to-I editing does not occur in every RNA molecule, cDNA sequencing normally shows a mixture of adenosine and guanosine at the respective site. To exclude the possibility that the mixed signal derives from a genomic SNP, gDNA sequencing should only identify an adenosine. Transcriptome-wide identification of A-to-I editing sites has been achieved by next-generation sequencing (NGS)-based methods comparing DNA and RNA sequencing, using RNA sequencing data alone or applying chemical labeling of inosine.⁴ So far, researchers have identified millions of editing sites.⁵⁻¹⁵ It has been shown that A-to-I editing is catalyzed by adenosine deaminases acting on RNA (ADARs) and regulates many cellular processes by manipulating protein functions, RNA splicing, immunogenicity and RNA interference.

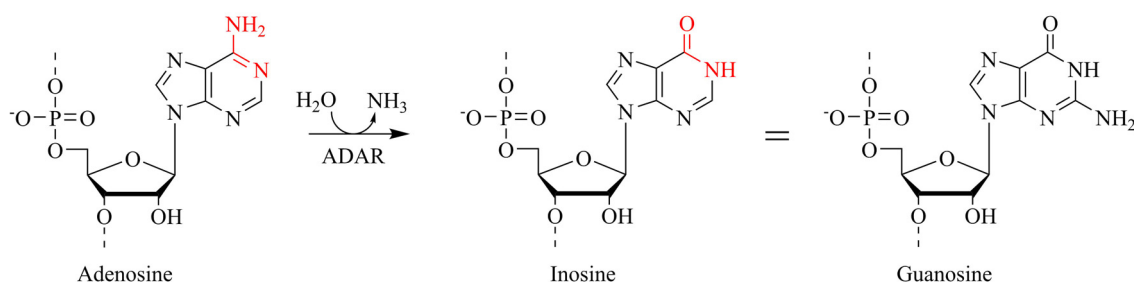
1.2 A-to-I RNA editing catalyzed by ADARs

1.2.1 ADAR enzymes

Three ADARs are encoded in the human genome and share common structural features including multiple N-terminal dsRNA-binding domains (RBDs) and a deaminase domain at the C-terminus (**Fig. 1b**).

2 INTRODUCTION

a



b

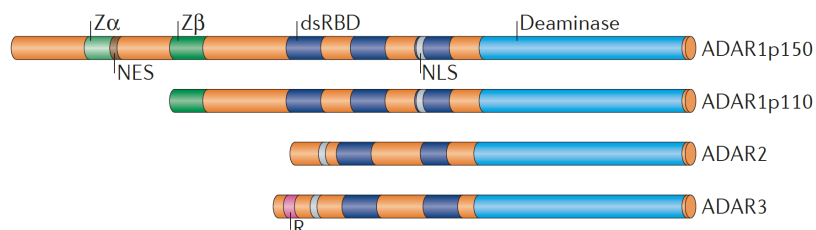


Figure 1 A-to-I RNA editing mediated by ADARs. **(a)** ADARs catalyze the hydrolytic conversion of adenosine to inosine which is biochemically interpreted as guanosine. **(b)** There are three ADAR enzymes existing in humans: ADAR1 expressed as two isoforms (p110 and p150), ADAR2 and ADAR3. All these enzymes have similar structural features. dsRBD, dsRNA-binding domain; NES, nuclear export signal; NLS, nuclear localization signal; R domain, arginine-rich domain; Z α and Z β , Z-DNA-binding domains. Adapted from ref. 2.

Constitutive gene expression results in the generation of ADAR1 as a 110-kDa protein (ADAR1p110) containing an N-terminal Z-DNA-binding domain together with three dsRBDs.^{16,17} However, ADAR1 can also be translated as a 150-kDa protein (ADAR1p150) initiated through an alternative promoter inducible by interferon (IFN).^{16,17} ADAR1p150 harbors an additional Z-DNA-binding domain and a nuclear export signal (NES) in the extended N-terminus compared to the shorter ADARp110. Both protein isoforms can shuttle between nucleus and cytoplasm, but ADAR1p110 is mainly localized in the nucleus and ADARp150 accumulates in the cytoplasm due to its NES.^{18,19} ADAR2 contains two dsRBDs and has been found to be restricted to the nucleus.²⁰⁻²² Similarly, ADAR3 has also two dsRBDs, but contains an arginine-rich (R) domain which enables binding of single-stranded RNA (ssRNA) and serves as a signal for nuclear import.^{23,24} Expression analyses revealed that *ADAR3* transcripts can only be detected in the brain.^{13,23} In contrast, *ADAR1* and *ADAR2* are expressed across multiple tissues.^{13,20,25} Whereas ADAR1 and ADAR2 have been proven to be catalytically active, ADAR3 lacks any evidence to mediate A-to-I RNA editing.

1.2.2 Physiological significance of ADARs

Only a very small fraction of editing sites causes amino acid substitutions in proteins. A prominent example is the transcript encoding the glutamate receptor subunit B (GRIA2). In *GRIA2* pre-mRNA, intron 11 and exon 11 form an imperfect RNA duplex, where editing results in the change of a glutamine (Q) codon (CAG) to an arginine (R) codon (CIG) in the coding sequence (**Fig. 2a**).²⁶ It has been shown that this site is only efficiently edited by ADAR2.²⁰ AMPA receptors containing GRIA2 with the Q/R substitution are impermeable to calcium.²⁷ *ADAR2*-null mice exhibit early onset epilepsy and die early after birth, but the phenotype can be fully rescued by the homozygous knock-in of the Q/R mutation at the edited site of *GRIA2*.²⁸ RNA editing of the Q/R site has been shown to be reduced in motor neurons from human patients with amyotrophic lateral sclerosis (ALS) and is therefore postulated to be necessary for neuronal integrity (**Fig. 2b**).²⁹ In mice, the death of motor neurons induced by the lack of *ADAR2* is indeed prevented by the expression of Q/R site-edited *GRIA2*.³⁰ *GRIA2* pre-mRNA harbors another recoding editing site which was also identified in the transcripts encoding two further AMPA receptor subunits (GRIA3 and GRIA4).³¹ In exon 13, which forms an imperfect hairpin structure with its adjacent intron, editing changes an arginine (R) codon (AGG) to a glycine (G) codon (IGG) and is efficiently catalyzed by both ADAR1 and ADAR2 (**Fig. 2a**).²⁰ The level of editing at the R/G site increases in rat brain during development and glutamate receptor subunits harboring the R/G substitution accelerate the recovery of AMPA receptors from desensitization.³¹

Further editing sites mediated by ADARs have been discovered and characterized in terms of their impact on protein properties. For instance, editing of 5 adenosines in the transcript encoding the serotonin receptor 2C (5-HT_{2C}R) changes the amino acid composition within the second intracellular loop and reduces the efficiency of the coupling between the receptor and G proteins (**Fig. 2a**).^{32,33} In *Kv1.1* (voltage-gated K⁺ channel subfamily A member 1) transcripts, editing leads to the recoding of isoleucine to valine in the ion-conducting pore and enables the channel to recover faster from inactivation (**Fig. 2a**).³⁴

Almost all editing sites identified in the human transcriptome have been found in *Alu* sequences which are highly repetitive elements generally located within introns and untranslated regions.⁵⁻¹⁵ *Alu* sequences are primate-specific retrotransposons of ~ 300 bp in length and comprise more than 10% of the human genome.³⁵ Two repetitive *Alu*

4 INTRODUCTION

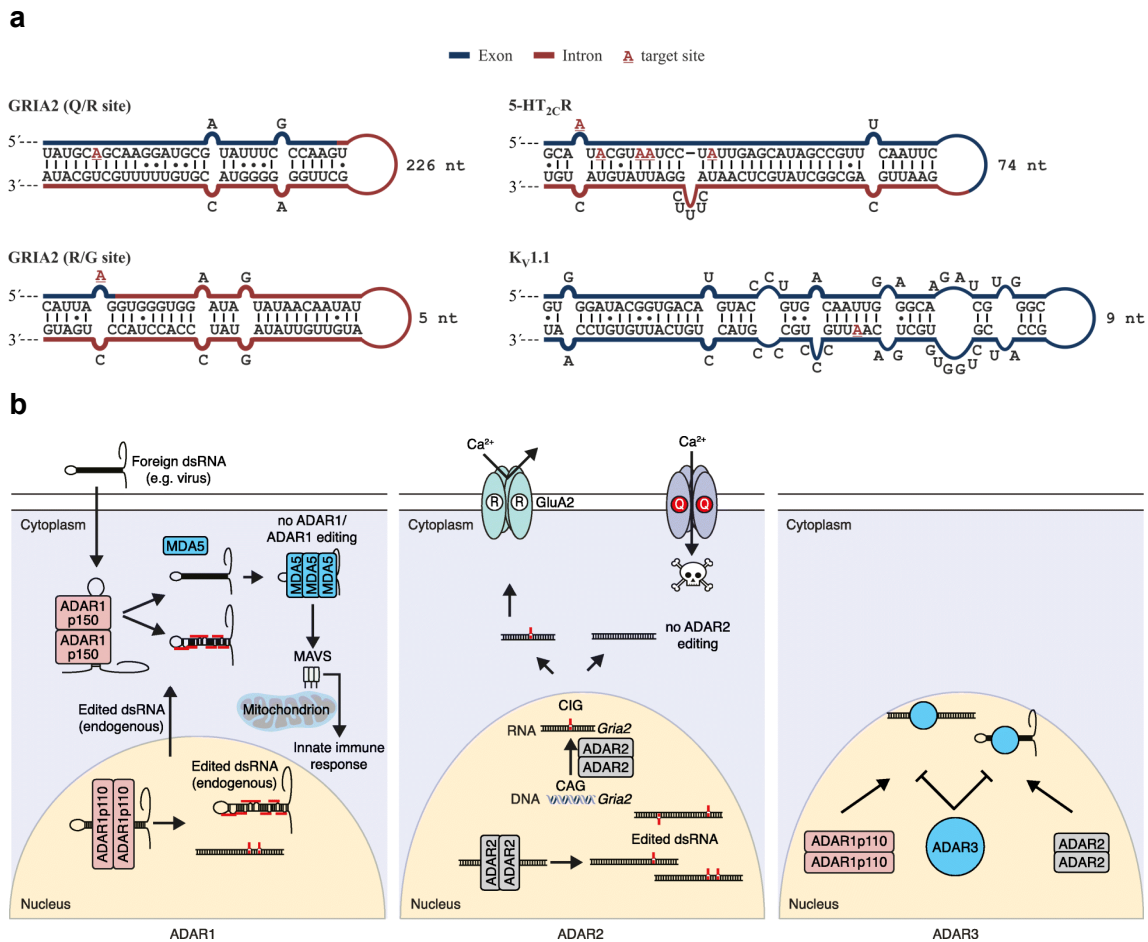


Figure 2 Substrates and functions of ADAR enzymes. **(a)** Predicted secondary structures of known ADAR substrates. Editing in the coding region of these transcripts (red underlined) has been shown to modulate the function of the respective proteins. The structures are typically formed by an exon and its adjacent intron within the pre-mRNA (*GRIA2*, 5-HT_{2c}R). **(b)** Physiological significance of ADARs. Primate-specific *Alu* sequences and other repetitive elements can form long dsRNA structures. Editing within such structures by ADAR1 (p110 and p150) is required to prevent endogenous RNA from triggering the MDA5/MAVS-mediated antiviral immune response. Additionally, IFN-inducible ADARp150 has been presumed to be primarily responsible to edit viral dsRNA in the cytoplasm³⁶ and has been found to uniquely regulate the MDA5/MAVS pathway in mice. ADAR2-mediated editing at the Q/R site of *GRIA2* pre-mRNA encoding an AMPA receptor subunit (also called GluA2) is essential for the neuronal integrity by preventing Ca²⁺ influx. Catalytically inactive ADAR3 inhibits editing by ADAR1 and ADAR2 via competitive binding to dsRNA substrates. Adapted from ref. 3.

sequences that are inversely orientated in the same transcript (inverted repeat *Alus*) can form a long double-stranded structure undergoing multiple editing.⁵ Recent findings have led to speculation that widespread RNA editing of *Alu* sequences in brain transcripts might contribute to increased cognitive capacity in humans compared to other primates.³⁷⁻³⁹ Additionally, there is evidence that RNA editing within *Alu* sequences prevents interferon response by regulating the innate immune system. Recently, ADAR1 has been

shown to be the primary editor of *Alu* sequences and other repetitive elements.^{15,40} Several *ADAR1* mutations have been identified in human patients with Aicardi-Goutières syndrome which is an autoimmune disease characterized by interferon overproduction.⁴¹ These gene mutations are speculated to impair ADAR1 function and to result in the accumulation of immune-stimulatory dsRNA formed by repetitive sequences.⁴¹ MDA5 is a pattern recognition receptor detecting viral dsRNA in the cytoplasm and mediating IFN response via its adapter protein MAVS.⁴² *ADAR1*-null mice die during the early stages of embryonic development with aberrant IFN production, liver disintegration, hematopoietic failure and widespread apoptosis.⁴³⁻⁴⁶ However, the embryonic lethality can be rescued by the loss of either *MDA5* or *MAVS*.^{45,46} Recent data indicate that ADAR1-catalyzed editing of endogenous dsRNA formed by inverted repeat *Alus* inhibits recognition by MDA5, leading to the prevention of IFN production.⁴⁷ Therefore, editing of *Alu* sequences by ADAR1 might serve as a mechanism for the innate immune system to discriminate self from non-self dsRNA (**Fig. 2b**). Additionally, ADAR1 appears to be involved in another mechanism ensuring the fine-tuning of the IFN-mediated immune response. It has been shown that ADAR1 inhibits translational shutdown in cells by preventing endogenous RNA from activating protein kinase R (PKR) during IFN treatment.⁴⁰ This might be accomplished by the action of ADAR1 on *Alu* dsRNAs.⁴⁰

Recent data demonstrated independent functions of the ADAR1 isoforms in mice.⁴⁶ Whereas ADAR1p150 regulates the MDA5/MAVS pathway and is necessary for intestinal homeostasis as well as B-cell development, ADAR1p110 is required for kidney patterning and contributes to the editing of *5-HT_{2c}R* transcripts.⁴⁶ It is still unknown whether the functions of the ADAR1 isoforms in multi-organ development and maintenance are dependent or independent of A-to-I editing. Editing independent functions of ADAR1 have been suggested since mice expressing catalytically inactive *ADAR1^{E861A}*, leading to the same embryonic-lethal phenotype as *ADAR1*-null mice, exhibit normal development and reach adulthood without any loss of viability when MDA5 is additionally deleted, while *ADAR1*-null mice with deleted *MDA5* or deleted *MAVS* survive only a few days after birth.^{45,46,48}

Sakurai *et al.* showed that ADAR1p110 is phosphorylated under cellular stress and is translocated to the cytoplasm to inhibit apoptosis independently of its catalytic activity.⁴⁹ ADAR1p110 binds anti-apoptotic gene transcripts containing dsRNA structures in the 3'-UTR and protects them from Staufen1-mediated decay.⁴⁹ Both ADAR1 isoforms have been reported to form a complex with Dicer, leading to enhanced

pre-siRNA and pre-miRNA processing without requiring the dsRNA-binding and deaminase activities of ADAR1.⁵⁰ Additionally, ADAR1 can interact with Ago2 via Dicer and promote RNA silencing.⁵⁰

ADARs can also influence RNA silencing by catalyzing A-to-I editing in miRNAs. A considerable number of pri-miRNAs undergoes A-to-I editing in the human brain and depending on the pri-miRNA, RNA editing can either suppress or enhance Drosha processing.⁵¹ In some cases, pri-miRNA editing results in inhibition of pre-miRNA cleavage by Dicer.^{51,52} The replacement of adenosine by inosine in the seed region of miRNAs can alter their affinities for target mRNAs. For instance, miR-376a-5p has been shown to target a different set of genes when edited at the +4 position.⁵³ In mice, edited miR-376a-5p has been found to be restricted to certain tissues and to contribute to decreased uric acid levels by silencing *PRPS1*, showing that tissue-specific regulation of the uric acid synthesis pathway can be managed by RNA editing-induced switching of miRNA target-specificity.⁵³ Several studies have reported that A-to-I editing in the 3'-UTR of transcripts can regulate mRNA levels by creating or disrupting miRNA-binding sites.⁵⁴⁻⁵⁷

The physiological role of the catalytically inactive ADAR3 is less understood. A recent NGS-based study has indicated that editing levels in the brain are negatively regulated by ADAR3 expression.¹⁵ This might be accomplished by the competitive binding of ADAR3 to RNA editing substrates (**Fig. 2b**).^{23,58}

1.2.3 Regulation of A-to-I RNA editing

RNA editing is dynamically regulated, but little is known about the underlying mechanisms determining which adenosine is edited at which rate. At the *cis*-regulatory level, the secondary structure of the dsRNA substrate has high influence on the editing outcome by modulating ADAR binding. *In vitro* studies have shown that multiple adenosines are converted within long stretches of dsRNA, such as *Alu* dsRNA.^{59,60} However, in shorter dsRNA regions interrupted by loops, bulges and mismatches, only one or few adenosines are edited. It is suggested that these structural features restrict the binding of ADARs to certain regions of the substrate, leading to selective A-to-I editing.^{61,62} Additionally, RNA structural elements distantly located from the editing site have been shown to be required for efficient deamination.⁶³⁻⁶⁶ The neighboring bases around the target adenosine have a major impact on the editing rate. *In vitro* data revealed that both ADAR1 and ADAR2 prefer uracil as 5'- and guanosine as 3'-nearest

neighboring base of the target adenosine.⁶⁷ Furthermore, 5'-GAN (N = G, C, A, U) triplets are highly disfavored for editing.⁶⁷ Several NGS studies have also confirmed that editing sites with G at the -1 position (upstream) are depleted, whereas editing sites with G at the +1 position (downstream) are enriched.⁵⁻¹⁴ It has been reported that the neighbor preferences are mainly determined by the catalytic domain.⁶⁷ In addition, an A:C mismatch at the editing site has been found to increase the editing level compared to a A-U base pair at the same site.⁶⁸ Consistent with this observation, an A:C mismatch at the Q/R site within *GRIA2* pre-mRNA lead to efficient conversion of the target adenosine by ADAR1 which is unable to edit the wild-type Q/R site (A-U).⁶⁹ Editing by both ADARs is remarkably inhibited when the target adenosine is mismatched with a purine base (A:A, A:G).⁶⁸

ADARs are important *trans*-regulators of A-to-I editing. As mentioned earlier, ADAR3 appears to inhibit ADAR1- and ADAR2-mediated editing in brain.¹⁵ Additionally, it has been shown that ADAR2 edits an intronic site of its own pre-mRNA to induce alternative splicing.⁷⁰ Consequently, the mature transcript contains 47 additional nucleotides causing a reading frameshift which ends in a premature translation stop and in a non-functional protein.⁷⁰ Recent data indicate that ADAR2 auto-editing represents a negative regulatory mechanism to adjust ADAR2 protein levels and substrate editing *in vivo*.⁷¹ In general, the total number of editing sites correlates with *ADAR* expression across human tissues as both increase from muscle to brain.^{13,15} In some cases, the levels of A-to-I editing in transcripts varies strongly across different tissues and does not correlate with *ADAR* expression.^{72,73} RNA editing at several sites increases during brain development, which can be explained by the enhancement of *ADAR* expression, but there are many sites which nevertheless exhibit stable editing yields.¹⁴ In addition, some of the identified sites are edited highly efficiently despite a low degree of double-strandedness.¹⁴ This implies that A-to-I editing is not only *trans*-regulated by ADARs alone. High-throughput screening has identified a small number of enhancers and repressors of ADAR2-catalyzed RNA editing, but the underlying mechanisms for regulating enzyme activity remains speculative.^{74,75} ADAR1 has been reported to be catalytically inactive when it is in complex with DICER.⁵⁰ The snoRNA HBII-52 promotes 2'-*O*-methylation at the C-site within *5-HT_{2c}R* pre-mRNA and inhibits editing of the target adenosine.⁷⁶ Additionally, the transcription factor CREB acts as negative regulator of *ADAR1* expression in metastatic melanoma cells.⁷⁷ Downregulation of ADAR1 inhibits editing of pri-miR-455, leading to the accumulation of mature wild-type

miR-455-5p which promotes melanoma growth and metastasis *in vivo*.⁷⁷ In contrast, activation of CREB results in increased ADAR2 protein levels and *GRIA2* Q/R site editing in rat brain.⁷⁸ The nuclear localization and stability of ADAR2 is ensured by the prolyl isomerase Pin1, leading to enhanced RNA editing.⁷⁹ In absence of PIN1, ADAR2 is released into the cytoplasm and is targeted by the E3 ubiquitin-protein ligase WWP2 promoting degradation.⁷⁹ Very recently, *ADAR2* has been found to be targeted by the transcription factor CLOCK inducing circadian gene expression.⁸⁰ Accordingly, several transcripts are rhythmically edited by ADAR2 *in vivo*.⁸⁰

1.2.4 Structural and mechanistic insights into ADAR catalysis

The crystal structure of the deaminase domain of ADAR2 (ADAR2-D) has revealed several important components including an inositol hexakisphosphate (IP₆) molecule which is located in the enzyme core and is necessary for proper protein folding.⁸¹ Within the catalytic site, a zinc ion and a glutamate residue (E396) activate water for the hydrolytic deamination of adenosine.⁸¹ The glutamate probably serves as a proton shuttle and its mutation to alanine has been shown to abolish editing activity of ADAR1 (E912A) and ADAR2 (E396A).^{82,83} NMR structural studies on both dsRBDs of ADAR2 in complex with the RNA hairpin, where the R/G site of *GRIA2* pre-mRNA is embedded, have revealed that the dsRBDs recognize both the sequence and the shape of the RNA substrate (**Fig. 3a**).⁸⁴ Whereas dsRBD1 interacts with the upper part of the hairpin, dsRBD2 binds near the R/G site and is supposed to bring the deaminase domain in close proximity to the target adenosine.⁸⁴ A yeast-based screen revealed that the substitution of a glutamate by a glutamine at position 488 in the deaminase domain enhances the catalytic rate of ADAR2.⁸⁵ It has been indicated that the glutamine facilitates base-flipping which is suggested to be used by ADARs to perform deamination.⁸⁵ In addition, the mutation has also been reported to increase the binding of ADAR2-D to a model substrate.⁸⁶ A screen for mutations at the corresponding position in ADAR1-D (E1008X) showed that, besides glutamine, several other amino acids can also enhance the catalytic activity.⁸⁷ Recently, Matthews *et al.* described crystal structures of ADAR2-D and ADAR2-D E488Q bound to RNA substrates and showed that ADARs indeed use base-flipping to catalyze A-to-I editing (**Fig. 3b,c**).⁸⁸ In this report, the authors used the adenosine analog 8-azanebularine (8-azaN) as target base since its hydration generates a mimic of the proposed high-energy intermediate during the deamination reaction, leading to tight

the NMR-based ADAR2 dsRBD2/dsRNA structure shows that ADAR2-D and dsRBD2 contact the same region, leading to a clash between both domains. Therefore, researchers have proposed new structural models showing the concurrent binding of dsRBD2 and ADAR2-D to dsRNA without steric hindrance.⁹¹ However, it cannot be excluded that both domains bind sequentially, meaning that dsRBD2 first contacts the region containing the editing site which is then released for the interaction with the deaminase domain.

1.3 Targeting RNA as therapeutic strategy

Classical approaches to treat diseases include targeting of enzymes and receptors with small molecule agonists and antagonists to modulate their activity. However, RNA is increasingly moving into the spotlight as a therapeutic target and/or agent. Therapeutic intervention on the RNA level is particularly useful when the targeting of the disease-related protein fails or when disorders are induced by RNA itself, such as dysregulated miRNAs or CAG-repeat RNAs. Furthermore, targeting RNA offers the possibility not only to up- or downregulate gene expression but also to generate new transcripts and protein products, for instance by creating or deleting splice sites. Antisense oligonucleotides (ASOs) interact with RNA via Watson-Crick base-pairing and regulate RNA function through various mechanisms.⁹²⁻⁹⁴

Medicinal chemists designing ASOs face many challenges concerning the pharmacokinetic and pharmacodynamic properties of ASOs. One typical modification is the replacement of a non-bridging phosphate oxygen atom in the nucleic acid backbone by a sulfur atom (**Fig. 4a**). The resulting phosphorothioate (PS) linkage enhances nuclease stability and prevents rapid renal filtration of ASOs by increasing their binding to plasma proteins. Another possibility to enhance nuclease resistance of oligonucleotides is the substitution of the sugar-phosphate backbone by phosphorodiamidate linkages, leading to so-called morpholinos (**Fig. 4a**). Furthermore, ASOs often contain sugar modifications at the 2'-position (**Fig. 4a**). Using modified nucleosides with 2'-fluoro (F), 2'-*O*-methyl (OMe) or 2'-*O*-methoxyethyl (MOE) groups increases the stability, potency and immunoresistance of nucleic acid-based drugs. The stability and potency are further improved by applying 2'-*O*,4'-*C*-methylene-bridged nucleic acids termed LNAs (locked nucleic acids; **Fig. 4a**). LNA and its analogs (cEt, ENA) are widely used for the new generation of ASO drug candidates.⁹⁵⁻⁹⁷ Recently, oligonucleotides have been conjugated with triantennary N-acetylgalactosamine (GalNAc₃) mediating liver-specific uptake.^{98,99}

Several concepts using ASOs have been successfully translated to the clinic. The majority of ASO drugs tested in clinical studies promotes RNase H-mediated target RNA cleavage (**Fig. 4b**). Interaction between short DNA oligomers and target mRNAs results in the formation of heteroduplexes, which are recognized by RNase H1. When the enzyme binds to such a heteroduplex, it catalyzes the cleavage of the RNA substrate.¹⁰⁰ The DNA oligomer remains intact and is released after RNA degradation. RNase H-dependent ASOs of the 1st generation contain a PS backbone, such as fomivirsen (Vitravene) which was the first FDA-approved ASO and was given to AIDS patients suffering from CMV retinitis.¹⁰¹⁻¹⁰³ 2nd generation ASOs are designed as so-called gapmers that are typically 20 nt long PS oligonucleotides with a central DNA gap allowing RNase H-mediated cleavage and with five flanking 2'-MOE-modified residues at both ends. A representative of this class is the FDA-approved mipomersen (Kynamro) targeting *apolipoprotein B-100* mRNA for the treatment of familial hypercholesterolemia.¹⁰⁴⁻¹⁰⁶ Recently developed gapmers are also terminally modified with an analogue of LNA (cEt) and/or conjugated with GalNAc₃.⁹⁸ Remarkably, the conjugation of GalNAc₃ has been shown in a Phase I/IIa study to increase the target affinity (potency) of an MOE gapmer more than 30-fold for lowering plasma levels of apolipoprotein(a).¹⁰⁷

RNA interference (RNAi) as a therapeutic strategy can be achieved by small interfering RNAs (siRNAs; **Fig. 4b**). Such ASOs are 21-23 nt RNA duplexes with 2 nt overhangs at their 3'-ends.^{108,109} The duplexes contain the target RNA sequence on the one strand (passenger strand) and the complementary sequence on the other (guide strand). Once delivered into the cell, siRNA is recognized by the RNA-induced silencing complex (RISC) which discards the passenger strand and retains the guide strand. To be preferably selected, the guide strand needs to fulfill several requirements, including that its 5'-terminus is located at the less stable end of the RNA duplex.¹¹⁰ The guide strand steers RISC to the target RNA which is subsequently cleaved by the RISC component Ago2. Therapeutic RNAi in the liver has benefited from the encapsulation of siRNA into lipid nanoparticles (LNPs) exhibiting hepatic accumulation. One of the most promising siRNA drugs is patisiran which contains 2'-OMe-modified nucleosides and 3'-terminal dT overhangs. The siRNA is delivered by LNPs to target *transthyretin* mRNA in patients suffering from hereditary ATTR amyloidosis.^{111,112} After successfully completing a Phase III trial, the drug is expected to be approved.¹¹³ Besides RNase H-dependent ASOs,

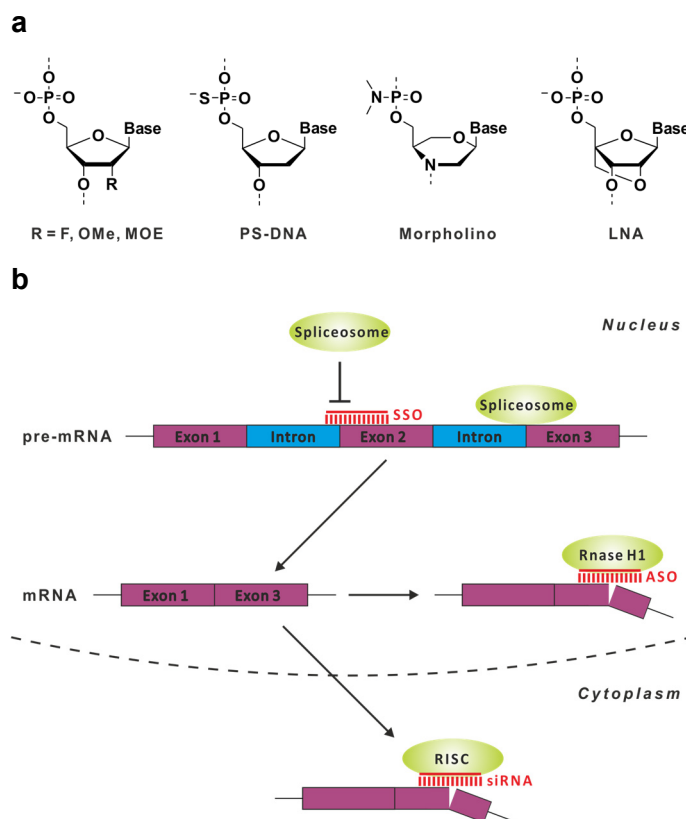


Figure 4 Chemical modification enables efficient RNA targeting by antisense oligonucleotides (ASOs). (a) Selection of nucleic acid modifications commonly used in ASOs. Adapted and modified from Manuscript (Man.) 6 (b) Classical antisense-mediated mechanisms of action. ASOs applied for the treatment of diseases modulate RNA splicing or promote RNA degradation via RNase H1 or RISC. RNase H-mediated RNA cleavage can be directed by ASOs in the nucleus and in the cytoplasm.¹¹⁴ Adapted and modified from Man. 6.

GalNac₃ has also been conjugated to siRNAs to enable specific liver uptake.⁹⁹ Various GalNac₃-siRNA conjugates are currently being evaluated in clinical trials.¹¹⁵

Another application field of therapeutic ASOs is the regulation of pre-mRNA splicing (**Fig. 4b**).¹¹⁶ Such ASOs are termed as splice-switching oligonucleotides (SSOs). Most recently, the FDA has approved two SSOs promoting exon skipping and exon retention in pre-mRNA to treat Duchenne muscular dystrophy (DMD) and spinal muscular atrophy (SMA), respectively. DMD is a rare muscle wasting disease mostly resulting from exonic out-of-frame deletions in the *DMD* gene encoding dystrophin. The morpholino-based oligomer eteplirsen (Exondys 51) binds to an internal splicing enhancer in *DMD* pre-mRNA and promote skipping of exon 51 to restore the ORF.¹¹⁷ Consequently, dystrophin is truncated, but can partially fulfill its role in the maintenance of muscle integrity, leading to the milder Becker muscular dystrophy phenotype. 14% of all DMD patients could benefit from the approved SSO-based drug.¹¹⁸ SMA is

characterized by motor neuron loss with subsequent muscle wasting. The neurodegenerative disease results from mutations in the *survival of motor neuron 1* (*SMN1*) gene. The closely related *SMN2* gene can potentially compensate the loss of *SMN1* function, but it contains a C-to-T point mutation in a splicing enhancer within exon 7. Due to this mutation, exon 7 is not included into most of the mature *SMN2* transcripts during splicing, leading to an insufficient amount of functional protein.¹¹⁹ Nusinersen (Spinraza) is a 2'-MOE PS-modified SSO recently approved for the treatment of SMA. The ASO drug targets an intronic splicing silencer within intron 7 of *SMN2* pre-mRNA and has been shown to increase functional SMN protein levels by promoting exon 7 inclusion.¹²⁰

1.4 Novel concepts for manipulating genetic information

1.4.1 Genome engineering

In recent years, tools to introduce specific changes in the genome of organisms have become very popular in basic research. Genome engineering approaches often rely on the induction of DNA double-strand breaks (DSBs) which are repaired by different mechanism, including nonhomologous end-joining (NHEJ) and homology-directed repair (HDR; **Fig. 5a**).¹²¹ NHEJ leads to random insertions or deletions (indels) at the site of the DSB. Since indels can create translational frameshifts, NHEJ is useful to disrupt gene functions. In contrast, HDR enables the insertion of single nucleotide substitutions or transgenes by an exogenously given DNA template that contains homologous sequences to the flanking regions of the DSB. However, NHEJ occurs concurrently and reduces the rate of HDR. To introduce site-specific DSBs, several programmable nucleases have been applied.

Meganucleases are obtained by protein engineering of naturally occurring DNA endonucleases (**Fig. 5b**).¹²² Since the DNA recognition activity and the cleavage activity coexist in a single domain, changing the recognition specificity of meganucleases for targeting different genomic loci without affecting its cleavage activity is a major challenge which limits the use of these enzymes for genome engineering.¹²³

Zinc-finger nucleases (ZFNs) are fusion proteins between an array of zinc-finger DNA-binding domains and the non-specific cleavage domain of the FokI endonuclease (**Fig. 5b**).¹²⁴ The zinc-finger DNA binding domains, each usually contacting three base

pairs of DNA, can be designed to target any chosen sequence, but the creation of a zinc-finger array with high binding affinity remains laboriously.¹²³

Transcription activator-like effector nucleases (TALENs) are constructed similar to ZFNs (**Fig. 5b**). DNA recognition is ensured by TALE repeats, each binding a single base pair.^{125,126} According to a user-defined sequence, an array of TALE repeats is constructed and fused to the FokI cleavage domain inducing a DSB at the chosen site.¹²⁷ In contrast to ZFNs, TALENs are easy to design, but their highly repetitive nature and their large size limit their potential to be delivered by viral gene delivery vehicles for *in vivo* applications.¹²³

A major drawback of all above-mentioned approaches is the need to design a new site-specific nuclease for each new target site. Recently, genome engineering has been revolutionized by using Cas9 (CRISPR-associated protein 9; **Fig. 5b**). The endonuclease originated from bacteria and is part of type II CRISPR-Cas immunity systems providing protection against invading MGEs (mobile genetic elements), such as viruses and plasmids.^{128,129} Foreign DNA is degraded and short fragments are integrated as new spacers into the CRISPR array. The array is transcribed and processed into CRISPR RNAs (crRNAs), each interacting with a second short RNA known as *trans*-activating CRISPR RNA (tracrRNA). The crRNA-tracrRNA duplex and Cas9 form a CRISPR ribonucleoprotein (crRNP) complex for targeting a region called protospacer in which one strand contains the complementary sequence to the crRNA. It has been revealed that DNA cleavage by Cas9 requires the presence of a specific protospacer adjacent motif (PAM) downstream the non-complementary strand of the protospacer.¹³⁰ Once the PAM is recognized, the endonuclease unwinds the DNA duplex for the hybridization between the crRNA and the complementary DNA strand.¹³¹ The subsequent cleavage within the protospacer by Cas9 produces a blunt DSB. The 1,368-amino acids Cas9 nuclease adapted from *Streptococcus pyogenes* is most widely used for genome engineering experiments and recognizes the simple PAM sequence of 5'-NGG.¹³⁰ In a typical experiment, genome engineering in cells is realized by the presence of the Cas9 nuclease and of a guide RNA (gRNA) which is a fused version of crRNA and tracrRNA.¹³² At the 5'-end, the gRNA consists of a variable region of 20 nucleotides that are designed to target to a chosen DNA sequence undergoing site-specific cleavage by Cas9. The number of genomic sites allowing the induction of a DSB by Cas9 is limited as the targeted sequences needs to be followed by a PAM sequence. To redirect Cas9 nuclease activity to a new appropriate target site, only the 20 nucleotides in the variable region of the gRNA

must be changed according to the DNA sequence, but the nuclease does not require any new protein design which would be necessary when using meganucleases, ZFNs or TALENs. Since the first report on genome engineering by CRISPR-Cas9, there has been an explosion of published work demonstrating successful gene targeting in various cell types and organisms. In contrast to the other genome engineering platforms, Cas9 does not require dimerization for its nuclease activity, leading to initial concerns that CRISPR-Cas9 exhibits higher off-target activity. However, the use of shorter gRNAs, paired Cas9 nickases, dimeric Cas9-FokI fusion proteins, high-fidelity Cas9 variants or small molecule-activated Cas9 nucleases has significantly improved the target specificity of this approach.¹³³

Due to its simplicity, the CRISPR-Cas9 technology is routinely used for efficient gene disruption by NHEJ-induced indels. However, the HDR strategy to introduce specific changes in the genome is still too inefficient to be widely used for gene repair. As mentioned before, the strategy relies on a DSB which can also concurrently lead to the formation of unintended indels by NHEJ. Recently, HDR efficiencies have been improved by several strategies, including rational template design and inhibition of NHEJ.¹³³ Another strategy for efficiently introducing point mutations applies CRISPR base-editing systems that do not rely on inducing DSBs and subsequent HDR-mediated genome engineering. Currently used editing systems are fusion proteins composed of a single-strand-breaking Cas9 nickase, a cytosine deaminase and an uracil glycosylase inhibitor (**Fig. 5c**).¹³⁴ An added gRNA steers the fusion protein to the protospacer and forms a duplex with the complementary strand, while the non-complementary strand containing the PAM undergoes C-to-U editing catalyzed by the cytosine deaminase. The glycosylase inhibitor prevents base excision repair of the resulting uracil, whereas the Cas9 nickase-induced break within the non-deaminated strand promotes mismatch repair, leading to the generation of an A-T base pair from the former G-C base pair. Recently, a highly engineered CRISPR base editing system has been presented which enables the targeted substitution of A-T base pairs by G-C base pairs in DNA.¹³⁵ This was accomplished by the fusion of a Cas9 nickase and a mutant derived from *E. Coli* t-RNA-specific adenosine deaminase (TadA). When the fusion protein is directed to the target DNA via an appropriate gRNA, the TadA mutant component catalyzes A-to-I editing within the displayed strand. After DNA repair or replication, an G-C base pair is present at the target site. Compared to HDR-mediated genome engineering, base editing systems

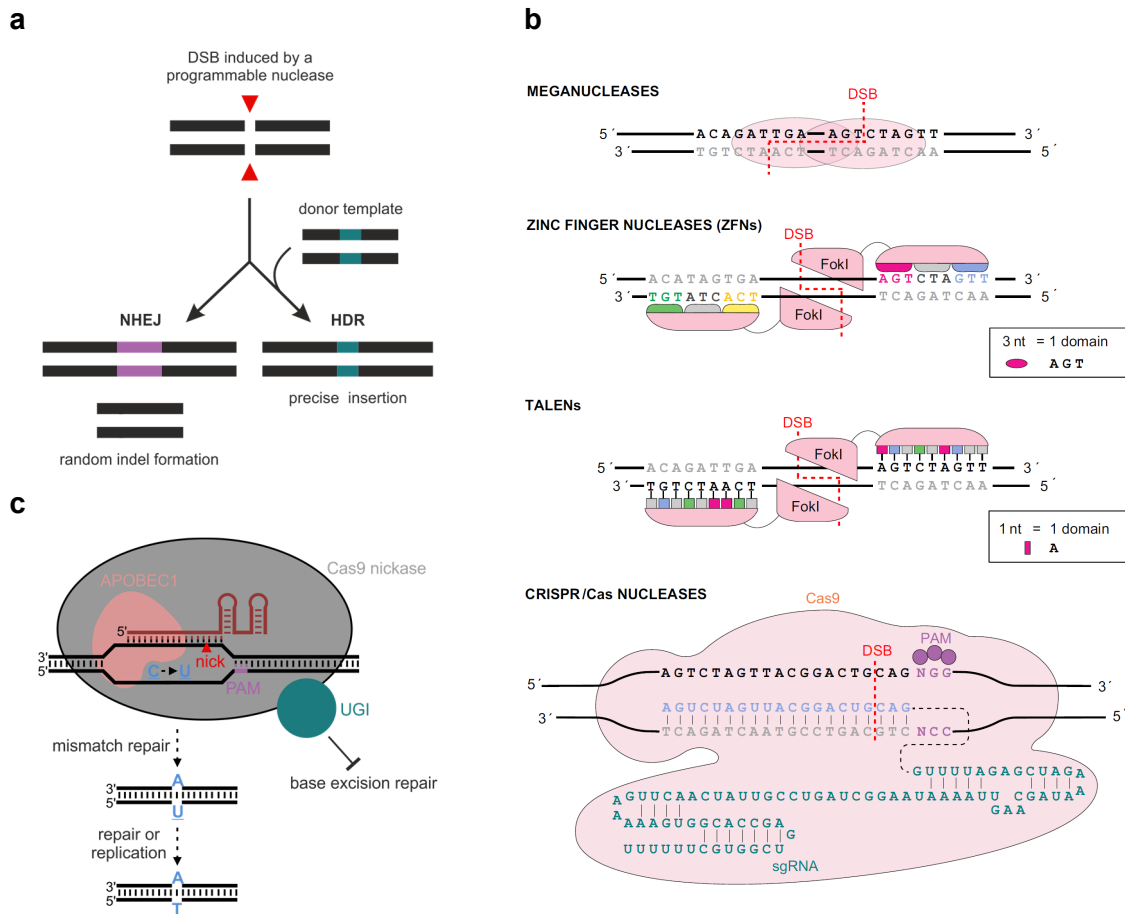


Figure 5 Genome engineering. (a) Altering genetic information often relies on the site-specific induction of a DNA double strand break (DSB). DSBs can be repaired by nonhomologous end-joining (NHEJ) resulting in random insertions or deletions (indels). Homology directed repair (HDR) is induced in the presence of a single- or double-stranded DNA donor template, leading to the specific insertion of single nucleotide substitutions or transgenes. (b) Site-specific DSBs have been introduced by four types of programmable nucleases. Meganucleases, ZFNs and TALENs need to be re-engineered to bind a new target site. This is simplified by the usage of Cas9 which only requires a PAM (Protospacer-adjacent motif) near the target site and the change of the antisense sequence within the single guide RNA (sgRNA). In contrast to the other nucleases, Cas9 induces blunt DSBs (red dashed line). Adapted from ref. 136. (c) CRISPR-base editing systems enable site-specific G-C to A-T base pair substitutions in genomic DNA without the induction of a DSB. This is accomplished by a fusion of a Cas9-nickase, a cytosine deaminase (here: APOBEC1) and an uracil glycosylase inhibitor (UGI).

offers higher efficiencies of single nucleotide changes while minimizing indel formation at the target site.^{134,135} However, these systems have been reported to cause off-target editing in the genome and are dependent on a PAM sequence which need to be located in a defined distance from the target base.^{135,137} Additionally, the systems are unable to distinguish between multiple copies of the target DNA base within the editing window (~ 5 nucleotides).^{134,135}

1.4.2 Site-directed A-to-I RNA editing

In 1995, Tod Woolf and co-workers introduced the concept of site-directed A-to-I RNA editing.¹³⁸ They could show that ADAR activity can be directed to catalyze RNA editing within a reporter mRNA inside *Xenopus* embryos when the mRNA was hybridized with a 52 nt oligonucleotide prior micro-injection.¹³⁸ Since then, several approaches have been developed enabling site-directed A-to-I editing within RNA to manipulate protein and RNA function without interfering at the DNA level.

In this PhD project, I applied so-called SNAP-ADAR (SA) enzymes. *O*⁶-alkylguanine-DNA alkyltransferase (AGT) is responsible for the repair of *O*⁶-alkylguanine in DNA by irreversibly transferring the alkyl group to a definite cysteine residue in the protein.¹³⁹ It has been reported that proteins fused to a mutated version of AGT termed SNAP-tag can be labelled by *O*⁶-benzylguanine (BG) derivatives *in vitro* and in cell culture.¹⁴⁰ Subsequently, the SNAP-tag has been further engineered to obtain increased activity towards BG derivatives, low affinity to DNA and improved expression in cells.¹⁴¹ Stafforst & Schneider reported the usage of the SNAP-tag for the design of the first published example of an engineered, RNA-guided deaminase performing site-directed RNA editing.¹⁴² The promiscuity of ADARs mainly derives from their dsRBDs binding to various dsRNA structures. Therefore, the dsRBDs of ADAR1 were replaced by the SNAP-tag, leading to the generation of SNAP-ADAR1 (SA1; **Fig. 6a,b**). The SNAP-ADAR (SA) approach also includes customized gRNAs that can be commercially acquired and contain a 5'-amino-C6 linker at the 5'-end for the coupling with BG. BG-modified gRNA (BG-gRNA) is covalently bound to the SNAP-tag and forms conjugates with the SA enzyme in a one-to-one ratio. The bound gRNA directs the enzyme to the complementary sequence of a chosen RNA and creates a duplex structure with a central A:C mismatch at the target site. Such a duplex structure is considered necessary for the efficient and selective editing of the target adenosine by the deaminase domain. *In vitro* data has shown that SA1 driven by 20 nt gRNAs is able to selectively correct nonsense and missense mutations in reporter transcripts, leading to the restoration of protein function.¹⁴² In the next report, our group showed *in vitro* that optimizing the gRNA results in the editing of all four 5'-NAG triplets by SA1 and SA2 (SNAP-ADAR2) with yields $\geq 80\%$ obtained for 5'-UAG, 5'-AAG and 5'-CAG.¹⁴³ Editing of the less preferred 5'-GAG triplet yielded in $\sim 50\%$ and required a gRNA with a mismatching A or G opposite the 5'-neighboring base to the target adenosine.¹⁴³ Besides an A:C mismatch, also an A-U

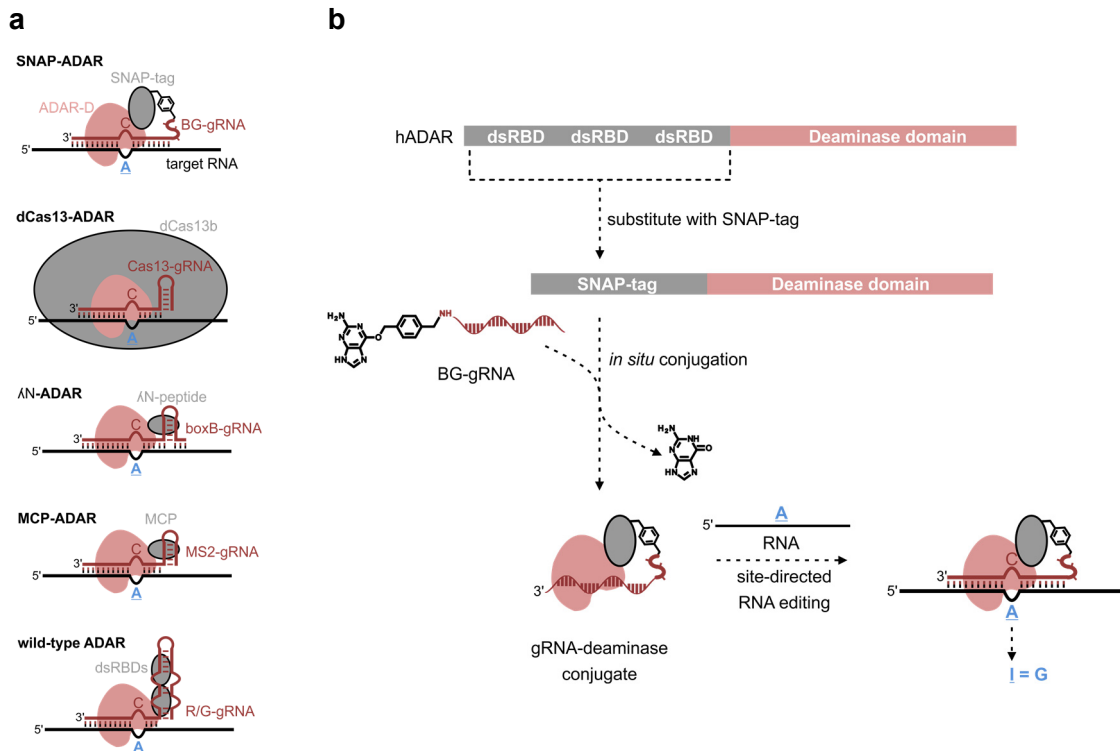


Figure 6 Approaches to perform site-directed A-to-I RNA editing. (a) ADAR-derived artificial enzymes or wild-type ADARs can be recruited by gRNAs to enable site-specific A-to-I substitutions in target RNAs. ADAR-D, ADAR deaminase domain; BG, benzylguanine; MCP, MS2 coat protein; dsRBD, dsRNA-binding domain. (b) The SNAP-ADAR (SA) approach. The deaminase domain of hADAR is fused to the so-called SNAP-tag which covalently binds benzylguanine (BG)-modified gRNA. The gRNA directs the SA enzyme to the target RNA and forms the duplex structure required for efficient A-to-I editing. Adapted and modified from Man. 5.

base pair at the target site can lead to quantitative editing yields *in vitro*.¹⁴³ In contrast, an A:G mismatch abolishes editing and can be used to prevent unintended off-target editing within the gRNA/mRNA duplex.¹⁴³

Cas13b enzymes have been previously identified as RNA-guided RNases belonging to type IV CRISPR-Cas systems.¹⁴⁴ Comprehensive screening of Cas13 family members revealed that the Cas13b ortholog from *Prevotella sp. P5-125* enables efficient and specific RNA knockdown in mammalian cells without evident sequence constraints.¹⁴⁵ The Cas13b enzyme was engineered to lack catalytic activity and fused to the ADAR2 deaminase domain carrying the hyperactive E488Q mutation to create an artificial RNA editing enzyme (dCas13b-ADAR2Q; **Fig. 6a**).¹⁴⁵ Expressing both dCas13b-ADAR2Q and ~85 nt gRNA, containing a 3'-terminal stem-loop and a targeting sequence at the 5'-end, enables site-directed RNA editing. Additionally, a highly-specific variant (dCas13b-ADAR2Q T375G) has been described which was

obtained from the substitution of threonine for glycine at position 375 within dCas13b-ADAR2Q.¹⁴⁵

The group of J. Rosenthal have replaced the dsRBDs of ADAR2 by a 22 amino acids long peptide (λ N) which is derived from the N protein of bacteriophage λ and binds with high affinity to *boxB* RNA which forms a hairpin structure (**Fig. 6a**).¹⁴⁶ Once the *boxB* hairpin is fused to an RNA oligonucleotide (*boxB*-gRNA), which is complementary to the sequence of interest, it can guide λ N-ADAR2 to the target site.

A similar system has been presented by Azad *et al.* who performed site-directed RNA editing by an artificial enzyme (MCP-ADAR1) derived from the fusion of the deaminase domain of ADAR1 and the bacteriophage MS2 coat protein (MCP) which exhibits high affinity binding to a specific stem-loop structure in MS2 RNA (**Fig. 6a**).¹⁴⁷ gRNAs containing six MS2 RNA stem-loops (MS2-gRNA) have been shown to direct MCP-ADAR1 for specific A-to-I RNA editing.¹⁴⁷

Our laboratory and others have developed a strategy enabling specific A-to-I substitutions in target RNAs by ADAR2.^{148,149} ADAR2 is steered by a guide RNA (gRNA) forming, together with the target RNA, a structure which mimics the hairpin where the *GRIA2* R/G editing site is embedded (**Fig. 6a**). At one end, the gRNA consists of the imperfect stem-loop (R/G motif) which is intended to be recognized by the dsRBDs of ADAR2. At the other end, the gRNA sequence is complementary to the user-defined target RNA. The gRNA design can also be used to recruit the ADAR1 isoforms p110 and p150.¹⁵⁰ So far, published studies have applied ectopically expressed ADAR enzymes. However, ongoing studies in our group show that even endogenous ADAR enzymes can be recruited for site-directed RNA editing. This is accomplished by applying chemically modified gRNAs instead of using gRNAs expressed from plasmids.

2 Aims of this study

Tools for the manipulation of genetic information have great potential for various applications in medicine and the life sciences. Genome engineering has been simplified by CRISPR-Cas9 tools which are applied in many laboratories to study gene function *in vitro* and *in vivo*. It appears promising that the CRISPR-Cas9 technology might be therapeutically used in the future to correct disease-causing mutations. However, changing the genome gives rise to urgent ethical questions associated with germline gene modification.¹⁵¹ There are also safety concerns regarding the therapeutic applicability of genome engineering tools which can induce unintended off-target mutations permanently persisting at the DNA level. In this regard, tools for manipulating genetic information at the RNA level are highly desired since ethical issues can be circumvented and potentially produced off-target mutations in RNA can be considered as reversible. The transient nature of RNA changes also allows to temporarily intervene in biological processes, such as inflammation or signal transduction, whose permanent alteration could have harmful effects. Additionally, the extend of specifically introduced RNA changes can be potentially adjusted from 0% to 100% to precisely regulate their biological outcome.

The SA approach differs from all other approaches for site-directed RNA editing due to the *in situ* covalent bond formation between gRNA and editing enzyme. Recent reports from our group successfully demonstrated that the unique assembly strategy enables SA enzymes to be recruited by small BG-gRNAs for efficient RNA editing in a PCR reaction tube.^{142,143} The study described here addressed the question of how powerful the SA approach is for future applications. Therefore, it was aimed at testing the performance of RNA-guided SA enzymes in mammalian cells. The editing system was extensively characterized regarding achievable editing yields, duration of RNA editing, the scope of editable triplets and gRNA potency. Additionally, possible applications in medicine and the life sciences were tested and discussed. This included the correction of disease-causing mutations, the manipulation of signal transduction and the light-driven translocation of proteins. The transcriptome-wide identification of off-target editing sites by NGS-based analysis provided insights into the specificity of SA enzymes. The performance of the BG-gRNA/SA editing system was compared with those of the other editing systems to assess whether the SA approach sets a new benchmark for site-directed RNA editing.

3 Results and discussion

This section contains the summary and the discussion of the results described in the research manuscripts (**Man. 1-4**).

3.1 SNAP-ADAR enzymes enable site-directed RNA editing in cell culture

The SA enzyme is genetically encodable and can be delivered via plasmid transfection for its expression in mammalian cells. In contrast, the gRNA is not encodable since it requires the modification with *O*⁶-benzylguanine (BG) to recruit the SA enzyme.¹⁴² Besides the BG modification at the 5'-end, the gRNAs lacked any other chemical modification in former *in vitro* experiments.^{142,143} However, to successfully perform site-directed RNA editing in cells, we anticipated that the gRNA requires chemical modification to enhance its nuclease resistance. Therefore, the BG-gRNA has been modified according to antagomirs which are ~ 22 nt RNA oligonucleotides antagonising miRNAs. Antagomirs, containing terminal PS linkages and global 2'-OMe-modified nucleosides, have been shown to ensure miRNA silencing in mice over three weeks.¹⁵² Similar to unmodified BG-gRNAs, the first antagomir-like BG-gRNA was designed to contain 17 nucleotides complementary to the chosen sequence, excepting a central mismatching C (counter base) facing the target A, and three non-complementary 5'-terminal nucleotides that serve as a linker between the gRNA and the SA enzyme. According to the antagomir design, the BG-gRNA contained two PS linkages at the 5'-end and four at the 3'-end. Additionally, the BG-gRNA was globally modified with 2'-OMe groups leaving a gap of three unmodified ribonucleosides that comprise the mismatching C and its two nearest neighbors. The impact of the antagomir design on the editing performance was first tested *in vitro* by M. F. Schneider (T. Stafforst group). When targeting a premature 5'-UAG stop codon at position 66 within *eCFP* mRNA (*eCFP W66X*), RNA editing was performed by SA1 to the same extent when using unmodified BG-gRNA or its antagomir-modified analog (**Man. 1, Fig. 1a,f**). The result was unexpected since oligonucleotides with such extensive modification typically impair the activity of nucleic-acid-driven enzymes. For example, RNase H activity has been shown to require gapmers containing a central region of at least five unmodified 2'-deoxy residues.¹⁵³ RISC-mediated RNA cleavage appears to be most efficient when using only modestly modified siRNAs. In particular, the siRNA guide strand is sensitive to base

modifications within its seed region (position 2-8) which is most important for target recognition.^{110,154} In contrast to SA1, SA2 seems not to tolerate extensively modified gRNAs, since the *in vitro* editing of *eCFP W66X* was reduced by ~40% when using antagomir-like BG-gRNA instead of the unmodified BG-gRNA (**Man. 1, Fig. 1a,f**). Considering the published crystal structure of the ADAR2 deaminase domain bound to dsRNA, the reduced editing activity of SA2 with the antagomir-like BG-gRNA may arise from decreased interaction between deaminase domain and gRNA.⁸⁸ The PS linkage between the nucleotides 16 and 17 within the gRNA could prevent the interaction with the arginine residue at position 474 in the ADAR2 deaminase domain (see **Introduction, Fig. 3c**) and could therefore lead to a destabilization of the duplex structure between gRNA and target mRNA during deamination.

Due to the results obtained from the *in vitro* editing experiments, SA1 was chosen to perform site-directed RNA editing in mammalian cells. For their CMV promoter-controlled gene expression, *SA1* and *CFP W66X* were cloned into pcDNA3.1 plasmids. 2×10^5 HEK 293T were seeded in a 24-well format and incubated for 24 hours before they were transiently transfected with 1800 ng of each plasmid + 14.4 μ l Lipofectamine 2000 (ratio = 1 μ g plasmid : 4 μ l Lipofectamine 2000). After 24 hours, 4×10^4 cells were transferred into a 96-well format. One day later, the cells were transfected with 50 pmol gRNA + 2.5 μ l Lipofectamine 2000 and incubated again for 24 hours before fluorescence microscopy. In case of RNA editing, the premature 5'-UAG stop codon is converted back to a tryptophan codon (5'-UIG) leading to restored eCFP fluorescence. Indeed, eCFP signal was detected in cells after transfecting them with the same antagomir-like BG-gRNA used in the *in vitro* experiment before (**Man. 1, Fig. 2e**). To demonstrate the restored fluorescence resulted from the RNA editing of the targeted stop codon at position 66, RNA was isolated from the cells, treated with DNaseI to remove DNA contaminations and reverse transcribed into cDNA which was then amplified by Taq-PCR for Sanger sequencing. Besides the signal for adenosine, Sanger sequencing revealed an additional signal for guanosine at the target site (**Man. 1, Fig. 2h**). This shows that the target adenosine was successfully converted to inosine indicated by the presence of the additional guanosine signal. Around 30% editing yield was calculated by dividing the height of the resulting guanosine peak by the sum of the peak heights of guanosine and adenosine. Additionally, *eCFP W66X/SA1*-coexpressing cells were transfected with either BG-gRNA without any chemical modification or antagomir-like NH₂-gRNA lacking the BG moiety. In both cases, the editing yield was $\leq 5\%$ (**Man. 1, Suppl. Fig.**

7b,e) and only a small fraction of the cells showed restored eCFP fluorescence compared to cells transfected with BG-antagomir-like gRNA (**Man. 1, Fig. 2d,g**). Therefore, efficient site-directed RNA editing with SA enzymes in cell culture requires a gRNA which is chemically modified and conjugated to BG. Chemical modification was necessary to enhance the nuclease resistance of the gRNA, but probably also to improve its potency.⁹²⁻⁹⁴ The BG conjugation was essential because BG enables the gRNA to recruit the SA enzyme as shown by former *in vitro* experiments¹⁴² and by all further studies described here (**Man. 2, Fig. 2-4; Man. 3, Fig. 1-4; Man. 4, Suppl. Fig. 14, Suppl. Note 1**).

After proving that site-directed RNA editing with SA enzymes can be indeed achieved in cell culture, several parameters of the cellular assay have been further optimized in collaboration with A. Hanswillemecke (T. Stafforst group) to improve the editing yield for the next study (**Man. 2**). We chose eGFP as a new reporter for editing since it is brighter than eCFP and can be made visible in cells even in case of a low concentration. According to this, the amount of plasmid with integrated reporter gene was reduced to 500 ng per 24-well format. For the transfection, the ratio of plasmid to Lipofectamine 2000 remained 1 µg : 4 µl. The gRNA transfection procedure was changed from forward to reverse transfection, meaning that 24 hours after plasmid transfection, the cells were detached and given to the prepared transfection mixture containing the gRNA. Similar protocols have been reported to be very efficient for the delivery of siRNA into mammalian cells.^{155,156} The amount of Lipofectamine 2000 used for the gRNA transfection has been reduced from 2.5 µl to 0.5 µl per 96-well. To maintain the viability of the cells, the plasmid and the gRNA transfections required cell confluences $\geq 80\%$ (2×10^5 cells/24-well) and $\geq 50\%$ (6×10^4 cells/96-well), respectively. A premature 5'-UAG stop codon was introduced at position 58 within *eGFP* (*eGFP W58X*) and served as a target site for RNA editing which was examined by Sanger sequencing and by fluorescence microscopy visualizing restored eGFP fluorescence. The chosen BG-gRNA was designed similarly to the BG-gRNA used for *eCFP W66X* editing, but the target-complementary sequence was extended by two additional nucleotides at the 3'-end (**Man. 2, Suppl. Table 9**). The new BG-gRNA design was applied in all subsequent studies and is shown in **Suppl. Fig. 1b** of **Man. 4**. With a constant BG-gRNA amount of 10 pmol/96-well, the yield of *eGFP W58X* editing ranged from $\sim 20\%$ with 25 ng SA1 plasmid to $\sim 45\%$ with 200 ng plasmid (**Man. 2, Suppl. Fig. 19**). With a constant SA1 plasmid amount of 100 ng, the editing yield increased from $\sim 20\%$ with 2 pmol to $\sim 45\%$ with 25

pmol or 50 pmol BG-gRNA (**Man. 2, Suppl. Fig. 20**). The results obtained from the new cellular assay indicate the possibility to adjust site-directed RNA editing by varying the expression of the SA enzyme and the amounts of the gRNA. To test the efficiency of the editing system under *in vivo* conditions, *eGFP W58X* editing was performed inside the annelid *Platynereis dumerilii* by T. Kuzdere (T. Stafforst group). For this, zygotes were micro-injected with the BG-gRNA and two mRNAs encoding SA1 and eGFP W58X. Indeed, restored GFP fluorescence was detected by microscopy one day after microinjection when the zygotes develop into trochophores (**Man. 2, Fig. 4**). Correspondingly, Sanger sequencing of the *eGFP W58X* cDNA revealed an editing yield of ~ 70%. The worms showed no obvious abnormalities during their early development. Taken together, the results show that the SA approach enables efficient RNA editing in *Platynereis dumerilii* and might therefore also be applicable in higher organisms.

Another study was conducted to demonstrate the efficiency of the editing system (**Man. 4**). For this, the SA plasmid was integrated as a single copy in the genome of 293 Flp-In T-REx cells under the control of a doxycycline-inducible CMV promoter. It was already known before that the expression of the SA enzyme in such cells is more homogeneous and much weaker than its transient expression in HEK 293T cells (**Man. 3, Suppl. Fig. 5,12,15**). I generated four cell lines expressing SA1, SA2, SA1Q (SNAP-ADAR1Q) or SA2Q (SNAP-ADAR2Q). The latter two harbor the hyperactive E/Q mutation in their deaminase domain.^{85,87} gRNAs were designed according to the gRNA used for *eGFP W58X* editing. For experiments, 3×10^5 cells were seeded in a 24-well format and incubated with 10 ng/ml doxycycline for 24 hours. To obtain editing at a respective site, 8×10^4 cells were reverse-transfected typically with 5 pmol BG-gRNA + 0.75 μ l Lipofectamine 2000 and incubated again with 10 ng/ml doxycycline in a 96-well format. After one day, RNA was isolated and processed as usual to determine the editing yield by Sanger sequencing. The SA enzymes were first directed to specific 5'-UAG triplets in the 3'UTRs of their own transcripts and of endogenous *GAPDH*, *ACTB* and *GUSB*. Editing yields of 40-80% (average 60%) were obtained when using SA1 and SA2 cells (**Man.4, Fig. 1a**). The editing yield was dependent on the targeted transcript (*GAPDH* \approx *ACTB* > *GUSB* > *SA*) rather than on the enzyme. According to the results obtained from the above-mentioned *in vitro* experiment, it was expected that the editing performance of SA2 may be impaired by using extensively modified gRNAs. However, it seems that the new gRNA design is well tolerated by the ADAR2 deaminase domain. Due to the extension of the gRNA, the four terminal PS linkages are shifted by two

nucleotides to the 3'-end, leading to additionally unmodified phosphodiester linkages between the nucleotides 16, 17 and 18. Regarding the previous discussion, it is tempting to speculate that the restored editing activity of SA2 resulted from the disappearance of the PS linkage between the nucleotides 16 and 17 in the gRNA. When BG-gRNA was transfected into SA1Q and SA2Q cells, editing was increased (65-90%, average 80%) compared to the cells expressing the wild-type SA enzymes (**Man. 4, Fig. 1a**). This was particularly true for *GUSB* and *SA* transcripts whose editing clearly profited from the activated SA enzymes which exhibit similar editing performances. The SA1Q and SA2Q cells were further used to investigate how long RNA editing stays stable after the transfection of the BG-gRNA targeting the 5'-UAG triplet in the 3'-UTR of *GAPDH*. It could be shown that 3 hours after gRNA transfection, maximum editing was achieved (**Man. 4, Fig. 1b**). The editing yield remained stable (80-90%) for three days and declined slightly afterwards, reaching ~ 60% at day 5 post-transfection. The reduction of editing may result from the dilution of the BG-gRNA-deaminase complex in the fast-growing cell line which was cultivated in medium supplemented with 10% FBS. Thus, the duration of stable editing may be even longer when using slowly dividing or non-dividing cells. M. Moschref (T. Stafforst group) showed that site-directed RNA editing was adjustable from 5% to 90% when targeting a 5'-UAG triplet in the ORF of *GAPDH* with increasing BG-gRNA amounts (39 fmol – 20 pmol/96 well; **Man. 4, Fig. 1c**). The activated SAQ enzymes reached the half-maximum editing yield with 0.15 pmol BG-gRNA and were up to 12-fold more potent than the wild-type SA enzymes with 1-2 pmol. High editing yields of $\geq 80\%$ were achieved with ≥ 1.25 pmol BG-gRNA when using SA1Q- or SA2Q-expressing cells. Given the high efficiency (up to 90% editing), potency (≥ 1.25 pmol BG-gRNA) and duration (several days), site-directed RNA editing with BG-gRNA-driven SA enzymes shows similar performance characteristics to siRNA-mediated RNAi in cell culture.¹⁵⁷ Editing of various 5'-UAG triplets in the 5'-UTRs and the ORFs of transcripts was shown to be more challenging than in the 3'-UTRs, especially in case of the wild-type SA enzymes which achieved editing yields of 25-50% (average 35%) in the 5'-UTRs and 15-60% (average 35%) in the ORFs (**Man. 4, Fig. 1d**). We assumed that the reduced editing in these areas resulted from the competition with the translation machinery which might disrupt the binding of the SA enzyme to the target site. In line with this, global translation inhibition by puromycin resulted in elevated editing within the ORF of *GUSB* mRNA in the wild-type SA cells with yields equal to those obtained from the 3'-UTR editing which was not notably influenced by the puromycin treatment

(**Man. 4, Suppl. Fig. 3**). The SAQ enzymes with the catalytic rate-enhancing E/Q mutation increased the editing yields in the 5'-UTRs (60-75%, average 70%) and the ORFs (50-85%, average 65%; **Man. 4, Fig. 1d**). The positive effect of the E/Q mutation on the editing performance has also been reported for the approaches based on λ N-ADAR2 and dCas13b-ADAR2.^{145,158,159} Concurrent editing of the 5'-UAG triplets in the 3'-UTRs of *SA*, *GAPDH*, *ACTB* and *GUSB* transcripts was examined in all four SA cell lines co-transfected with the four corresponding BG-gRNAs. The editing yields were equal to those obtained from the single transcript editings (**Man. 4, Fig. 1a**). Furthermore, three BG-gRNAs were co-transfected for the concurrent editing of three 5'-UAG triplets (2×ORF, 1×3'-UTR) within the *GAPDH* mRNA. Also in this case, the editing yields remained unchanged compared to the editings at the single sites (**Man. 4, Suppl. Fig. 2**). All 16 triplets containing an adenosine at the middle position (5'-NAN) were selected in the ORF of the *GAPDH* transcript to be edited in SA1Q and SA2Q cells (**Man. 4, Suppl. Note 4**). The editing experiments were conducted by M. Moschref (T. Stafforst group). The obtained results were in accordance with the reported preferences of ADAR enzymes with the 5'-GAN triplets being most challenging to edit (< 30%; **Man. 4, Fig. 1e**).^{67,85} Nevertheless, editing yields > 50% were obtained for 10/16 triplets and even > 70% for 7/16 triplets (5'-CAA, CAC, CAG, AAC, UAU, UAC, UAG). SA1Q and SA2Q achieved similar editing yields, excepting for the triplets GAU (SA1Q < SA2Q), CAU (SA1Q > SA2Q) and CAG (SA1Q < SA2Q). Due to the large number of decently editable triplets, the application of the SA approach holds promise for recoding many functionally important amino acid residues, including those involved in signal transduction.

Since the SA enzyme covalently binds to the BG-gRNA, both components merge into one. The number of required components, which need to interact during the deamination reaction, is reduced to two, the mRNA and the gRNA-SA conjugate. In contrast, all other approaches are based on non-covalent interactions between editase and gRNA. These editing systems require the simultaneous interaction between three components (mRNA + gRNA + editase) during deamination. Therefore, the probability for successful deamination might be lower compared to the SA/BG-gRNA system. This should be reflected in higher editing yields when using the SA approach. As mentioned earlier, over-expressed SA1 achieved ~ 45% editing of *eGFP W58X* in cell culture (**Man. 2, Suppl. Fig. 19,20**). Compared to this, MCP-ADAR1 directed by MS2-gRNA edits the same transcript with a much lower efficiency of ~ 5%.¹⁴⁷ The boxB-gRNA-driven λ N-ADAR2 editing system was also less efficient in catalyzing *eGFP W58X* (~ 20%) as

shown by the first report.¹⁴⁶ However, ~70% editing of *eGFP W58X* was recently achieved by an optimized version of the system which consists of the ADAR2 deaminase domain with the activating E/Q mutation and four λ N peptides at the N-terminus (4 λ N-ADAR2Q) and of a gRNA containing two *boxB* hairpins (2boxB-gRNA).¹⁵⁸ Additionally, a similar variant (NLS₃- λ N-ADAR2Q), containing three nuclear localization signal (NLS) copies and only one λ N peptide at the N-terminus, was able to correct a Rett syndrome-causing mutation within endogenous *Mecp2* mRNA in mouse neurons.¹⁵⁹ The targeted 5'-CAA triplet was edited with ~70% efficiency. In comparison, when targeting a CAA triplet within the endogenous *GAPDH* message, the genomically expressed SAQ mutants achieved even higher editing yields (SA1Q: ~80%, SA2Q: ~90%; **Man. 4, Fig. 1e**). However, since the editing of the CAA triplet was tested on two distinct transcripts, it is difficult to judge which editing system performs better. Besides the applied system, the editing efficiency also depends on the target accessibility which may be hampered by RNA secondary structures, RNA-binding proteins, low mRNA levels and short transcript half-lives. So far, the λ N-ADAR2Q enzymes have been highly over-expressed from plasmids in mammalian cells. It remains to be determined how efficiently the editases perform under genomic expression which generally lead to lower protein levels compared to transient expression. In contrast, the approach based on steering human ADARs by gRNA containing the R/G-motif of the *GRIA2* transcript (R/G-gRNA) has already been applied upon stable integration of the ADAR-expressing plasmid into the genome.¹⁴⁸⁻¹⁵⁰ It could be shown that although the genomic expression of ADAR2 was 20-fold lower than the transient expression, *eGFP W58X* editing was even enhanced from ~45% to ~65%.¹⁴⁸ Despite the promising yields obtained from the editing of the exogenously expressed marker construct, editing of endogenous transcripts was much weaker. For instance, a 5'-UAG triplet within the *ACTB* transcript was edited with ~15% efficiency in case of the genomic expression of ADAR2 and ~25% efficiency in case of transient expression.¹⁴⁸ In comparison, the same 5'-UAG triplet within *ACTB* was targeted in SA-expressing cells and edited with much higher efficiencies ranging from ~70% (SA1 cells) to ~85% (SA2Q cells; **Man. 4, Fig. 1e**). dCas13b-ADAR2Q has been reported to achieve 89% editing of a premature 5'-UAG stop codon in a luciferase reporter transcript.¹⁴⁵ One has to be skeptical about the high editing yield since the editase reached much smaller yields in case of all other targeted transcripts. dCas13b-ADAR2Q were tested for the correction of over 30 disease-relevant 5'-UAG nonsense mutations in a reporter construct (for further discussion, see **section 3.2.1**). In contrast to the highly efficient 5'-UAG

editing in the luciferase transcript, the enzyme edited the 5'-UAG triplets with only low efficiencies between ~ 0% and 30%.¹⁴⁵ 5'-UAG triplets within endogenous transcripts were edited by dCas13b-ADAR2Q with 15-45% efficiency. SA1Q and SA2Q achieved editing yields between 50% and 90% when testing several 5'-UAG triplets within endogenous transcripts (**Man. 4, Fig. 1d**). Even SA1 and SA2 showed a higher overall efficiency (15-85%) in the editing of 5'-UAG triplets compared to dCas13b-ADAR2Q (**Man. 4, Fig. 1d**). dCas13b-ADAR2Q was also tested for the editing of *eGFPW58X* by T. Merkle (T. Stafforst group; **Man. 4, Suppl. Fig. 11, Suppl. Note 2**). The experiments were conducted according to the conditions applied by Cox *et al.*¹⁴⁵ It could be revealed that dCas13b-ADAR2Q edits *eGFPW58X* with ~ 30% efficiency which was lower than that achieved by wild-type SA1 (~ 45%, **Man. 2, Suppl. Fig. 19,20**). As described later in **section 3.2.2**, further evidence was provided that the SA approach achieves higher editing efficiencies than the dCas13-ADAR approach. The T375G mutation in dCas13b-ADAR2Q has been described to enhance the specificity of the enzyme without the substantial loss of editing efficiency.¹⁴⁵ However, when tested for the editing of the luciferase reporter transcript, the yield of 89% obtained by Cas13b-ADAR2Q decreased to ~ 45% when Cas13b-ADAR2Q T375G was used.¹⁴⁵ As shown by S. Flad (T. Stafforst group), the editing of five tested triplets (5'-UAG, CAA, CAG, AAG, GAU) within the ORF of *GAPDH* mRNA was greatly impaired when the T375G mutation was introduced into SA2Q (SA2QG) (**Man. 4, Suppl. Fig. 10**). Even the wild-type SA enzymes edited the triplets with higher efficiencies than SA2QG. Since the results indicate that this mutation substantially decreases the editing activity, the claim, which was made for the T375G mutation, seems overstated.

3.2 SNAP-ADAR enzymes as a promising tool for medicine and life sciences

3.2.1 Correcting disease-causing mutations

The potential of SNAP-ADAR enzymes to be applied for correcting disease-causing mutations has been demonstrated *in vitro* by T. Stafforst (T. Stafforst group). The single G-to-A substitution at position 1746 in *coagulation factor V* mRNA lead to the replacement of an arginine residue by a glutamine residue (R534Q). The mutated variant is called Factor V Leiden which is resistant to the cleavage by protein C and increases the risk of thrombosis 8-fold in heterozygous individuals and even 80-fold in homozygous

individuals.¹⁶⁰ The *in vitro* experiment was conducted with a 1000-nt long fragment of the *factor V* transcript containing the disease-causing 5'-CAA glutamine codon. SA2 was shown to edit the codon with ~70% efficiency (**Man. 1, Fig. 3c**). Such a high editing efficiency would be enough to dramatically alleviate the disease phenotype.¹⁶⁰ Other editing systems have also been used to repair disease-related mutations within RNA. The λ N-ADAR2 approach was applied to correct a disease-related nonsense mutation within *CFTR* mRNA in *Xenopus* oocytes, leading to restored CFTR protein activity.¹⁴⁶ The rescue of mitophagy after repairing a loss-of-function mutation in the *PINK1* transcript has been achieved in mammalian cells by R/G-gRNA-directed ADAR2.¹⁴⁸ As mentioned earlier, dCas13b-ADAR2Q was able to edit over 30 disease-relevant UAG sites.¹⁴⁵ However, only small fragments (200 bp) of the original sequences were used to be expressed within a reporter cassette. It is therefore difficult to estimate whether the endogenous transcripts are also editable with the same efficiencies. Furthermore, it is unclear if the achieved editing yields ($\leq 30\%$) would be sufficient to change the disease phenotypes. In cell culture experiments, the correction of disease-causing mutations by site-directed RNA editing has so far been largely tested on over-expressed transcripts in standard cell lines. Given the current delivery methods, it remains challenging to efficiently introduce an artificial editing enzyme or gRNA into primary cells or higher organisms to perform RNA editing of endogenous transcripts. Recently, the boxB-gRNA/ λ N-ADAR2Q editing system was successfully delivered into primary mice neurons via adeno-associated virus (AAV) transduction to perform the above-mentioned repair of endogenous *Mecp2* mRNA, leading to increased levels and restored heterochromatin enrichment of Mecp2 protein.¹⁵⁹ Viral delivery vehicles have also been successfully applied to enable *in vivo* gene transfer.¹⁶¹ Regarding the SA approach, only the SA enzyme can be delivered by viral carriers. The uptake of the BG-gRNA into primary cells or *in vivo* may be achieved by already existing oligonucleotide delivery vehicles including lipoplexes, cell-penetrating peptides, cholesterol and GalNAc₃.^{98,99,152,162,163} Since the BG-gRNA/SA system was shown to achieve high editing yields in a standard cell line and in a simple organism, it has the potential to attenuate a disease phenotype *ex vivo* or *in vivo* when efficiently delivered.

3.2.2 Reversible manipulation of signaling networks

KRAS is an important molecular on/off switch regulating signal transduction.¹⁶⁴ Recently, the dCas13b-ADAR approach was applied for the editing at two 5'-UAG sites

(#1, #2) within the ORF of endogenous *KRAS* mRNA. dCas13b-ADAR2Q achieved 22% editing at site #1 and 32% at site #2.¹⁴⁵ The editing at the two sites was repeated with the SA approach using SA1Q cells. It could be shown that SA1Q substantially increased the editing yields at both sites (#1: ~ 55%, #2: ~ 46%; **Man. 4, Fig. 2e**). Additionally, the editing remained similar when both sites were targeted concurrently in SA1Q cells. SA1 was also tested for editing at site #1 and achieved 18% editing (**Man. 4, Fig. 2e**). In comparison, the specific variant dCas13b-ADAR2Q T375G edits the same site with 13% efficiency.¹⁴⁵ Signal transduction is mediated by the phosphorylation of serine, threonine and tyrosine residues. The triplets 5'-UAU and 5'-UAC code for tyrosine and were edited up to 90% within endogenous GAPDH mRNA by the activated SAQ variants (**Man. 4, Fig. 1e**). Thus, it was assumed that the SA approach enables the efficient recoding of tyrosine residues within signaling proteins. The phosphorylation of Tyr 701 in the transcription factor STAT1 is essential for signal transduction and appeared therefore as an appropriate target.¹⁶⁵ The editing of the corresponding 5'-UAU codon was examined in SA1 and SA1Q cells under the assistance of K. D. Selvasaravanan (T. Stafforst group). It could be shown that SA1 edited the codon with ~ 30% efficiency while SA1Q reached ~ 80% (**Man. 4, Fig. 2e**). This shows that it is readily feasible to recode tyrosine residues which are relevant for signaling. Additionally, *KRAS* (site #1) and *STAT1* (Tyr 701) mRNAs were simultaneously targeted in SA1Q cells. Both transcripts were edited with similar efficiencies compared to their single editings (**Man. 4, Fig. 2e**). Thus, concurrent editing of signaling transcripts can be achieved which opens the possibility of manipulating entire signaling networks.

3.2.3 Switching protein localization by light

We were wondering if SA enzymes can expand the synthetic biology toolbox for manipulating protein localization. To achieve spatiotemporal control of protein translocation, I applied the concept of light-triggered RNA-editing that was developed in the previous study by A. Hanswillemenke (T. Stafforst group; **Man. 2**). The 6-nitropiperonyloxymethyl (NPOM) caging group has recently been shown to be stable under physiological conditions and to enable efficient decaging of aromatic *N*-heterocyclic compounds upon irradiation with UV-light of 365 nm.¹⁶⁶ The NPOM group was installed on BG before its coupling to NH₂-gRNA. NPOM-caged BG-gRNA (^{NPOM}BG-gRNA) fails to act as a substrate for the SA enzyme. However, UV-light irradiation cleaves the NPOM group from the BG moiety and allows the formation of the

gRNA-deaminase complex ready for catalyzing site-directed RNA editing (**Man. 2, Fig. 1, 2a,b**). The decaging of ^{NPOM}BG-gRNAs has been proven to be very efficient to mediate light-triggered RNA editing in a reaction tube (**Man. 2, Fig. 2c**), a mammalian cell (HEK 293T cell; **Man. 2, Fig. 3**) and in a simple organism (*Platynereis dumerilii*; **Man. 2, Fig. 5**). The editing yield was adjustable by the duration of UV-light exposure. However, minor background editing at the target site was observed in non-irradiated cells and worms after the transfection/injection of ^{NPOM}BG-gRNA. This has led to the speculation that the NPOM group does not entirely block the complex formation between SA enzyme and gRNA. Therefore, when applying ^{NPOM}BG-gRNA in an experiment, optimization of the SA and gRNA amounts is required to reduce the background editing levels to a minimum.

The strategy was ready to be used for the photo-induced switching of protein localization within a mammalian cell. Short peptide sequences typically define the subcellular localization of a protein. For instance, proteins containing a nuclear localization signal (NLS) are transported into the nucleus.¹⁶⁷ In this study, site-directed RNA editing by SA enzymes was exploited to introduce three copies of the SV40 large T antigen NLS (NLS₃) into the N- or C-terminus of a protein to induce its nuclear transport. SA2 is localized in the cytoplasm without NLS inclusion and its localization can be easily visualized under a fluorescence microscope when stained with BG-fluorescein (BG-FITC; **Man. 4, Suppl. Note 3**). Therefore, the *SA2* transcript was chosen as a target for NLS₃ inclusion mediating the switch in the localization of the SA2 protein from the cytoplasm to the nucleus. Since SA2 can be used for site-directed RNA editing, only one gene construct was required for the editing experiments. For C-terminal NLS₃ inclusion by site-directed RNA editing, a gene construct was designed containing a premature 5'-TAG stop codon between the *SA2* sequence and the downstream NLS₃ sequence (SA2-TAG-NLS₃; **Man. 3, Scheme 1**). At the RNA level, the conversion of the stop codon (5'-UAG) to a tryptophan codon (5'-UIG) by RNA editing results in the translation of SA2 harboring the NLS₃ at its C-terminus which redirects the enzyme from the cytoplasm to the nucleus. The *SA2-TAG-NLS₃* construct was first transiently expressed from a pcDNA3.1 plasmid in 293T cells. The editing experiment was again conducted according to the new cellular assay described in **section 3.1**. Four hours after the reverse transfection with BG-gRNA, the cells were detached and seeded on cover slips. 24 hours later, the cells were stained with BG-FITC or harvested for the Sanger sequencing of *SA2-TAG-NLS₃* cDNA. Without BG-gRNA, the stained SA2 protein was

strictly located in the cytoplasm (**Man. 3, Fig. 1a**). However, when *SA2-TAG-NLS₃*-expressing cells were additionally transfected with BG-gRNA, ~50% of these cells showed a new phenotype representing a mixture of cytoplasmic and nuclear SA2 protein. The editing yield was ~75%. The mixed phenotype differed from the phenotype of the positive control (*SA2-TGG-NLS₃*), where SA2 protein could only be found in the nucleus. It was reasoned that the mixture resulted from the incomplete editing and old SA2 protein which was already produced before BG-gRNA transfection. The ability of the BG-gRNA/SA system to mediate C-terminal NLS₃ inclusion was also tested when *SA2-TAG-NLS₃* was weaker expressed under genomic control in 293 Flp-In T-REx cells. Again, transfecting BG-gRNA induces the phenotype change from strict cytoplasmic to cytoplasmic/nuclear localization (**Man. 3, Fig. 1b**). The mixed phenotype was observed in ~35% of the cells and the target mRNA was edited with ~50% efficiency. Taken together, C-terminal NLS₃ inclusion to redirect a protein from the cytoplasm to the nucleus can be indeed achieved by SA-mediated RNA editing under transient or genomic expression. Next, NLS₃ inclusion at the N-terminus was examined under both expression conditions. Translation of mRNA is initiated by an AUG start codon typically embedded in a Kozak sequence.¹⁶⁸ The gene construct designed for this study contained a Kozak sequence (5'-CCACC-ATG-G) between the *NLS₃* sequence and the downstream *SA2* sequence (**Man. 3, Scheme 1**). In addition, another Kozak sequence harboring a single G-to-A substitution within the start codon (5'-CCACC-ATA-G) was placed in front of the *NLS₃* sequence (*ATAG-NLS-SA2*). Without the editing of the defective 5'-AUA start codon at the RNA level, translation was assumed to start in front of the *SA2* sequence, leading to cytoplasmic SA2 protein. In contrast, the conversion to 5'-AUI would initiate translation in front of the *NLS₃* sequence which results in the translocation of the protein to the nucleus. When the gene construct harboring the defect start codon in front of *NLS₃* (*ATAG-NLS-SA2*) was transiently expressed in 293T cells, SA2 was found almost exclusively in the cytoplasm (**Man. 3, Fig. 2a**). However, in ~10% of the transfected cells, a small portion of SA2 protein was also located in the nucleus as indicated by a faint nuclear staining (**Man. 3, Suppl. Fig. 10**). This indicates that translation may be slightly initiated in front of the *NLS₃* sequence despite the defective start codon. In contrast, the same sequence (5'-CCACC-AUA-G) has been shown not to initiate the translation of plasmid-encoded preproinsulin in transfected COS cells.¹⁶⁹ The small amount of nuclear SA2 protein observed in a minor fraction of cells transiently expressing *ATAG-NLS-SA2* did not result from gRNA-independent editing of the defect start codon

since the editing yield was negligible ($\leq 2\%$). When the cells were additionally transfected with BG-gRNA, $\sim 60\%$ editing was obtained and $\sim 55\%$ of the cells showed a mixed cytoplasm/nuclear phenotype similar to that described after C-terminal NLS₃ inclusion (**Man. 3, Fig. 2a**). The fluorescence of nuclear SA2 protein in cells transfected with BG-gRNA was much stronger than that of nuclear SA2 protein occasionally observed in the negative control (without gRNA; **Man. 3, Suppl. Fig. 10**). The *ATAG-NLS-SA2* gene construct was also genomically expressed in 293 Flp-In T-REx cells. Less than 3% of the cells contained a small portion of nuclear SA2 protein, probably because of the reduced expression of the construct. When transfected with BG-gRNA, only $\sim 10\%$ of the cells contained nuclear SA2 protein. According to this, Sanger sequencing revealed an editing yield of $\sim 15\%$ (**Man. 3, Fig. 2b**). Thus, reducing the expression of target mRNA and SA enzyme appears to substantially impair the editing within the 5'-UTR for N-terminal NLS₃ inclusion. The difference between the editing yields under genomic and transient construct expression was much more pronounced than obtained for the C-terminal NLS₃ inclusion where editing was catalyzed within the ORF. To increase the editing within the 5'-UTR under genomic expression, the gene construct was re-engineered using the sequence of the hyperactive SA2Q variant instead of wild-type SA2. Similar to the old gene construct, the genomic expression of the new version (*ATAG-NLS-SA2Q*) in 293 Flp-In T-REx cells led to a small percentage of cells ($< 4\%$) containing a minor fraction of nuclear SA protein (**Man. 3, Fig. 2b**). After BG-gRNA transfection, $\sim 45\%$ of the cells showed the cytoplasmic/nuclear phenotype and $\sim 40\%$ editing was obtained. Thus, choosing the hyperactive SA2Q mutant greatly increases the editing in the 5'-UTR under genomic expression, leading to an improved efficiency of N-terminal NLS₃ inclusion comparable with that of C-terminal NLS₃ inclusion.

As mentioned earlier, decaging of ^{NPOM}BG-gRNAs were tested for light-induced protein translocation. For this, the medium was changed four hours after gRNA-transfection and the cells were irradiated at 365 nm for 5 s (UV-LED) before they were seeded on cover slips. ^{NPOM}BG-gRNAs were first applied for N-terminal NLS₃ inclusion in HEK 293T cells transiently expressing *ATAG-NLS-SA2*. Without UV-light irradiation, $\sim 20\%$ of the cells contained a small portion of nuclear SA2 after ^{NPOM}BG-gRNA transfection (**Man. 3, Fig. 3a**). However, after UV-light irradiation, the mixed phenotype with clearly visible amount of nuclear SA2 protein was observed in $\sim 55\%$ of the cells. Accordingly, the editing in ^{NPOM}BG-gRNA-transfected cells increased from $\sim 15\%$ before to $\sim 40\%$ after irradiation. Photo-induced N-terminal NLS₃ inclusion was also tested

under genomic expression. As expected, the editing yield and phenotype switch induced by light-triggered RNA editing were inefficient (< 20%) in Flp-In T-REx cells expressing the wild-type enzyme (ATAG-NLS-SA2, **Man. 3, Fig. 3b**). Compared to this, when transfecting ^{NPOM}BG-gRNA into Flp-In T-REx cells expressing the hyperactive SAQ version (ATAG-NLS-SA2Q), the editing was substantially increased from ~ 15% in non-irradiated cells to ~ 35% in irradiated cells. Similarly, the percentage of the cells showing the cytoplasm/nuclear phenotype was enhanced from ~ 10% before to ~ 40% after irradiation. Therefore, nuclear import of a protein, whether expressed transiently or genomically, can be indeed controlled by light-induced ^{NPOM}BG-gRNA decaging which enables complex formation between SA enzyme and gRNA. There are several tools inducing protein translocation into the nucleus to control gene expression. For instance, the small molecule tamoxifen is administered to induce nuclear translocation of the engineered CreER recombinase which mediates on/off switching of gene expression *in vivo*.¹⁷⁰ Furthermore, inducing the NLS-directed nuclear transport of transcription factors has been shown to trigger gene expression.^{171,172} In this regard, the inclusion of an NLS into the C- or N-terminus of a transcription factor by site-directed RNA editing might be also used for the induction of gene expression. When additionally applying Npom-caged BG-gRNAs, the SNAP-ADARs might enable light-dependent activation of gene expression and could be an alternative to current TALE-based or CRISPR-Cas9-based optogenetic tools which require substantial engineering due to the high number of components.¹⁷³⁻¹⁷⁶ Light-induced protein translocation from the cytoplasm to the nucleus can also be achieved by other strategies including photo-decaging of amino acids or small molecules, and photo-controlled protein unfolding.^{171,172,177,178} The strategies have been successfully applied since nuclear protein import is mediated post-translationally. However, these strategies are only applicable at the protein level and fail to manipulate co-translational protein translocation into the ER. Membrane or secretory proteins typically contain an N-terminal signal sequence which is recognized by the signal recognition particle (SRP) during translation.¹⁷⁹ The SRP recruits the ribosome to the ER where the elongating polypeptide chain is translocated across the ER membrane. The signal sequence is usually cleaved off and after translocation, the mature protein can fold before being transported towards the plasma membrane. Since RNA editing is catalyzed before translation, SA enzymes might be used to control the co-translational protein transport to the ER. To test this, a gene construct was designed containing a defective 5'-ATA start codon followed by the IgK-chain leader sequence which functions as signal

sequence for ER-mediated protein secretion (**Man. 3, Scheme 1**).¹⁸⁰ An alternative start codon together with the sequence of *HA-tagged eGFP (HA-eGFP)* was placed after the signal sequence. To attach the resulting protein to the plasma membrane, the gene construct additionally contained the sequence of the transmembrane domain (TMD) of the PDGF receptor at its C-terminus. Repairing the defective 5'-AUA start codon at the RNA level allows the translation of the IgK-chain leader sequence which mediates the transport of the protein to the plasma membrane. When integrated into the membrane via the TMD, the N-terminus of the protein (HA-eGFP) is located extracellularly. An antibody against the HA-tag was used to visualize the protein at the plasma membrane via red immunofluorescence. The gene construct (*ATAG-IgK-HA-eGFP-TMD*) was transiently expressed in HEK 293T cells together with *BFP-tagged SA2 (BFP-SA2)*. The co-transfected cells contained only intracellular BFP and eGFP fluorescence (**Man. 3, Suppl. Fig. 19, Suppl. Table 7**). When transfecting these cells with BG-gRNA, ~45% of the cells additionally exhibit red fluorescence at the plasma membrane (**Man. 3, Fig. 4c**). The editing yield was ~65%. Protein translocation to the membrane was also mediated by light-induced RNA editing. When transfecting ^{NPOM}BG-gRNA into the cells expressing *ATAG-IgK-HA-eGFP-TMD* and *BFP-SA2*, only ~10% RNA editing was detected and the percentage of cells showing red fluorescence at the membrane was negligible (<2%; **Man. 3, Fig. 4d**). In contrast, the irradiation of the ^{NPOM}BG-gRNA-transfected cells with UV-light resulted in ~45% RNA editing and ~30% of the cells exhibiting red membrane staining (**Man. 3, Fig. 4e**). Taken together, site-directed RNA editing can indeed be used to translocate a protein from the cytoplasm to the plasma membrane, even under the control of light. So far, no other strategy has been exploited to steer a protein to the membrane. RNA editing by SNAP-ADARs therefore expands the synthetic biology toolbox for directing protein localization within a cell. Inducing protein translocation to the plasma membrane might be useful to modulate cell signaling or cell-to-cell/matrix interactions. In addition, the use of Npom-caged BG-gRNA offers the possibility to manipulate such biological processes in a spatiotemporal manner.

3.3 SNAP-ADAR enzymes have a promising safety profile

A major hurdle for the application of RNA editing systems is maintaining the integrity of the A-to-I RNA editome in the cell. Altered A-to-I editing is associated with several human diseases, including cancer, neurodegeneration and autoimmunity.^{181,182} Therefore, an optimal RNA editing system does not interfere with the function of endogenous

ADARs and fails to cause any off-target editing while being very efficient in catalyzing editing at target sites.

To determine global off-target A-to-I editing events caused by the SA approach, the poly(A)⁺ transcriptome of the SA cell lines (SA1, SA2, SA1Q, SA2Q) was analyzed using NGS (Illumina HiSeq 4000 platform, 50 million 2×100 bp paired end reads per sample, two independent experiments). For the study, RNA was isolated from the cells which were transfected with BG-gRNA targeting a 5'-UAG triplet in the 3'-UTR of endogenous *ACTB*. To distinguish between gRNA-dependent and -independent off-target editing, RNA samples from SA1Q and SA2Q cells lacking the BG-gRNA were additionally prepared. The analysis of the NGS data was done by Q. Li (J. B. Li group, Stanford University, USA). The RNA-seq reads were aligned to the human (hg19) genome and exonic sequences around known splice junctions. The rate of uniquely mapped reads was typically ~ 95%. Around 25,000 distinct transcripts with FPKM ≥ 2 (both replicates combined) were detectable (**Man. 4, Suppl. Note 4**). Editing sites were identified through a pipeline described by Ramaswami *et al.* (separate samples method).¹⁰ The pipeline can be applied when using RNA-seq data alone and removes known human SNPs to reduce the false positive rate to usually less than 3%. Around 50,000 editing sites with ≥ 50 reads coverage (both replicates combined) were identified in every sample (**Man. 4, Suppl. Note 4**). The editing levels of these sites were compared between the samples obtained from the SA cell lines and the control obtained from 293 Flp-In T-REx cells stably transfected with the empty vector lacking any SA gene. In case of changes in editing > 10%, Fisher's exact test was performed to identify significantly differently edited sites ($p < 0.01$). Almost all (> 99.8%) significantly differently edited sites were found to exhibit higher editing levels in the SA samples compared to the control. These sites were considered as off-target editing sites. The analysis also identified the targeted editing of the 5'-UAG triplet within *ACTB* in all four SA+gRNA samples (**Man. 4, Fig 2b**). Both wild-type SA enzymes were shown to be highly precise editing machines since very few off-target sites (SA1+gRNA: 6 sites, SA2+gRNA: 30 sites) could be found (**Man. 4, Fig. 2b**). Most of the off-target sites were edited < 25%, were known from the RADAR database¹⁸³ and were located within the 3'-UTRs of transcripts (**Man. 4, Table 1, Suppl. Fig. 5a**). Edits, which happen in the coding regions and lead to amino acid substitutions, require special attention. Only one of such so-called nonsynonymous editings was found in the SA1+gRNA sample. However, the site (in *TMX3*) was edited with an efficiency of only ~ 10% (**Man. 4, Fig. 2d**). The region around the off-target site shares high sequence

similarity to the gRNA-targeted region within *ACTB*, indicating that the found nonsynonymous editing was gRNA-dependent and would therefore disappear when omitting the BG-gRNA (**Man. 4, Suppl. Table 1**). Indeed, this site was also identified to be edited in the SA1Q+gRNA sample but not in the SA1Q sample without gRNA. Two nonsynonymous off-target sites (*AAGAB*: 42%, *CHFR*: 32%) were identified in SA2 cells. One of them (*AAGAB*) was also located in a region similar to the gRNA-targeted region and was found to be edited in a gRNA-dependent manner in the SA2Q samples (**Man. 4, Suppl. Table 1**). In contrast to the wild-type SA enzymes, the hyperactive versions generated a considerable number of off-target edits in the transcriptome. 835 off-target edits were identified in the SA1Q+gRNA sample and 1310 in the SA2Q+gRNA sample (**Man. 4, Table 1, Fig. 2b, Suppl. Fig. 5a,8c**). Most of these edits were found at novel sites (SA1Q+gRNA: 706/835 sites, SA2Q+gRNA: 972/1310 sites) and located either in the ORF (SA1Q+gRNA: 347/835, SA2Q+gRNA: 496/1310) or in the 3'-UTR (SA1Q+gRNA: 402/835, SA2Q+gRNA: 637/1310). The number of nonsynonymous off-target sites was 230 for the SA1Q+gRNA sample and 347 for the SA2Q+gRNA sample (**Man. 4, Fig. 2d, Suppl. Fig. 8d**). However, the majority of these sites were edited < 25% in both samples (SA1Q+gRNA: 167/230, SA2Q+gRNA: 240/347). Only 4/230 nonsynonymous off-target sites in the SA1Q+gRNA sample and 20/347 in the SA2Q+gRNA sample were found to be edited > 50% but below the level of the on-target editing (SA1Q+gRNA: 67%, SA2Q+gRNA: 77%). The off-target editing was generated by the editing enzyme alone rather than by the misguiding of the enzyme through the gRNA. This was concluded because compared to the SAQ samples without gRNA, only a few off-target sites (SA1Q: 29, SA2Q: 30) were additionally identified or edited with higher efficiency in the SAQ samples with gRNA (**Man. 4, Fig. 2c**). Almost all of them (58/59) were novel and edited < 50% (57/59). Approximately half of the off-target edits occurred at nonsynonymous sites (27/59). Furthermore, 414 off-target sites identified in the SA1Q+gRNA sample were also found to be edited in the SA2Q+gRNA sample. Therefore, most of the off-target edits (SA1Q+gRNA: 421/835, SA2Q+gRNA: 896/1310) were generated in an enzyme-dependent manner. Analysis of the sequence around the off-target sites revealed that the off-target editing produced by SA1Q is almost restricted to sites with an U or an A at the -1 position (5'-UAN or 5'-AAN; **Man. 4, Suppl. Fig. 5b**). In contrast, SA2Q seems to be more promiscuous as it additionally produced off-target edits at sites with a C at the -1 position to a noticeable extent. The weak neighbor preference at the -1 position may explain the higher off-target activity of SA2Q compared

to SA1Q. Off-target editing identified in this study had no obvious effects on cell physiology since the SA cell lines behaved identically to the cell line containing the empty vector in terms of morphology and proliferation. Accordingly, no difference in global gene expression was found in the SAQ cells transfected with gRNA compared to cells containing the empty vector (**Man. 4, Suppl. Fig. 6**). Nevertheless, attempts were made to lower the off-target editing in the SAQ cells without affecting the on-target editing. For this, site-directed RNA editing was performed in SA1Q and SA2Q cells which were incubated with doxycycline for different periods of time (0 h, 4 h, 8 h, 24 h, 48 h). Sanger sequencing was applied to determine the editing yield at the target site (GAPDH ORF#2) and at six top-ranked off-target sites. The amount of the SAQ proteins was reduced when decreasing the incubation time (**Man. 4, Suppl. Fig. 9a**). At four hours of doxycycline incubation, the protein amounts were less than 10% of those at 48 h of incubation. It could be shown that off-target editing was typically more affected by the SAQ protein reduction than the on-target editing (**Man. 4, Suppl. Fig. 9b,c**). Whereas most of the off-target editing yields were reduced by 2 to 3-fold when decreasing the incubation time from 48 h to 4 h, the loss of the on-target editing yield was only 28% in SA1Q cells (from 78% to 50%) and 18% in SA2Q cells (from 80% to 62%). Taken together, reducing the SAQ protein expression offers a possibility to dramatically decrease off-target editing in SAQ cells while maintaining high on-target editing.

The extend of global off-target editing has been also described for the dCAS13-ADAR approach and the λ N-ADAR approach. In case of dCas13b-ADAR2Q, NGS analysis identified more than 18,000 off-target events in the transcriptome.¹⁴⁵ In comparison with that, ~ 18-fold less off-target edits were found in SA1Q and SA2Q. dCas13b-ADAR2Q T375G caused only 20 off-target editing events and therefore provides a high specificity similar to those of the wild-type SA enzymes. One should note that the NGS analysis was performed with samples from cells transfected with only 10 ng of the respective dCas13-ADAR construct instead of 150 ng which was typically used for editing experiments including those conducted for the editing of endogenous *KRAS*. Therefore, it is reasonable to assume that both dCas13b-ADAR2Q and dCas13b-ADAR2Q T375G caused more off-target edits in the editing experiments than observed in the analysis of the NGS experiments. Additionally, it is hard to believe that 15-fold less plasmid, as used in the NGS experiments, can lead to the same editing yields achieved with 150 ng of the dCas13-ADAR constructs, especially in cases when endogenous transcripts are targeted. 4λ N-ADAR2Q has been reported to produce editing at ~ 77,000

off-target sites.¹⁸⁴ The off-target activity was reduced by approximately half when the editing enzyme was translocated from the cytoplasm to the nucleus (NLS₃-4λN-ADAR2Q). Wild-type 4λN-ADAR2 exhibits higher specificity as ~ 20,000 off-target edits were found. Also in this case, nuclear localization reduced the off-target activity of the enzyme (NLS₃-4λN-ADAR2, ~ 8,800 off-target sites). The authors claimed that sequestering the enzyme in the nucleus decreases off-target editing without affecting on-target editing. The NGS data was reanalyzed by Q. Li (J. B. Li group, Stanford University, USA) using the same pipeline that was applied for the SA approach. Contrary to the claim, the reanalysis revealed that nuclear translocation of the 4λN-ADAR enzymes results in a reduction of the on-target editing by ~ 20% (**Man. 4, non-included data a**). In accordance, redirecting SA2 or SA2Q from the cytoplasm to the nucleus affected editing at target sites (**Man. 4, non-included data b,c**). Additionally, the reanalysis showed that the wild-type 4λN-ADAR enzymes (4λN-ADAR2, NLS₃-4λN-ADAR2) were several 100-fold less specific than the wild-type SA enzymes (**Man. 4, Suppl. Fig. 8**). Around 14-fold more off-target edits were identified for the hyperactive 4λN-ADARQ versions (4λN-ADAR2Q, NLS₃-4λN-ADAR2Q) compared to the SAQ enzymes.

Off-target editing produced by an editing system can occur on the target RNA. Since the gRNA and the target RNA form a nearly perfect double-stranded RNA duplex, not only the target adenosine can be edited but also other adenosines which are located within the duplex structure. SA enzymes are steered by BG-gRNAs which cannot be expressed from plasmids and need to be stabilized by chemical modification to be effective in cell culture. Besides stabilizing the BG-gRNA, chemical modification also prevents the off-target editing within gRNA/target RNA duplexes. It could be shown by M. F. Schneider (T. Stafforst group) that the *in vitro* editing of *eCFP W66X* can be blocked by a gRNA containing several 2'-OMe-modified nucleosides next to the editing site (**Man. 1, Fig. 1b**). This knowledge was applied by T. Stafforst (T. Stafforst group) to inhibit the off-target editing when repairing the disease-causing 5'-CAA glutamine codon within the *factor V* transcript *in vitro*. In case of the unmodified gRNA, the target A (5'-CAA) was edited by SA2 with ~ 70% efficiency while the neighboring A (5'-CAA) suffered from strong off-target editing (~ 50%; **Man. 1, Fig. 3b**). The off-target editing could be eliminated by using a gRNA, which were modified with two 2'-OMe-modified ribonucleosides opposite the off-site A and its 3'-neighboring base, while maintaining editing at the target site (~ 70%; **Man. 1, Fig. 3c**). When performing site-directed RNA editing of endogenous transcripts in SA cells, almost all 29 tested BG-gRNAs did not

produce any off-target edits within the gRNA/target RNA duplexes. Only four out of the 29 BG-gRNAs led to off-target editing within the gRNA/target RNA duplexes. Off-target editing was observed when targeting the triplets 5'-CAA, AAA, AAC and UAA within *GAPDH* mRNA in SA1Q or SA2Q cells (**Man. 4, Fig. 2a**). In all four cases, the adenosine adjacent to the target adenosine suffered from off-target editing (5'-CAA, AAA, AAC, UAA). The levels of editing at these off-target sites were very low in SA1Q cells (0-10%, average 6%), but noticeable in SA2Q cells (5-75%, average 30%). In particular, the off-site A within the 5'-CAA codon was found to be strongly edited with ~ 75% efficiency in SA2Q cells. The unintended edits within the four triplets happened due to the structure of the BG-gRNA which contains a gap of three unmodified ribonucleosides opposite to the targeted triplet (**Man. 4, Suppl. Fig. 1b**). The ribonucleoside facing the off-site adenosine was additionally modified with a 2'-OMe or a 2'-F group to avoid off-target editing (**Man. 4, Fig. 2a, Suppl. Fig. 4**). Indeed, the editing of the off-site adenosines could be reduced by these modifications, whereby the 2'-OMe group had a stronger inhibitory effect on the off-target editing than the 2'-F group. In case of SA2Q cells, the off-target editing within the triplets 5'-AAA (3%), AAC (2%) and UAA (~ 10%) almost disappeared when using BG-gRNAs carrying the additional 2'-OMe modification. The new BG-gRNA variant was also able to substantially decrease the off-target editing within the 5'-CAA triplet from ~ 75% to 25%. Remarkably, the additional modifications in the BG-gRNA did not affect the on-target editing within all four triplets. In case of the 5'-AAA triplet, the 2'-F-modified BG-gRNA even increased the on-target editing from 40% to 55%. In contrast to the SA approach, all other approaches rely on genetically encodable gRNAs expressed from plasmids. Since such gRNAs lack any chemical modification, they make gRNA/target RNA duplexes vulnerable to off-target editing. Indeed, off-target edits within gRNA/target RNA duplexes were observed when using R/G-gRNAs directing wild-type ADAR enzymes.^{148,150} Furthermore, applying the λ N-ADAR2 approach to repair *Mecp2 R106Q* produced up to 60% off-target editing at several sites within the transcript.¹⁵⁹ Four out of the five off-target sites were located in the gRNA/target RNA duplex. Massive off-target editing within gRNA/target RNA duplexes is also generated by dCas13b-ADAR2Q.¹⁴⁵ For instance, multiple off-target edits were detected within the gRNA/target RNA duplexes when the dCas13b-ADAR2Q was applied for the targeting of the KRAS sites #1 and #2. In contrast, SA enzymes edited these sites without any detectable off-target editing (**Man. 4, Suppl. Fig. 7, Suppl. Note**

1). The off-target editing within gRNA/target RNA duplexes by dCas13b-ADAR2Q was found to be reduced when using the more specific dCas13b-ADAR2Q T375G.¹⁴⁵

The editing of off-site adenosines on the target RNA can also occur outside of a gRNA/target RNA duplex. When testing the SNAP-ADAR approach for the editing of endogenous transcripts, off-target editing outside the gRNA/target RNA duplexes was only found within *GAPDH* mRNA when using SA1Q and SA2Q cells (**Man. 4, Suppl. Fig. 15a**). In wild-type SA1 or SA2 cells, such off-target edits were absent. In the SAQ cells, six off-target sites were detected, whereby only one site in each cell line (SA1Q: up to 50% editing of a 5'-AAG triplet in the 3'-UTR; SA2Q: up to 70% of a 5'-CAG triplet in the ORF) were strongly edited (**Man. 4, Suppl. Fig. 15c-d**). The RNA secondary structures around these two off-target sites were predicted using Mfold. The analysis predicted that both sites are located within dsRNA regions (**Man. 4, Suppl. Fig. 15b**). Additionally, off-target editing was analyzed using three BG-gRNAs binding to different locations of the transcript (against ORF #1, ORF #2 and 3'-UTR; **Man. 4, Suppl. Fig. 15a**). In the SA1Q and SA2Q cells, the editing at an off-target site within *GAPDH* mRNA increased when the SA enzyme was directed in the proximity to that site (**Man. 4, Suppl. Fig. 15c-d**). Therefore, gRNA-dependent editing of an off-site adenosine outside the gRNA/target duplex appears to rely on the secondary structure around that site and on the proximity of the editase. This observation is also confirmed by the results obtained from a recently introduced method called TRIBE which applies fusions of the ADAR2 deaminase domain and RNA-binding proteins (RBPs) to identify RBP targets by A-to-I editing.¹⁸⁵ The editing reaction has been reported to take place preferentially within dsRNA regions near the binding site of the fusion protein. Furthermore, the hyperactive E/Q mutation in the deaminase domain increased the sensitivity of the TRIBE method as more edited sites and more edited RBP targets were identified.¹⁸⁶ With regard to site-directed editing approaches, possible off-target edits within dsRNA regions in proximity of a gRNA/target duplex need to be considered, especially when using systems with the hyperactive ADAR deaminase domain. To avoid such off-target edits, the application of 2'-OMe/LNA mixmer-based ASOs might be helpful since they have been shown to be potent inhibitors of RNA editing by remodeling the RNA secondary structure.¹⁸⁷ 4λN-ADAR2 has been reported to edit several off-site adenosines outside of a duplex formed by the gRNA and an over-expressed target mRNA with efficiencies reaching that of the target adenosine.¹⁵⁸ The off-target editing was even more pronounced when using the hyperactive 4λN-ADAR2Q variant. It could be shown that the editing of the off-site

adenosines can be reduced by lowering the gRNA amount. However, reducing the gRNA amount similarly impaired the on-target editing. gRNA-dependent off-target editing outside the gRNA/target RNA duplex was also found within targeted endogenous *Mecp2* mRNA when applying the λ N-ADAR approach.¹⁵⁹ This type of off-target edits has also been shown to be generated by the dCas13-ADAR approach.¹⁴⁵

Endogenous ADAR function might be influenced by the applied gRNA. Using R/G-gRNAs for directing wild-type ADAR enzymes might prevent the editing at important sites in the transcriptome. Similarly, wild-type ADAR enzymes might also recognize the dsRNA duplex formed by the target RNA and the gRNA used for the recruitment of engineered editing machines. Such a duplex potentially sequesters wild-type ADAR enzymes and therefore might lead to decreased levels of editing within endogenous substrates. It has been shown that dsRNA with a minimum length of ~ 20 bp can be recognized as editing substrates, whereby dsRNA with a length < 40 bp are edited with less than 10% efficiency.⁵⁹ BG-gRNAs directing SA enzymes have a typical length of 22 nt (**Man. 4, Suppl. Table 4**) and form an RNA duplex of 19 bp with the target RNA. A former study tested the recruitment of transiently or genomically expressed wild-type ADAR2 by such a gRNA targeting *eGFP W58W*.¹⁴⁸ Since the editing yields stayed $< 10\%$, the gRNA seems to insufficiently recruit wild-type ADAR enzymes. The same result was obtained for a gRNA which led to the formation of a 21 bp RNA duplex. As indicated above, NGS analysis revealed that applying BG-gRNA does not reduce naturally occurring A-to-I edits in the transcriptome of 293 Flp-In T-REx cells (**Man. 4, Fig. 2b**). The gRNAs used for the dCas13-ADAR approach form a 50 bp duplex with the target RNA.¹⁴⁵ The recognition of such a duplex by wild-type ADAR enzymes was tested by T. Merkle (T. Stafforst group). For the experiments, a Cas13-gRNA was designed to target *eGFP W58X* mRNA. The plasmid containing the gRNA sequence was transfected together with *wild-type ADAR2* and *eGFP W58X* into HEK 293T cells according to the protocol applied in the original study by Cox *et al.*¹⁴⁵ It could be shown that the Cas13-gRNA indeed directed ADAR2 to the target site which was edited with $\sim 25\%$ efficiency similar to that obtained by dCas13b-ADAR2Q ($\sim 30\%$; **Man. 4, Suppl. Fig. 11**). The gRNA even recruited SA2Q which also achieved 25% editing. Therefore, the results indicate that, under the applied conditions, the ADAR2 deaminase domain alone is sufficient to be recruited by Cas13-gRNAs which lack specificity for dCas13b-ADAR2Q. Cas13-gRNAs might also be able to recruit endogenously expressed ADAR enzymes, potentially leading to a reduction of naturally occurring A-to-I edits in the transcriptome.

This might also be true for 2boxB-gRNAs (~ 85 nt) directing λ N-ADAR2Q enzymes since they form gRNA/target RNA duplexes with similar length (~ 50 bp).^{158,159,184} Despite of the large size of the MS2-gRNA used to steer MCP-ADAR1 (~ 400 nt), it forms smaller gRNA/target RNA duplexes (21 bp) as compared to Cas13-gRNAs or 2boxB-gRNAs.¹⁴⁷ Therefore, a substantial recruitment of wild-type ADARs by such a gRNA seems rather unlikely.

4 Conclusion

Six approaches (SA, dCas13-ADAR, λ N-ADAR, MCP-ADAR, wild-type ADAR; see **Introduction, Fig. 6**) have been developed enabling specific A-to-I substitutions in target RNAs. So far, the approaches applying SA, dCas13-ADAR or λ N-ADAR enzymes have been best characterized with respect to editing efficiency and specificity. Endogenous transcripts were highly efficiently edited by SAQ enzymes which achieved 50-90% editing within 11 out of the 16 adenosine-containing 5'-NAN triplets. The wild-type SA enzymes allowed decent editing within preferred triplets, especially in the 3'-UTR. Editing of endogenous transcripts by dCas13b-ADAR2Q has only been tested within highly preferred 5'-UAG triplets and enabled yields up to 45%. dCas13b-ADAR2Q edited *KRAS* mRNA with less efficiency (up to 2.2-fold) as compared to SA1Q.¹⁴⁵ NLS₃- λ N-ADAR2Q achieved a promising yield of ~ 70% when editing a 5'-CAA triplet within endogenous *Mecp2* mRNA.¹⁵⁹ However, more work is needed to determine the efficiency of the system in editing various endogenous substrates and triplets. Highly precise RNA editing was provided by the wild-type SA enzymes whose genes were integrated into the cell genome as a single copy. The more efficient SAQ variants showed some off-target activity, but there is evidence that reducing the amounts of the SAQ enzymes can inhibit off-target editing while maintaining high on-target editing. In contrast, dCas13b-ADAR2Q and (NLS₃)-4 λ N-ADAR2Q were transiently over-expressed and produced one order of magnitude more off-target edits in the transcriptome.^{145,184} Wild-type (NLS₃)-4 λ N-ADAR2 enzymes were even two orders of magnitude less specific than wild-type SA enzymes. The Cas13-ADAR and λ N-ADAR approaches struggled with off-target edits within a duplex formed by the gRNA and the target RNA.^{145,158,159} This specificity problem is fixed in the SA approach due to the chemical modification of the BG-gRNAs, even when A-rich triplets are targeted. The dCas13b-ADAR2Q T375G reach a similar specificity as provided by the wild-type SA enzymes.¹⁴⁵ However, it remains unclear to which extent dCas13b-ADAR2Q T375G can be used to efficiently catalyze site-directed RNA editing of transcripts since the T375G mutation within the ADAR2 deaminase domain seems to substantially decrease the editing efficiency of an editase to a level which is lower than that achieved by a wild-type SA enzyme. When considering the achieved editing yields and the extent of produced off-target edits together, the SA approach provides the best balance between editing efficiency and specificity compared to the dCas13-ADAR and λ N-ADAR approaches. Therefore, the SA approach, which

relies on the unique assembly mechanism of gRNA-deaminase complexes, sets a new benchmark for site-directed RNA editing. The BG-gRNA/SA editing system has been well characterized regarding maximum editing yields, editing duration, the scope editable triplets and gRNA potency. The results obtained from the studies using SA enzymes to correct disease-causing mutations, to manipulate signaling networks and to switch protein localization by light are promising for future applications of the SA approach in medicine and the life sciences.

References

- 1 Boccaletto, P. *et al.* MODOMICS: A database of RNA modification pathways. 2017 update. *Nucleic Acids Res.* **46**, D303-D307 (2018).
- 2 Nishikura, K. A-to-I editing of coding and non-coding RNAs by ADARs. *Nat. Rev. Mol. Cell Biol.* **17**, 83-96 (2016).
- 3 Walkley, C. R. & Li, J. B. Rewriting the transcriptome: Adenosine-to-inosine RNA editing by ADARs. *Genome Biol.* **18**, 205 (2017).
- 4 Ramaswami, G. & Li, J. B. Identification of human RNA editing sites: A historical perspective. *Methods* **107**, 42-47 (2016).
- 5 Levanon, E. Y. *et al.* Systematic identification of abundant A-to-I editing sites in the human transcriptome. *Nat. Biotech.* **22**, 1001-1005 (2004).
- 6 Li, J. B. *et al.* Genome-wide identification of human RNA editing sites by parallel DNA capturing and sequencing. *Science* **324**, 1210 (2009).
- 7 Sakurai, M., Yano, T., Kawabata, H., Ueda, H. & Suzuki, T. Inosine cyanoethylation identifies A-to-I RNA editing sites in the human transcriptome. *Nat. Chem. Biol.* **6**, 733-740 (2010).
- 8 Ramaswami, G. *et al.* Accurate identification of human Alu and non-Alu RNA editing sites. *Nat. Methods* **9**, 579-581 (2012).
- 9 Peng, Z. *et al.* Comprehensive analysis of RNA-Seq data reveals extensive RNA editing in a human transcriptome. *Nat. Biotech.* **30**, 253-260 (2012).
- 10 Ramaswami, G. *et al.* Identifying RNA editing sites using RNA sequencing data alone. *Nat. Methods* **10**, 128-132 (2013).
- 11 Bazak, L. *et al.* A-to-I RNA editing occurs at over a hundred million genomic sites, located in a majority of human genes. *Genome Res.* **24**, 365-376 (2014).
- 12 Zhang, Q. & Xiao, X. Genome sequence-independent identification of RNA editing sites. *Nat. Methods* **12**, 347-350 (2015).
- 13 Picardi, E. *et al.* Profiling RNA editing in human tissues: Towards the inosinome Atlas. *Sci. Rep.* **5**, 14941 (2015).
- 14 Hwang, T. *et al.* Dynamic regulation of RNA editing in human brain development and disease. *Nat. Neurosci.* **19**, 1093-1099 (2016).
- 15 Tan, M. H. *et al.* Dynamic landscape and regulation of RNA editing in mammals. *Nature* **550**, 249 (2017).
- 16 Patterson, J. B. & Samuel, C. E. 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 (1995).
- 17 George, C. X. & Samuel, C. E. Human RNA-specific adenosine deaminase ADAR1 transcripts possess alternative exon 1 structures that initiate from different promoters, one constitutively active and the other interferon inducible. *Proc. Natl. Acad. Sci. USA* **96**, 4621-4626 (1999).
- 18 Strehblow, A., Hallegger, M. & Jantsch, M. F. Nucleocytoplasmic distribution of human RNA-editing enzyme ADAR1 is modulated by double-stranded RNA-binding domains, a leucine-rich export signal, and a putative dimerization domain. *Mol. Biol. Cell* **13**, 3822-3835 (2002).
- 19 Fritz, J. *et al.* RNA-regulated interaction of transportin-1 and exportin-5 with the double-stranded RNA-binding domain regulates nucleocytoplasmic shuttling of ADAR1. *Mol. Cell. Biol.* **29**, 1487-1497 (2009).
- 20 Melcher, T. *et al.* A mammalian RNA editing enzyme. *Nature* **379**, 460-464 (1996)

- 21 Desterro, J. M. P. *et al.* Dynamic association of RNA-editing enzymes with the nucleolus. *J. Cell Sci.* **116**, 1805-1818 (2003).
- 22 Sansam, C. L., Wells, K. S. & Emeson, R. B. Modulation of RNA editing by functional nucleolar sequestration of ADAR2. *Proc. Natl. Acad. Sci. USA* **100**, 14018-14023 (2003).
- 23 Chen, C.-X. *et al.* A third member of the RNA-specific adenosine deaminase gene family, ADAR3, contains both single- and double-stranded RNA binding domains. *RNA* **6**, 755-767 (2000).
- 24 Maas, S. & Gommans, W. M. Identification of a selective nuclear import signal in adenosine deaminases acting on RNA. *Nucleic Acids Res.* **37**, 5822-5829 (2009).
- 25 Kim, U., Wang, Y., Sanford, T., Zeng, Y. & Nishikura, K. Molecular cloning of cDNA for double-stranded RNA adenosine deaminase, a candidate enzyme for nuclear RNA editing. *Proc. Natl. Acad. Sci. USA* **91**, 11457-11461 (1994).
- 26 Higuchi, M. *et al.* RNA editing of AMPA receptor subunit GluR-B: A base-paired intron-exon structure determines position and efficiency. *Cell* **75**, 1361-1370 (1993).
- 27 Sommer, B., Kohler, M., Sprengel, R. & Seeburg, P. H. RNA editing in brain controls a determinant of ion flow in glutamate-gated channels. *Cell* **67**, 11-19 (1991).
- 28 Higuchi, M. *et al.* Point mutation in an AMPA receptor gene rescues lethality in mice deficient in the RNA-editing enzyme ADAR2. *Nature* **406**, 78-81 (2000).
- 29 Kawahara, Y. *et al.* Glutamate receptors: RNA editing and death of motor neurons. *Nature* **427**, 801-801 (2004).
- 30 Hideyama, T. *et al.* Induced loss of ADAR2 engenders slow death of motor neurons from Q/R site-unedited GluR2. *J. Neurosci.* **30**, 11917 (2010).
- 31 Lomeli, H. *et al.* Control of kinetic properties of AMPA receptor channels by nuclear RNA editing. *Science* **266**, 1709 (1994).
- 32 Burns, C. M. *et al.* Regulation of serotonin-2C receptor G-protein coupling by RNA editing. *Nature* **387**, 303-308 (1997).
- 33 Niswender, C., Sanders-Bush, E. & Emeson, R. Identification and characterization of RNA editing events within the 5-HT_{2C} receptor. *Ann. N. Y. Acad. Sci.* **861**, 38-48 (1998).
- 34 Bhalla, T., Rosenthal, J. J. C., Holmgren, M. & Reenan, R. Control of human potassium channel inactivation by editing of a small mRNA hairpin. *Nat. Struct. Mol. Biol.* **11**, 950-956 (2004).
- 35 International Human Genome Sequencing Consortium. Initial sequencing and analysis of the human genome. *Nature* **409**, 860-921 (2001).
- 36 Samuel, C. E. Adenosine deaminases acting on RNA (ADARs) are both antiviral and proviral. *Virology* **411**, 180-193 (2011).
- 37 Mattick, J. S. & Mehler, M. F. RNA editing, DNA recoding and the evolution of human cognition. *Trends Neurosci.* **31**, 227-233 (2008).
- 38 Paz-Yaacov, N. *et al.* Adenosine-to-inosine RNA editing shapes transcriptome diversity in primates. *Proc. Natl. Acad. Sci. USA* **107**, 12174-12179 (2010).
- 39 Li, J. B. & Church, G. M. Deciphering the functions and regulation of brain-enriched A-to-I RNA editing. *Nat. Neurosci.* **16**, 1518-1522 (2013).
- 40 Chung, H. *et al.* Human ADAR1 prevents endogenous RNA from triggering translational shutdown. *Cell* **172**, 811-824 (2018).
- 41 Rice, G. I. *et al.* Mutations in ADAR1 cause Aicardi-Goutieres syndrome associated with a type I interferon signature. *Nat. Genet.* **44**, 1243-1248 (2012).

- 42 Wu, J. & Chen, Z. J. Innate immune sensing and signaling of cytosolic nucleic acids. *Annu. Rev. Immunol.* **32**, 461-488 (2014).
- 43 Wang, Q. *et al.* Stress-induced apoptosis associated with null mutation of ADAR1 RNA editing deaminase gene. *J. Biol. Chem.* **279**, 4952-4961 (2004).
- 44 Hartner, J. C. *et al.* Liver disintegration in the mouse embryo caused by deficiency in the RNA-editing enzyme ADAR1. *J. Biol. Chem.* **279**, 4894-4902 (2004).
- 45 Mannion, Niamh M. *et al.* The RNA-editing enzyme ADAR1 controls innate immune responses to RNA. *Cell Rep.* **9**, 1482-1494 (2014).
- 46 Pestal, K. *et al.* Isoforms of RNA-editing enzyme ADAR1 independently control nucleic acid sensor MDA5-driven autoimmunity and multi-organ development. *Immunity* **43**, 933-944 (2015).
- 47 Ahmad, S. *et al.* Breaching self-tolerance to Alu duplex RNA underlies MDA5-mediated inflammation. *Cell* **172**, 797-810 (2018).
- 48 Liddicoat, B. J. *et al.* RNA editing by ADAR1 prevents MDA5 sensing of endogenous dsRNA as nonself. *Science* **349**, 1115-1120 (2015).
- 49 Sakurai, M. *et al.* ADAR1 controls apoptosis of stressed cells by inhibiting Staufen1-mediated mRNA decay. *Nat. Struct. Mol. Biol.* **24**, 534-543 (2017).
- 50 Ota, H. *et al.* ADAR1 forms a complex with Dicer to promote microRNA processing and RNA-induced gene silencing. *Cell* **153**, 575-589 (2013).
- 51 Kawahara, Y. *et al.* Frequency and fate of microRNA editing in human brain. *Nucleic Acids Res.* **36**, 5270-5280 (2008).
- 52 Kawahara, Y., Zinshteyn, B., Chendrimada, T. P., Shiekhattar, R. & Nishikura, K. RNA editing of the microRNA-151 precursor blocks cleavage by the Dicer-TRBP complex. *EMBO Rep.* **8**, 763-769 (2007).
- 53 Kawahara, Y. *et al.* Redirection of silencing targets by adenosine-to-inosine editing of miRNAs. *Science* **315**, 1137-1140 (2007).
- 54 Borchert, G. M. *et al.* Adenosine deamination in human transcripts generates novel microRNA binding sites. *Hum. Mol. Genet.* **18**, 4801-4807 (2009).
- 55 Wang, Q. *et al.* ADAR1 regulates ARHGAP26 gene expression through RNA editing by disrupting miR-30b-3p and miR-573 binding. *RNA* **19**, 1525-1536 (2013).
- 56 Nakano, M. *et al.* RNA editing modulates human hepatic aryl hydrocarbon receptor expression by creating microRNA recognition sequence. *J. Biol. Chem.* **291**, 894-903 (2016).
- 57 Nakano, M., Fukami, T., Gotoh, S. & Nakajima, M. A-to-I RNA Editing Up-regulates Human Dihydrofolate Reductase in Breast Cancer. *J. Biol. Chem.* **292**, 4873-4884 (2017).
- 58 Oakes, E., Anderson, A., Cohen-Gadol, A. & Hundley, H. A. Adenosine deaminase that acts on RNA 3 (ADAR3) binding to glutamate receptor subunit B pre-mRNA inhibits RNA editing in glioblastoma. *J. Biol. Chem.* **292**, 4326-4335 (2017).
- 59 Nishikura, K. *et al.* Substrate specificity of the dsRNA unwinding/modifying activity. *EMBO J.* **10**, 3523-3532 (1991).
- 60 Polson, A. G. & Bass, B. L. Preferential selection of adenosines for modification by double-stranded RNA adenosine deaminase. *EMBO J.* **13**, 5701-5711 (1994).
- 61 Lehmann, K. A. & Bass, B. L. The importance of internal loops within RNA substrates of ADAR1. *J. Mol. Biol.* **291**, 1-13 (1999).
- 62 Ohman, M., Källman, A. M. & Bass, B. L. In vitro analysis of the binding of ADAR2 to the pre-mRNA encoding the GluR-B R/G site. *RNA* **6**, 687-697 (2000).

- 63 Daniel, C., Venø, M. T., Ekdahl, Y., Kjems, J. & Öhman, M. A distant cis acting intronic element induces site-selective RNA editing. *Nucleic Acids Res.* **40**, 9876-9886 (2012).
- 64 Rieder, L. E., Staber, C. J., Hoopengardner, B. & Reenan, R. A. Tertiary structural elements determine the extent and specificity of messenger RNA editing. *Nat. Commun.* **4**, 2232 (2013).
- 65 Daniel, C., Silberberg, G., Behm, M. & Ohman, M. Alu elements shape the primate transcriptome by cis-regulation of RNA editing. *Genome Biol.* **15**, R28 (2014).
- 66 Daniel, C., Widmark, A., Rigardt, D. & Öhman, M. Editing inducer elements increases A-to-I editing efficiency in the mammalian transcriptome. *Genome Biol.* **18**, 195 (2017).
- 67 Eggington, J. M., Greene, T. & Bass, B. L. Predicting sites of ADAR editing in double-stranded RNA. *Nat. Commun.* **2**, 319 (2011).
- 68 Wong, S. K., Sato, S. & Lazinski, D. W. Substrate recognition by ADAR1 and ADAR2. *RNA* **7**, 846-858 (2001).
- 69 Maas, S. *et al.* Structural requirements for RNA editing in glutamate receptor pre-mRNAs by recombinant double-stranded RNA adenosine deaminase. *J. Biol. Chem.* **271**, 12221-12226 (1996).
- 70 Rueter, S. M., Dawson, T. R. & Emeson, R. B. Regulation of alternative splicing by RNA editing. *Nature* **399**, 75-80 (1999).
- 71 Feng, Y., Sansam, C. L., Singh, M. & Emeson, R. B. Altered RNA editing in mice lacking ADAR2 autoregulation. *Mol. Cell. Biol.* **26**, 480-488 (2006).
- 72 Stulić, M. & Jantsch, M. F. Spatio-temporal profiling of filamin A RNA-editing reveals ADAR preferences and high editing levels outside neuronal tissues. *RNA Biol.* **10**, 1611-1617 (2013).
- 73 Huntley, M. A. *et al.* Complex regulation of ADAR-mediated RNA-editing across tissues. *BMC Genomics* **17**, 61 (2016).
- 74 Garnarcz, W., Tariq, A., Handl, C., Pusch, O. & Jantsch, M. F. A high-throughput screen to identify enhancers of ADAR-mediated RNA-editing. *RNA Biol.* **10**, 192-204 (2013).
- 75 Tariq, A. *et al.* RNA-interacting proteins act as site-specific repressors of ADAR2-mediated RNA editing and fluctuate upon neuronal stimulation. *Nucleic Acids Res.* **41**, 2581-2593 (2013).
- 76 Vitali, P. *et al.* ADAR2-mediated editing of RNA substrates in the nucleolus is inhibited by C/D small nucleolar RNAs. *J. Cell Biol.* **169**, 745-753 (2005).
- 77 Shoshan, E. *et al.* Reduced adenosine-to-inosine miR-455-5p editing promotes melanoma growth and metastasis. *Nat. Cell Biol.* **17**, 311-321 (2015).
- 78 Peng, P. L. *et al.* ADAR2-dependent RNA editing of AMPA receptor subunit GluR2 determines vulnerability of neurons in forebrain ischemia. *Neuron* **49**, 719-733 (2006).
- 79 Marcucci, R. *et al.* Pin1 and WWP2 regulate GluR2 Q/R site RNA editing by ADAR2 with opposing effects. *EMBO J.* **30**, 4211 (2011).
- 80 Terajima, H. *et al.* ADARB1 catalyzes circadian A-to-I editing and regulates RNA rhythm. *Nat. Genet.* **49**, 146-151 (2017).
- 81 Macbeth, M. R. *et al.* Inositol hexakisphosphate is bound in the ADAR2 core and required for RNA editing. *Science* **309**, 1534-1539 (2005).
- 82 Lai, F., Drakas, R. & Nishikura, K. Mutagenic analysis of double-stranded RNA adenosine deaminase, a candidate enzyme for RNA editing of glutamate-gated ion channel transcripts. *J. Biol. Chem.* **270**, 17098-17105 (1995).

- 83 Cho, D.-S. C. *et al.* Requirement of dimerization for RNA editing activity of adenosine deaminases acting on RNA. *J. Biol. Chem.* **278**, 17093-17102 (2003).
- 84 Stefl, R. *et al.* The solution structure of the ADAR2 dsRBM-RNA complex reveals a sequence-specific readout of the minor groove. *Cell* **143**, 225-237 (2010).
- 85 Kuttan, A. & Bass, B. L. Mechanistic insights into editing-site specificity of ADARs. *Proc. Natl. Acad. Sci. USA* **109**, E3295-E3304 (2012).
- 86 Phelps, K. J. *et al.* Recognition of duplex RNA by the deaminase domain of the RNA editing enzyme ADAR2. *Nucleic Acids Res.* **43**, 1123-1132 (2015).
- 87 Wang, Y., Havel, J. & Beal, P. A. A phenotypic screen for functional mutants of human adenosine deaminase acting on RNA 1. *ACS Chem. Biol.* **10**, 2512-2519 (2015).
- 88 Matthews, M. M. *et al.* Structures of human ADAR2 bound to dsRNA reveal base-flipping mechanism and basis for site selectivity. *Nat. Struct. Mol. Biol.* **23**, 426-433 (2016).
- 89 Haudenschild, B. L. *et al.* A transition state analogue for an RNA-editing reaction. *J. Am. Chem. Soc.* **126**, 11213-11219 (2004).
- 90 Wang, Y. & Beal, P. A. Probing RNA recognition by human ADAR2 using a high-throughput mutagenesis method. *Nucleic Acids Res.* **44**, 9872-9880 (2016).
- 91 Thomas, J. M. & Beal, P. A. How do ADARs bind RNA? New protein-RNA structures illuminate substrate recognition by the RNA editing ADARs. *Bioessays* **39**, doi: 10.1002/bies.201600187 (2017).
- 92 Burnett, John C. & Rossi, John J. RNA-based therapeutics: Current progress and future prospects. *Chem. Biol.* **19**, 60-71 (2012).
- 93 Kole, R., Krainer, A. R. & Altman, S. RNA therapeutics: Beyond RNA interference and antisense oligonucleotides. *Nat. Rev. Drug Discov.* **11**, 125-140 (2012).
- 94 Bennett, C. F., Baker, B. F., Pham, N., Swayze, E. & Geary, R. S. Pharmacology of antisense drugs. *Annu. Rev. Pharmacol. Toxicol.* **57**, 81-105 (2017).
- 95 Janssen, H. L. A. *et al.* Treatment of HCV infection by targeting microRNA. *New Engl. J. Med.* **368**, 1685-1694 (2013).
- 96 Hong, D. *et al.* AZD9150, a next-generation antisense oligonucleotide inhibitor of STAT3 with early evidence of clinical activity in lymphoma and lung cancer. *Sci. Transl. Med.* **7**, 314ra185 (2015).
- 97 Koizumi, M. 2'-O,4'-ethylene-bridged nucleic acids (ENA) as next-generation antisense and antigene agents. *Biol. Pharm. Bull.* **27**, 453-456 (2004).
- 98 Prakash, T. P. *et al.* Targeted delivery of antisense oligonucleotides to hepatocytes using triantennary N-acetyl galactosamine improves potency 10-fold in mice. *Nucleic Acids Res.* **42**, 8796-8807 (2014).
- 99 Nair, J. K. *et al.* Multivalent N-acetylgalactosamine-conjugated siRNA localizes in hepatocytes and elicits robust RNAi-mediated gene silencing. *J. Am. Chem. Soc.* **136**, 16958-16961 (2014).
- 100 Wu, H. *et al.* Determination of the role of the human RNase H1 in the pharmacology of DNA-like antisense drugs. *J. Biol. Chem.* **279**, 17181-17189 (2004).
- 101 Vitravene Study Group. A randomized controlled clinical trial of intravitreal fomivirsen for treatment of newly diagnosed peripheral cytomegalovirus retinitis in patients with AIDS. *Am. J. Ophthalmol.* **133**, 467-474 (2002).
- 102 Vitravene Study Group. Randomized dose-comparison studies of intravitreal fomivirsen for treatment of cytomegalovirus retinitis that has reactivated or is

- persistently active despite other therapies in patients with AIDS. *Am. J. Ophthalmol.* **133**, 475-483 (2002).
- 103 Vitravene Study Group. Safety of intravitreal fomivirsen for treatment of cytomegalovirus retinitis in patients with AIDS. *Am. J. Ophthalmol.* **133**, 484-498 (2002).
- 104 Raal, F. J. *et al.* Mipomersen, an apolipoprotein B synthesis inhibitor, for lowering of LDL cholesterol concentrations in patients with homozygous familial hypercholesterolaemia: A randomised, double-blind, placebo-controlled trial. *Lancet* **375**, 998-1006 (2010).
- 105 Stein, E. A. *et al.* Apolipoprotein B synthesis inhibition with mipomersen in heterozygous familial hypercholesterolemia: Results of a randomized, double-blind, placebo-controlled trial to assess efficacy and safety as add-on therapy in patients with coronary artery disease. *Circulation* **126**, 2283-2292 (2012).
- 106 Santos, R. D. *et al.* Long-term efficacy and safety of mipomersen in patients with familial hypercholesterolaemia: 2-year interim results of an open-label extension. *Eur. Heart J.* **36**, 566-575 (2015).
- 107 Viney, N. J. *et al.* Antisense oligonucleotides targeting apolipoprotein(a) in people with raised lipoprotein(a): Two randomised, double-blind, placebo-controlled, dose-ranging trials. *Lancet* **388**, 2239-2253 (2016).
- 108 Lam, J. K. W., Chow, M. Y. T., Zhang, Y. & Leung, S. W. S. siRNA versus miRNA as therapeutics for gene silencing. *Mol. Ther. Nucleic Acids* **4**, e252 (2015).
- 109 Wittrup, A. & Lieberman, J. Knocking down disease: a progress report on siRNA therapeutics. *Nat. Rev. Genet.* **16**, 543-552 (2015).
- 110 Deleavey, G. F. & Damha, M. J. Designing chemically modified oligonucleotides for targeted gene silencing. *Chem. Biol.* **19**, 937-954 (2012).
- 111 Coelho, T. *et al.* Safety and efficacy of RNAi therapy for transthyretin amyloidosis. *New Engl. J. Med.* **369**, 819-829 (2013).
- 112 Suhr, O. B. *et al.* Efficacy and safety of patisiran for familial amyloidotic polyneuropathy: A phase II multi-dose study. *Orphanet J. Rare Dis.* **10**, 109 (2015).
- 113 APOLLO: The study of an investigational drug, Patisiran (ALN-TTR02), for the treatment of transthyretin (TTR)-mediated amyloidosis (NCT01960348), <https://clinicaltrials.gov> (access on 15.03.2018).
- 114 Liang, X.-H., Sun, H., Nichols, J. G. & Crooke, S. T. RNase H1-dependent antisense oligonucleotides are robustly active in directing RNA cleavage in both the cytoplasm and the nucleus. *Mol. Ther.* **25**, 2075-2092 (2017).
- 115 Bobbin, M. L. & Rossi, J. J. RNA Interference (RNAi)-based therapeutics: Delivering on the promise? *Annu. Rev. Pharmacol. Toxicol.* **56**, 103-122 (2016).
- 116 Disterer, P. *et al.* Development of therapeutic splice-switching oligonucleotides. *Hum. Gene Ther.* **25**, 587-598 (2014).
- 117 Lim, K. R. Q., Maruyama, R. & Yokota, T. Eteplirsen in the treatment of Duchenne muscular dystrophy. *Drug Des. Devel. Ther.* **11**, 533-545 (2017).
- 118 Bladen, C. L. *et al.* The TREAT-NMD DMD global database: Analysis of more than 7,000 Duchenne muscular dystrophy mutations. *Hum. Mutat.* **36**, 395-402 (2015).
- 119 Lorson, C. L., Hahnen, E., Androphy, E. J. & Wirth, B. A single nucleotide in the SMN gene regulates splicing and is responsible for spinal muscular atrophy. *Proc. Natl. Acad. Sci. USA* **96**, 6307-6311 (1999).

- 120 Finkel, R. S. *et al.* Treatment of infantile-onset spinal muscular atrophy with nusinersen: A phase 2, open-label, dose-escalation study. *Lancet* **388**, 3017-3026 (2016).
- 121 Ceccaldi, R., Rondinelli, B. & D'Andrea, A. D. Repair pathway choices and consequences at the double-strand break. *Trends Cell Biol.* **26**, 52-64 (2016).
- 122 Smith, J. *et al.* A combinatorial approach to create artificial homing endonucleases cleaving chosen sequences. *Nucleic Acids Res.* **34**, e149-e149 (2006).
- 123 Maeder, M. L. & Gersbach, C. A. Genome-editing technologies for gene and cell therapy. *Mol. Ther.* **24**, 430-446 (2016).
- 124 Kim, Y. G., Cha, J. & Chandrasegaran, S. Hybrid restriction enzymes: Zinc finger fusions to Fok I cleavage domain. *Proc. Natl. Acad. Sci. USA* **93**, 1156-1160 (1996).
- 125 Mak, A. N.-S., Bradley, P., Cernadas, R. A., Bogdanove, A. J. & Stoddard, B. L. The crystal structure of TAL effector PthXo1 bound to its DNA target. *Science* **335**, 716 (2012).
- 126 Deng, D. *et al.* Structural basis for sequence-specific recognition of DNA by TAL effectors. *Science* **335**, 720 (2012).
- 127 Miller, J. C. *et al.* A TALE nuclease architecture for efficient genome editing. *Nat. Biotechnol.* **29**, 143 (2010).
- 128 van der Oost, J., Westra, E. R., Jackson, R. N. & Wiedenheft, B. Unravelling the structural and mechanistic basis of CRISPR-Cas systems. *Nat. Rev. Microbiol.* **12**, 479 (2014).
- 129 Makarova, K. S. *et al.* An updated evolutionary classification of CRISPR-Cas systems. *Nat. Rev. Microbiol.* **13**, 722-736 (2015).
- 130 Jinek, M. *et al.* A Programmable dual-RNA-guided DNA endonuclease in adaptive bacterial immunity. *Science* **337**, 816 (2012).
- 131 Anders, C., Niewoehner, O., Duerst, A. & Jinek, M. Structural basis of PAM-dependent target DNA recognition by the Cas9 endonuclease. *Nature* **513**, 569 (2014).
- 132 Ran, F. A. *et al.* Genome engineering using the CRISPR-Cas9 system. *Nat. Protoc.* **8**, 2281 (2013).
- 133 Komor, A. C., Badran, A. H. & Liu, D. R. CRISPR-based technologies for the manipulation of eukaryotic genomes. *Cell* **168**, 20-36.
- 134 Hess, G. T., Tycko, J., Yao, D. & Bassik, M. C. Methods and applications of CRISPR-mediated base editing in eukaryotic genomes. *Mol. Cell* **68**, 26-43 (2017).
- 135 Gaudelli, N. M. *et al.* Programmable base editing of A•T to G•C in genomic DNA without DNA cleavage. *Nature* **551**, 464-471 (2017).
- 136 Ruiz de Galarreta, M. & Lujambio, A. Therapeutic editing of hepatocyte genome in vivo. *J. Hepatol.* **67**, 818-828 (2017).
- 137 Kim, D. *et al.* Genome-wide target specificities of CRISPR RNA-guided programmable deaminases. *Nat. Biotechnol.* **35**, 475 (2017).
- 138 Woolf, T. M., Chase, J. M. & Stinchcomb, D. T. Toward the therapeutic editing of mutated RNA sequences. *Proc. Natl. Acad. Sci. USA* **92**, 8298-8302 (1995).
- 139 Pegg, A. E. Repair of O6-alkylguanine by alkyltransferases. *Mutat. Res.* **462**, 83-100 (2000).
- 140 Keppler, A. *et al.* A general method for the covalent labeling of fusion proteins with small molecules in vivo. *Nat. Biotech.* **21**, 86-89 (2003).

- 141 Gronemeyer, T., Chidley, C., Juillerat, A., Heinis, C. & Johnsson, K. Directed evolution of O6-alkylguanine-DNA alkyltransferase for applications in protein labeling. *Protein Eng. Des. Sel.* **19**, 309-316 (2006).
- 142 Stafforst, T. & Schneider, M. F. An RNA-deaminase conjugate selectively repairs point mutations. *Angew. Chem. Int. Ed.* **51**, 11166-11169 (2012).
- 143 Schneider, M. F., Wettengel, J., Hoffmann, P. C. & Stafforst, T. Optimal guideRNAs for re-directing deaminase activity of hADAR1 and hADAR2 in trans. *Nucleic Acids Res.* **42**, e87 (2014).
- 144 Smargon, A. A. *et al.* Cas13b is a type VI-B CRISPR-associated RNA-guided RNase differentially regulated by accessory proteins Csx27 and Csx28. *Mol. Cell* **65**, 618-630.e617 (2017).
- 145 Cox, D. B. T. *et al.* RNA editing with CRISPR-Cas13. *Science* **358**, 1019-1027 (2017).
- 146 Montiel-Gonzalez, M. F., Vallecillo-Viejo, I., Yudowski, G. A. & Rosenthal, J. J. C. Correction of mutations within the cystic fibrosis transmembrane conductance regulator by site-directed RNA editing. *Proc. Natl. Acad. Sci. USA* **110**, 18285-18290 (2013).
- 147 Azad, M. T. A., Bhakta, S. & Tsukahara, T. Site-directed RNA editing by adenosine deaminase acting on RNA for correction of the genetic code in gene therapy. *Gene Ther.* **24**, 779 (2017).
- 148 Wettengel, J., Reautschnig, P., Geisler, S., Kahle, P. J. & Stafforst, T. Harnessing human ADAR2 for RNA repair - Recoding a PINK1 mutation rescues mitophagy. *Nucleic Acids Res.* **45**, 2797-2808 (2017).
- 149 Fukuda, M. *et al.* Construction of a guide-RNA for site-directed RNA mutagenesis utilising intracellular A-to-I RNA editing. *Sci. Rep.* **7**, 41478 (2017).
- 150 Heep, M., Mach, P., Reautschnig, P., Wettengel, J. & Stafforst, T. Applying human ADAR1p110 and ADAR1p150 for site-directed RNA editing - G/C Substitution Stabilizes GuideRNAs against Editing. *Genes* **8**, 34 (2017).
- 151 Bosley, K. S. *et al.* CRISPR germline engineering - The community speaks. *Nat. Biotechnol.* **33**, 478 (2015).
- 152 Krützfeldt, J. *et al.* Silencing of microRNAs in vivo with 'antagomirs'. *Nature* **438**, 685-689 (2005).
- 153 Monia, B. P. *et al.* Evaluation of 2'-modified oligonucleotides containing 2'-deoxy gaps as antisense inhibitors of gene expression. *J. Biol. Chem.* **268**, 14514-14522 (1993).
- 154 Prakash, T. P. *et al.* Positional effect of chemical modifications on short interference RNA activity in mammalian cells. *J. Med. Chem.* **48**, 4247-4253 (2005).
- 155 Whitehurst, A. W. *et al.* Synthetic lethal screen identification of chemosensitizer loci in cancer cells. *Nature* **446**, 815 (2007).
- 156 Fujita, S. *et al.* Highly efficient reverse transfection with siRNA in multiple wells of microtiter plates. *J. Biosci. Bioeng.* **104**, 329-333 (2007).
- 157 Kim, D.-H. *et al.* Synthetic dsRNA Dicer substrates enhance RNAi potency and efficacy. *Nat. Biotechnol.* **23**, 222 (2004).
- 158 Montiel-González, M. F., Vallecillo-Viejo, I. C. & Rosenthal, Joshua J. C. An efficient system for selectively altering genetic information within mRNAs. *Nucleic Acids Res.* **44**, e157-e157 (2016).
- 159 Sinnamon, J. R. *et al.* Site-directed RNA repair of endogenous Mecp2 RNA in neurons. *Proc. Natl. Acad. Sci. USA* **114**, E9395-E9402 (2017).
- 160 Khan, S. & Dickerman, J. D. Hereditary thrombophilia. *Thromb. J.* **4**, 15 (2006).

- 161 Kotterman, M. A., Chalberg, T. W. & Schaffer, D. V. Viral vectors for gene therapy: Translational and clinical outlook. *Annu. Rev. Biomed. Eng.* **17**, 63-89 (2015).
- 162 Akinc, A. *et al.* A combinatorial library of lipid-like materials for delivery of RNAi therapeutics. *Nat. Biotechnol.* **26**, 561 (2008).
- 163 Boisguérin, P. *et al.* Delivery of therapeutic oligonucleotides with cell penetrating peptides. *Adv. Drug Del. Rev.* **87**, 52-67 (2015).
- 164 Simanshu, D. K., Nissley, D. V. & McCormick, F. RAS proteins and their regulators in human disease. *Cell* **170**, 17-33 (2017).
- 165 Miklossy, G., Hilliard, T. S. & Turkson, J. Therapeutic modulators of STAT signalling for human diseases. *Nat. Rev. Drug Discov.* **12**, 611 (2013).
- 166 Lusic, H. & Deiters, A. A New photocaging group for aromatic N-heterocycles. *Synthesis* **2006**, 2147-2150 (2006).
- 167 Cautain, B., Hill, R., de Pedro, N. & Link, W. Components and regulation of nuclear transport processes. *FEBS J.* **282**, 445-462 (2015).
- 168 Kozak, M. An analysis of 5'-noncoding sequences from 699 vertebrate messenger RNAs. *Nucleic Acids Res.* **15**, 8125-8148 (1987).
- 169 Kozak, M. Context effects and inefficient initiation at non-AUG codons in eucaryotic cell-free translation systems. *Mol. Cell. Biol.* **9**, 5073-5080 (1989).
- 170 Feil, S., Valtcheva, N. & Feil, R. Inducible Cre mice. *Methods Mol. Biol.* **530**, 343-363 (2009).
- 171 Engelke, H., Chou, C., Uprety, R., Jess, P. & Deiters, A. Control of protein function through optochemical translocation. *ACS Synth. Biol.* **3**, 731-736 (2014).
- 172 Niopek, D. *et al.* Engineering light-inducible nuclear localization signals for precise spatiotemporal control of protein dynamics in living cells. *Nat. Commun.* **5**, 4404 (2014).
- 173 Konermann, S. *et al.* Optical control of mammalian endogenous transcription and epigenetic states. *Nature* **500**, 472 (2013).
- 174 Polstein, L. R. & Gersbach, C. A. A light-inducible CRISPR-Cas9 system for control of endogenous gene activation. *Nat. Chem. Biol.* **11**, 198 (2015).
- 175 Nihongaki, Y., Yamamoto, S., Kawano, F., Suzuki, H. & Sato, M. CRISPR-Cas9-based photoactivatable transcription system. *Chem. Biol.* **22**, 169-174 (2015).
- 176 Nihongaki, Y. *et al.* CRISPR-Cas9-based photoactivatable transcription systems to induce neuronal differentiation. *Nat. Methods* **14**, 963 (2017).
- 177 Gautier, A. *et al.* Genetically encoded photocontrol of protein localization in mammalian cells. *J. Am. Chem. Soc.* **132**, 4086-4088 (2010).
- 178 Inlay, M. A. *et al.* Synthesis of a photocaged tamoxifen for light-dependent activation of Cre-ER recombinase-driven gene modification. *Chem. Commun.* **49**, 4971-4973 (2013).
- 179 Park, E. & Rapoport, T. A. Mechanisms of Sec61/SecY-mediated protein translocation across membranes. *Annu. Rev. Biophys.* **41**, 21-40 (2012).
- 180 Coloma, M. J., Hastings, A., Wims, L. A. & Morrison, S. L. Novel vectors for the expression of antibody molecules using variable regions generated by polymerase chain reaction. *J. Immunol. Methods* **152**, 89-104 (1992).
- 181 Gallo, A. & Locatelli, F. ADARs: allies or enemies? The importance of A-to-I RNA editing in human disease: From cancer to HIV-1. *Biol. Rev.* **87**, 95-110 (2012).
- 182 Bajad, P., Jantsch, M. F., Keegan, L. & O'Connell, M. A to I editing in disease is not fake news. *RNA Biol.* **14**, 1223-1231 (2017).

- 183 Ramaswami, G. & Li, J. B. RADAR: A rigorously annotated database of A-to-I
RNA editing. *Nucleic Acids Res.* **42**, D109-D113 (2014).
- 184 Vallecillo-Viejo, I. C., Liscovitch-Brauer, N., Montiel-Gonzalez, M. F.,
Eisenberg, E. & Rosenthal, J. J. C. Abundant off-target edits from site-directed
RNA editing can be reduced by nuclear localization of the editing enzyme. *RNA
Biol.* **15**, 104-114 (2018).
- 185 McMahon, Aoife C. *et al.* TRIBE: Hijacking an RNA-editing enzyme to identify
cell-specific targets of RNA-binding proteins. *Cell* **165**, 742-753 (2016).
- 186 Xu, W., Rahman, R. & Rosbash, M. Mechanistic implications of enhanced editing
by a hyperTRIBE RNA-binding protein. *RNA* **24**, 173-182 (2018).
- 187 Mizrahi, R. A., Schirle, N. T. & Beal, P. A. Potent and selective inhibition of A-
to-I RNA editing with 2'-O-methyl/locked nucleic acid-containing antisense
oligoribonucleotides. *ACS Chem. Biol.* **8**, 832-839 (2013).

Curriculum vitae

PERSONAL DETAILS

- Name: Paul Vogel
- Date of birth: 22.09.1987
- Place of birth: Dresden

EDUCATION

- PhD in Biochemistry, University of Tuebingen** since 09/2013
- M.Sc. in Chemistry, Technical University Dresden** 10/2010 - 01/2013
- B.Sc. in Chemistry, Technical University Dresden** 10/2007 - 09/2010

WORK EXPERIENCE

- PhD Thesis, Interfaculty Institute of Biochemistry, University of Tuebingen** since 09/2013
 - Research group: Thorsten Stafforst - Chemical biology of nucleic acids
 - Title: Establishing site-directed A-to-I RNA editing in cell culture
- Master's Thesis, Institute of Radiopharmaceutical Cancer Research, Helmholtz-Zentrum Dresden-Rossendorf** 03/2012 - 01/2013
 - Research group: Jens Pietzsch - Radiopharmaceutical and chemical biology
 - Title: The impact of the EphA2 receptor on the radiosensitivity of melanoma cells
- Bachelor's Thesis, Institute of Radiopharmaceutical Cancer Research, Helmholtz-Zentrum Dresden-Rossendorf** 06/2010 - 09/2010
 - Research group: Jens Pietzsch - Radiopharmaceutical and chemical biology
 - Title: COX-2 inhibitors as potential radiosensitizers

Conference contributions

Talk

Vogel, P., Moschref, M., Hanswillemenke, A., Li, Q., Li, J. B. & Stafforst, T. SNAP-tagged ADARs - A tool for precise and efficient RNA editing. **GASB I Conference**, Marburg, Germany, November 24-25, 2017.

Poster presentations

Vogel, P., Moschref, M., Li, Q., Li, J. B. & Stafforst, T. Efficient and precise RNA editing of endogenous transcripts with SNAP-tagged ADARs. **1st Symposium on Nucleic Acid Modifications**, Mainz, Germany, September 4-6, 2017.

Vogel, P., Hanswillemenke, A. & Stafforst, T. SNAP-tagged deaminases enable highly efficient site-directed RNA editing. **Gordon Research Conference on RNA Editing**, Ventura, USA, March 12-17, 2017.

Vogel, P., Hanswillemenke, A., Bhatt, N. & Stafforst, T. Engineering artificial riboproteins for site-directed RNA editing. **67. Mosbacher Kolloquium**, Mosbach, Germany, March 31-April 2, 2016.

Vogel, P., Hanswillemenke, A., Bhatt, N. & Stafforst, T. Changing protein localization by site-directed RNA editing. **VII. Nukleinsäurechemietreffen**, Berlin, Germany, September 17-18, 2015.

Vogel, P., Schneider, M. F., Wettengel, J. & Stafforst, T. Site-directed RNA editing as a tool for medicine and basic research. **Gordon Research Conference on RNA Editing**, Lucca, Italy, March 8-13, 2015.

Manuscripts

Man. 1: accepted

Vogel, P., Schneider, M. F., Wettengel, J. & Stafforst, T. Improving site-directed RNA editing in vitro and in cell culture by chemical modification of the guideRNA. *Angew. Chem. Int. Ed.* **53**, 6267-6271 (2014).

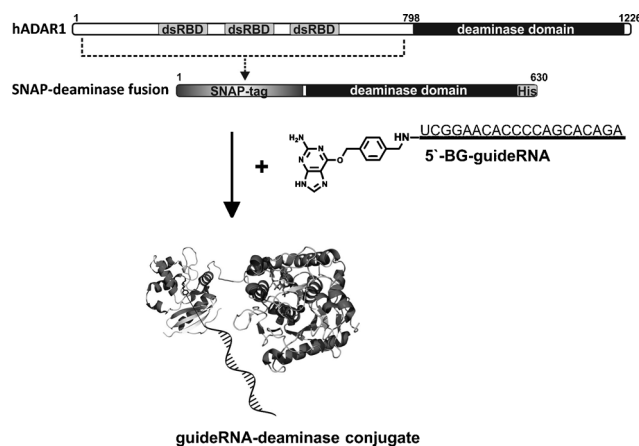
Improving Site-Directed RNA Editing In Vitro and in Cell Culture by Chemical Modification of the GuideRNA**

Paul Vogel, Marius F. Schneider, Jacqueline Wettengel, and Thorsten Stafforst*

Abstract: Adenosine-to-inosine deamination can be re-addressed to user-defined mRNAs by applying phosphothioate/2'-methoxy-modified guideRNAs. Dense chemical modification of the guideRNA clearly improves performance of the covalent conjugates inside the living cell. Furthermore, careful positioning of a few modifications controls editing selectivity in vitro and was exploited for the challenging repair of the Factor 5 Leiden missense mutation.

RNA editing has the power to reprogram genetic information on the RNA level.^[1] The outcome depends on the site at which a single adenosine-to-inosine (A-to-I) conversion occurs. If it happens in the open reading frame, the substitution of a single amino acid residue results; if it happens in an untranslated region, RNA processing is altered. Thus, directing RNA editing activity to a user-defined site of a (pre)-mRNA makes it possible to manipulate RNA and protein function with high potential for application in basic biology research and medicine (transcript repair). We recently reported a strategy for the assembly of an artificial, guideRNA-dependent RNA editing machinery that allows the application of simple Watson–Crick binding rules for the site-selective and highly rational targeting of any arbitrary codon.^[2] Here, we report several improvements of the tool including a proof-of-principle for its cellular application.

To direct RNA editing, we have engineered the protein-guided human ADAR1 (adenosine deaminase acting on RNA) into a guideRNA-dependent enzyme by fusing its C-terminal catalytic domain to a SNAP-tag (Scheme 1).^[2] The SNAP-tag^[3] enables formation of fully defined one-to-one conjugates with guideRNAs that carry a 5'-terminal O₆-benzylguanine (BG) modification. The RNA part of such a tool fulfills two tasks. First, it steers the deaminase domain to the target site at a user-defined mRNA, and second, it forms the secondary structure required for the highly efficient and selective editing of a single adenosine residue without overediting of nearby off-target adenosines. Most appealing is the modular nature of our approach which allows us to



Scheme 1. Engineering of a guideRNA-dependent deaminase. The N-terminal RNA substrate binding protein domains (dsRBD) are substituted with a SNAP-tag, which allows for the formation of defined 1-to-1 covalent conjugates with short guideRNAs 5'-terminally modified with O-benzylguanine (BG). The guideRNA steers editing activity towards new mRNA substrates.

program the machinery to target virtually any given codon by designing a respective guideRNA.

Single nucleotide polymorphisms (SNP) can have a profound effect on the processing of an RNA transcript or the function of the derived protein product.^[4] Hence, many SNP are directly linked to diseases. By reprogramming adenosine formally into guanosine, site-directed A-to-I RNA editing has the potential to either model, repair, or attenuate genetic disease phenotypes.^[5] However, for application in research or therapy, the RNA–protein conjugate has to become effective and specific inside the cell. The direct transduction of the protein–RNA conjugate could be difficult.^[6] An alternative would be the expression of the editing machinery in a genetically engineered animal or tissue culture. Whereas the SNAP-ADAR fusion is genetically encodable, the guideRNA strictly requires the chemical modification with the BG group for its proper functioning and thus is genetically not encodable.^[3] Hence, with respect to in vivo application, we got interested in the stabilization of the guideRNA component by chemical modification.

Antisense oligomers are often globally modified at the ribose backbone.^[7] A typical example is the antagomir^[8] that contains global 2'-O-methyl groups and terminal phosphothioate groups in combination with a single cholesterol modification. Such modifications make the probe resistant against nucleases; thus, long-lasting microRNA knockdown of more than one week is typically obtained by a single transfection.^[8] The increased lipophilicity supports their

[*] P. Vogel, M. F. Schneider, J. Wettengel, Dr. T. Stafforst
Interfaculty Institute of Biochemistry, University of Tübingen
Auf der Morgenstelle 15, 72076 Tübingen (Germany)
E-mail: thorsten.stafforst@uni-tuebingen.de
Homepage: <http://www.ifib.uni-tuebingen.de/research/stafforst.html>

[**] We thank Prof. Dr. Alfred Nordheim for assistance with cell culture and fluorescence microscopy. We thank the DFG (STA 1053/3-1), the Fonds der Chemischen Industrie and the University of Tübingen for generous financial support.



Supporting information for this article is available on the WWW under <http://dx.doi.org/10.1002/anie.201402634>.

penetration into the cell membrane and also facilitates their trafficking between cyto- and nucleoplasm.^[9] The cholesterol modification facilitates the receptor-mediated uptake in various tissues without the need for additional transfection agents.^[8] Thus, the action of such probes has been reported after systemic administration to mice with remarkably low toxicity and immunogenicity.^[10]

The situation is more complex for RNA probes that address a catalytically active protein complex such as the antisense strand of a siRNA after loading into the RNA-induced silencing complex (RISC). Since such probes have to fulfill their tasks inside the active site of the respective protein complex they are less receptive for modification.^[11] However, it was demonstrated that chemical modification at selected sites^[12] can improve pharmacokinetics, target selectivity,^[13] and also toxicity and immunogenicity.^[14] Various modifications including 2'-*O*-methyl, 2'-fluorine, and LNA have been reported.^[11,14]

If ADAR's deaminase domain would accept chemically modified guideRNAs as substrates for editing mRNAs, an important hurdle towards *in vivo* application would be overcome. To not entirely block editing, we had to carefully optimize the site and degree of modification. As modifications we chose 2'-*O*-methylation and terminal phosphothioate since they are particularly simple to synthesize and commercially affordable, and have proven their utility in various applications.^[7–15] To test the acceptance of 2'-*O*-methyl groups in BG-guideRNA-dependent A-to-I RNA editing, we systematically studied the effect of such modifications in our *in vitro* assay^[2] on a nonsense Stop66 eCFP mRNA with our optimized 17nt guideRNA. As one may expect, editing of the eCFP mRNA substrate was fully inhibited by a protein-RNA conjugate containing a guideRNA that carried three 2'-methoxy groups centered at the counter base of the targeted adenosine (Figure 1B versus 1A). It was shown before that densely modified nucleic analogues inhibit RNA editing.^[16] However, modifying only the counter base of the targeted adenosine with a single 2'-methoxy group gave not complete but rather roughly 50% inhibition (Figure 1C). Notably, a single modification one nucleotide up- or downstream from the counter base had nearly no or only very little influence on the overall editing yield (Figure 1D,E). Thus, full inhibition can only be achieved by cooperative action of several proximate 2'-modifications. We then tested a guideRNA that was densely modified with 2'-methoxy groups leaving only a little gap of three natural ribonucleotides centered around the counter base, and that further carried two phosphothioate modifications at the 5'-terminus and four at the 3'-terminus, as antagomirs^[8] typically do. Even though this 17nt duplex was heavily chemically modified on the guide strand, a very good editing yield of $\geq 80\%$ was obtained with SNAP-ADAR1 (Figure 1F). Thus, antagomir-like guideRNAs are well accepted by the deaminase domain of hADAR1 if they contain a minimal gap (3nt) of unmodified RNA at the target site. Nucleic acid dependent enzymes typically tolerate less extensive chemical modification on the guide. For instance, gapmers that redirect RNaseH activity require a gap of 7 to 10 unmodified deoxyribonucleotides,^[17] whereas the antisense strand of siRNAs can tolerate 2'-

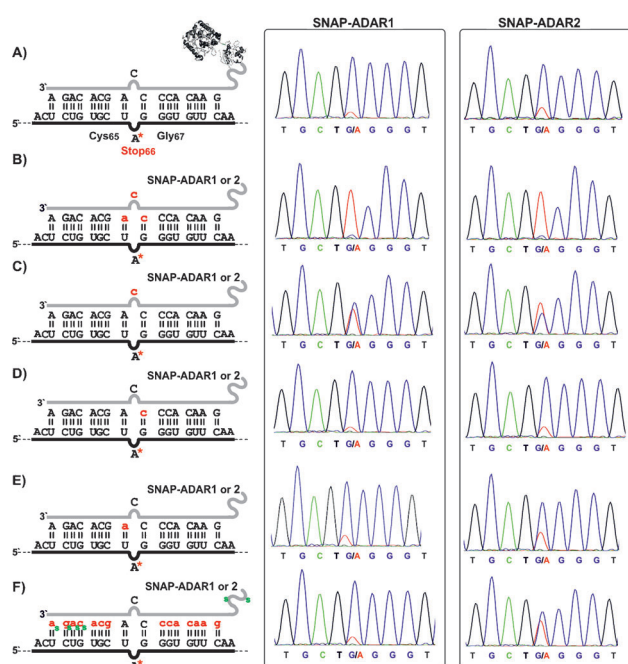


Figure 1. Effect of chemically modified guideRNAs on editing of the Stop66 codon in eCFP mRNA (mRNA black; guideRNA gray; the targeted adenosine is marked with an asterisk; 2'-OMe-modified bases are indicated by small red letters, phosphothioate linkages by green "s"). Editing yields are estimated from the areas for adenosine (red) versus guanosine (blue) in the respective sequencing traces with SNAP-ADAR1 (left) and SNAP-ADAR2 (right). Editing conditions: 3 h at 30/37 °C; 50 nM mRNA, 500 nM BG-guideRNA, 650 nM SNAP-ADAR1 or -ADAR2, 75 mM KCl, 25 mM Tris-HCl, 10 mM DTT, 0.75 mM MgCl₂, 2 μM heparin, pH 8.3). One full sequencing trace is given in the Supporting Information (Figure S3).

fluorine,^[12] but is disabled in the presence of global 2'-methoxylation.^[7,8,11] Thus applying a very limited number of modified bases is usually recommended.^[11] Our finding was unexpected but is highly important: in the future it may allow us to express SNAP-ADAR fusions inside an animal or a tissue and to then manipulate protein or RNA function under conditional control by administration of a chemically stabilized guideRNA.

To further illustrate that approach we transiently expressed SNAP_F-ADAR1 and the fluorogenic reporter gene under control of a CMV promoter in HEK 293T cells and stimulated transcript repair by transfection of the guideRNA. For this, SNAP_F-ADAR1 and Stop66 eCFP were subcloned into the pcDNA3.1 vector. Wild-type (Trp66) eCFP served as a positive control. Cotransfection of both plasmids in a 1-to-1 ratio was achieved with lipofectamine 2000 in a 24-well format. After 24 h, transfected cells were detached with trypsin, distributed evenly over several wells in a 96-well plate, and incubated for 24 h prior to lipofection with various guideRNAs (50 pMol in 150 μL/well). After one day of incubation, the fluorescence phenotype was analyzed by microscopy.

Whereas the positive control (Figure 2A) gave a clearly visible fluorescence signal, no fluorescence was observed when the nonfunctional Stop66 eCFP was transfected in the

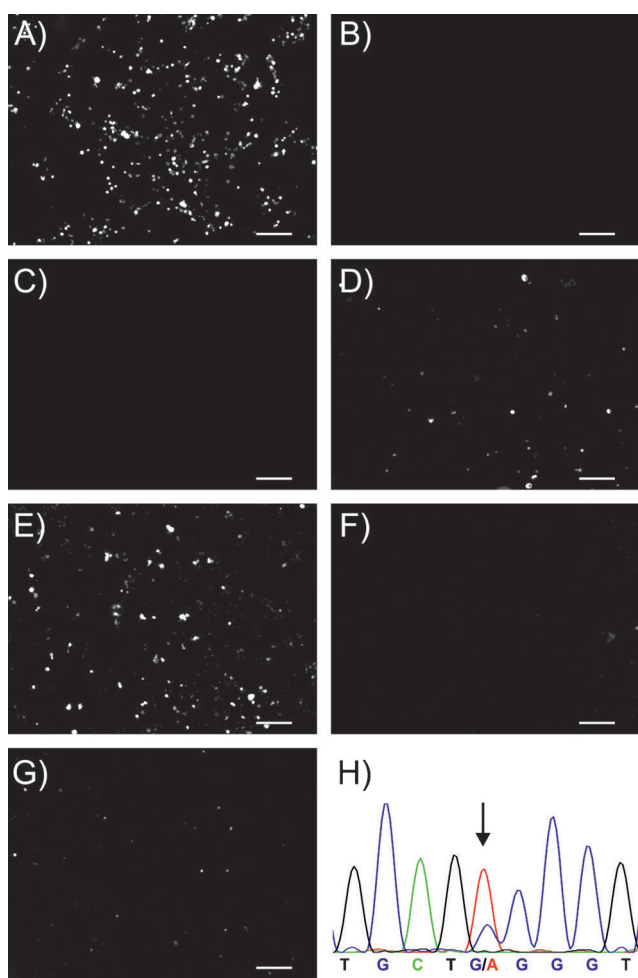


Figure 2. Directed RNA editing in 293T cells. Cells were cotransfected with equal amounts of reporter gene (Stop66 eCFP or wild-type) and SNAP_F-ADAR1 (or empty pcDNA3.1). After 24 h guideRNAs were transfected and 24 h later the fluorescence phenotype was analyzed by fluorescence microscopy. A) Positive control (wild-type eCFP with SNAP_F-ADAR1 and BG-antagomir-guideRNA); B) negative control 1 (Stop66 eCFP/empty pcDNA3.1/BG-antagomir-guideRNA); C) negative control 2 (Stop66 eCFP/SNAP_F-ADAR1/no guideRNA); D)–G) experiments with Stop66 eCFP/SNAP_F-ADAR1 and various guideRNAs: D) standard BG-guideRNA; E) BG-antagomir-guideRNA; F) standard NH₂-guideRNA; G) NH₂-antagomir-guideRNA. The scale bar represents 200 μm. H) Sanger sequencing of the RNA extracted from cells treated as in experiment (E). For details on cell culturing see the Supporting Information.

presence of either SNAP_F-ADAR1 or the BG-modified antagomir-like guideRNA alone (Figure 2B,C). For further controls see Figure S6. In contrast, CFP fluorescence was detectable when Stop66 eCFP was coexpressed with SNAP_F-ADAR1 and subsequently transfected with our standard BG-modified guideRNA lacking 2'-methoxy and phosphothioate modifications (Figure 2D). However, compared to the positive control, the number of bright fluorescing cells was low. In order to estimate the effect of further chemical modification we transfected Stop66 eCFP/SNAP_F-ADAR1-coexpressing cells with the BG-modified antagomir-like guideRNA. Compared to the standard guideRNA (Figure 2D), the chemically stabilized guideRNA displayed a markedly increased number

of fluorescent cells (Figure 2E), hence clearly demonstrating the expected beneficial effects of chemical modification in a cellular environment.

In vitro, the editing reaction benefits strongly from the covalent conjugation between guideRNA and deaminase. Applying 5'-amino (NH₂)-modified guideRNAs instead of the 5'-BG-modified guideRNAs typically results in a reduced or even abolished editing activity depending on the in vitro editing conditions.^[2] Thus we tested the cellular transcript repair in the presence of the corresponding NH₂-guideRNAs. As expected, no (Figure 2F), or only very little CFP fluorescence (Figure 2G) was restored by applying guideRNAs that cannot address the deaminase domain by means of covalent attachment. Apparently covalent conjugation is fast enough inside the cell and is required for efficient transcript repair. The conjugation kinetics of our applied SNAP_F-tag variant for BG-tagged moieties was reported to be $2.8 \times 10^4 \text{ M}^{-1} \text{ s}^{-1}$ and was previously shown to be sufficient for fluorophore conjugation in the cytosol of living mammalian cells.^[18]

To clearly demonstrate that the fluorescence phenotype was due to RNA editing at the targeted Stop66 codon, we extracted the total RNA from the cells, removed possible DNA contaminations with DNaseI, reversely transcribed the eCFP mRNA with a specific backward primer, and amplified the cDNA through Taq-PCR with CFP-specific primers. In agreement with the fluorescence imaging, Sanger sequencing revealed the highest editing yield for the antagomir-like BG-guideRNA (Figure 2H) and no detectable editing for the control experiment expressing SNAP_F-ADAR1 but lacking the guideRNA (Figure S7A). Natural editing enzymes achieve selective and up to quantitative editing, similar to our optimized in vitro directed editing approach. Since many parameters remain to be optimized for the cellular assay, including the length, sequence, concentration, modification, and transfection of the guideRNA, as well as the ratio and transfection of the reporter and SNAP-ADAR genes, we expect further improvements in the future.

As a last control, we transfected Stop66 eCFP/SNAP_F-ADAR1 co-expressing cells with a 2'-methoxy-phosphothioate modified BG-guideRNA that binds around Trp codon 58, thus 24 nt upstream of the targeted codon. Even though this BG-guideRNA is capable of repairing a Stop58 eGFP mutation there was no detectable repair of the Stop66 eCFP mutation (Figures S6H and S7F). This is in very accordance with our experimental experience. Re-directed RNA editing is highly selective and requires the positioning of the targeted adenosine in a well-defined secondary structure.^[2]

Chemical modification of the guideRNA may also improve substrate specificity in a cellular application, because random binding of the probe to a partially complementary sequence in the transcriptome may less often lead to unwanted off-site editing.^[15] Furthermore, chemical modification may not only improve the pharmacological properties including lifetime and off-site effects, but could also improve the editing selectivity. We demonstrate this for the repair of the Factor 5 Leiden polymorphism. This disease-causing single point mutation (G¹⁷⁴⁶→A) represents the most abun-

dant genetic risk factor in heritable multifactorial thrombophilia in the Caucasian population.^[19] Due to the point mutation, a single amino acid substitution (R534→Q) appears at the Protein C dependent proteolytic cleavage site (R533R534) of the blood coagulation factor F5. Whereas the heterozygous defect is accompanied by an only minor increase in thrombosis risk (ca. 8-fold), the homozygous defect has a much more pronounced effect (>80-fold increased risk).^[19] Directed RNA editing has the potential to compensate for this genetic defect by its repair at the RNA level. However, a look into the gene revealed a very adenosine-rich and thus highly challenging target site (Figure 3 A). The glutamine codon (5'-CAA) is known to be more difficult to activate than the amber stop codon.^[20, 2b]

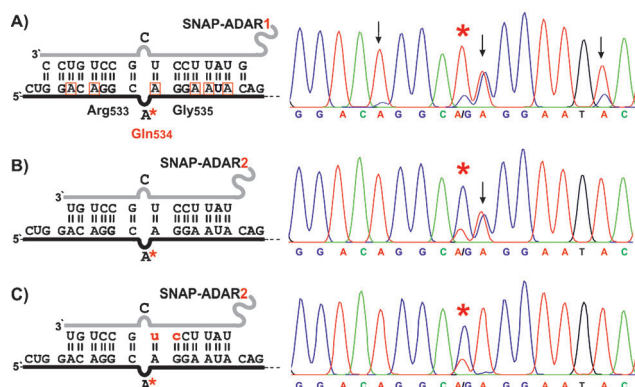


Figure 3. Repair of the F5 Leiden polymorphism. Potential off-target adenosines are highlighted with red boxes. Overedited sites are marked with black arrows, the targeted site with an asterisk. Editing conditions: 50 nM R534Q F5 mRNA, 200 nM BG-guideRNA, 350 nM SNAP-ADAR1/2; 3 h at 30/37 °C in 75 mM KCl, 25 mM Tris-HCl, 10 mM DTT, 0.70 mM MgCl₂, pH 8.3. An overview of the editing results for all guideRNAs with SNAP-ADAR1 and SNAP-ADAR2 is shown in the Supporting Information together with one complete sequencing trace (Figures S4 and S5).

The repair reaction was studied on a 1000 nt long piece of the F5 transcript containing the F5 Leiden mutation centrally. This was required since both the F5 pre-mRNA (> 70 kb) and the mature mRNA (7 kb) are too large^[19d] for our standard editing assay.^[2] We started the repair study with a 17 nt long guideRNA that puts the targeted adenosine into an A/C mismatch in the middle of the guideRNA/mRNA helix. Even though none of the 330 adenosines outside of the helix have been edited, we found massive overediting at neighboring off-target adenosines inside the guideRNA/mRNA duplex (Figure 3 A). Specifically, editing was observed at four sites, with the targeted site being barely activated ($\leq 20\%$ yield). The highest editing yield (ca. 50%) was obtained at the direct neighbor of the targeted base. Since ADAR2 is known to edit the Q/R site in the glutamate receptor with up to quantitative yield,^[1] we made use of our modular approach and changed the deaminase domain in our SNAP fusion protein from that of hADAR1 to that of hADAR2 (see the Supporting Information). Indeed, this helped to activate the 5'-CAA codon considerably (ca. 70% yield, Figure 3 B). Shortening

the guideRNA (from 17 to 14 nt) was sufficient to suppress overediting at the distal sites; however, the strong overediting (ca. 50%) at the adenosine directly neighboring the targeted base stayed unaffected (Figure 3 B). Our work on 2'-O-methylated guideRNAs (Figure 1 A–G) suggested that incorporation of 2'-methoxy groups around the off-targeted adenosine may provide a means to suppress overediting in this very delicate codon context. To avoid affecting the targeted adenosine, we placed only two modifications on the guideRNA, one opposite the off-site adenosine and one opposite its neighboring guanosine. Indeed, overediting was completely abolished without affecting editing at the target site for which the overall yield remained roughly 70% (Figure 3 C). Such repair yields would be more than sufficient to attenuate the disease phenotype.^[19] Thus chemical modification of the guideRNA is not only useful to increase nuclease resistance but provides a means to finetune editing selectivity.

In summary, we have demonstrated the strength and applicability of site-directed RNA editing as a rational approach for RNA repair. Chemical modification turned out to be a reliable means to suppress overreaction and steer selectivity. Importantly, we found that even massive chemical modification including global 2'-methoxy and terminal phosphothioate groups is well accepted as long as a small gap of three nonmodified ribonucleotides is maintained. Furthermore, we have demonstrated the functioning of the tool inside the cell. To our knowledge this is the first example of the assembly of an enzymatically active, covalent protein nucleic acid (analogue) conjugate inside the living cell and may represent an attractive strategy for the spatially or temporally controlled assembly of protein arrays in general.^[21] In contrast to competing strategies that apply the protein-RNA recognition by means of the MS2 or λ N phage system for assembly,^[22, 21] our guideRNAs are particularly short (20 nt) and lack additional BoxB RNA hairpin (19 nt) and linker (10–20 nt) motifs required for recognition and spacing that make other guideRNAs ≥ 60 nt long.^[22] Our stabilized 20 nt short guideRNAs are particularly suitable for transfection and are supposed to be too small to elicit an immune response.^[7–15] The necessity to transfect our guideRNAs instead of expressing them encourages us to think about introducing further chemical modifications that endow our probes with additional layers of control, such as photoactivation. Our findings dramatically improve the prospect of directed RNA editing for in vivo applications in basic biology research and medicine. Along those lines, we demonstrated the first repair of a disease-causing missense point mutation in vitro. This required the exchange of the deaminase domain of the SNAP fusion protein from hADAR1 to that of hADAR2 and the careful chemical modification of the guideRNA to target the codon in its highly challenging adenosine-rich sequence context. The high modularity of the SNAP-deaminases in combination with their useful tolerance for chemical modification highlight the potential of this approach for site-directed RNA repair.

Received: February 21, 2014

Keywords: antagomirs · genetic diseases · nucleic acid analogues · RNA editing · transcript repair

- [1] a) K. Nishikura, *Annu. Rev. Biochem.* **2010**, *79*, 321–349; b) B. L. Bass, *Annu. Rev. Biochem.* **2002**, *71*, 817–846; c) S. M. Rueter, C. M. Burns, S. A. Coode, P. Mookherjee, R. B. Emeson, *Science* **1995**, *267*, 1491–1494.
- [2] a) T. Stafforst, M. F. Schneider, *Angew. Chem.* **2012**, *124*, 11329–11332; *Angew. Chem. Int. Ed.* **2012**, *51*, 11166–11169; b) M. F. Schneider, J. Wettengel, P. C. Hoffmann, T. Stafforst, *Nucl. Acids Res.* **2014**, DOI: 10.1093/nar/gku272.
- [3] A. Keppler, S. Gendreizig, T. Gronemeyer, H. Pick, H. Vogel, K. Johnsson, *Nat. Biotechnol.* **2002**, *21*, 86–89.
- [4] The International SNP Map Working Group, *Nature* **2001**, *409*, 928–933.
- [5] a) T. W. Woolf, J. M. Chase, D. Stinchcomb, *Proc. Natl. Acad. Sci. USA* **1995**, *92*, 8298–8302; b) T. M. Woolf, *Nat. Biotechnol.* **1998**, *16*, 341.
- [6] a) M. Rapoport, H. Lorberboum-Galski, *Expert Opin. Drug Delivery* **2009**, *6*, 453–463; b) S. R. Schwarze, A. Ho, A. Vocero-Akbani, S. F. Dowdy, *Science* **1999**, *285*, 1569–1572; c) Y. Yang, N. Ballatori, H. C. Smith, *Mol. Pharmaceutics* **2002**, *61*, 269–276.
- [7] a) M. D. Horwich, P. D. Zamore, *Nat. Protoc.* **2008**, *3*, 1537; b) R. Kole, A. R. Krainer, S. Altman, *Nat. Rev. Drug Discovery* **2012**, *11*, 125–140.
- [8] J. Krützfeldt, N. Rajewsky, R. Braich, K. G. Rajeev, T. Tuschl, M. Manoharan, M. Stoffel, *Nature* **2005**, *438*, 685–689.
- [9] a) R. L. Juliano, X. Ming, O. Nakagawa, *Bioconjugate Chem.* **2012**, *23*, 147–157; b) P. Lorenz, T. Misteli, B. F. Baker, C. F. Bennett, D. L. Spector, *Nucleic Acids Res.* **2000**, *28*, 582–592.
- [10] a) A. Bonauer, G. Carmona, M. Iwasaki, M. Mione, M. Koyanagi, A. Fischer, J. Burchfield, H. Fox, C. Doebele, K. Ohtani, E. Chavakis, M. Potente, M. Tjwa, C. Urbich, A. M. Zeiher, S. Dimmeler, *Science* **2009**, *324*, 1710–1713; b) M. Trajkovski, J. Haussler, J. Soutschek, B. Bhat, A. Akin, M. Zavalan, M. H. Heim, M. Stoffel, *Nature* **2011**, *474*, 649–654; c) R. A. Boon, K. Iekushi, S. Lechner, T. Seeger, A. Fischer, S. Heydt, D. Kaluza, K. Tréguer, G. Carmona, A. Bonauer, A. J. G. Horrevoets, N. Didier, Z. Girmatsion, P. Biliczki, J. R. Ehrlich, H. A. Katus, O. J. Müller, M. Potente, A. M. Zeiher, H. Hermeking, S. Dimmeler, *Nature* **2013**, *416*, 107–111.
- [11] a) A. Birmingham, E. Anderson, K. Sullivan, A. Reynolds, Q. Boese, D. Leake, J. Karpilow, A. Khvorova, *Nat. Protoc.* **2007**, *2*, 2068–2078; b) J. B. Bramsen, M. B. Laursen, A. F. Nielsen, T. B. Hansen, C. Bus, N. Langkjaer, B. R. Babu, T. Hojland, M. Abramov, A. Van Aerschot, D. Odadzic, R. Smiccius, J. Haas, C. Andree, J. Barman, M. Wenska, P. Srivastava, C. Zhou, D. Honcharenko, S. Hess, E. Müller, G. V. Bobkov, S. N. Mikhailov, E. Fava, T. F. Meyer, J. Chattopadhyaya, M. Zerial, J. W. Engels, P. Herdewijn, J. Wengel, J. Kjems, *Nucleic Acids Res.* **2009**, *37*, 2867–2881.
- [12] T. P. Prakash, *J. Med. Chem.* **2005**, *48*, 4247–4253.
- [13] A. L. Jackson, J. Burchard, D. Leake, A. Reynolds, J. Schelter, J. Guo, J. M. Johnson, L. Lim, J. Karpilow, K. Nichols, W. Marshall, A. Khvorova, P. S. Linsley, *RNA* **2006**, *12*, 1197–1205.
- [14] a) V. Hornung, M. Guenther-Biller, C. Bourquin, A. Ablasser, M. Schlee, S. Uematsu, A. Noronha, M. Manoharan, S. Akira, A. de Fougerolles, S. Endres, G. Hartmann, *Nat. Med.* **2005**, *11*, 263–270; b) J. Soutschek, A. Akinc, B. Bramlage, K. Charisse, R. Constien, M. Donoghue, S. Elbashir, A. Geick, P. Hadwiger, J. Harborth, M. John, V. Kesavan, G. Lavine, R. K. Pandey, T. Racie, K. G. Rajeev, I. Rohl, I. Toudjarska, G. Wang, S. Wuschko, D. Bumcrot, V. Koteliansky, S. Limmer, M. Manoharan, H. P. Vornlocher, *Nature* **2004**, *432*, 173–178.
- [15] a) Y.-L. Chiu, T. M. Rana, *RNA* **2003**, *9*, 1034–1048; b) M. M. Fabani, M. J. Gait, *RNA* **2008**, *14*, 336–346; c) J. Elmén, M. Lindow, S. Schütz, M. Lawrence, A. Petri, S. Obad, M. Lindholm, M. Hedtjärn, H. F. Hansen, U. Berger, S. Gullans, P. Kearney, P. Sarnow, E. M. Straarup, S. Kauppinen, *Nature* **2008**, *452*, 896–900; d) J. Soutschek, *Nature* **2004**, *432*, 173–178; W. P. Kloosterman, A. K. Lagendijk, R. F. Ketting, J. D. Moulton, R. H. A. Plasterk, *PLoS Biol.* **2007**, *5*, 1738–1749; e) aktuelle Übersicht: N. M. Snead, J. J. Rossi, *Nucleic Acid Ther.* **2012**, *22*, 139.
- [16] R. A. Mizrahi, N. T. Schirle, P. A. Beal, *ACS Chem. Biol.* **2013**, *8*, 832–839.
- [17] J. E. Lee, C. F. Bennett, T. A. Cooper, *Proc. Natl. Acad. Sci. USA* **2012**, *109*, 4221–4226.
- [18] A. Gautier, A. Juillerat, C. Heinis, I. R. Correa, Jr., M. Kindermann, F. Beaufils, K. Johnsson, *Chem. Biol.* **2008**, *15*, 128–136.
- [19] a) S. Khan, J. D. Dickerman, *Thromb. J.* **2006**, *4*, 15–32; b) E. G. Bovill, S. J. Hasstedt, M. F. Leppert, G. L. Long, *Thromb. Haemostasis* **1999**, *82*, 662–666; c) B. Zöller, P. G. de Frutos, A. Hillarp, B. Dahlbäck, *Haematologica* **1999**, *84*, 59–70; d) J. L. Kujovich, *GeneReviews* <http://www.ncbi.nlm.nih.gov/books/NBK1368/>; e) D. C. Rees, M. Cox, J. B. Clegg, *Lancet* **1995**, *346*, 1133–1134.
- [20] J. M. Eggington, T. Greene, B. L. Bass, *Nat. Commun.* **2011**, *2*, DOI: 10.1038/ncomms1324.
- [21] a) C. J. Delebecque, A. B. Lindner, P. A. Silver, F. A. Aldaye, *Science* **2011**, *333*, 470–474; b) F. C. Simmel, *Curr. Opin. Biotechnol.* **2012**, *23*, 516–521; c) C. M. Agapakis, P. M. Boyle, P. A. Silver, *Nat. Chem. Biol.* **2012**, *8*, 527–535.
- [22] M. F. Montiel-Gonzales, I. Vallecillo-Viejo, G. A. Yudowski, J. C. Rosenthal, *Proc. Natl. Acad. Sci. USA* **2013**, DOI: 10.1073/pnas.1306243110.

Supporting Information

© Wiley-VCH 2014

69451 Weinheim, Germany

**Improving Site-Directed RNA Editing In Vitro and in Cell Culture by
Chemical Modification of the GuideRNA****

*Paul Vogel, Marius F. Schneider, Jacqueline Wettengel, and Thorsten Stafforst**

anie_201402634_sm_miscellaneous_information.pdf

Protein and gene constructs used in this study

One part of the work was done with SNAP-ADAR1, the same construct described in a previous publication (*Angew. Chem. Int Ed.* **2012**, *51*, 11166). Another part of the work was done with SNAP-ADAR2 as described below.

Cloning of SNAP-ADAR2-His6 into pRS426 under control of the Gal1-10 promoter

The gene of the deaminase domain (aa281-701) of human adar2 (BC065545, with a sequenced clone from a commercial cDNA library as template) was subcloned via the *Ascl* and *BamHI* restriction sites into the SNAP-ADAR1-His6 construct replacing the ADAR1 domain. The plasmid is based on the yeast/*E. coli* shuttle vector pRS426 (P. Hieter et al., *Gene* **1992**, *110*, 119.). The fusion protein is under control of a Gal1-10 promoter, adding a C-terminal 6xHis-tag. To allow the usage of *BamHI*, a natural *BamHI* site in the ADAR2 gene was disrupted by a silent point mutation. Thus the insert was cloned via overlap PCR. 5'-Site: fw primer: 5'-TAGGCGCGCC AGGGTCTGGC GCGCGCAGTA AGAAGCTTGC CAAGGCCCGG; bw primer: 5'-CTGAGCAGGG AACCCCTGGAT GCC; 3'-site: fw primer: 5'-GGCATCCAGG GTTCCCTGCT CAG; bw primer: 5'-GCGGATCCTA TTAATGGTGA TGGTGATGGT GGGGCGTGAG TGAGAACTGG TC; overlap with the 5'-site fw and the 3'-site bw primer. Phusion polymerase (NEB) was used for all cloning steps. PCR products were always cleaned by 1.2-1.5% agarose gel electrophoresis in 1fold TAE buffer. Ligation products were transformed into *xl1blue E.coli*. After minipreparation, plasmids were sequenced over the whole promoter and ORF.

Protein production and purification

The fusion protein was produced on a 1L scale in YVH10 (*Saccharomyces cerevisiae*, K. D. Wittrup et al., *Nat. Biotech.* **1998**, *16*, 773) very similar to a well described literature protocol (B. L. Bass et al., *Meth. Enzym.* **2007**, *424*, 319). For this the SNAP-ADAR2-His6 gene in pRS426 was transformed into chemically competent YVH10 using the Frozen-EZ Yeast Transformation II Kit (Zymo Research) according to the recommendations of the manufacturer. Cells were grown in SD-CAA+W for 4 days at 25°C, 270 rpm in culture flasks and switched to SG-CAA+W for induction. After 5 days at 20°C, 270 rpm, cells were harvested, lysed with a French press (20 000 Psi, 3 runs) and clarified from the debris by centrifugation (40 000 g, 1h). The lysate was subjected to pre-equilibrated Ni-NTA gel via a 0.4 μ M PES-sterile syringe filter. The Ni-NTA gel was washed with 10 mM imidazole, the protein was eluted into 15 ml 400 mM imidazole, 100 mM NaCl, 20 mM Tris-HCl pH 8.0, 5% glycerol. The protein containing fractions were subjected to a 1 ml HiTrap Heparin column (GE-Healthcare) equilibrated with 100 mM NaCl, 20 mM Tris-HCl pH 8.0, 5% glycerol. The protein was manually eluted by step-wise increasing salt concentration. The protein was typically found between 250 and 350 mM NaCl and was already >90% clean as judged from the SDS-PAGE (Fig. S1). The presence of the SNAP-tag was confirmed by staining the protein with O-benzylguanine-modified fluoresceine (Fig. S1). Since yeast is not containing an endogenous editing machinery that could interfere with our construct, no further purification was done. The protein was concentrated to a final volume of 250 μ l with a 15 ml 10kDa MWCO amicon centrifugal filter and was changed to 150 mM NaCl without changing the other buffer conditions. The concentration of the protein solution was estimated by UV-spectroscopy with extinction coefficients of 600 $\text{cm}^{-1} \text{mM}^{-1}$ (230 nm), 85 $\text{cm}^{-1} \text{mM}^{-1}$ (260 nm), 120 $\text{cm}^{-1} \text{mM}^{-1}$ (280 nm) prior to addition of DTT. The protein stock solution was filled up with 86% glycerol to a total concentration of 20% glycerol and DTT was added to a final concentration of 2 mM, and was stored at -20°C for short-term and -80°C for long-term. A typical 1 L production resulted in about 10-15 nmol (600-1000 μ g) >90% pure protein. Protein properties: 622 aa (190 aa SNAP-tag, 420 aa deaminase); $pI \approx 9.4$; MW = 67.8 kDa.

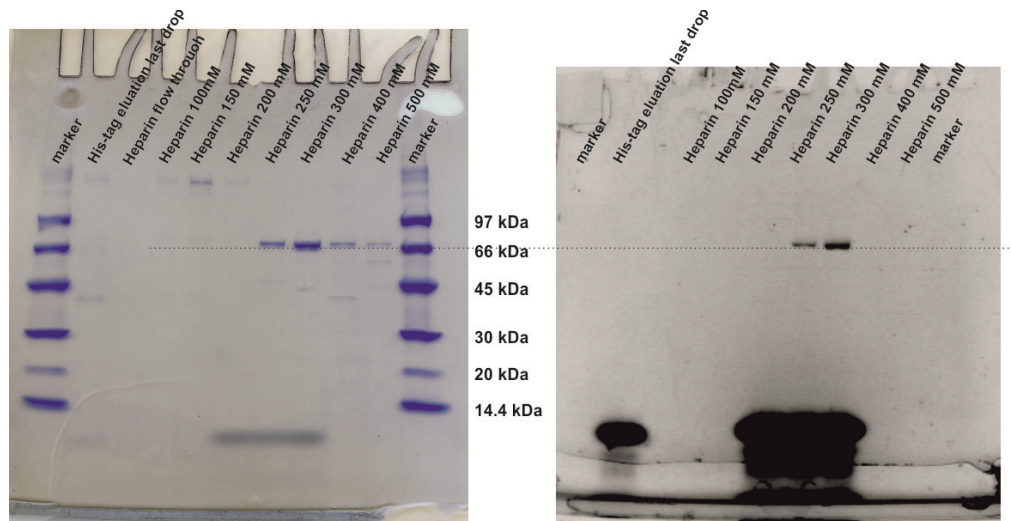
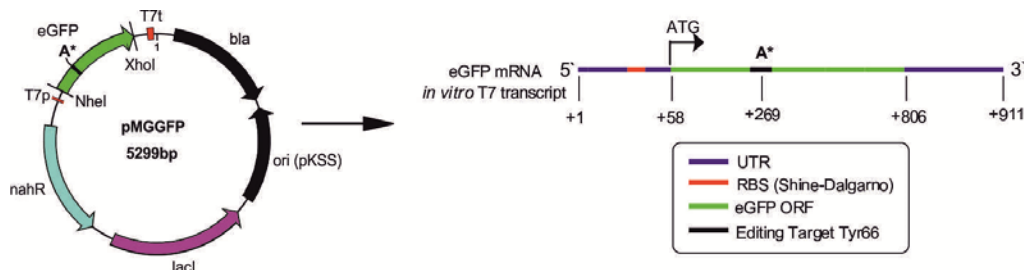


Figure S1. SDS-Page of SNAP-ADAR2 protein production after purification on heparin. Left: Coumassie blue stain, right FITC-BG staining. Shown are the fractions of the step-wise heparin column elution with 100-500 mM NaCl. Protein samples were incubated in presence of ca. 1 μ M *O*-benzylguanine-fluoresceine for highly selective fluorescence staining of SNAP-tagged proteins. The 250 mM and 300 mM fractions were unified and concentrated and used without further purification. The expected size of the protein (68 kDa) fits well to the LMW protein marker (GE Healthcare).

Editing substrates

Editing was investigated on mRNA substrates of Stop66 eCFP and R⁵³⁴Q F5 mRNA. The Stop66 eCFP gene was subcloned into the pMG211 *E. coli* vector (M. Gamper et al., *FEBS J.* **2005**, 272, 375.) and was produced from XI1Blue with a Miniprep kit (Macherey&Nagel).



Scheme S1. pMG211 plasmids carrying the Stop66 eCFP gene used in this study.

Stop66 eCFP

The repair of the Stop66 nonsense mutation to functional tryptophan was investigated on eCFP. As compared to eGFP, eCFP contains the typical mutations N146I, M153T, V163A, and N164H which are required for the proper folding and maturation of the tryptophan-containing chromophore (Tsien, et al., *Current Biology* **6**, 178-82 (1996)). A gene containing the non-functional Stop66 codon (UAG) at position of Trp66 (UGG) is the starting material for the editing-mediated repair. Point mutations were introduced by overlap-extension PCR via the XhoI/NheI restriction sites. The product genes were approved by sequencing over the complete ORF.

Stop66 eCFP protein sequence

```

1      MASKGEELFT  GVVPILVELD
21     GDVNGHKFSV  SGEGEGDATY
41     GKLTLKFICT  TGKLPVPWPT
61     LVTTLC*GVQ  CFSRYPDHMK
81     RHDFFKSAMP  EGYVQERTIF
101    FKDDGNYKTR  AEVKFEGDTL
121    VNRIELKGID  FKEDGNILGH
141    KLEYNYSISHN  VYITADKQKN
161    GIKAHFKTRH  NIEDGVSQLA
181    DHYQQNTPIG  DGPVLLPDNH
201    YLSTQSALSK  DPNEKRDHNV
221    LLEFVTAAGI  THGMDELYKS
241    GGSMALE

```

Stop66 eCFP gene sequence

```

1      ATGGCTAGCA  AAGGAGAAGA  ACTCTTCACT  GGAGTTGTCC  CAATTCTTGT  TGAATTAGAT
61     GGTGATGTTA  ACGGCCACAA  GTTCTCTGTC  AGTGGAGAGG  GTGAAGGTGA  TGCAACATAC
121    GGAAACTTTA  CCCTGAAGTT  CATCTGCACT  ACTGGCAAAC  TGCCTGTTCC  GTGGCCAACA
181    CTAGTCACTA  CTCTGTGCTA  GGGTGTTCAA  TGCTTTTCAA  GATACCCGGA  TCACATGAAA
241    CGGCATGACT  TTTTCAAGAG  TGCCATGCCC  GAAGGTTATG  TACAGGAAAAG  GACCATCTTC
301    TTCAAAGATG  ACGGCAACTA  CAAGACACGT  GCTGAAGTCA  AGTTTGAAGG  TGATACCCTT
361    GTTAATAGAA  TCGAGTTAAA  AGGTATTGAC  TTCAAGGAAG  ATGGCAACAT  TCTGGGACAC
421    AAATTGGAAT  ACAACTATAT  CTCACACAAT  GTATACATCA  CCGCAGACAA  ACAAAAAGAA
481    GGAATCAAAG  CCCACTTCAA  GACCCGCCAC  AACATTGAAG  ATGGAAGCGT  TCAACTAGCA
541    GACCATTATC  AACAAAATAC  TCCAATTGGC  GATGGCCCTG  TCCTTTTACC  AGACAACCAT
601    TACCTGTCCA  CACAATCTGC  CCTTTCGAAA  GATCCCAACG  AAAAGAGAGA  CCACATGGTC
661    CTTCTTGAGT  TTGTAACAGC  TGCTGGGATT  ACACATGGCA  TGGATGAACT  ATACAAATCC
721    GGGCGCTCCA  TGGCGCTCGA  G

```

The F5 gene covers roughly >70 kb on the genome and results in a 7 kb mRNA, spliced together from 25 exons. For T7 in vitro transcription, we have used a plasmid from a human ORFeome collection database that contains the full mature cDNA and the F5 Leiden mutation ([BC111588](#)). To run the F5 editing assay efficiently, we have copied a 1000 bp section of this cDNA downstream of a T7 promoter via a 2-step PCR (Phusion polymerase, 1. reaction: template ORF clone BC111588 after miniprep, forward primer: 5`-GGG ACT ATG CAC CTG TAA TAC CAG; backward primer: 5`-CCG TGT AGC CAT GAC TGT AGA TTC; 2. reaction: template PCR reaction 1 after agarose gel purification, forward primer: 5`- GCT AAT ACG ACT CAC TAT AGG GAG AGG GAC TAT GCA CCT GTA ATA CCA G; backward primer: 5`-CCG TGT AGC CAT GAC TGT AGA TTC). After agarose gel purification, a 1000 nt long transcript containing the F5L mutation right in the middle was formed via in-vitro-T7 transcription from this template. The F5VL mutation is located in exon 10 (220 bp) and sits 10 bp upstream of the border to exon 11. Thus, all guideRNAs tested so far can target both, the mature and the pre-mRNA.

R534Q F5 protein sequence

```

1      DYAPVIPANM  DKKYRSQHLD
21     NFSNQIGKHY  KKVMYTQYED
41     ESFTKHTVNP  NMKEDGILGP
61     IIRAQVRDTL  KIVFKNMASR

```

```

81      PYSIYPHGVT FSPYEDEVNS
101     SFTSGRNNTM IRAVQPGETY
121     TYKWNILEFD EPTENDAQCL
141     TRPYYSVDVI MRDIASGLIG
161     LLLICKSRSL DRQGIQRAAD
181     IEQQAVFAVF DENKSWYLED
201     NINKFCENPD EVKRDDPKFY
221     ESNIMSTING YVPESITTLG
241     FCFDDTVQWH FCSVGTQNEI
261     LTIHFTGHSF IYGKRHEDTL
281     TLFPMRGESV TVTMDNVGTW
301     MLTSMNSSPR SKKLRLKFRD
321     VKCIPDDEDED SYEIFEPPES
341     TVMATR

```

R534Q F5 gene sequence

```

1      GGGACTATGC ACCTGTAATA CCAGCGAATA TGGACAAAAA ATACAGGTCT CAGCATTTGG
61     ATAATTTCTC AAACCAAATT GGAAAAACATT ATAAGAAAAGT TATGTACACA CAGTACGAAG
121    ATGAGTCCTT CACCAAACAT ACAGTGAATC CCAATATGAA AGAAGATGGG ATTTTGGGTC
181    CTATTATCAG AGCCCAGGTC AGAGACACAC TCAAAATCGT GTTCAAAAAT ATGGCCAGCC
241    GCCCCTATAG CATTTACCCT CATGGAGTGA CCTTCTCGCC TTATGAAGAT GAAGTCAACT
301    CTTCTTTTAC CTCAGGCAGG AACAACACCA TGATCAGAGC AGTTCAACCA GGGGAAACCT
361    ATACTTATAA GTGGAACATC TTAGAGTTTG ATGAACCCAC AGAAAATGAT GCCCAGTGCT
421    TAACAAGACC ATACTACAGT GACGTGGACA TCATGAGAGA CATCGCCTCT GGGCTAATAG
481    GACTACTTCT AATCTGTAAG AGCAGATCCC TGGACAGGCA AGGAATACAG AGGGCAGCAG
541    ACATCGAACA GCAGGCTGTG TTTGCTGTGT TTGATGAGAA CAAAAGCTGG TACCTTGAGG
601    ACAACATCAA CAAGTTTTGT GAAAATCCTG ATGAGGTGAA ACGTGATGAC CCCAAGTTTT
661    ATGAATCAAA CATCATGAGC ACTATCAATG GCTATGTGCC TGAGAGCATA ACTACTCTTG
721    GATTCTGCTT TGATGACACT GTCCAGTGGC ACTTCTGTAG TGTGGGGACC CAGAATGAAA
781    TTTTGACCAT CCACTTCACT GGGCACTCAT TCATCTATGG AAAGAGGCAT GAGGACACCT
841    TGACCCTCTT CCCCATGCGT GGAGAATCTG TGACGGTCAC AATGGATAAT GTTGGAACTT
901    GGATGTAAAC TTCCATGAAT TCTAGTCCAA GAAGCAAAAA GCTGAGGCTG AAATTCAGGG
961    ATGTTAAATG TATCCCAGAT GATGATGAAG ACTCATATGA GATTTTGTAA CCTCCAGAAT
1021   CTACAGTCAT GGCTACACGG

```

mRNA-Synthesis for RNA editing

Editing was investigated on mRNA substrates of Stop66 eCFP and R534Q F5. Those were generated from the respective genes under control of the T7 promoter with T7 RNA polymerase. To avoid contamination with RNaseA, mRNA production was templated with PCR products rather than plasmids. For this, each gene was amplified from the respective pMG211 plasmids with Phusion Pol, T7forw primer and a respective backward primer, yielding 900 or 1050 nt long ssRNA transcripts in very good quality after spin column clean-up with the RNeasy Minelute kit (Qiagen). To remove even traces of DNA template, mRNA transcripts were subsequently treated with DNaseI (Qiagen, 30 min at 37°C), and were finally cleaned by another RNeasy spin column work-up. Average yields of DNA-free mRNA were 1.0 OD/50µl reaction mix. The absence of DNA template was always proven by PCR prior to cDNA synthesis.

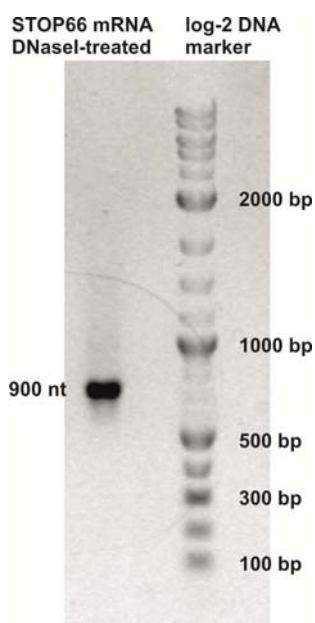


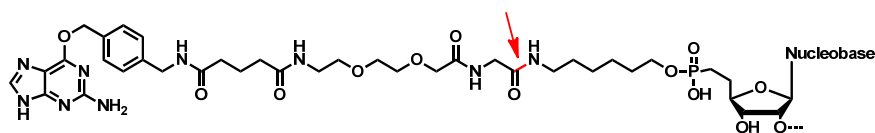
Figure S2. Analytical agarose gel (1x TAE, 1.5% agarose) of a representative mRNA (STOP66 *ecfp*) obtained via T7 RNA polymerase-dependent transcription from a 930 bp PCR transcript as template, after DNaseI treatment and spin column workup. As expected, the 900 nt ssRNA runs slightly faster as compared to the 1000 bp dsDNA marker and is of very good quality (no degradation detectable).

Properties of the mRNA transcripts: **Stop66 eCFP**: length 903 nt; molecular weight = 290.3 kDa, 287 adenosines (32%); 207 cytosines (23%); 194 guanosines (22%); 215 uridines (24%); 5'-UTR: length 70 nt; 3'-UTR: length 68 nt, ORF: length 765 nt: 245 A (32%); 179 C(23.4%); 161 G (21.0%); 180 U (23.5%); **R534Q F5**: length 1040 nt; molecular weight = 334.6 kDa, 334 adenosines (32%); 222 cytosines (21%); 228 guanosines (22%); 256 uridines (25%).

Syntheses

Synthesis of 5'-benzylguanine-modified gRNAs

gRNA oligos with no, one and up to three 2'-methoxy groups were obtained from Eurofins Germany as HPLC-purified, MALDI-TOF-confirmed ssRNA oligos carrying a 5'-C6-aminolinker (NH₂-gRNA). The phosphothioate / 2'-OMe modified oligomer was obtained from Biospring GmbH (Frankfurt) in HPLC-clean, desalted purity. In a typical procedure (*Angew. Chem. Int. Ed.* **2012**, *51*, 11166), 150 µg NH₂-gRNA (20-25 nMol) were first cleaned by ethanol precipitation, washed, dried, taken up in 25 µl hepes buffer (75 mM hepes, 50 mM NaCl, pH 8.1) and were given to a pre-activated OSu-ester of BG-linker-COOH (10 equiv. BG-linker-COOH, 10 equiv. EDCI*HCl, 14 equiv. N-hydroxysuccinimide, 40 equiv. Hünig base, all in 25 µl DMSO, 1h preactivation). After 1h at 30°C, another lot of 10 equiv. preactivated (2h) BG-linker-OSu in 25 µl DMSO were added and incubated for 1h at 30°C. The raw gRNA was precipitated with ethanol, taken up in 1xTBE, 7M urea and purified on a 20% 19:1 1xTBE-7M urea PAGE mini gel (10x10 cm, 1mm thick), cut out on a TLC-plate under low-intensity 254 nm UV-light and was isolated by the crush-soak method into RNase-free water at 4°C overnight. To remove urea and buffer salts, the BG-gRNA was precipitated again with ethanol, washed, dried and dissolved into 80 µl RNase-free water. Typically, 80-100% conversion was observed and around 40% pure BG-modified gRNA was obtained after crush soak / precipitation. The concentrations were determined by UV-absorbance at 260 nm. Special care must be taken in the preparation of the PTO/2'-OMe-modified guideRNA with regard to ethanol precipitation and PAGE purification. Even though clean oligomers can be precipitated, precipitation from a reaction mixture or urea containing solutions turned out to be difficult. The RNA pellet also tends to dissolve in 70% ethanol (take short incubation of 5 min with -20°C pre-cooled ethanol). PTO/OMe oligomers run slower on the PAGE, thus cyanol blue should be added as a dye front rather than bromphenol blue. In our experience the 20 mer oligomer with 6xPTO and 17x2'-OMe runs close to the cyanol blue band in 18% PAGE (19:1). To reduce problems, we directly loaded the reaction mixture on the PAGE skipping the precipitation step. However, BG-linker-COOH should also be loaded as a marker to clearly distinguish the modified guideRNA from the excess reagent that would block conjugation if it was not removed.



Scheme S3. Structure of the BG-modified guideRNAs. BG was introduced via peptide bond formation (red bond, see arrow) into C6-aminolinker-modified guideRNAs.

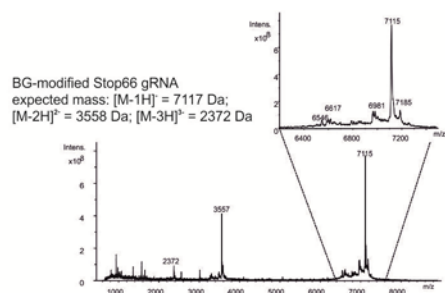
Stop66 gRNA	5'-r(UCG GAA CAC CC <u>C</u> AGC ACA GA)	$\epsilon_{260}^a = 230$
Stop66 gRNA-me	5'-r(UCG GAA CAC CC <u>c</u> AGC ACA GA)	$\epsilon_{260}^a = 230$
Stop66 gRNA-3' me	5'-r(UCG GAA CAC CC <u>C</u> aGC ACA GA)	$\epsilon_{260}^a = 230$
Stop66 gRNA-5' me	5'-r(UCG GAA CAC Cc <u>C</u> AGC ACA GA)	$\epsilon_{260}^a = 230$
Stop66 gRNA-3me	5'-r(UCG GAA CAC Ccc <u>c</u> aGC ACA GA)	$\epsilon_{260}^a = 230$
Stop66 gRNA-PTO/OMe	5'-r(u _s c _s g gaa cac c CC <u>C</u> A gc a _s c _s a _s g _s a)	$\epsilon_{260}^a = 230$
F5 gRNA-17nt	5'-r(CAC GUA UUC CU <u>C</u> GCC UGU CC)	$\epsilon_{260}^a = 189$
F5 gRNA-14nt	5'-r(UGC UAU UCC UC <u>G</u> CCU GU)	$\epsilon_{260}^a = 160$
F5 gRNA-14nt-2me	5'-r(UGC UAU UCc u <u>CG</u> CCU GU)	$\epsilon_{260}^a = 160$

Table S1. Commercially obtained 5'-NH₂-gRNAs that were transformed into 5'-BG-modified gRNA; the nucleotide opposite of the targeted adenosine is unlined and in bold; 2'-methoxylation is indicated by small letters; phosphothioate linkages are indicated by "s" subscripts. The first three nucleotides 5' of each guideRNA are not binding to the mRNA substrate but link the guideRNA to the SNAP-tag.

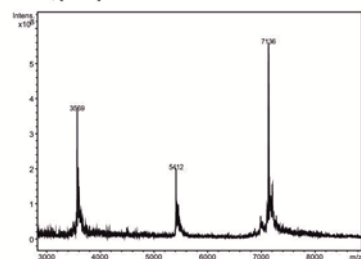
a) extinction coefficients of NH₂-gRNAs are given in mM⁻¹ cm⁻¹, after modification with BG, 12 mM⁻¹ cm⁻¹ were added to the respective extinction coefficients listed here

MALDI-TOF-MS-analysis of BG-modified gRNAs

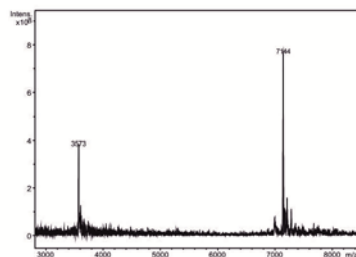
PAGE-purified BG-gRNAs were precipitated with 0.5 volumes 7.5 M ammonium acetate / 2.5 volumes ethanol, washed with 70% ethanol and were dissolved in nanopure water. Samples were then mixed with a matrix of 2,4,6-trihydroxyacetophenone monohydrate (0.3 M in EtOH) / diammonium citrate (0.1 M in water) 2:1. Spectra were collected by the Richert group (University of Stuttgart) on a Bruker REFLEX-IV spectrometer (see SI in C. Richert et al., *Nat. Chem.* **2011**, 3, 603) in the negative ion mode. Importantly, all BG-gRNAs contain the BG-modification and are free from impurities by the respective non-modified amino-gRNAs (-568 Da).



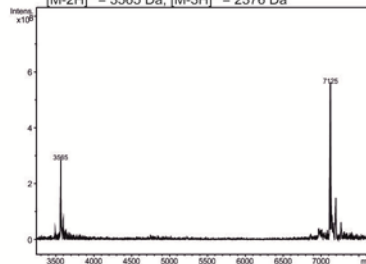
BG-modified Stop66 gRNA-me
expected mass: $[M-1H]^+ = 7131$ Da;
 $[M-2H]^+ = 3565$ Da; $[M-3H]^+ = 2376$ Da



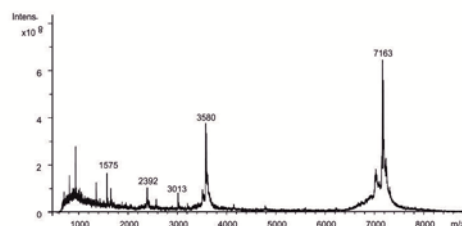
BG-modified Stop66 gRNA-3' me
expected mass: $[M-1H]^+ = 7131$ Da;
 $[M-2H]^+ = 3565$ Da; $[M-3H]^+ = 2376$ Da



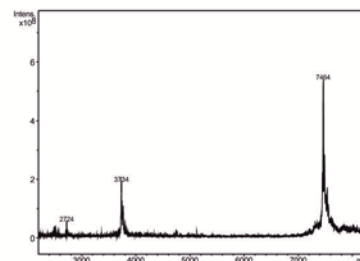
BG-modified Stop66 gRNA-5' me
expected mass: $[M-1H]^+ = 7131$ Da;
 $[M-2H]^+ = 3565$ Da; $[M-3H]^+ = 2376$ Da



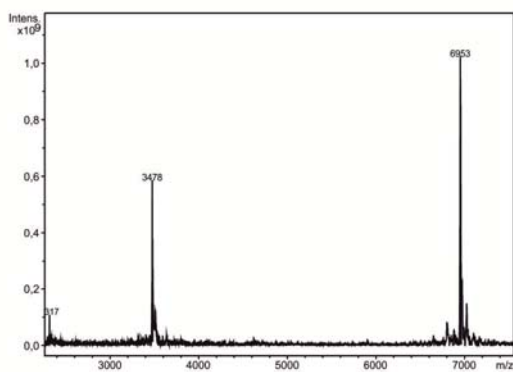
BG-modified Stop66 gRNA-3me
expected mass: $[M-1H]^+ = 7159$ Da;
 $[M-2H]^+ = 3579$ Da; $[M-3H]^+ = 2386$ Da



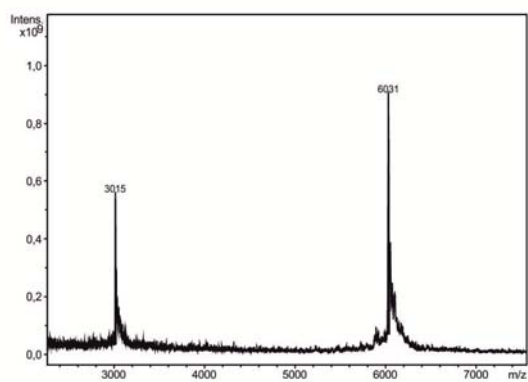
BG-modified Stop66 gRNA-PTO/OME
expected mass: $[M-1H]^+ = 7453$ Da;
 $[M-2H]^+ = 3727$ Da



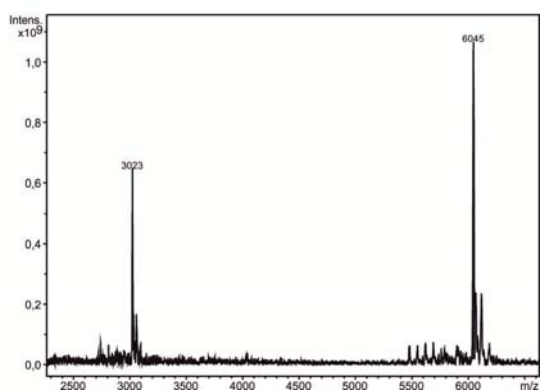
BG-modified F5 gRNA-17nt
expected mass: $[M-1H]^- = 6960$ Da;
 $[M-2H]^{2-} = 3480$ Da; $[M-3H]^{3-} = 2319$ Da



BG-modified F5 gRNA-14nt
expected mass: $[M-1H]^- = 6023$ Da;
 $[M-2H]^{2-} = 3011$ Da; $[M-3H]^{3-} = 2007$ Da



BG-modified F5 gRNA-14nt-2me
expected mass: $[M-1H]^- = 6051$ Da;
 $[M-2H]^{2-} = 3025$ Da; $[M-3H]^{3-} = 2016$ Da



Editing of mRNA substrates with gRNA-ADAR-conjugates

General considerations

Editing reactions were performed in 1x reverse transcription (RT) buffer from New England Biolabs but with a magnesium concentration reduced to physiological conditions (25 mM Tris-HCl, 75 mM KCl, 0.75 mM MgCl₂, 10 mM DTT, pH 8.3). To our experience editing reactions can be run between magnesium concentrations of 0 to 5 mM, however, some codons are better edited at low magnesium content (≤ 1 mM). To be as close as possible to later in vivo conditions, reactions are generally run at 0.75 mM magnesium concentration. eCFP mRNA has a short self-pairing region around nt 400. Due to the formation of this dsRNA motif little over-editing can occur. However, adding supplements as heparin (0.5-2 μ M), spermidine or BSA can completely block this off-target editing. Thus, editing reactions with eCFP mRNA (not F5 mRNA) had been supplemented with heparin (2 μ M, calculated on an average molar mass of 20 kDa). SNAP-ADAR proteins were always given in at least 1.3 equiv. excess over the respective gRNA to allow full in situ conversion of the latter and thus to ensure the presence of the deaminase at every gRNA/mRNA pairing complex. To minimize mRNA degradation the reaction mixture was complemented with 0.5 u/ μ l murine RNase inhibitor (NEB). Typically, reactions were carried out at 30°C 30 min, 37°C 30 min for 3 cycles (overall reaction time = 3 h), however, >90% of the final A to G conversion was usually observed after 30 min. Editing reactions were run at the given concentrations of the components on a 25 μ l scale according to our recently published protocol (Stafforst, Schneider, *Angew. Chem. Int Ed.* **2012**, *51*, 11166).

Post-editing work-up / PCR amplification

After editing, the edited mRNA was reverse transcribed into cDNA. For this, the editing reaction was stopped by addition of a reverse transcription mix containing an excess of a short ssDNA that is fully reverse complementary to the respective gRNA used (see Table 1.) and heating to 70°C for 3 min. The magnesium concentration was adjusted to the requirements of the reverse transcriptase (3 mM final Mg concentration). Heating enables to break possible RNA secondary structure and to allow primer binding. RT was started by adding 50 u M MuLV RT (NEB) in 0.5 μ l. After 2 h at 42 °C the cDNA was worked-up and concentrated using the Nucleospin PCR work-up kit (Macherey & Nagel).

Reverse complementary strands to stop editing reactions:

Stop66 ecfp ssDNA: 5'-d(ATCTGTGCTGGGGTGTTCGAT)

R534Q F5 ssDNA: 5'-d(CGGACAGGCGAGGAATACGTGT)

For subsequent analysis, a PCR amplification was required. For this, the respective cDNA from edited mRNA was used as template (typically 1 μ l OD260 = 0.2, ca. 5 ng) in a Taq-PCR reaction on a 50 μ l scale. Negative controls were performed with mRNA substrate (1 μ l OD260 = 0.2), to exclude that traces of DNA template in the mRNA originate the PCR product, and without adding any template to ensure that none of the mixed PCR components was contaminated with any DNA template. A PCR product containing the gene served as positive control. PCR products were separated by 1.5 % 1xTAE agarose gel electrophoresis containing the Rotisafe stain (Carl Roth GmbH), product bands were cut out and DNA was extracted with the Nucleospin PCR & Gel work-up kit (Macherey & Nagel).

Taq PCR primers used:

eCFP: fw: 5'-d(GCGGATAACAATTCCCCTCTAG)

bw: 5'-d(CAGCGGTGGCAGCAGCCAAC)

F5: fw: 5'-d(GGGACTATGCACCTGTAATACCAG)

bw: 5'-d(GAGTCTTCATCATCATCTGGGATAC)

Sequencing

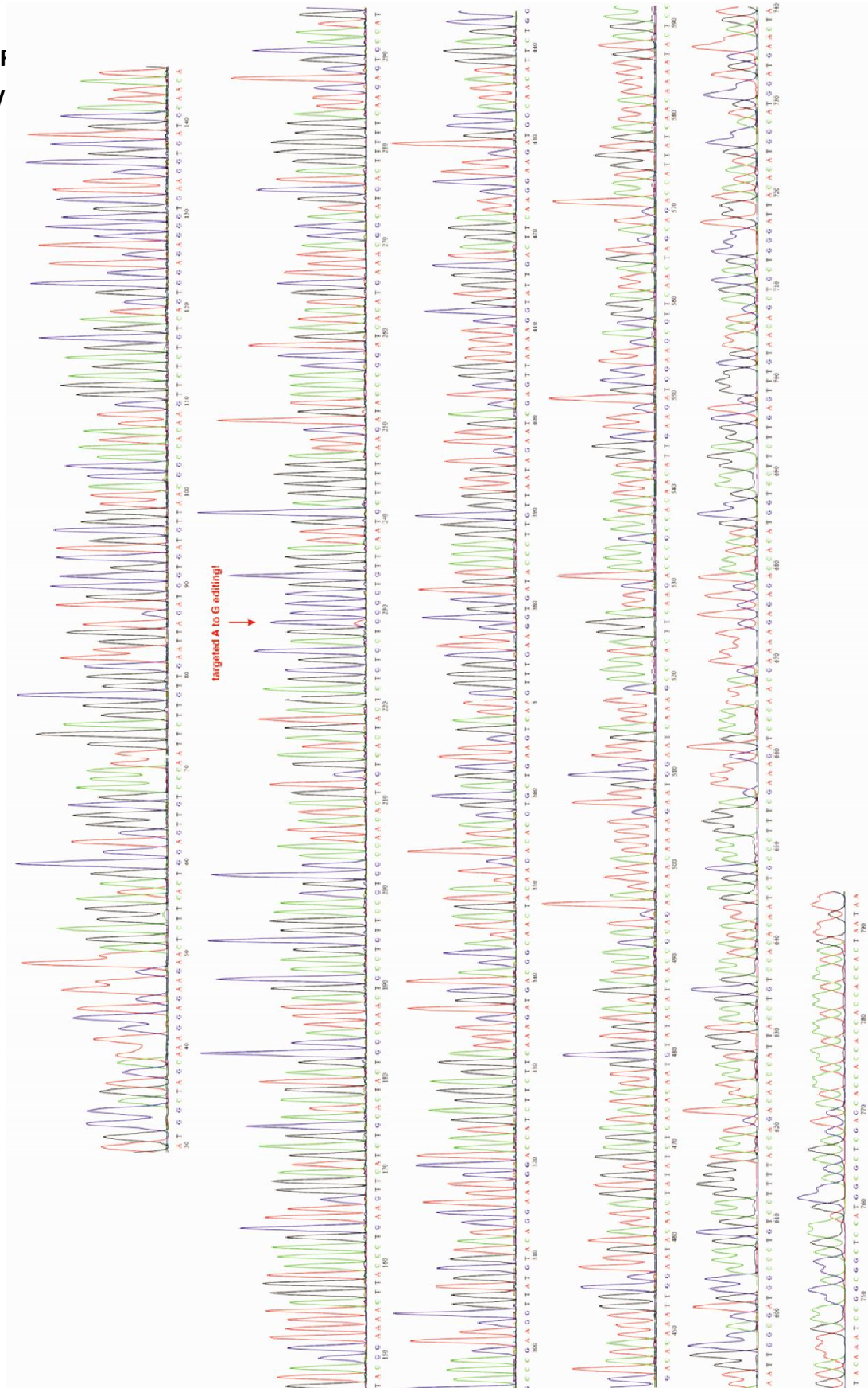
For sequencing, PCR products were sent to either LGC-Genomics, Berlin (F5) or to Eurofins MWG, Ebersberg (eCFP). The F5 samples were sequenced using the Taq PCR backward primer, the eCFP samples were sequenced using the Taq PCR forward primer reported above. The areas in the abi-traces of the respective adenosine versus guanosine peaks at a specific site were used to estimate the degree of editing.

$$\text{editing efficiency} = \frac{\text{area G}}{(\text{area A} + \text{area G})}$$

Results

Figure S3. One complete Sanger sequencing trace that covers the full ORF of eCFP is given to exemplify the high editing selectivity achieved with directed RNA editing. The given sequence is of the experiment shown in Figure S3 F): the editing with Stop66 BG-gRNA-ant in presence of SNAP-ADAR1. No second site (off-target) editing was observed for any of the 14 editing settings shown in Figure S3.

Editing of I chemically



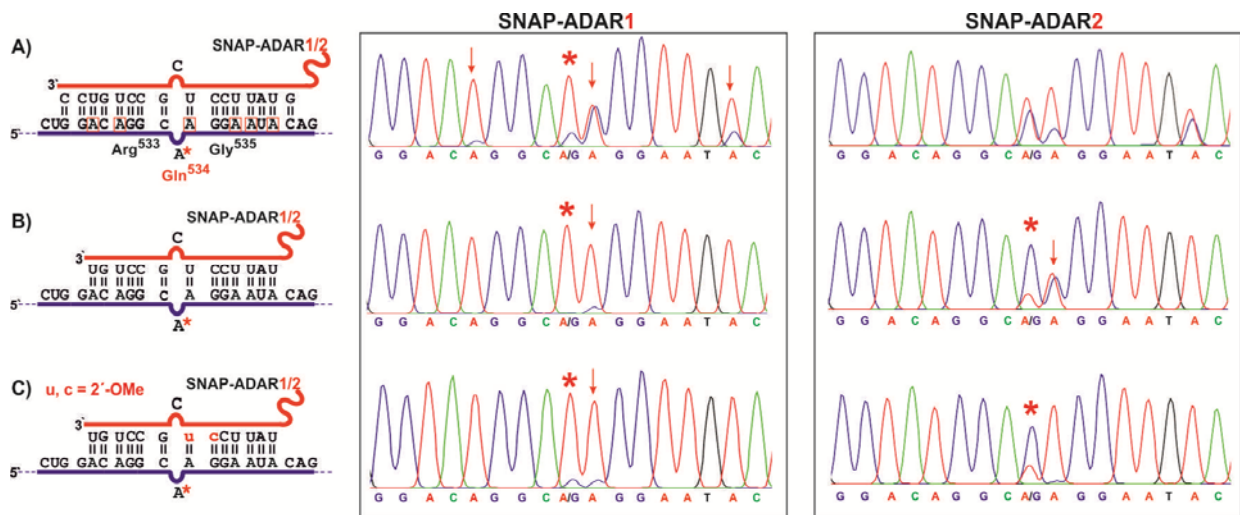


Figure S4. Effect of chemically modified guideRNAs on editing of R⁵³⁴Q F5 mRNA. As compared to ADAR2, ADAR1 is not at all able to sufficiently activate the 5'-CAA codon. ADAR2 efficiently and selectively edits the targeted adenosine (*). However, over-editing at the direct 3'-neighboring adenosine base is problematic. Selective methylation at two sites (C) enable to control over-editing in closest vicinity. The mRNA is shown in blue, the guideRNA in red. Bases modified with 2'-OMe are colored red, phosphothioate linkages are indicated with a green 's'. Editing yields can be estimated from the areas for adenosine (red) versus guanosine (blue) in the respective sequencing trace after reverse transcription and PCR amplification. Editing conditions: 3h at 30/37°C; 25 nM mRNA, 200 nM BG-guideRNA, 350 nM SNAP-ADAR1/2, 75 mM KCl, 50 mM Tris-HCl, 10 mM DTT, 0.70 mM MgCl₂, pH 8.3.

Figure S5. One complete Sanger sequencing trace that covers the full 1000 nt part of the F5 mRNA which was be used for editing. The F5 leiden mutations lies centrally in this transcript. The full sequencing trace is given to exemplify the high editing selectivity achieved with directed RNA editing (no other adenosine of 330 possible off-targets is edited). The given sequencing trace is that of the experiment shown in Figure S5 C): the editing with F5 BG-gRNA-14nt-2me in presence of SNAP-ADAR2.



Cellular editing assay

Constructs

SNAP-ADAR1, Stop66 eCFP and wt eCFP were subcloned into pcDNA3.1 (Life Technologies) at the BamH1 / Xba1 site using the following primers: CFP: fw: 5'-GCGGATCCAC CATGGCTAGC AAAGGAGAAG AACTC; bw: 5'-CCTCTAGAGC CGGATTTGTA TAGTTCATCC ATGCC; SNAP-ADAR1: fw: 5'-CGGATCCACC ATGGACAAAG ATTGCGAAAT GAAAC; bw: 5'-CCTCTAGATA CTGGGCAGAG ATAAAAGTTC TTTTC. The genes are under control of the CMV promotor and contain a 5'-CCACCATGG Kozak sequence. In contrast to the in vitro experiments, we have used a new, improved SNAP-tag variant called SNAP_f (cloned from the SNAP_f vector, N9183S, New England Biolabs into the pcDNA3.1 vector using the BamH1/Asc1 restriction sites with fw primer: 5'-GCGGATCCAC CATGGACAAA GACTGCG; bw primer: 5'-CGGCGCGCC TCCGCCTGCA GGACCCAGC). Compared to the older SNAP-tag used in the in-vitro-assays (pSNAP-tag (T7)-2 Vector, N9182S, NEB) this new SNAP_f variant possess 10 mutations that are important for faster labeling kinetics^[1] and are sequence-optimized for translation in mammalian cells. (Protein-Sequence of SNAP_f: MDKDCMKRRTTLDSPKLELSGCEQGLHRIIFLGKGTSAADAVEVPAPAAVLGGPEPLMQATAWLNAYFHQPEAIEEFPVPALHHPVFQQESFTRQVLWKLKVVVKFGEVISYSLAALAGNPAATAAVKTALSGNPVPIIPCHRVVQGDLDVGGYEGGLAVKEWLLAHEGHRGKPLG).

Cell culture & transfection

Fresh HEK293T cells (DSMZ Braunschweig, Germany, ACC-635) were cultured in DMEM (Life Technologies) supplemented with 10% FBS (Life Technologies) and 1% penicillin/streptomycin (Life Technologies) under standard conditions (37°C and 5% CO₂ in a water saturated steam atmosphere). The transfection experiments were performed with Lipofectamine 2000 reagent (Life Technologies) according to the manufacture's protocol. For this, cells were seeded in full media (DMEM+FBS+ antibiotics) 24hrs prior to the first transfection with plasmid-DNA. Subsequently, the media was changed to DMEM+FBS lacking antibiotics. Plasmid-DNA and Lipofectamine 2000 reagent in a ratio of 1 µg to 4 µl were separately diluted in equal amounts of OptiMEM (Life Technologies) and incubated for 5 min. Then the solutions were mixed and incubated for 20 min prior to addition to the cells. 24 hrs later, the cells were harvested by trypsination and were seeded again. After incubation overnight the cells were transfected with the respective guideRNA applying basically the same procedure as for plasmid transfection (20:1 ratio of guideRNA (pmol) to Lipofectamine (µl)). If no RNA was transfected (negative controls, f.i.) the cells were treated with the same amount of Lipofectamine 2000 instead.

Fluorescence imaging of Stop66 eCFP editing

2×10⁵ Cells were seeded in a 24-well plate in 500 µl media and were co-transfected with 1,8 µg of each plasmid in 50 µl OptiMEM + Lipofectamine in 50 µl OptiMEM on the following day. 24hrs after transfection, 4×10⁴ cells were given in each 96-well plate in 100 µl media. guideRNA (50 pmol) in 25 µl OptiMEM + Lipofectamine in 25 µl OptiMEM were used for the second transfection of the cells. 24hrs later, the eCFP fluorescence was imaged at 50× magnification using an Axiovert 200M microscope (Zeiss) equipped with an AxioCam MRm camera (Zeiss). Access to the fluorescence microscope was kindly provided by Prof. Dr. Alfred Nordheim (Interfaculty Institute for Cell Biology, University of Tübingen).

[1] A. Gautier, A. Juillerat, C. Heinis, I. Reis Correa Jr., M. Kindermann, F. Beaufils, K. Johnsson, *Chem. Biol.* **2008**, *15*, 128.

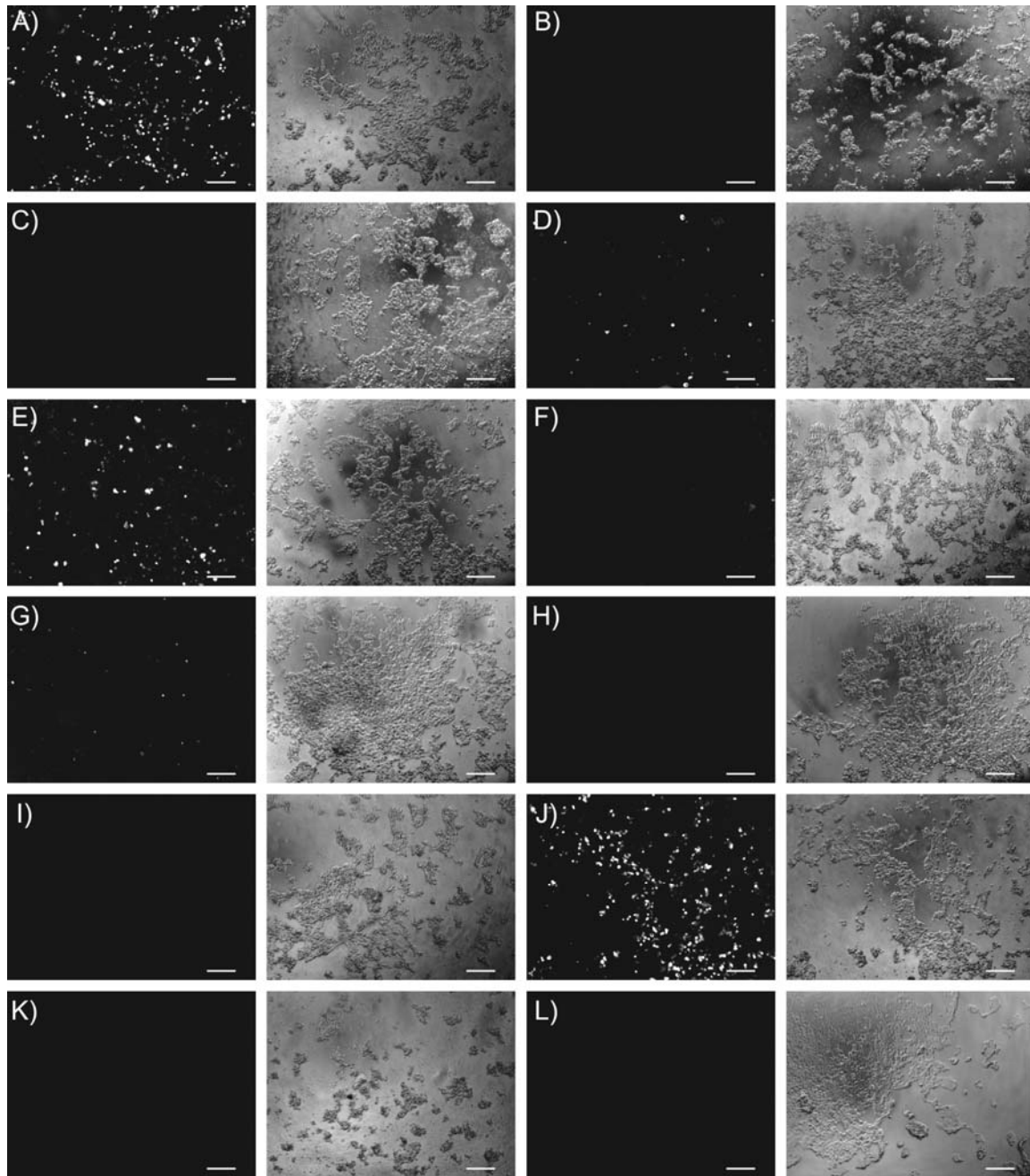


Figure S6. Fluorescence (left) and phase contrast pictures (right) for analyzing Stop66 eCFP editing in HEK293T cells (scale bar: 200 μ m). Cells were co-transfected with equal amounts of reporter gene (Stop66 eCFP or wildtype) and SNAP-ADAR1 (or empty pcDNA3.1). After 24hrs the guideRNAs were transfected and 24hrs later, the fluorescence phenotype was analyzed by fluorescence microscopy.

A) Positive control 1 (wt eCFP with SNAP-ADAR1 and BG-antagomir-guideRNA);

B) Negative control 1 (Stop66 eCFP / empty pcDNA3.1 / BG-antagomir-guideRNA);

C) Negative control 2 (Stop66 eCFP / SNAP-ADAR1 / no guideRNA);

D)-H) experiments with Stop66 eCFP / SNAP-ADAR1 and various guideRNAs:

D) standard BG-guideRNA;

E) BG-antagomir-guideRNA;

F) standard NH₂-guideRNA;

G) NH₂-antagomir-guideRNA;

H) Stop58 eGFP BG-antagomir-guideRNA (5'-BG-UsAsUGUGUCGG CCA_CGGAAsCsAsGsG-3', italic letters: 2'-methoxylation, s: phosphothioate modification, ribonucleotides: underlined);

I)-L) further control experiments:

I) Negative control 3 (Stop66 eCFP / empty pcDNA3.1 / no guideRNA);

J) Positive control 2 (wt eCFP / SNAP-ADAR1 / BG-antagomir-guideRNA);

K) Negative control 4 (empty pcDNA3.1 / empty pcDNA3.1 / BG-antagomir-guideRNA);

L) Negative control 5 (untreated cells).

Determination of in-vivo-editing yield by RNA extraction and Sanger sequencing

In comparison to the fluorescence imaging protocol, the protocol for determining the editing efficiency in HEK293T cells was 10-fold up-scaled from 24 well plates to 6-cm dishes to obtain enough material for RNA sequence analysis. 2×10^6 cells were plated on a 6-cm dish in 5 ml media and the first transfection was performed with 18 μ g of both Stop66 eCFP and SNAP-ADAR1 plasmids in 500 μ l OptiMEM + Lipofectamine in 500 μ l OptiMEM. After 24hrs 2×10^5 harvested cells were seeded in each 24-well plate in 500 μ l media, and after 24hrs incubation 250 pmol guideRNA in 50 μ l OptiMEM were transfected per well with Lipofectamine 2000 in 50 μ l OptiMEM. RNA was extracted 24hrs after the second transfection using the RNeasy MinElute Cleanup Kit (Qiagen). The samples were purified from DNA by DNase I digest (RNase-Free DNase Set, Qiagen). Afterwards, M-MuLV reverse transcriptase (New England Biolabs) was applied for synthesizing cDNA according to the manufacture's protocol (RT Primer: 5'-CTAGAAGGCACAGTCGAGGC). The resulting cDNA was amplified by Taq-PCR (fwd Primer: 5'-GCGGATCCACCATGGCTAGCAAAGGAGAAGAAGTCT, rev Primer: 5'-CCTCTAGAGCCGGATTTGTATAGTTCATCCATGCC) and analyzed through Sanger Sequencing (Eurofins MWG Operon, Germany) applying the forward primer from the PCR amplification.

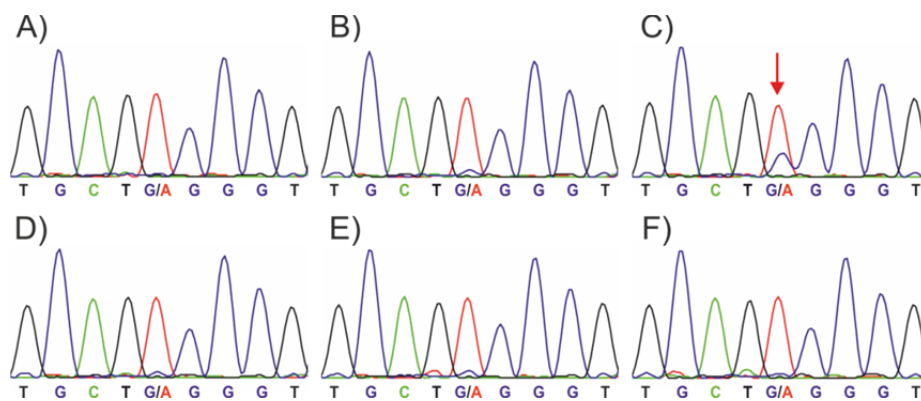


Figure S7. RNA sequence analysis for determining the editing yield in HEK293T cells. Cells were co-transfected with equal amounts of Stop66 eCFP and SNAP-ADAR1. After 24hrs various guideRNAs were transfected and 24hrs later, mRNA was extracted. After DNase I digestion cDNA was prepared by reverse transcription, amplified by Taq-PCR and subsequently analyzed by Sanger Sequencing. The following guideRNAs were applied, respectively:

- A) no guideRNA;
- B) standard BG-guideRNA;
- C) BG-antagomir-guideRNA;
- D) standard NH₂-guideRNA;
- E) NH₂-antagomir-guideRNA;
- F) Stop58 eGFP BG-antagomir-guideRNA.

Man. 2: accepted

Hanswillemenke, A., Kuzdere, T., Vogel, P., Jékely, G. & Stafforst, T. Site-directed RNA editing in vivo can be triggered by the light-driven assembly of an artificial riboprotein. *J. Am. Chem. Soc.* **137**, 15875-15881 (2015).

Site-Directed RNA Editing in Vivo Can Be Triggered by the Light-Driven Assembly of an Artificial Riboprotein

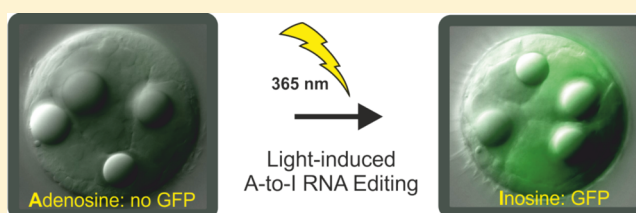
Alfred Hanswillemenke,[†] Tahsin Kuzdere,[†] Paul Vogel,[†] Gáspár Jékely,[‡] and Thorsten Stafforst^{*,†}

[†]Interfaculty Institute of Biochemistry, University of Tübingen, Auf der Morgenstelle 15, 72076 Tübingen, Germany

[‡]Max-Planck-Institute for Developmental Biology, Spemannstraße 35, 72076 Tübingen, Germany

S Supporting Information

ABSTRACT: Site-directed RNA editing allows for the manipulation of RNA and protein function by reprogramming genetic information at the RNA level. For this we assemble artificial RNA-guided editases and demonstrate their transcript repair activity in cells and in developing embryos of the annelid *Platynereis dumerilii*. A hallmark of our assembly strategy is the covalent attachment of guideRNA and editing enzyme by applying the SNAP-tag technology, a process that we demonstrate here to be readily triggered by light in vitro, in mammalian cell culture, and also in *P. dumerilii*. Lacking both sophisticated chemistry and extensive genetic engineering, this technology provides a convenient route for the light-dependent switching of protein isoforms. The presented strategy may also serve as a blue-print for the engineering of addressable machineries that apply tailored nucleic acid analogues to manipulate RNA or DNA site-specifically in living organisms.



INTRODUCTION

RNA-guided machineries provide highly selective and rationally programmable tools for the site-specific manipulation of nucleic acids. Several endogenous riboproteins are known that are steered toward their endogenous targets by nucleic acid hybridization and that are readily re-addressed toward new targets by expression or administration of artificial external guideRNAs. Those include the snoRNA-guided 2'-O-methylation¹ and pseudo-uridylation² machineries and the microRNA-guided RNA-induced silencing complex. The harnessing of the latter machinery, better known as RNA interference,³ has developed into a standard tool in cell biology. Besides harnessing endogenous eukaryotic machineries, the engineering of artificial riboproteins for the site-specific manipulation of nucleic acids comes more and more into focus now. Tools are highly desired that simplify genetic engineering⁴ and that help to elucidate the role of point mutations and RNA modifications.^{5–7} Besides their application in basic biology research, such tools have potential for translation into individualized medicine. A highly topical example is the re-engineering of the bacterial CRISPR-Cas9 system for site-selective genome editing in eukaryotic cells.⁸

Endogenous riboproteins are typically assembled by molecular recognition between specific protein and RNA structures.⁹ The formation of a single covalent bond between an RNA and a protein component, however, is virtually unknown for that purpose. Nevertheless, we could recently demonstrate the assembly and functioning of highly selective adenosine (A)-to-inosine RNA editing machineries inside living cells following the latter approach.^{10,11} Since inosine is biochemically read as guanosine (G), editing formally creates A-to-G point mutations at the RNA level. If RNA editing is

directed to the open reading frame, 12 out of the 20 canonical amino acids can be substituted,¹² including most of the polar residues essential for enzyme catalysis, post-translational protein modification, or signaling. Furthermore, editing in the non-coding part of the RNA can interfere with translation initiation (start codon), translation stop, microRNA action, and splicing among others.^{13,14} Thus, the potential of site-directed RNA editing for application in basic biology research and medicine is evident.^{15–18}

We apply the SNAP-tag technology¹⁹ to assemble the editing machinery via covalent bond formation. This technology requires the fusion of a SNAP-tag domain (an evolved O⁶-alkylguanine-DNA alkyltransferase) with the C-terminal catalytic domain of a human ADAR enzyme (adenosine deaminases acting on RNA).¹⁰ At the RNA component, the incorporation of a small chemical moiety, O⁶-benzylguanine (BG), is necessary. The covalent bond is then formed in situ in a single-turnover enzymatic reaction between the SNAP-tag and the BG moiety with very fast kinetics ($k_{\text{conjugation}} = 2.8 \times 10^4 \text{ M}^{-1} \text{ s}^{-1}$)²⁰ and high specificity (Figure 1a). Recently, we demonstrated the repair of a premature stop codon (UAG) into a tryptophan codon (UG) in a fluorescent reporter gene in human cells (293T).¹¹ Notably, the repair reaction was strongly dependent on the covalent attachment of the guideRNA to the deaminase. This opens the appealing possibility of controlling the editing reaction by triggering the assembly of the covalent RNA–protein conjugate (Figure 1a). We decided to apply light as a trigger, as it allows for the very precise and fast control in time, space, and dosage.²¹

Received: October 3, 2015

Published: November 23, 2015

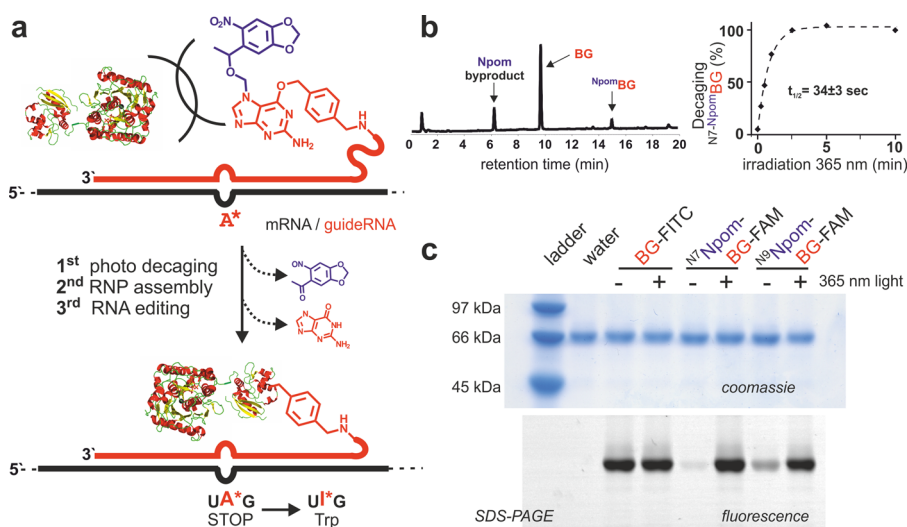


Figure 1. (a) Concept of light-triggered site-directed RNA editing. Assembly of the guideRNA–deaminase conjugate requires release of the Npom-protected benzylguanine (BG) moiety and is a prerequisite for the editing reaction. (b) First-order kinetic analysis (via HPLC) of the photodeprotection of N7-NpomBG at the small-molecule level. The HPLC trace shows the product mixture after 60 s of 365 nm irradiation (75% conversion). The respective analysis for N9-NpomBG can be found in Figure S13. (c) Light-triggered conjugation reaction of fluorescein-labeled Npom-BG with SNAP-ADAR1 protein (SDS-PAGE coomassie versus fluorescein stain). BG-FITC refers to the conjugate of BG with fluorescein isothiocyanate, and BG-FAM refers to the conjugate with 6-carboxyfluorescein.

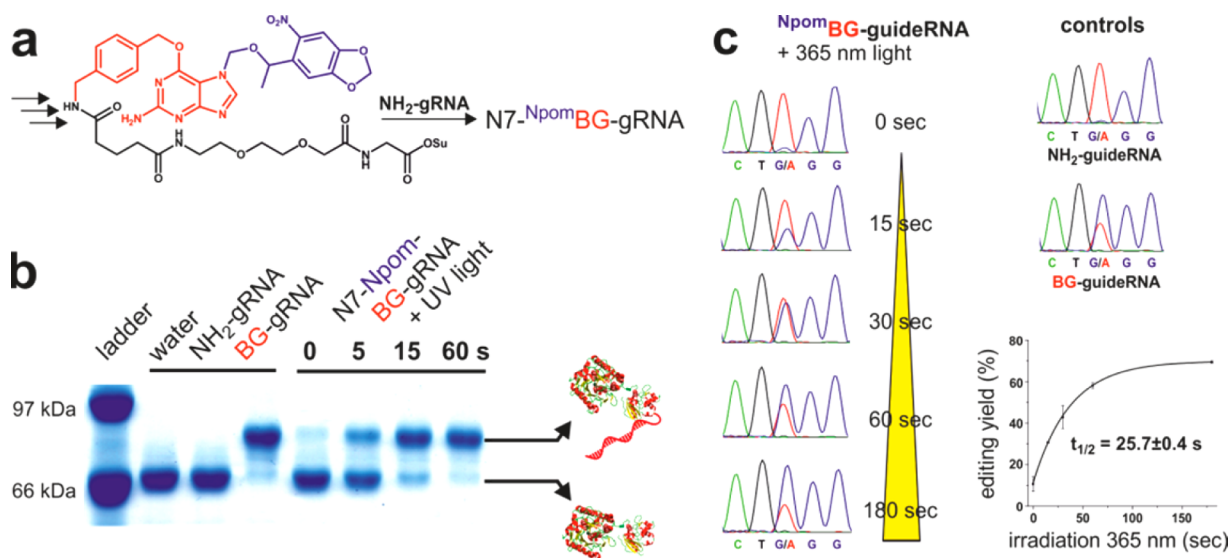


Figure 2. Light-dependent assembly of the editase and in vitro RNA editing. (a) N7-NpomBG is included into an activated linker that readily reacts with the aminolinker of commercially available RNA analogues to obtain the N7-NpomBG-gRNAs. (b) The light-driven conjugation reaction between Npom-BG-gRNA and SNAP-ADAR1 is easily monitored by SDS-PAGE (coomassie stain). (c) In vitro site-directed RNA editing of the amber stop codon at position 66 in the eCFP gene. The editing yield is clearly light-dose-dependent obeying first-order kinetics. Sanger sequencing of the entire ORF of the eCFP gene shows no off-target editing (Figure S15). The respective editing applying SNAP-ADAR2 instead of SNAP-ADAR1 is given in Figure S16. For further details, see the Supporting Information.

RESULTS AND DISCUSSION

Synthesis and Decaging of Npom-Protected O6-Benzylguanine. To achieve the light-dependent assembly of the covalent RNA–protein conjugate, we masked the BG moiety chemically by installment of a light-sensitive 6-nitropiperonyloxymethyl (Npom) protection group^{22,23} which absorbs broadly in the 330–420 nm range. During synthesis we obtained a separable 1:2 mixture of regioisomers containing the Npom group either at N7 or N9 position of the guanine base. Upon irradiation with 365 nm light on a common UV-light table, both isomers, N7 and N9, decay efficiently into free BG and the respective nitroso acetophenone byproduct with similar

kinetics (N7 isomer, Figure 1b, $t_{1/2} = 34 \pm 3$ s; N9 isomer, Figure S13, $t_{1/2} = 47 \pm 4$ s). The decaging efficiency $\epsilon\phi$ was determined by comparison with a commercial standard (DMNB-cAMP) to be ~ 2000 and ~ 1500 M⁻¹ cm⁻¹ for the N7 and N9 isomers, respectively, giving quantum yields $\phi \approx 0.5$ and 0.36 (for details, see the Supporting Information).

To determine the reactivity of the Npom-protected BG with SNAP-deaminases, we modified the aminomethyl linker of the BG moiety with fluorescein. Such probes were incubated with sub-stoichiometric amounts of SNAP-ADAR1 either in the dark or in the presence of 365 nm light. The conjugate formation was then determined by SDS-PAGE and fluorescence analysis

(Figure 1c). It was clearly shown that full fluorescence labeling of the SNAP-deaminase was readily accessible upon irradiation. However, some background reactivity of the protected BG in particular of the N9 isomer with SNAP-ADAR1 was visible. The latter is coherent with the requirement of the natural ancestor of the SNAP-tag to accept the desoxyribose at the N9 position of the nucleobase.²⁴ We did not expect the low-level residual activity to play a role under dilute conditions inside the living cell; nevertheless, we continued all further work with the pure N7 isomer of Npom-BG.

Light-Triggered Assembly of Protein–RNA Conjugates Controls RNA Editing in Vitro. To study the assembly of the guideRNA–deaminase conjugate and its effect on in vitro RNA editing, we attached the N7-NpomBG via a short linker¹⁰ to the 5'-terminal aminolinker of a 17 nt guideRNA that directs the conjugate to codon 66 of the eGFP transcript in order to stimulate the repair of a premature amber stop codon (UAG) back to tryptophan (Figure 2a). Via SDS-PAGE we first characterized the light-dependent assembly of the riboprotein (Figure 2b). Conjugation results in a readily detectable shift of the SNAP-deaminase toward higher molecular weight. Indeed, excellent control of the conjugate assembly was achieved in a clearly light-dose-dependent manner, shifting the SNAP-deaminase from non-conjugated to nearly complete conversion following kinetics agreeing with the decaging kinetics of the NpomBG precursor described above. To study the light-dependent in vitro RNA editing reaction, a master mix containing all components was aliquoted in the dark into PCR tubes and aliquots were irradiated individually with 365 nm light for 0, 15, 30, 60, or 180 s, respectively, prior to starting the editing reaction by incubation at 37 °C. A guideRNA lacking the BG moiety served as a negative, and a guideRNA modified with authentic BG served as a positive control. No editing was observed in the negative control. Similarly, only very minor editing above background was detectable in the non-irradiated sample with the Npom-caged BG-guideRNA. However, upon irradiation editing was restored in a light-dose-dependent manner regaining an editing level comparable to that of the positive control (Figure 2c, 75% with SNAP-ADAR1; Figure S16, 60% with SNAP-ADAR2). Plotting the intermediate editing levels against the irradiation time resulted in first-order kinetics (Figure 2c, $t_{1/2} = 26 \pm 0.5$ s) very similar to those obtained with the small-molecule precursor (Figure 1b).

RNA Editing Is Controllable by Light in Living Cells.

For the study of intracellular light-activated RNA editing, we incorporated the N7-NpomBG into a 19 nt Antagomir-like²⁵ chemically stabilized nucleic acid analogue¹¹ that contained a gap of three natural ribonucleotides around the editing site. We applied Antagomir-like chemistry to improve the stability of the guideRNA and the selectivity of the editing reaction;¹¹ this has been shown for RNA interference before.²⁶ The guideRNA targets a premature amber stop codon (UAG) at an eGFP reporter (W58amber), and successful editing is indicated by turn-on of eGFP fluorescence. A guideRNA with authentic BG served as a positive and the same guideRNA lacking the BG moiety as a negative control. Further controls were done to test the necessity of all components of the machinery. SNAP-ADAR1 and the reporter gene were transiently overexpressed from plasmids in HEK293T cells. One day after transfection of the plasmids, the respective guideRNA was lipofected into the cells. Four hours after lipofection, the medium was changed and cells were irradiated with 365 nm light under high control of

dosage and wavelength by using a fluorescence microscope equipped with a LED light source. One day later, the eGFP fluorescence was analyzed by microscopy before the RNA was isolated, and the editing yield was determined by Sanger sequencing. Compared to our previous protocol, we had to optimize the amounts and stoichiometry of SNAP-ADAR1 and guideRNA in order to suppress some low-level (~10%) editing caused by the Npom-protected guideRNA in the dark (for details, see Figures S19 and S20). The optimal amount of guideRNA used was 10 pmol/150 μ L and is in a range typical for siRNA duplexes. Applying the NH₂-guideRNA (negative editing control), only a very few cells developed a low-level GFP fluorescence and no editing was detectable in the sequencing trace (<5%). However, transfecting BG-guideRNA (positive editing control) gave brightly fluorescent cells, similar to the transfection of functional wt eGFP, and an editing yield of typically 45% was determined (Figures 3 and S17). Notably,

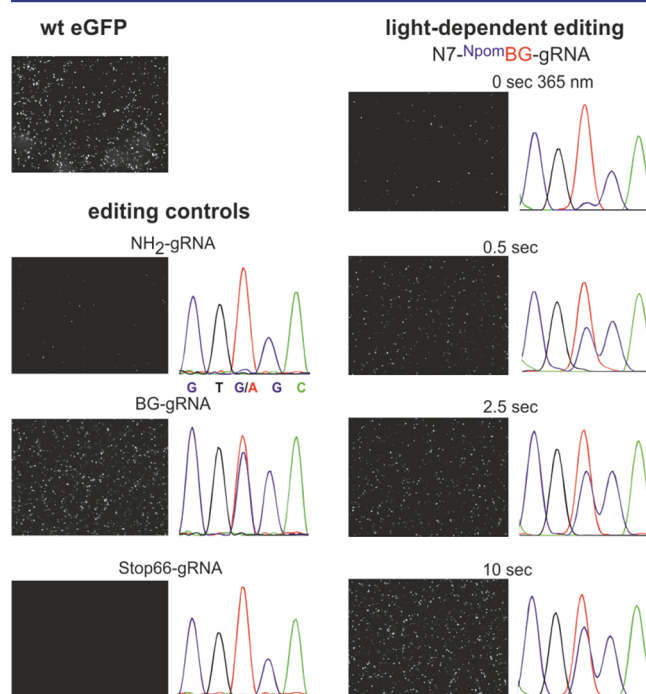


Figure 3. Light-controlled RNA editing in living 293T cells. SNAP-ADAR1 and the reporter gene (W58X eGFP, or wt eGFP) are provided on plasmids, the guideRNAs are reverse-transfected, all as described in the Supporting Information. Shown is the fluorescence microscopy analysis together with the respective Sanger sequencing traces 24 h post-transfection of the respective guideRNA. 5'-Terminal, the guideRNAs are either carrying an aminolinker (NH₂), the BG moiety (BG), or the Npom-protected BG moiety. The Stop66-guideRNA is a negative control BG-guideRNA targeting the GFP gene around codon 66 instead of codon 58.

no other edited site was observed in the reporter transcript. Furthermore, absolutely no editing was obtained at codon 58 by a chemically stabilized, negative control guideRNA that directs repair to codon 66. Thus, the formation of the RNA secondary structure directly at the targeted codon is strictly required for site-directed RNA editing and is the major determinant of specificity. The Npom-protected guideRNA gave only very low editing yield over background (<5%) and only a small number of low-intensity fluorescent cells. However, following illumination, editing was switched on to a level similar to that of the positive editing control, as indicated

by fluorescence microscopy but also by RNA sequencing (45% editing yield). Intermediate editing levels have been accessible by varying the light dose (Figure 3; more details can be found in Figure S17). The light dose applied to photoactivate editing was well tolerated by the cells. In comparison to the Npom-guideRNA, the editing yield of neither the positive nor the negative editing controls was dependent on light (Figure S17). No unspecific off-target editing was observed in the reporter gene, as indicated by Sanger sequencing (Figure S18).

Site-Directed RNA Editing in *Platynereis dumerilii*. As it requires massive genetic manipulation to switch the expression of one isoform to another that differs only in a single point mutation, site-directed RNA editing might offer a practical alternative.¹⁵ A light-triggered variant would be particularly attractive for developmental biology, as early stages are often transparent,²¹ and the spatiotemporal control of gene expression is of particular interest.^{27,28} An emerging model system for developmental and neurobiology is the marine annelid *Platynereis dumerilii*²⁹ that is readily cultivated³⁰ and easily manipulated at the one-cell zygote by microinjection.²⁷ To test site-directed RNA editing inside the worm, we injected two mRNA transcripts encoding SNAP-ADAR1 and eGFP together with chemically stabilized 21 nt guideRNAs. One day after microinjection, when the zygotes were developing into trochophore larvae, the fluorescence phenotype was analyzed by microscopy (Figures 4 and S22). A GFP-positive phenotype was only detectable in the positive control (wt GFP) and in the

editing sample (Figure 4a,f). All negative controls lacking parts of the machinery, such as the guideRNA, SNAP-ADAR1, or both, showed no green fluorescence (Figure 4b–d). In the editing sample as well as in the positive control, there was some heterogeneity of fluorescence intensity that may result from the difficulty of precisely controlling the injection volume. To determine the editing yield, a cohort of trochophores (each 80–100) were lysed, and RNA was extracted and analyzed by Sanger sequencing (Figure 4; for detailed analysis of all larvae, see Table S23). Editing was observed only when all components were included and achieved 60–70% over the entire population. No off-target editing was observed in the targeted transcript (Figure S24). The worms seem to develop and behave normally over the first days and stages of development.

Controlling Site-Directed RNA Editing in Living *P. dumerilii* by Light. Also in *Platynereis*, efficient editing requires assembly of the covalent guideRNA–deaminase conjugate and fails when using the NH₂-guideRNA lacking the BG moiety (Figure 4e). This encouraged us to test light-activated RNA editing inside the worm. For this, a guideRNA containing the Npom-protected BG was microinjected. In contrast to using the NH₂-guideRNA (Figure 4e) lacking the BG moiety, microinjection of the Npom-protected guideRNA resulted in a small but significant number of faintly fluorescing trochophores (18%) besides a large number of dark ones (>80%, Figure 5a, Table S23). In accordance with this, RNA sequencing of a cohort of 80–100 animals revealed a low but significant residual editing at the targeted stop codon (~10%). In faintly fluorescent trochophore larvae, the rhodamine signal was typically stronger (Figure 5a), indicating that the low-level editing might be due to an undesirably high injection volume of the editing components. This low-level residual editing activity is reminiscent of the situation described above for the light-dependent editing in cell culture. However, when the microinjected trochophores were treated with 365 nm light on a UV trans-illuminator (5 min), half of the trochophores developed a bright eGFP signal (Figure 5b, Table S23). The fluorescence imaging was in agreement with an editing yield of ~60%, as determined by RNA sequencing of 80–100 animals (Figure 5b). Thus, irradiation allows for activating RNA editing to a yield nearly identical to that of the positive editing control with an unprotected BG moiety (Figure 4f).

CONCLUSION

RNA-guided enzymes represent rationally programmable tools that allow for the efficient and precise manipulation of nucleic acids at specific sites in living organisms. Here, we further elaborate a novel strategy for site-directed adenosine-to-inosine RNA editing (a) by introducing photocontrol and (b) by applying the tool in developing *Platynereis dumerilii*.

The presented approach is unique in that the artificial RNA-guided editing enzyme is assembled via the formation of a single covalent bond.^{10,15} As covalent bond formation is essential for the functioning, photocontrol is feasible by blocking the SNAP-tag-mediated bond formation via installation of a single photoprotection group at the O6-benzyl-guanine moiety. Specifically, we demonstrate the ready synthesis of Npom-protected BG and its convenient introduction into diversely chemically modified antisense oligomers after their solid-phase synthesis. This is in contrast to other strategies that require the site-specific incorporation of (often several) photoprotected nucleosides during solid-phase

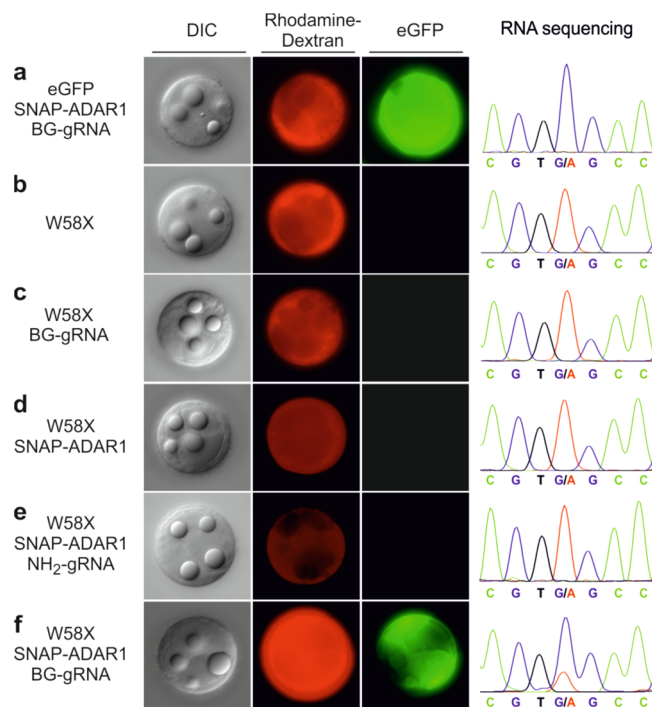


Figure 4. Site-directed RNA editing in *P. dumerilii*. Reporter mRNA (eGFP) and SNAP-ADAR1 mRNA were microinjected into one-cell zygotes, together with the respective BG/NH₂-guideRNA and rhodamine–dextran as an injection control. Shown are the fluorescence images of one representative embryo 24 hours post fertilization (hpf) for each experiment and the sequencing trace obtained from the RNA of 80–100 animals per experiment: (a) positive control, (b–e) negative controls lacking single components of the editing machinery, and (f) editing experiment. For details, see the Supporting Information. DIC = differential interference contrast.

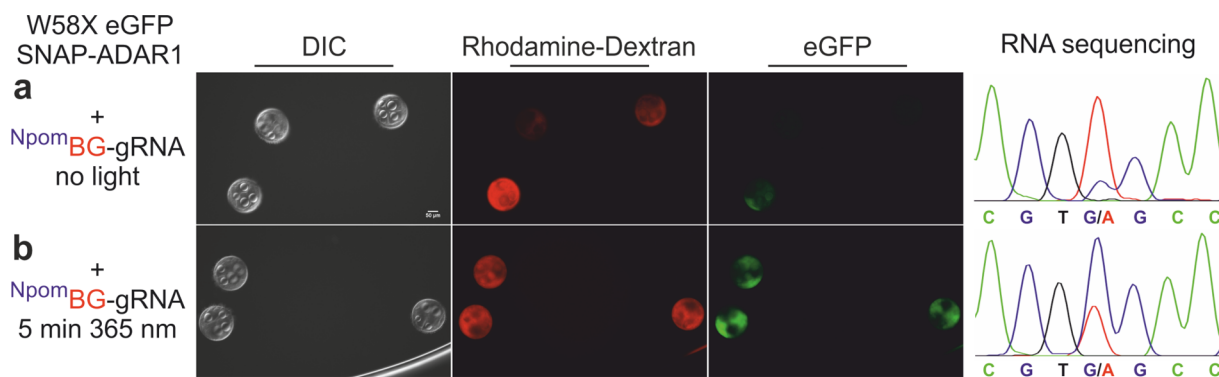


Figure 5. Light-dependent editing in *P. dumerilii*. Reporter gene and SNAP-ADAR1 have been microinjected into one-cell *Platyneris* zygotes as described in Figure 4, but now with a photoprotected chemically stabilized ^{Npm}BG-guideRNA. Within 1 h after microinjection, zygotes have been (a) kept in the dark or (b) treated with 365 nm light (5 min). Fluorescence images and RNA sequences (80–100 animals/experiment) are taken 24 hpf. The sequence of the guideRNA is the same as in Figure 4. For further details, see the Supporting Information.

oligonucleotide synthesis to achieve photocontrol of biochemical processes.^{21,23,31} Furthermore, our strategy needs less genetic engineering compared to the introduction of photo-responsive groups into enzymes by means of amber suppression or related strategies.^{32,33}

In vitro we could show that the attachment of the Npom group at N7, but not N9, of the BG moiety blocks the conjugation reaction with the SNAP-tag. However, reactive benzyl guanine is readily released upon 365 nm irradiation with high efficiency ($\epsilon\varphi \approx 2000 \text{ M}^{-1} \text{ cm}^{-1}$) and allows for the light-dose-dependent assembly of guideRNA–deaminase conjugates. Besides editing, the Npom-protected BG will be applicable in other approaches that rely on the SNAP-tag, like chemical inducers of dimerization.³⁴ By controlling the assembly of the editase, we could trigger the in vitro editing of a purified mRNA in a light-dose-dependent manner covering the whole dynamic range from absence of editing in the absence of light until full editing in the presence of light. The desired action of our tool could be directly translated into mammalian cell culture; however, optimization was required to control low-level residual editing by the photoprotected guideRNA. Again, a similarly high dynamic range was achieved. Furthermore, we established site-directed RNA editing for the first time in a living organism. Specifically, we achieved the efficient and highly selective switch of a premature stop into a tryptophan codon in developing *Platyneris dumerilii* zygotes. Notably, no genetic engineering and livestock breeding is required, thus circumventing time-consuming and cost-intensive laboratory work. As our editing tool is independent of any host-specific factors, the technology should be transferable to any other organism. In *Platyneris*, the covalent assembly of the guideRNA–deaminase conjugate was again essential, and our simple photocontrol strategy for site-directed RNA editing was directly transferable. The tool could now be further elaborated to achieve precise spatiotemporal control of protein isoforms in cellular networks or in developing *Platyneris*.

MATERIALS AND METHODS

Synthesis of Npom-Caged O6-Benzylguanine. Trifluoroacetamide protected O6-benzylguanine (BG, 120 mg, 0.33 mmol)¹⁹ was solved in dry DMF (1.2 mL) under argon. Diazabicycloundecene (150 μL , 153 mg) was added at room temperature, and the solution was stirred for 30 min. Npom chloride (0.5 mmol, ~ 1.5 equiv, dissolved in 1.6 mL of DMF) prepared *in situ* as described²² was added dropwise. After 2.5 h, the reaction mixture was diluted with EtOAc, washed with 1% citric acid (3 \times) and brine (1 \times), and dried over Na_2SO_4 . The

evaporated crude product was cleaned via silica chromatography (2–4% MeOH in DCM) and yielded 24 mg (21%) of ^{N7}Npom-BG-TFA and 50 mg (42%) of ^{N9}Npom-BG-TFA. For full characterization and assignment of the isomers and downstream synthesis, see the Supporting Information.

N7 Isomer. ¹H NMR (600 MHz, DMSO-*d*₆): δ = 9.99 (t, *J* = 5.9 Hz, 1H), 8.07 (s, 1H), 7.48 (d, *J* = 8.1 Hz, 2H), 7.35 (s, 1H), 7.28 (d, *J* = 8.1 Hz, 2H), 6.84 (s, 1H), 6.22 (s, 2H), 6.16 (s, 1H), 6.04 (s, 1H), 5.54 (m, 2H), 5.46 (m, 2H), 5.13 (q, *J* = 6.2 Hz, 1H), 4.39 (d, *J* = 5.9 Hz, 2H), 1.33 (d, *J* = 6.2 Hz, 3H). ¹³C NMR (151 MHz, DMSO-*d*₆): δ = 164.1, 159.8, 156.4 (q, ²*J*(C,F) = 36 Hz), 156.3, 151.8, 146.5, 145.9, 140.5, 137.2, 136.2, 135.5, 127.9, 127.5, 116.0 (q, ¹*J*(C,F) = 288 Hz), 105.6, 105.2, 104.3, 103.2, 74.9, 72.0, 66.6, 42.4, 23.3. HR-ESI-MS: [M + H]⁺(_{theoretical}) = 590.16056 for C₂₅H₂₃F₃N₇O₇; found 590.16118. R_f(DCM/MeOH, 98:2) = 0.08. R_f(DCM/MeOH, 95:5) = 0.50.

N9 Isomer. ¹H NMR (600 MHz, DMSO-*d*₆): δ = 10.00 (t, *J* = 6.0 Hz, 1H), 7.80 (s, 1H), 7.49 (d, *J* = 8.1 Hz, 2H), 7.46 (s, 1H), 7.31 (d, *J* = 8.1 Hz, 2H), 6.97 (s, 1H), 6.34 (s, 2H), 6.15 (s, 1H), 6.03 (s, 1H), 5.43–5.49 (m, 2H), 5.40 (d, ²*J* = 11.4 Hz), 5.32 (d, ²*J* = 11.4 Hz), 5.21 (q, *J* = 6.3 Hz, 1H), 4.41 (d, *J* = 6.0 Hz, 2H), 1.38 (d, *J* = 6.3 Hz, 3H). ¹³C NMR (151 MHz, DMSO-*d*₆): δ = 160.8, 157.2 (q, ²*J*(C,F) = 36 Hz), 155.3, 152.7, 147.5, 142.1, 140.5, 138.2, 137.0, 136.6, 129.6, 128.3, 116.9 (q, ¹*J*(C,F) = 288 Hz), 114.3, 106.7, 105.2, 104.1, 73.1, 71.7, 67.5, 43.3, 24.1. HR-ESI-MS: [M + Na]⁺(_{theoretical}) = 612.14250 for C₂₅H₂₂F₃N₇O₇Na; found 612.14262. R_f(DCM/MeOH, 98:2) = 0.32. R_f(DCM/MeOH, 95:5) = 0.55.

Light-Triggered In Vitro RNA Editing. Purified SNAP-ADAR1 (170 nM), purified eCFP mRNA (10 nM), and one of the respective guideRNAs (50 nM) were prepared in buffer (25 mM Tris-HCl, 0.75 mM MgCl₂, 75 mM KCl, 2 μM heparin, and 640 u/mL murine RNase inhibitor, 10 mM DTT, pH 8.3) in PCR tubes. Irradiation with 365 nm light was performed on a UV trans-illuminator (UVP TFL-40V, 25 W, intensity high) for the indicated amount of time at room temperature. Subsequent editing was performed by incubation for 120 min while cycling between 30 and 37 °C. Reactions were stopped by heating to 70 °C for 3 min and subsequent reverse transcription. After PCR amplification of the cDNA, editing yields were estimated by the relative height of the guanosine versus adenosine traces by Sanger sequencing. All experiments were done in at least two replicates. Sequence of the guideRNAs: (Npom)BG/NH₂-UCG-GAACACCCC-AGCACAGA-3' (natural ribonucleotides; 5'-terminal modifications were introduced via amino-linker, the 5'-terminal three nucleotides serve as linker and do not base-pair with the target).

Light-Triggered Cellular RNA Editing. Cells (293T: DSMZ code ACC-635; 200 000 cells/well) were seeded on 24-well plates in full media (DMEM, 10% FBS, 1% penicillin/streptomycin, grown in 5% CO₂, 37 °C). At 60–80% confluency, plasmid pcDNA3.1 vector (Life Technologies) carrying SNAP-ADAR1 (100 ng/well) and pcDNA3.1 vector carrying the respective eGFP variant (500 ng/well)¹¹ were co-transfected with Lipofectamine 2000 (4 $\mu\text{L}/\mu\text{g}$).¹¹

After 24 h, the cells were reverse transfected into 96-well plates (60 000 cells/well) containing the respective guideRNAs (10 pmol/well) pretreated with Lipofectamine 2000 (0.5 μ L/well). Four hours after reverse transfection, media was replaced with DMEM without FBS and phenol red, containing HEPES (25 mM). Irradiation (365 nm) was performed in a fluorescence microscope (Zeiss CellObserverZ.1, equipped with a 365 nm Colibri.2 LED) at 100% LED power for the indicated amount of time. Twenty-four hours later, the fluorescence phenotype was analyzed by fluorescence microscopy (Zeiss CellObserverZ.1), and RNA was extracted using the RNeasy MinElute Cleanup Kit (Qiagen). After reverse transcription and PCR amplification, the editing yield was estimated by Sanger sequencing. All experiments were done in at least two replicates. The sequence of the W58X guideRNAs was (Npom)BG/NH₂-UsAsU-GUGUCGG-CCA-CGGAAsCsAsGsG-3'; the sequence of the Stop66-guideRNA was BG-UsCsG-GAACACC-CCA-GCAsCsAsGsA-3' (s = phosphothioate linkage; plain font indicates 2'-methoxyribonucleotides, and italic underlined indicates unmodified ribonucleotides; the three 5'-terminal nucleotides serve as a linker and do not base-pair with the target).

Editing in *Platynereis dumerilii*. For the in vitro transcription of stabilized mRNAs of SNAP-ADAR1 and eGFP variants with the mMACHINE mMACHINE T7 Ultra Kit (Life Technologies), the respective genes were subcloned into the pUC57-T7-RPP2 vector, resulting in 5'-capped and 3'-polyadenylated transcripts additionally stabilized by a *Platynereis*-specific RPP2 5'-UTR, as described before.³⁵ mRNA transcripts were cleaned by the RNeasy MinElute Cleanup Kit (Qiagen). GuideRNAs were precipitated with potassium acetate prior to use. Fertilized zygotes were obtained from an in-house breeding culture and were incubated at 14.8 °C for 55 min. Prior to microinjection, the egg jelly was removed by rinsing the zygotes with natural seawater (NSW) in a 100 μ m sieve. To soften the vitellin envelope, a 1-min-long proteinase K treatment (70 μ g/mL) was performed as described before.²⁷ Around 100 zygotes were embedded in the injection stage (2% agarose in NSW). Samples were injected using Femtotips II microcapillaries with a Femtojet express microinjector (700 hPa injection pressure, 0.1 s injection time, 35 hPa compensation pressure) in a cooled (14.8 °C) Zeiss Axiovert 40 CL microscope equipped with a Luigs and Neumann micro-manipulator as described before.²⁷ Injection started 1 hours post fertilization (hpf) and was stopped when the first cleavage was detected (ca. 2 hpf). Irradiation at 365 nm was performed immediately after microinjection for the indicated amount of time on a UV transilluminator (UVP TFL-40V, 25 W, intensity high). Microinjected zygotes were bred at 19 °C in Nunclon six-well plates containing 6 mL NSW. Twenty-four hpf, healthy larvae (early trochophore) were separated from unhealthy ones. The fluorescence phenotype was analyzed by microscopy (Axio Imager Z1). RNA from 80–100 healthy larvae (two injection sessions) was isolated 25 hpf by shock freezing (liquid nitrogen), shear forces (passing through 0.6 mm needle), vortexing (10 s), and subsequent use of the RNeasy MinElute Cleanup Kit (Qiagen). After reverse transcription and PCR amplification, the editing yield was determined by Sanger sequencing. Injection samples contained 1.5 μ g/ μ L rhodamine-dextran (10 kDa MW, Sigma) for injection control, 250 ng/ μ L of the respective reporter mRNA, 450 ng/ μ L SNAP-ADAR1 mRNA, and 25 μ M of the respective guideRNA. Sequence of the guideRNA: BG/NH₂-UsAsU-GUGUCGG-CCA-CGGAACAsGsCsA-3' (s = phosphothioate linkage; plain font indicates 2'-methoxy ribonucleotides, and italic underlined indicates unmodified ribonucleotides; the 5'-terminal three nucleotides serve as linker and do not base-pair with the target).

■ ASSOCIATED CONTENT

Supporting Information

The Supporting Information is available free of charge on the ACS Publications website at DOI: 10.1021/jacs.5b10216.

Chemical synthesis, compound characterization, molecular biology, and editing experiments, including Figures S1–S26 and Tables S1–S23 (PDF)

■ AUTHOR INFORMATION

Corresponding Author

*thorsten.stafforst@uni-tuebingen.de

Notes

The authors declare no competing financial interest.

■ ACKNOWLEDGMENTS

We gratefully acknowledge support from the University of Tübingen, the Deutsche Forschungsgemeinschaft (STA 1053/3-2, STA 1053/4-1), and the Max Planck Society. This work has received funding from the European Research Council (ERC) under the European Union's Horizon 2020 research and innovation program (grant agreement no. 647328). The research leading to these results received funding from the ERC under the European Union's Seventh Framework Programme (FP7/2007-2013)/ERC grant agreement no. 260821. The authors thank Aurora Panzera for technical assistance during microinjection.

■ REFERENCES

- (1) Zhao, X.; Yu, Y.-T. *Nat. Methods* **2008**, *5*, 95–100.
- (2) Karijolic, J.; Yu, Y.-T. *Nature* **2011**, *474*, 395–398.
- (3) Dorsett, Y.; Tuschl, T. *Nat. Rev. Drug Discovery* **2004**, *3*, 318–329.
- (4) Kim, H.; Kim, J.-S. *Nat. Rev. Genet.* **2014**, *15*, 321–334.
- (5) Machnicka, M. A.; Milanowska, K.; Oglou, O. O.; Purta, E.; Kurkowska, M.; Olchowik, A.; Januszewski, W.; Kalinowski, S.; Dunin-Horkawicz, S.; Rother, K. M.; Helm, M.; Bujnicki, J. M.; Grosjean, H. *Nucleic Acids Res.* **2013**, *41*, D262–D267.
- (6) Liu, N.; Pan, T. *Transl. Res.* **2015**, *165*, 28–35.
- (7) Li, J. B.; Church, G. M. *Nat. Neurosci.* **2013**, *16*, 1518–22.
- (8) Jinek, M.; Chylinski, K.; Fonfara, I.; Hauer, M.; Doudna, J. A.; Charpentier, E. *Science* **2012**, *337*, 816–821.
- (9) Watkins, N. J.; Bohnsack, M. T. *WIREs RNA* **2012**, *3*, 397–414.
- (10) Stafforst, T.; Schneider, M. F. *Angew. Chem., Int. Ed.* **2012**, *51*, 11166–11169.
- (11) Vogel, P.; Schneider, M. F.; Wettengel, J.; Stafforst, T. *Angew. Chem., Int. Ed.* **2014**, *53*, 6267–6271.
- (12) Schneider, M. F.; Wettengel, J.; Hoffmann, P. C.; Stafforst, T. *Nucleic Acids Res.* **2014**, *42*, e87.
- (13) Nishikura, K. *Annu. Rev. Biochem.* **2010**, *79*, 321–349.
- (14) Bass, B. L. *Annu. Rev. Biochem.* **2002**, *71*, 817–846.
- (15) Vogel, P.; Stafforst, T. *ChemMedChem* **2014**, *9*, 2021–2025.
- (16) Reenan, R. N. *Engl. J. Med.* **2014**, *370*, 172–174.
- (17) Montiel-Gonzalez, M. F.; Vallecillo-Viejo, I.; Yudowski, G. A.; Rosenthal, J. J. C. *Proc. Natl. Acad. Sci. U. S. A.* **2013**, *110*, 18285–18290.
- (18) Kole, R.; Krainer, A. R.; Altman, S. *Nat. Rev. Drug Discovery* **2012**, *11*, 125–140.
- (19) Keppler, A.; Gendreizig, S.; Gronemeyer, T.; Pick, H.; Vogel, H.; Johnsson, K. *Nat. Biotechnol.* **2003**, *21*, 86–89.
- (20) Gautier, A.; Juillerat, A.; Heinis, C.; Corrêa, I. R., Jr.; Kindermann, M.; Beaufils, F.; Johnsson, K. *Chem. Biol.* **2008**, *15*, 128–136.
- (21) Brieke, C.; Rohrbach, F.; Gottschalk, A.; Mayer, G.; Heckel, A. *Angew. Chem., Int. Ed.* **2012**, *51*, 8446–8476.
- (22) Lusic, H.; Deiters, A. *Synthesis* **2006**, 2147–2150.
- (23) Connelly, C. M.; Uprety, R.; Hemphill, J.; Deiters, A. *Mol. Biosyst.* **2012**, *8*, 2987–2993.
- (24) Banala, S.; Arnold, A.; Johnsson, K. *ChemBioChem* **2008**, *9*, 38–41.
- (25) Krützfeldt, J.; Rajewsky, N.; Braich, R.; Rajeev, K. G.; Tuschl, T.; Manoharan, M.; Stoffel, M. *Nature* **2005**, *438*, 685–689.
- (26) Bramsen, J. B.; Laursen, M. B.; Nielsen, A. F.; Hansen, T. B.; Bus, C.; Langkjær, N.; Babu, B. R.; Højland, T.; Abramov, M.; Van Aerschot, A.; Odadzic, D.; Smcius, R.; Haas, J.; Andree, C.; Barman, J.;

Wenska, M.; Srivastava, P.; Zhou, C.; Honcharenko, D.; Hess, S.; Müller, E.; Bobkov, G. V.; Mikhailov, S. N.; Fava, E.; Meyer, T. F.; Chattopadhyaya, J.; Zerial, M.; Engels, J. W.; Herdewijn, P.; Wengel, J.; Kjemis, J. *Nucleic Acids Res.* **2009**, *37*, 2867–2881.

(27) Conzelmann, M.; Williams, E. A.; Tunaru, S.; Randel, N.; Shahidi, R.; Asadulina, A.; Berger, J.; Offermanns, S.; Jékely, G. *Proc. Natl. Acad. Sci. U. S. A.* **2013**, *110*, 8224–8229.

(28) Williams, E. A.; Conzelmann, M.; Jékely, G. *Front. Zool.* **2015**, *12*, 1–15.

(29) Fischer, A.; Dorresteijn, A. *BioEssays* **2004**, *26*, 314–325.

(30) Garcia-Alonso, J.; Smith, B. D.; Rainbow, P. S. *Pan-Am. J. Aquat. Sci.* **2013**, *8*, 142–146.

(31) Hemphill, J.; Liu, Q.; Uprety, R.; Samanta, S.; Tsang, M.; Juliano, R. L.; Deiters, A. *J. Am. Chem. Soc.* **2015**, *137*, 3656–62.

(32) Gautier, A.; Gauron, C.; Volovitch, M.; Bensimon, D.; Jullien, L.; Vríz, S. *Nat. Chem. Biol.* **2014**, *10*, 533–541.

(33) Hemphill, J.; Borchardt, E. K.; Brown, K.; Asokan, A.; Deiters, A. *J. Am. Chem. Soc.* **2015**, *137*, 5642–45.

(34) Erhart, D.; Zimmermann, M.; Jacques, O.; Wittwer, M. B.; Ernst, B.; Constable, E.; Zvelebil, M.; Beaufils, F.; Wymann, M. P. *Chem. Biol.* **2013**, *20*, 549–57.

(35) Randel, N.; Asadulina, A.; Bezares-Calderón, L. A.; Verasztó, C.; Williams, E. A.; Conzelmann, M.; Shahidi, R.; Jékely, G. *eLife* **2014**, *3*, e02730.

Supporting Information

Site-directed RNA editing in vivo can be triggered by the light-driven assembly of an artificial riboprotein

Alfred Hanswillemenke¹, Tahsin Kuzdere¹, Paul Vogel¹, Gaspar Jekely², Thorsten Stafforst¹

¹Interfaculty Institute of Biochemistry; University of Tübingen
Auf der Morgenstelle 15; 72076 Tübingen (Germany)

²Max-Planck-Institute for Developmental Biology
Spemannstraße 35; 72076 Tübingen (Germany)

Supplementary Information on chemical synthesis, compound characterization and additional editing data is available in the online version of the paper. Correspondence and requests for materials should be addressed to TS (thorsten.stafforst@uni-tuebingen.de).

Content:

Chemical Synthesis	2
Synthesis of Npom-BG derivatives	2
Synthesis and MALDI characterization of Npom-guideRNAs	13
Experiments with Npom-caged O6-Benzylguanine	18
Photo-deprotection kinetics of ^{N7} Npom-BG-TFA and ^{N9} Npom-BG-TFA	18
BG-FITC/FAM assay and SNAP-ADAR1 band-shift assay	20
Editing <i>in vitro</i>	22
Primers & Sequences for <i>in vitro</i> and in cell culture experiments	22
Methods	29
Additional experimental data	32
Editing in cell culture	33
Methods	33
Additional experimental data	38
Editing in <i>Platynereis Dumerilii</i>	40
Gene sequences	40
Methods	44
Additional experimental data	48
Supporting Literature	53

Chemical Synthesis

Chemicals

If not stated otherwise, all substrates and reagents required for synthesis and biochemical studies were purchased from commercial providers and used without further purification.

General Methods

All column chromatographic purifications were carried out on self-packed columns of silica gel (0.04-0.063 mm/230-240 mesh). Thin-layer chromatography (TLC) was performed on silica gel sheets (60 F254, 0.2 mm, 5 x 10 cm, Merck) and visualized under UV light (254 nm). All analytical and preparative HPLC runs were performed on a Shimadzu system (SCL-10A VP, SPD-20AV, LC-20AT) running with 0.1% TFA in water (Eluent A) and 0.1% TFA in acetonitrile/water (9:1, Eluent B). Analytical HPLC was performed using an EC 125/4 Nucleodur C18 column by Machery + Nagel and preparative HPLC was performed using a VP 250/10 Nucleodur C18 column by Machery + Nagel. ¹H and ¹³C spectra were recorded on a Bruker ARX 250 or a Bruker Avance 400. ¹H, ¹³C, HCCH-Cosy, HSQC and HMBC experiments for the discrimination of Npom-BG isomers (**1a/1b**) were performed using a Bruker AMX-600. High resolution mass spectrometry was performed on a maXis4G ESI-TOF-MS by Bruker Daltonics.

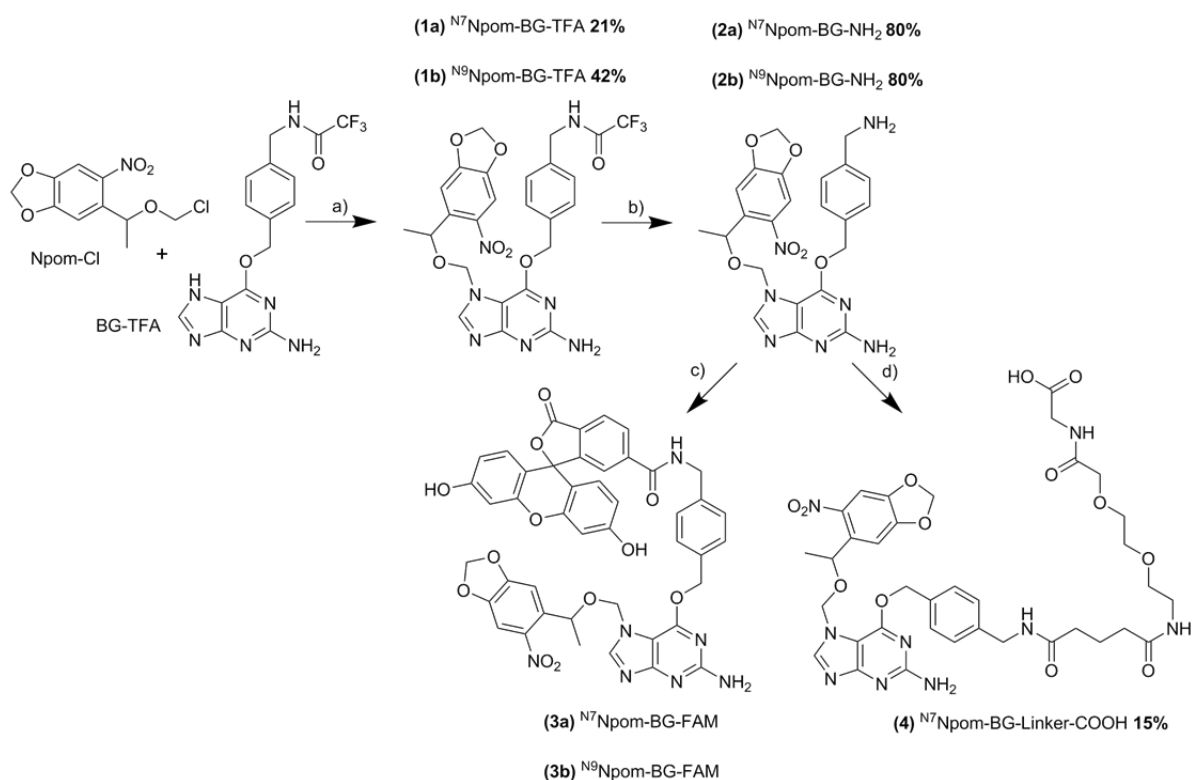


Figure S1. Synthesis of Npom-caged BG- derivatives. **a)** DBU, DMF, r.t., 2.5 h **b)** K₂CO₃, MeOH, 50°C, 4 h **c)** OSu-FAM, TEA, DMF, r.t., 1 h **d)** Solid-phase peptide synthesis (see page S12).

Synthesis of ^{N7/N9}Npom-BG-TFA (1a/1b)

BG-TFA (120 mg, 0.33 mmol, 1 eq; synthesized according to Keppler et al.^[S1]) was solved in DMF (1.2 ml) in a dry flask under argon atmosphere. Diazabicycloundecen (DBU, 150 µl, 153 mg, 1 mmol, 3 eq) was added at r.t. and the solution was stirred for 30 min. Npom-Cl (0.5 mmol, ~1.5 eq, *in situ*; synthesized according to Lusic et al.^[S2]) was solved in DMF (1.6 ml) and added drop wise. After 2.5 h stirring at r.t., the reaction mixture was diluted with EtOAc, washed three times with 1% citric acid and Brine and dried over Na₂SO₄. Evaporation of the organic phase resulted in 180 mg crude product. Silica gel column chromatography (2-4% MeOH in DCM) yielded 120 mg (63%) total product. The isomers were cleanly separable and were obtained in a 1:2 ratio (21% yield of ^{N7}Npom-BG-TFA (**1a**) and 42% yield of ^{N9}Npom-BG-TFA (**1b**)).

R_f(DCM/MeOH, 98:2): N7 = 0.08, N9 = 0.32

R_f(DCM/MeOH, 95:5): N7 = 0.50, N9 = 0.55

^{N7}Npom-BG-TFA, characterization:

¹H NMR (600 MHz, DMSO-d6): δ = 9.99 (t, J = 5.9 Hz, 1H), 8.07 (s, 1H), 7.48 (d, J = 8.1 Hz, 2H), 7.35 (s, 1H), 7.28 (d, J = 8.1 Hz, 2H), 6.84 (s, 1H), 6.22 (s, 2H), 6.16 (s, 1H), 6.04 (s, 1H), 5.54 (m, 2H), 5.46 (m, 2H), 5.13 (q, J = 6.2 Hz, 1H), 4.39 (d, J = 5.9 Hz, 2H), 1.33 (d, J = 6.2 Hz, 3H).

¹³C NMR (151 MHz, DMSO-d6): 164.1, 159.8, 156.4 (q, $^2J(C,F)$ = 36.2 Hz), 156.3, 151.8, 146.5, 145.9, 140.5, 137.2, 136.2, 135.5, 127.9, 127.5, 116.0 (q, $^1J(C, F)$ = 288.4 Hz), 105.6, 105.2, 104.3, 103.2, 74.9, 72.0, 66.6, 42.4, 23.3.

HR-ESI-MS: $[M+H]^+$ (theoretical) = 590.16056 for C₂₅H₂₃F₃N₇O₇;
 $[M+H]^+$ (found) = 590.16118

^{N9}Npom-BG-TFA, characterization:

¹H NMR (600 MHz, DMSO-d6): 10.00 (t, J = 6.0 Hz, 1H), 7.80 (s, 1H), 7.49 (d, J = 8.1 Hz, 2H), 7.46 (s, 1H), 7.31 (d, J = 8.1 Hz, 2H), 6.97 (s, 1H), 6.34 (s, 2H), 6.15 (s, 1H), 6.03 (s, 1H), 5.43-5.49 (m, 2H), 5.40 (d, 2J = 11.4 Hz), 5.32 (d, 2J = 11.4 Hz), 5.21 (q, J = 6.3 Hz, 1H), 4.41 (d, J = 6.0 Hz, 2H), 1.38 (d, J = 6.3 Hz, 3H).

¹³C NMR (151 MHz, DMSO-d6): 160.8, 157.2 (q, $^2J(C,F)$ = 36.2 Hz), 155.3, 152.7, 147.5, 142.1, 140.5, 138.2, 137.0, 136.6, 129.6, 128.3, 116.9 (q, $^1J(C,F)$ = 288.4 Hz), 114.3, 106.7, 105.2, 104.1, 73.1, 71.7, 67.5, 43.3, 24.1.

HR-ESI-MS: $[M+Na]^+$ (theoretical) = 612.14250 for C₂₅H₂₂F₃N₇O₇Na;
 $[M+Na]^+$ (found) = 612.14262

The N7- and N9- isomer of Npom-BG-TFA were characterized by 1D- and 2D-NMR spectroscopy (see Figure S2-5). All peaks in the ¹H and ¹³C spectra were assigned supported by the data from heteronuclear multiple-bond correlation spectroscopy (HMBC) experiments. The carbon atoms of the TFA-protection group show characteristic coupling to the ¹⁹F atoms.

In the HMBC spectrum of ^{N7}Npom-BG-TFA a crosspeak between the protons of the oxymethyl bridge of the Npom-group and C5 of guanosine was observed, while a cross-peak between the oxymethyl bridge and C4 was observed for ^{N9}Npom-BG-TFA. Figure S5 shows the HMBC spectra of both isomers and signals relevant for the assignment of the regioisomers are highlighted. The Npom substitution has a strong effect on chemical shifts of C4 (N7: 165.14 ppm, N9: 155.25 ppm) and C5 (N7: 105.23 ppm, N9: 114.28 ppm). Furthermore, the oxymethyl bridge of the Npom group and the oxygen and O6 methylen group of BG are in close proximity in the case of ^{N7}Npom-BG-TFA, shifting the 1H signal of the oxymethyl bridge from 5.36 ppm (N9) to 5.53 ppm (N7).

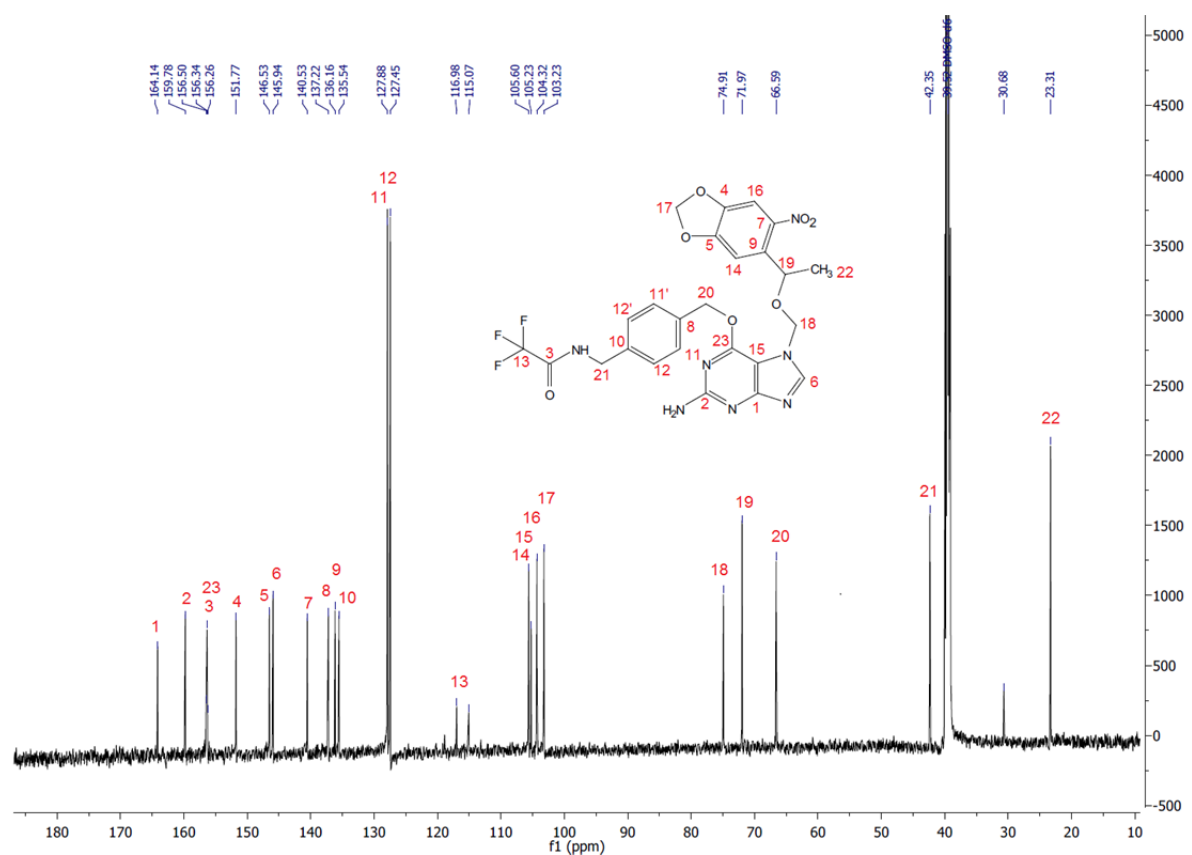
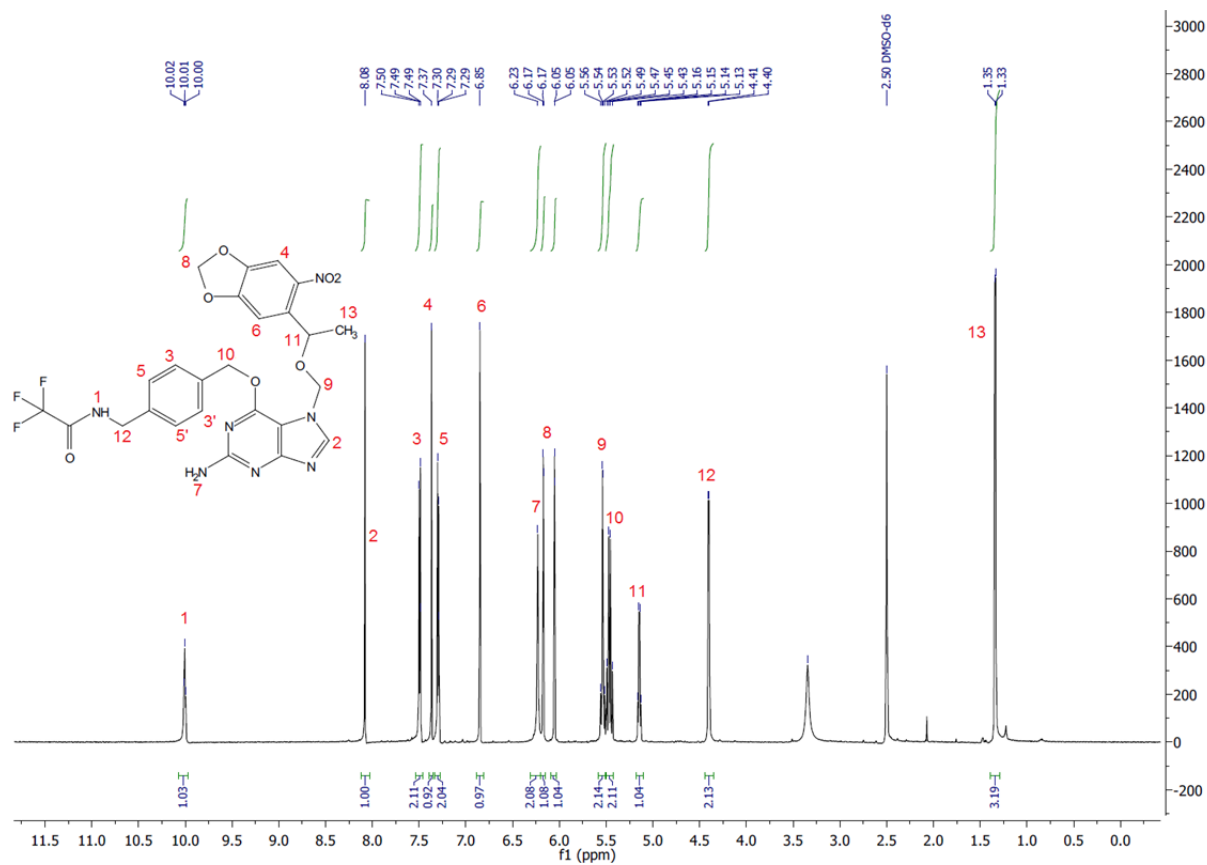


Figure S2. 1H- & 13C-NMR spectra of ^{N7}Npom-BG-TFA (**1a**).

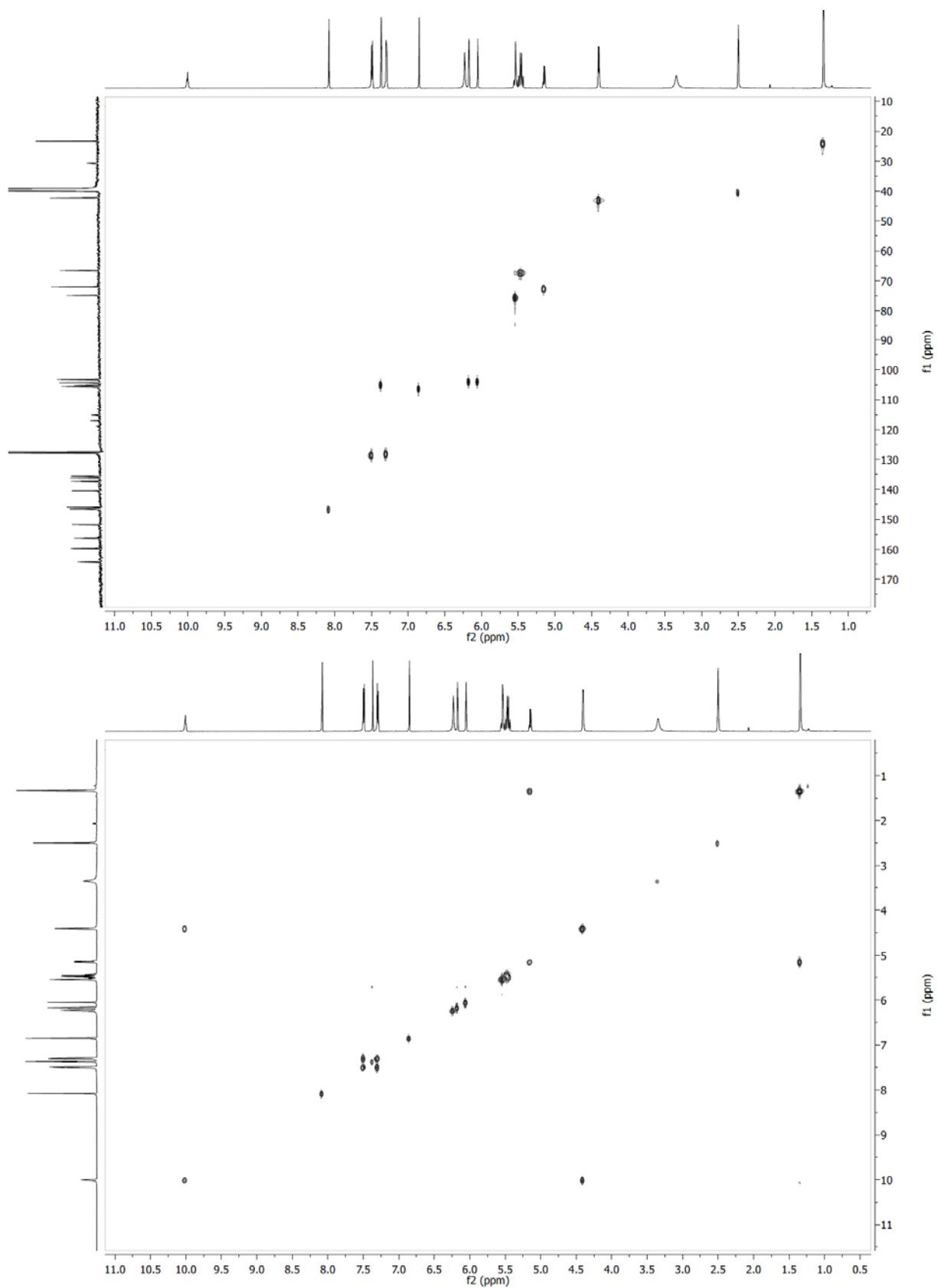


Figure S3. HSQC- and H,H-COSY of ^{15}N pom-BG-TFA (**1a**).

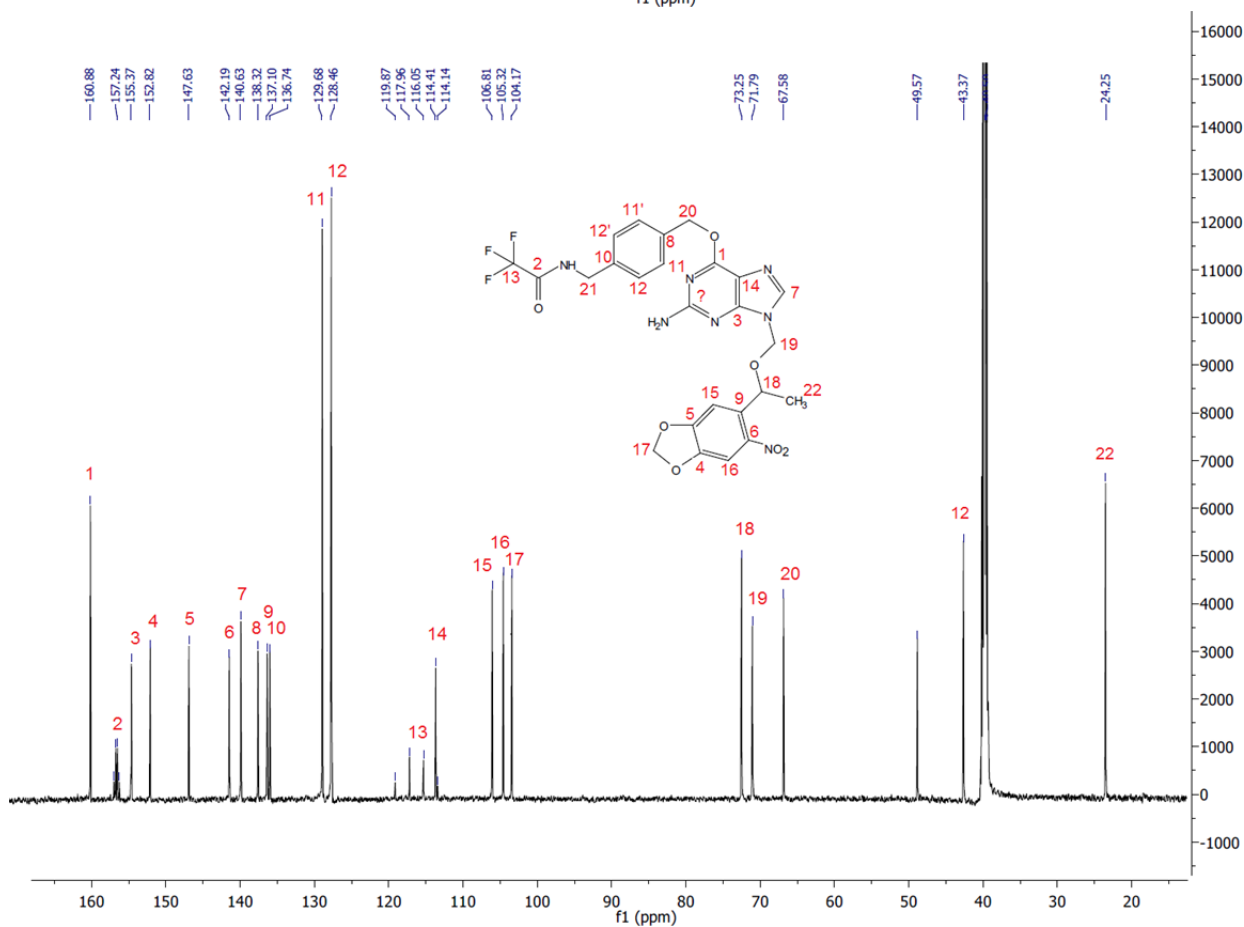
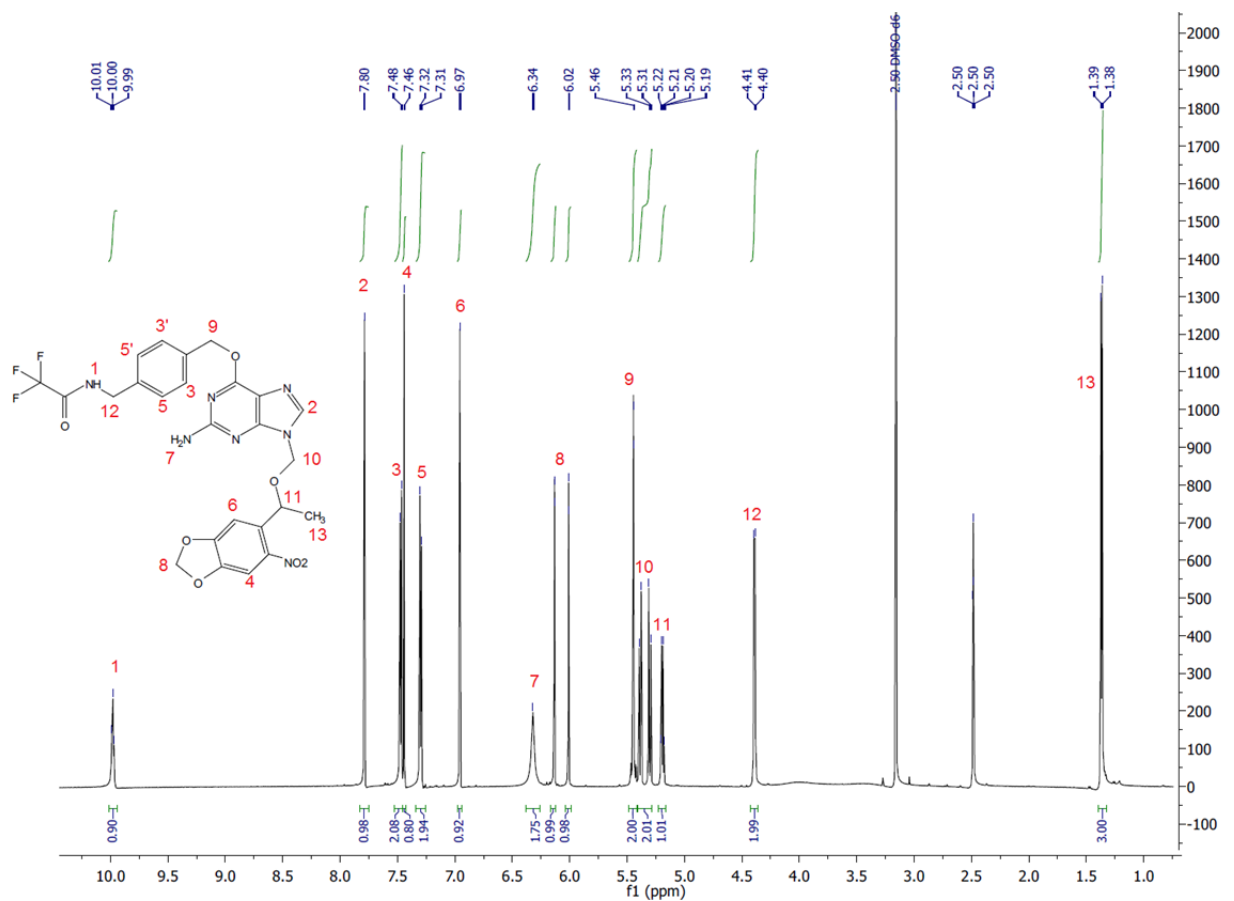


Figure S4. 1H- & 13C-NMR spectra of N⁹Npom-BG-TFA (**1b**).

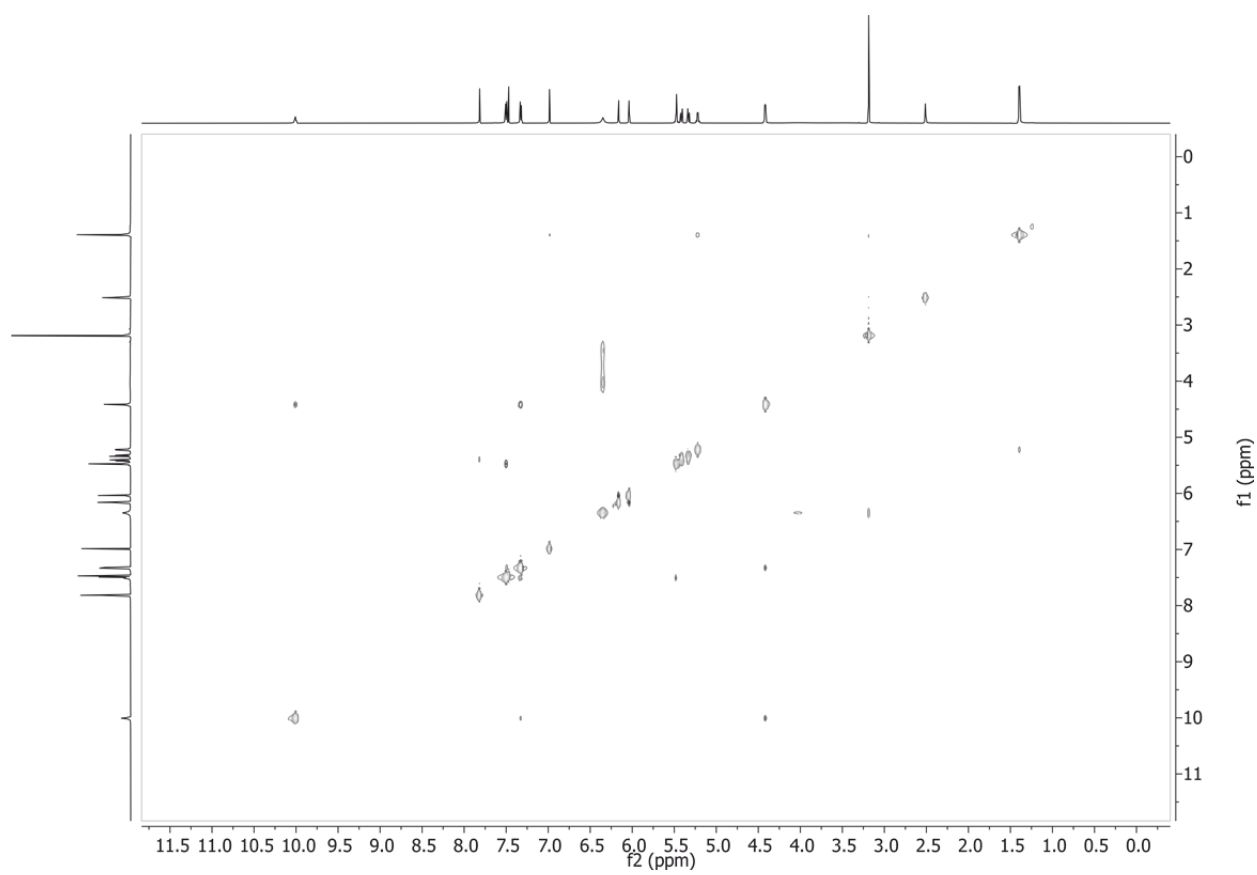


Figure S4 continued H,H-COSY of N^9 Npom-BG-TFA (**1b**).

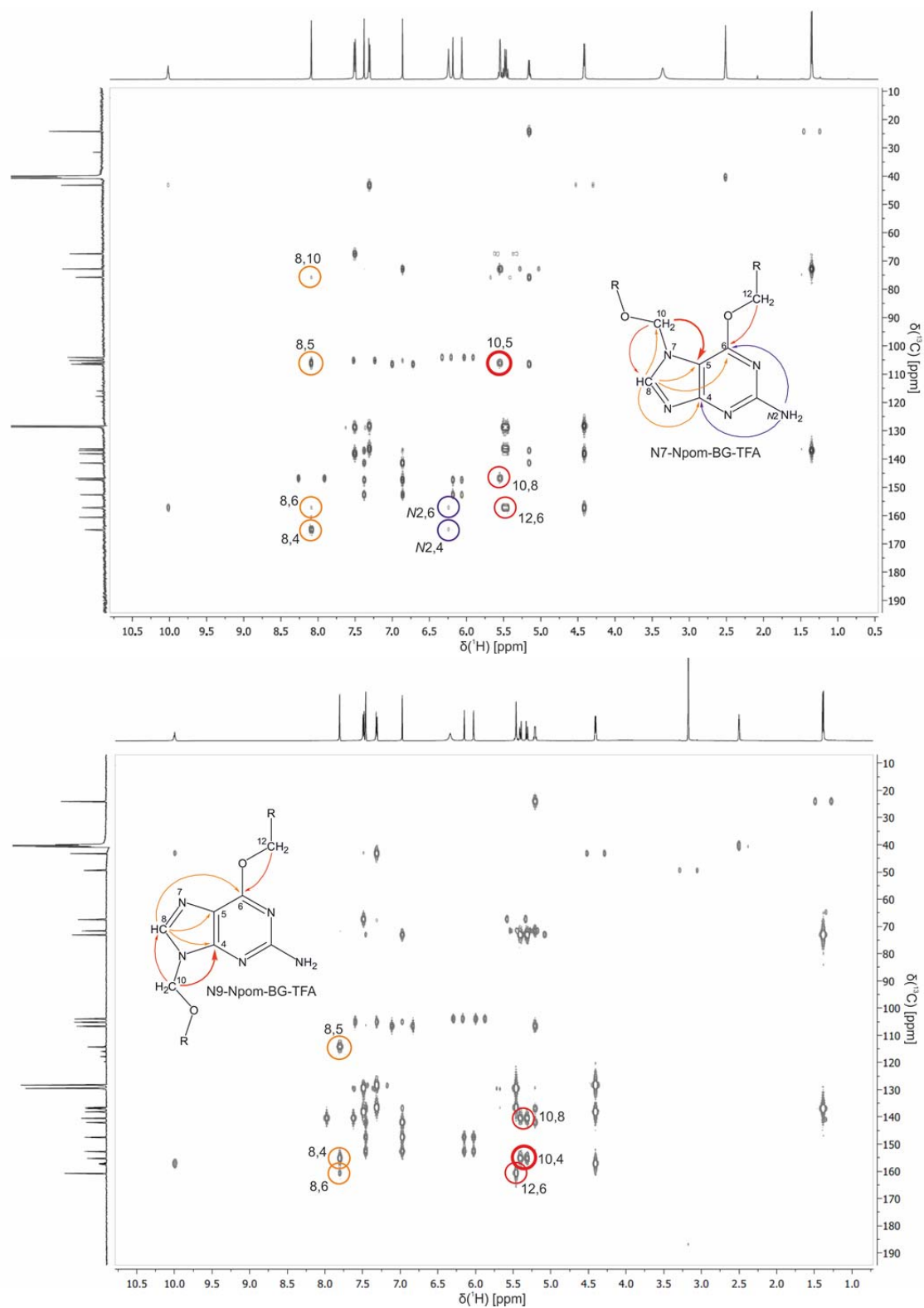


Figure S5. Assignment of the regio-isomers via HMBC NMR experiments of compounds **1a** (top, N7) versus **1b** (bottom, N9). Signals significant for the assignment are marked.

^{N7/N9}Npom-BG-NH₂ (2a/2b)

For deprotection of the trifluoroacetamide group, 50 mg of the respective isomer of Npom-BG-TFA (**1a/1b**, 90 μmol, 1 eq) was solved in 1 ml MeOH and 150 μl water. Potassium carbonate (K₂CO₃, 100 mg, 720 μmol, 8 eq) was added and the reaction mixture was heated to 50°C for 4 h. HPLC analysis showed full conversion of the starting material. No hydrolysis of the Npom-group was observed. 1 M HCl was added to the solution until a neutral pH was reached and all solvent was evaporated. When Npom-BG-NH₂ was used for solid phase synthesis of the Npom-BG-Linker for the later incorporation into guideRNAs, then Npom-BG-NH₂ was purified via preparative HPLC. For this, the crude reaction mixture was neutralized to pH 7 with 1 M HCl, all solvents were evaporated, the crude product was solved in water/ACN/TFA (50:50:0.1), filtered and applied to preparative HPLC. 38 mg (90%) HPLC-pure product was recovered (see Figure S6).

For workup of larger scale reactions MeOH was removed by evaporation and the resulting precipitate was resolved in EtOAc by ultrasonification. The organic phase was washed three times with 1M NaOH/Brine (1:9) and dried over Na₂SO₄. Evaporation yielded ~80% of the respective Npom-BG-NH₂ isomer that was used without further purification for the synthesis of ^{N7/N9}Npom-BG-FAM (**3a/3b**).

^{N7}Isomer: HR-ESI-MS: [M+H]⁺ (theoretical) = 494.1783 for C₂₃H₂₄N₇O₆;
[M+H]⁺ (found) = 494.1781

^{N9}Isomer: HR-ESI-MS: [M+H]⁺ (theoretical) = 494.1783 for C₂₃H₂₄N₇O₆;
[M+H]⁺ (found) = 494.1784

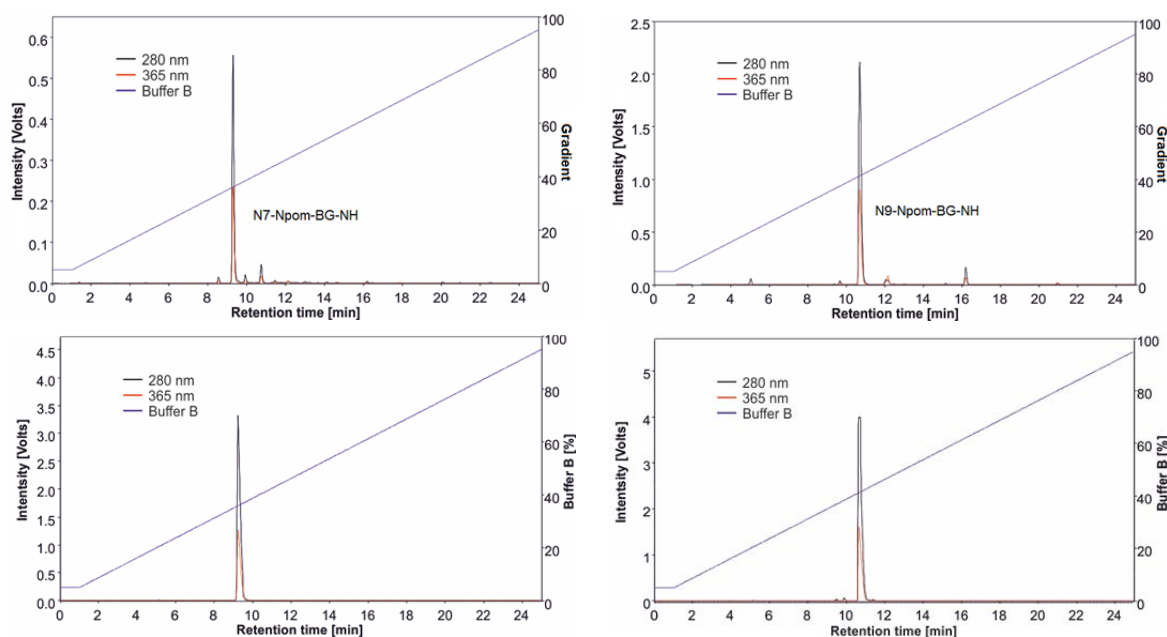


Figure S6. Analytical HPLC-analyses of N7 and N9 Npom-BG-NH₂ (**2a/2b**), top: unpurified compounds for NpomBG-FAM synthesis; bottom: HPLC-purified compounds for incorporation into guideRNAs.

^{N7/N9}Npom-BG-FAM (3a/3b)

3-4 mg of the respective isomer of Npom-BG-NH₂ (**2a/2b**, ~7 μmol, 1 eq) were solved in DMF (300 μl) in an amber eppendorf cup (1.5 ml) and trimethylamine (TEA, 5 μl) was added. Then, 5(6)-carboxyfluorescein succinimidyl ester (~3.5 mg, 7.4 μmol, 1.05 eq) was solved in DMF (50 μl) and added to the Npom-BG-NH₂. The reaction cup was incubated at r.t. for 1 h. DMF and TEA were removed under high vacuum and the crude product was solved in water/ACN/TFA (60:40:0.1), filtered and applied to preparative HPLC. For both isomers, the product fractions contained two products referring to the C5 and C6 isomers of carboxyfluorescein (see Figure S7). The main fractions were lyophilized and resolved in dimethylsulfoxide (DMSO, 50 μl). The concentration was adjusted to 1 mM by spectroscopic measurement of the carboxyfluorescein absorbance at a wavelength of 500 nm (assuming an extinction coefficient of FAM at 500 nm to equal 93.000 M⁻¹cm⁻¹). The extinction was determined in a buffer containing 20 mM Tris-HCl (pH 8), 100 mM NaCl and 5% glycerol.

HR-ESI-MS (^{N7}Npom-BG-FAM): [M+H]⁺ (theoretical) = 852.22600;
[M+H]⁺ (found) = 852.22538

HR-ESI-MS (^{N9}Npom-BG-FAM): [M+H]⁺ (theoretical) = 852.22600;
[M+H]⁺ (found) = 852.22499

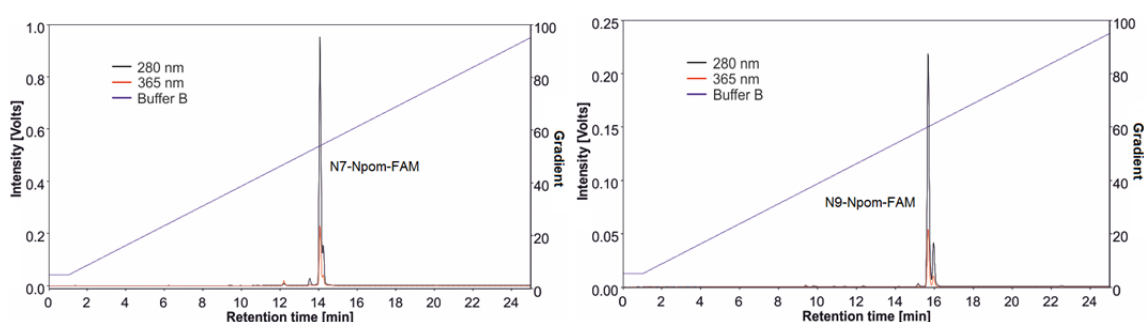


Figure S7. Analytical HPLC of N7- and N9-Npom-BG-FAM (**3a/3b**)

Synthesis of ^{15}N -Npom-BG-Linker-COOH (**4**) via solid phase peptide synthesis

Synthesis of ^{15}N -Npom-BG-Linker-COOH was performed in a syringe and if not indicated otherwise the resin was washed with N-Methyl-2-pyrrolidone (NMP)/DCM (4 x 1:1, 4 x DCM and 4 x NMP) after every coupling or deprotection step. Fmoc-EEG-COOH (Fmoc-aminoethoxy-ethoxy-acetic acid, synthesized according to Visintin et al. [S3], 179 mg, 460 μmol , 4.14 eq) was pre-activated with HBTU (162 mg, 427 μmol , 3.85 eq) and HOBT (54 mg, 400 μmol , 3.6 eq) for 10 min and coupled to the H-Gly-2-CITrt resin (EMD Millipore, 178 mg, 0.63 mmol/g, 111 μmol , 1 eq) in DIPEA/NMP (1:8). After 50 min of coupling, the resin was washed and capped with 5 ml NMP/DIPEA/acetic anhydride (10:1:1, 3x). Fmoc-deprotection was performed using 20% piperidine in NMP (3x5 ml, 10 min each). Glutaric anhydride (125 mg, 1.1 mmol, 10 eq) was coupled in NMP (1.2 ml) with DIPEA (200 μl) for 25 min. An additional washing step with 1% NaOH in Dioxan/ H_2O (1:1, 1 min) was performed. Afterwards, the resin was washed with NMP/DCM thoroughly to ensure that all water was removed. The glutaric acid was activated on the solid phase two times with pentafluorophenyl trifluoroacetate (Tokyo chemical industries, 155 mg, 95 μl , 555 μmol , 5 eq) in 2 ml pyrimidine/DCM (1:1) and the resin was washed with dry NMP (4 x). ^{15}N -Npom-BG- NH_2 (**2a**, 15 mg, 30 μmol , 0.27 eq) was coupled to the resin in NMP/DMSO/DIPEA (100:15:8, 1.2 ml) at r.t. overnight. The syringe was protected from light. Cleavage from the trityl-resin was performed with 0.5% TFA in DCM/hexafluoro-2-propanol (9:1, 50 ml). Evaporation of the cleavage solution yielded 80 mg of crude product which was solved in ACN/water, filtered and applied to preparative HPLC. Product fractions were identified by ESI-MS and analyzed by analytical HPLC (Figure S8). A stock solution of ^{15}N -Npom-BG-Linker-COOH in DMSO was prepared and the concentration was determined photometrically by observing Npom absorption ($\epsilon_{360\text{nm}} = 4300 \text{ M}^{-1}\text{cm}^{-1}$, as determined for (R,S)-1-(6-Nitro-1,3-benzodioxol-5-yl)ethan-1-ol) (Npom-OH)). Overall, 3.05 μmol (10% with respect to **2a**) of clean product were obtained.

HR-ESI-MS: $[\text{M}+\text{H}]^+$ (theoretical) = 810.30531 for $\text{C}_{36}\text{H}_{44}\text{N}_9\text{O}_{13}$;
 $[\text{M}+\text{H}]^+$ (found) = 810.30417

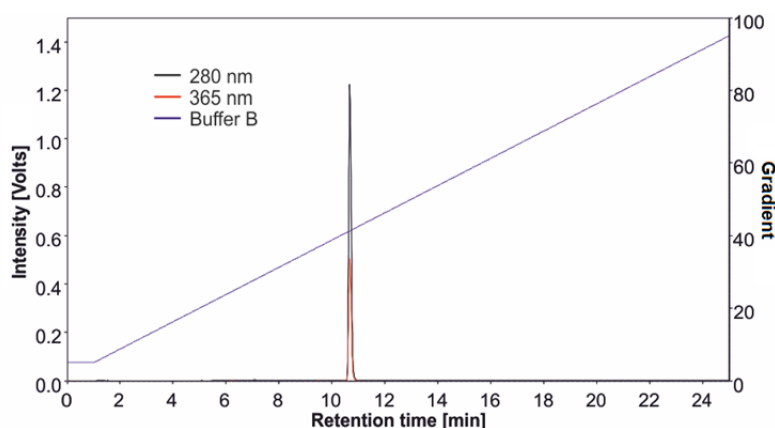


Figure S8. Analytical HPLC of ^{15}N -Npom-BG-Linker-COOH (**4**).

Synthesis of Npom-guideRNAs

guideRNAs that carry only the 5'-C6-aminolinker but no further chemical modifications were obtained from Eurofins Germany in HPLC-purified, MALDI-TOF-confirmed quality. The phosphothioate / 2'-OMe modified oligomers were obtained from Biospring GmbH (Frankfurt) in HPLC-clean quality with sodium as counter ion. Npom-guideRNAs were synthesized on a 100 µg scale (~12-15 nmol depending on sequence, length and modification of the guideRNAs). Antogamir-like modified guideRNAs were solved in RNase free water to a concentration of 6 µg/µl. Stop66 NH₂-guideRNAs were precipitated with 0.1 volumes of 3 M NaCl and 3 volumes of 100% EtOH, washed with 70% EtOH and dissolved in RNase free water (6 µg/µl) prior to coupling.

A 35 mM stock solution of N7-Npom-BG-Linker-COOH (**4**) in DMSO was prepared. For activation ^{N7}Npom-BG-Linker-COOH (10 µl, 0.350 µmol, ~25 eq) was incubated with 1-ethyl-3-(3-dimethylaminopropyl)carbodiimide hydrochloride (EDCI*HCl, 0.72 µmol, ~50 eq), N-hydroxysuccinimide (NHS, 1.23 µmol, ~90 eq) in 24 µl DMSO. After 15 min. incubation at 30°C, half of the pre-activated linker was added to 17 µl of the NH₂-guideRNA solution (6 µg/µl, 100 µg, ~14 µmol, 1 eq) together with DIPEA (2.35 µmol, 170 eq, in 8 µl DMSO) and incubated for another 30 min at 37°C. Then, the remaining pre-activation mix was added and the reaction was continued for 60-90 min.

The crude BG-guideRNA was mixed with 0.1 volumes of 10 x TBE-7M urea containing two dyes (bromophenol blue and xylene cyanol) and purified on a 20% urea-PAGE (four lanes per reaction, 1 x TBE-7 M urea, 120-140 V, 300 min). The PAGE was performed in the dark to prevent Npom-deprotection. While the main part of the gel was stored in the dark, one lane containing ca. 10% of the crude Npom-BG-guideRNA was cut from the gel and analyzed on a TLC plate under low intensity 254 nm UV light. The migration of NH₂-guideRNA and Npom-BG-Linker were noted and the gel slice was discarded. The region corresponding to Npom-BG-guideRNA migration was cut out from the remaining lanes and was transferred to a 1.5 ml reaction tube. In the 20% urea PAGE the xylene cyanol band migrated between the NH₂- and the Npom-BG-guideRNA and served as additional orientation.

600 µl of RNAase free water were added to the gel slices and the tube was shaken at 4°C overnight to allow the NpomBG-guideRNA to diffuse out of the gel. To remove urea and buffer salts, the NpomBG-guideRNA was precipitated with 0.1 volumes sodium acetate (NaOAc, 3 M) and 3 volumes of EtOH 100% (incubation at -80°C for >4 h). The precipitated RNA was centrifuged (45 min, 14.000 rpm, -4°C), washed with 70% EtOH, dried and dissolved in 50 µl RNase free water. Concentrations were determined photometrically by UV-absorbance of RNA at 260nm and typically around 40% of clean Npom-BG-guideRNA was recovered (compare Table S9).

Table S9. Sequences and extinction coefficients of guideRNAs synthesized in this study. The 5'-BG-modified and 5'-(^{N7}Npom-BG)-guideRNAs were prepared from the commercially obtained 5'-NH₂-guideRNAs as described above. The nucleotides opposite of the targeted adenosine are unlined and bold, they are normal ribonucleotides. 2'-Methoxylation is indicated by italic characters. Phosphothioate linkages are indicated by "s" subscripts. The 5'-terminal three nucleotides 5' of each guideRNA are not binding to the mRNA substrate but link the guideRNA to the SNAP-tag. BG-modification adds 2.5 mM⁻¹ cm⁻¹, Npom-BG modification adds 6.5 mM⁻¹ cm⁻¹ to the extinction coefficient of the NH₂-guideRNAs.

Short name	Sequence	$\epsilon_{260\text{nm}}$ (mM ⁻¹ cm ⁻¹)
NH ₂ -Stop66	UCG GAA CAC CCC AGC ACA GA	230
NH ₂ -W58X 19nt	<i>UsAsU GUG UCG GCC ACG GAAs CsAsGs G</i>	226
NH ₂ -W58X 21nt	<i>UsAsU GUG UCG GCC ACG GAA CAsGs GsCsA</i>	236
BG-Stop66	UCG GAA CAC CCC AGC ACA GA	232.5
BG-W58X 19nt	<i>UsAsU GUG UCG GCC ACG GAAs CsAsGs G</i>	228.5
BG-W58X 21nt	<i>UsAsU GUG UCG GCC ACG GAA CAsGs GsCsA</i>	238.5
BG-Stop66 PTO	<i>UsCsG GAA CAC CCC AGC AsCsAs GsA</i>	232.5
Npom-Stop66	UCG GAA CAC CCC AGC ACA GA	236.5
Npom-W58X 19nt	<i>UsAsU GUG UCG GCC ACG GAAs CsAsGs G</i>	232.5
Npom-W58X 21nt	<i>UsAsU GUG UCG GCC ACG GAA CAsGs GsCsA</i>	242.5

MALDI-TOF characterization of guideRNAs

150-300 pmol of each guideRNA were precipitated by adding 0.5 volumes ammonium acetate (7.5 M) and 2.5 volumes EtOH 100% and subsequent incubation at -80°C for >4 h. The precipitated RNA was centrifuged (45 min, 14 000 rpm, -4°C), washed with 70% EtOH, dried and dissolved in 5 µl RNase free water. The samples were mixed with a matrix of [2,4,6]-trihydroxyacetophenone monohydrate (0.3 M in EtOH) / diammonium citrate (0.1 M in water) (2:1) and the mixture applied on a ground steel target and analyzed using a Bruker Reflex MALDI-TOF mass spectrometer (linear mode, negative mode). All spectra were processed using mMass (Martin Strohm, Germany, see Figure S10).

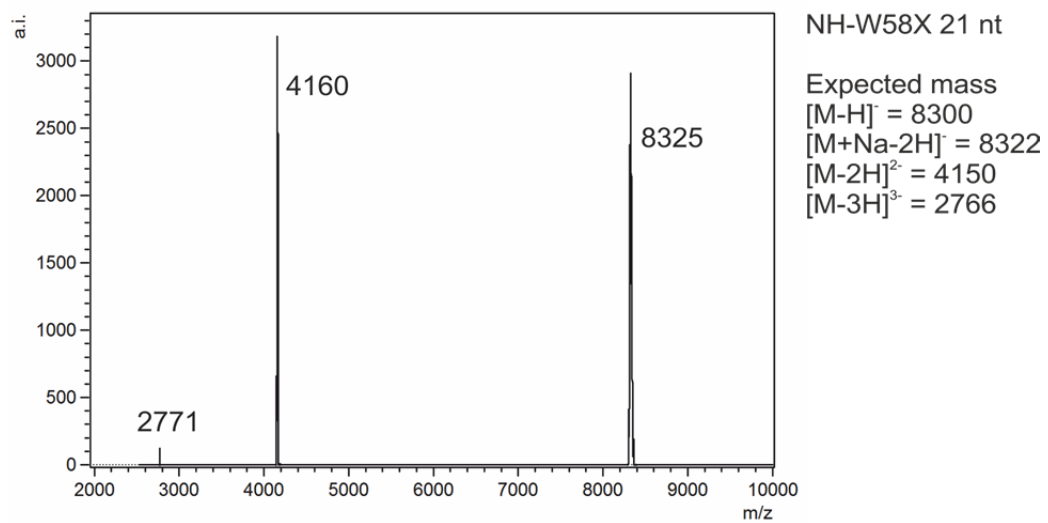
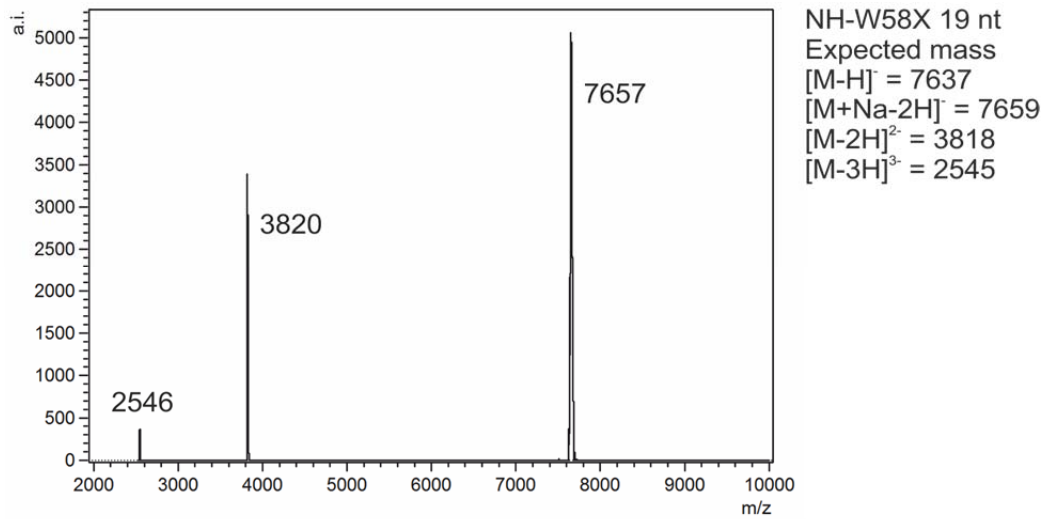
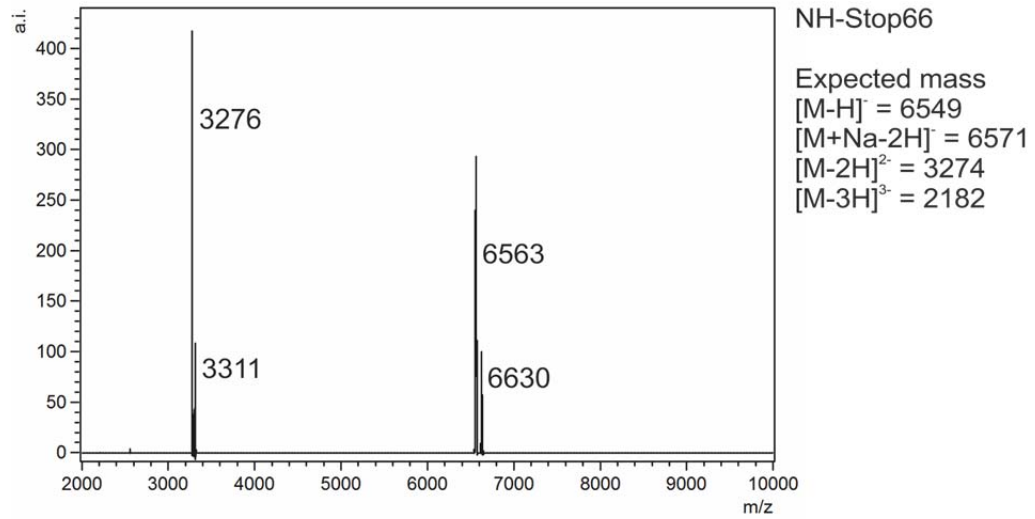
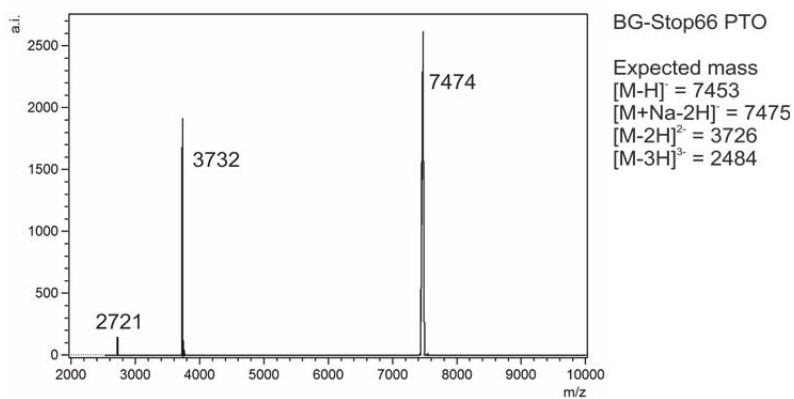
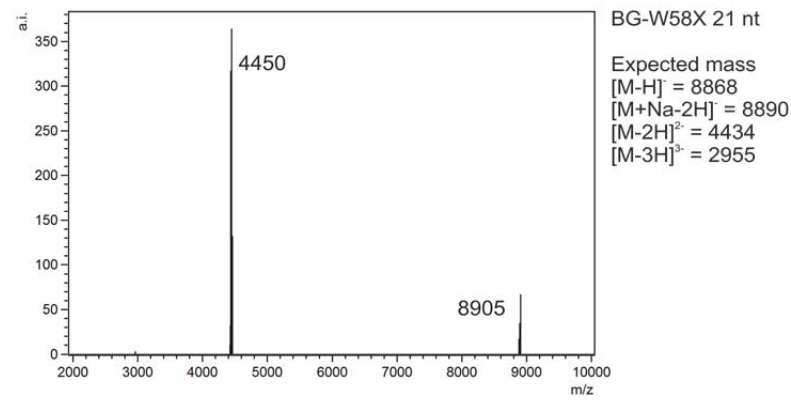
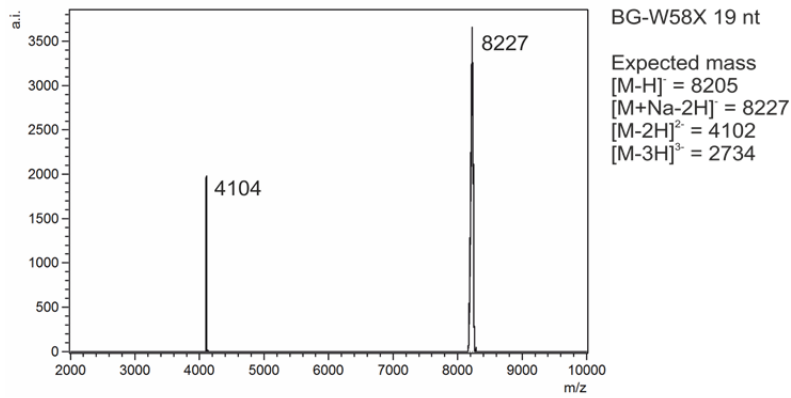
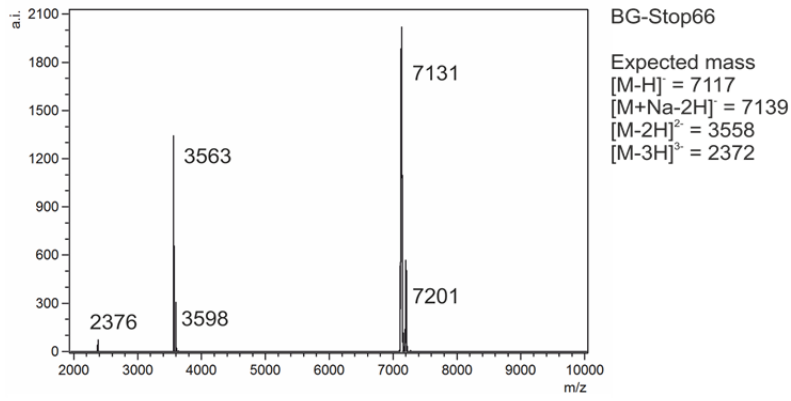
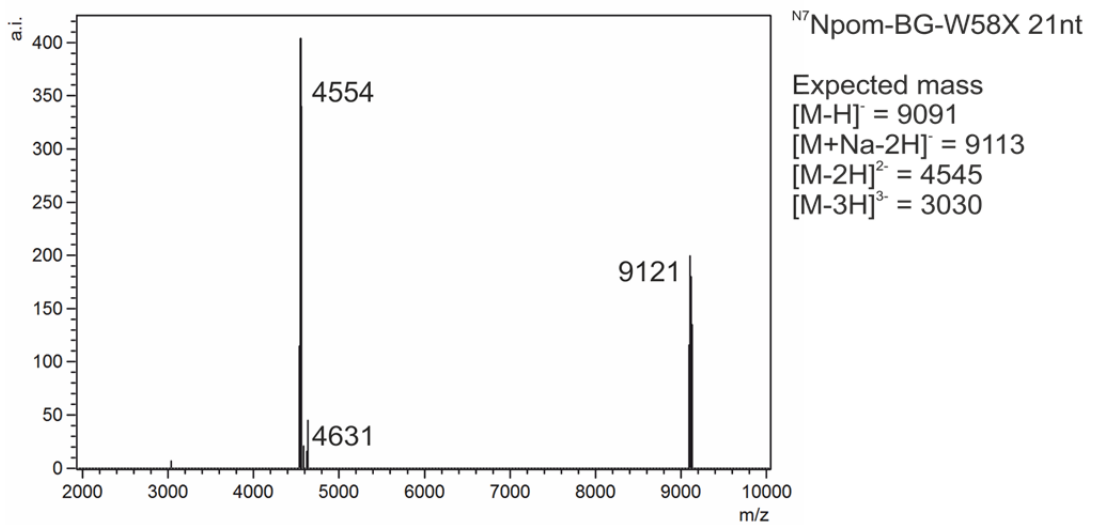
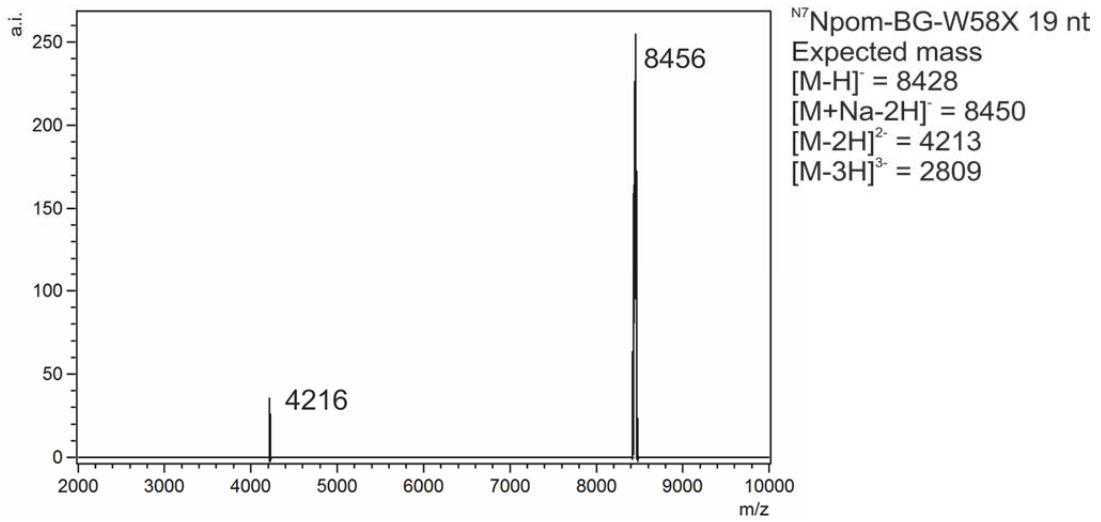
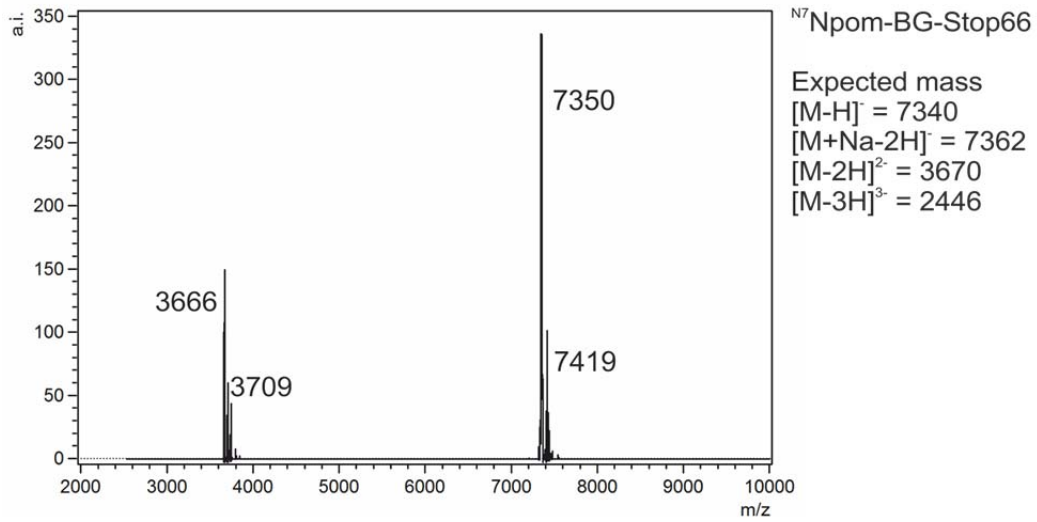


Fig S10. MALDI-TOF spectra of NH-guideRNAs



cont. Fig S10. MALDI-TOF spectra of BG-guideRNAs



cont. Fig S10 MALDI-TOF spectra of Npom-guideRNAs

Experiments with Npom-caged O6-Benzylguanine

Photo-deprotection kinetics of ^{N7/N9}Npom-BG-TFA

To determine the decaging efficiency $\epsilon\Phi$, the decaging of our compounds **1a** and **1b** was compared to the decay of a commercially available compound DMNB-cAMP (4,5-dimethoxy-2-nitrobenzyl adenosine 3',5'-cyclic monophosphate; Life Technologies). The latter was chosen because of the very similar absorbance properties with our compounds. First, the absorption coefficient ϵ of the Npom-BG-TFA was determined. Therefore, the absorption spectra of ^{N7}Npom-BG-TFA, BG-NH₂ and Npom-OH were recorded (Fig S11). The extinction coefficient of Npom-OH and the caged product ^{N7}Npom-BG-TFA are very similar around 360 nm. $\epsilon_{365\text{nm}}$ of Npom-OH was determined to be 4300 M⁻¹cm⁻¹. In the range of the methoxy-nitrobenzyl group the presence of the BG moiety didn't change the UV spectra significantly. Stock solutions of ^{N7}Npom-BG-TFA and DMNB-cAMP in DMSO were prepared and diluted in sodium phosphate buffer (NaCl 100 mM, KH₂PO₄/K₂HPO₄ 10mM, pH 7.4) to final concentrations of 10 μ M. Decaging was performed in PCR tubes (60 μ L scale) by irradiation with 365 nm light on a UV transilluminator (UVP TFL-40V, 25 W, intensity high) for the indicated amount of time at room temperature. Decaging was performed at low concentrations of the compounds making inner-filter effects negligible. Taking the extinction coefficient and the maximal diameter of the PCR tube (diameter of 5 mm) into consideration a transmission of >95% is expected. Photodecomposition of the caged substances was monitored in analytical HPLC with UV detection at 280nm and 365 nm. For both substances an exemplary chromatogram is shown in Figure S12. Decomposition of ^{N7}Npom-BG-TFA results in two clean products, BG-TFA which shows high absorption at 280nm but now absorption at 365nm and the released photocaging group that shows high absorbance at 365nm. DMNB-cAMP shows two starting materials that refer to the two isomers (axial versus equatorial). The peak areas of the emerging products, BG-TFA and cAMP, were determined and plotted against irradiation time (see Figure S13). By 1st-order fitting, the half-life was determined to be 5.53 \pm 0.63 min for DMNB-cAMP decay, 0.57 \pm 0.04 min for ^{N7}Npom-BG-TFA and 0.78 \pm 0.06 min for ^{N9}Npom-BG-TFA.

Literature reports a $\epsilon\Phi$ value of 250 M⁻¹cm⁻¹ for of cAMP formation at 350 nm with $\epsilon = 5.0 \text{ mM}^{-1}\text{cm}^{-1}$ and $\Phi = 5\%$.^[S6] $\epsilon_{365\text{nm}}$ of DMNB-cAMP was determined to be 4.0 mM⁻¹cm⁻¹ and from this $\epsilon\phi$ at 365nm was calculated to be 200 M⁻¹cm⁻¹. By comparing the half-lives, an $\epsilon\Phi_{365\text{nm}}$ for N7-Npom-BG-TFA of 2000 M⁻¹cm⁻¹ results. Taking a ϵ of ca. 4 mM⁻¹cm⁻¹ into account, the quantum yields can be estimated to be $\approx 50\%$ for the N7- and $\approx 36\%$ for the N9-isomer.

To test the stability of the Npom-protected BG in the dark, compounds **1a** and **1b** were dissolved at 10 μ M in phosphate buffer (NaCl 100 mM, KH₂PO₄/K₂HPO₄ 10mM, pH 7.4) and kept at r.t. for 3 day. Per HPLC (Figure S14), no release of BG is detectable.

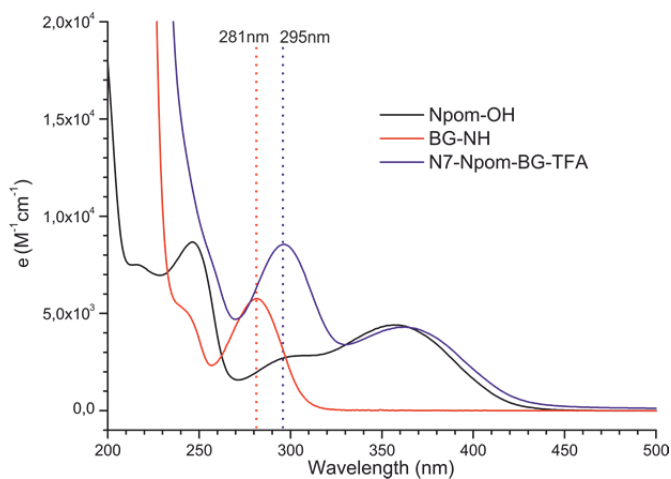


Figure S11. Determination of the extinction coefficients of Npom-OH, N7-Npom-BG-TFA and BG-NH₂.

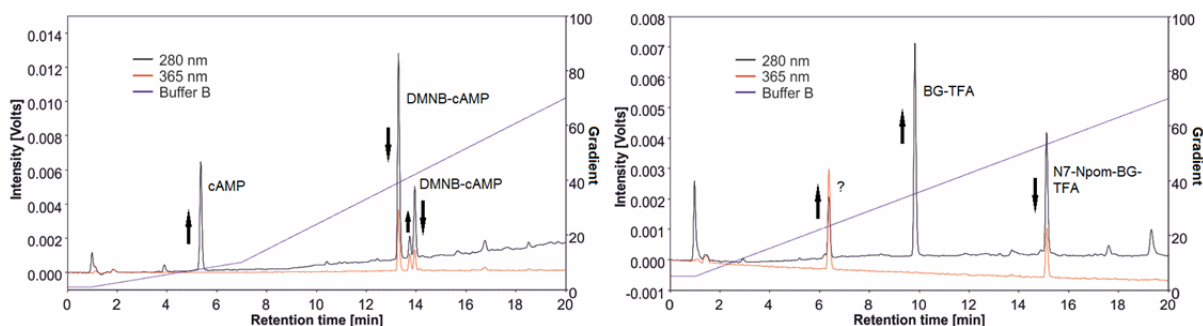


Figure S12. Analytical HPLC traces of the photo-decaging of reference DMNB-cAMP versus compound N7-Npom-BG-TFA. Shown are the HPLC trace after 5 min (left) and 0.5 min (right) irradiation with 365 nm.

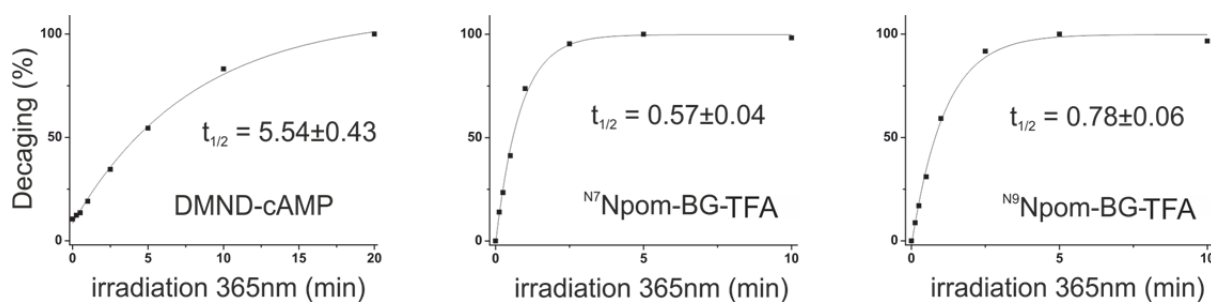


Figure S13. Photodecaging kinetics at 365 nm irradiation for N7-Npom-BG-TFA and N9-Npom-BG-TFA in comparison to commercial reference DMNB-cAMP for the determination of the quantum yield.

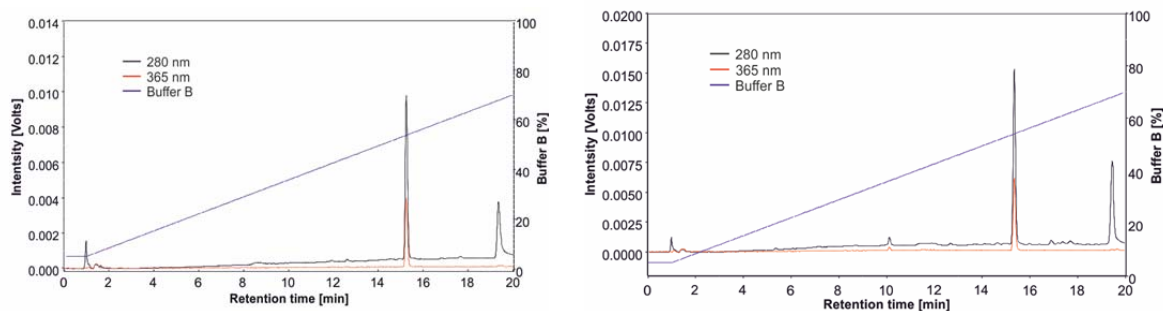


Figure S14. Analytical HPLC for N7-Npom-BG-TFA before and after 3 days at r.t. in the dark (NaCl 100 mM, $\text{KH}_2\text{PO}_4/\text{K}_2\text{HPO}_4$ 10mM, pH 7.4). No release of BG-TFA ($t_R = 10$ min) was observed. The little impurity at $t_R = 10$ min after 3 days is not BG-TFA as it shows absorption at 365 nm (compare Figure S12).

BG-FITC/FAM assay

The BG-FITC/FAM assay (Figure 1c in the manuscript) was performed in 20mM Tris-HCl (pH 8), 100mM NaCl and 5% glycerol. The reaction was performed on a 8 μl scale in PCR reaction tubes. The concentration of SNAP-ADAR1 was adjusted to 1 μM and the respective BG derivatives (BG-FITC, $^{\text{N}7}\text{Npom-BG-FAM}$ and $^{\text{N}9}\text{Npom-BG-FAM}$) were used in a concentration of 7.5 μM . Irradiation was performed on a UV-table (365 nm, power=high). The reaction tubes were incubated at 30°C for 20 min in the dark. The reaction was stopped by adding 0.5 volumes of 4 x SDS-PAGE loading buffer and heating to 95°C for 4 min. 10 μl of each sample were applied to SDS-PAGE (4% stacking gel, 12% separation gel, 100 V, 2 h). Furthermore, GE Healthcare LMW protein marker was applied to one lane. To determine FITC/FAM fluorescence, the gel was scanned in a Fujifilm FLA-5100 fluorescence scanner using an excitation wavelength of 473nm (Intensity = 500V) and recording the emission at 557nm (Cy3 filter set). To verify consistent loading, coomassie staining was performed. The staining solution was composed of Coomassie Brilliant Blue G-250 (0.02% w/v), $\text{Al}_2(\text{SO}_4)_3$ (5% w/v), EtOH (10% v/v) and phosphoric acid (2% v/v).

SNAP-ADAR1 band-shift assay

The SNAP-ADAR1 band-shift assay (Figure 2b in the manuscript) was performed in 10 ml Tris-HCl (pH 8), 50 mM NaCl and 2.5% glycerol, 2 mM DTT. The reaction was performed on an 8 μ l scale in PCR reaction tubes. The concentration of SNAP-ADAR1 was adjusted to 1 μ M and the respective guideRNAs (NH2-Stop66, BG-Stop66, Npom-Stop66) were used in a concentration of 7.5 μ M. Irradiation was performed on a UV-table (365 nm, power=high). The reaction tubes were incubated at 30°C in the dark for 40 min. The reaction was stopped by addition of 0.5 volumes of 4 x SDS-PAGE loading buffer and heating to 95°C for 4 min. 10 μ l of each sample were applied to SDS-PAGE (4% stacking gel, 12% separation gel, 100 V, 3.5 h). Furthermore, GE Healthcare LMW protein marker was applied to one lane. The proteins were visualized by coomassie staining. The staining solution was composed of Coomassie Brilliant Blue G-250 (0.02% w/v), $\text{Al}_2(\text{SO}_4)_3$ (5% w/v), EtOH (10% v/v) and phosphoric acid (2% v/v).

Primer

All primers used in this study were purchased from Sigma-Aldrich GmbH (Muenchen) or MWG Eurofins (Ebersberg).

Stop66 fw: GCGGATAACA ATTCCCCTCT AG
Stop66 rv: CAGCGGTGGC AGCAGCCAAC

W58X RT: CTAGAAGGCA CAGTCGAGGC
W58X fw: GCGGATCCAC CATGGCTAGC AAAGGAGAAG AACTC
W58X rv: CCTCTAGAGC CGGATTTGTA TAGTTCATCC ATGCC
W58Xpos327 fw: GACACGTGCT GAAGTCAAGT TTGAAGGTG

W58X PD fw: ATGGCGCGCC TAGCTAGCAA AGGAGAAGAA CTC
W58X PD bw: TAACCGGTTT TGTATAGTTC ATCCATGCCA TG

Gene sequences

PCR-template for the generation of W66amber eCFP mRNA

The premature stop codon (TAG) at position 66 is underlined and highlighted.

```

          10          20          30          40          50          60
1  TAATACGACTCACTATAGGGGAATTGTGAGCGGATAACAATTCCCCTCTAGAAAATAATTT
   T7 promoter
          70          80          90          100         110         120
61  TGTTTAACTTTAAGAAGGAGATATACATATGGCTAGCAAAGGAGAAGAACTCTTCACTGG
           M A S K G E E L F T G
          130         140         150         160         170         180
121  AGTTGTCCCAATTCTTGTGTAATTAGATGGTGATGTTAACGGCCACAAGTTCTCTGTCTAG
41  V V P I L V E L D G D V N G H K F S V S
          190         200         210         220         230         240
181  TGGAGAGGGTGAAGGTGATGCAACATACGAAAACTTACCCTGAAGTTCATCTGCACTAC
61  G E G E G D A T Y G K L T L K F I C T T
          250         260         270         280         290         300
241  TGGCAAACCTGCCTGTTCCGTGGCCAACACTAGTCACTACTCTGTGCTTAGGGTGTTCAAATG
81  G K L P V P W P T L V T T L C * G V Q C
          310         320         330         340         350         360
301  CTTTTCAAGATACCCGGATCACATGAAACGGCATGACTTTTTTCAAGAGTGCCATGCCCGA
101  F S R Y P D H M K R H D F F K S A M P E
          370         380         390         400         410         420
361  AGGTTATGTACAGGAAAGGACCATCTTCTTCAAAGATGACGGCAACTACAAGACACGTGC
121  G Y V Q E R T I F F K D D G N Y K T R A
          430         440         450         460         470         480
421  TGAAGTCAAGTTTGAAGGTGATAACCTTGTTAATAGAATCGAGTTAAAAGGTATTGACTT
141  E V K F E G D T L V N R I E L K G I D F
```

```

          490      500      510      520      530      540
481  CAAGGAAGATGGCAACATTCTGGGACACAAATTTGGAATACAACCTATATCTCACACAATGT
161      K E D G N I L G H K L E Y N Y I S H N V

          550      560      570      580      590      600
541  ATACATCACCGCAGACAAAACAAAAGAATGGAATCAAAGCCCCTTCAAGACCCGCCACAA
181      Y I T A D K Q K N G I K A H F K T R H N

          610      620      630      640      650      660
601  CATTGAAGATGGAAGCGTTCAACTAGCAGACCATTATCAACAAAATACTCCAATTGGCGA
201      I E D G S V Q L A D H Y Q Q N T P I G D

          670      680      690      700      710      720
661  TGGCCCTGTCTTTTTACCAGACAACCATTACCTGTCCACACAATCTGCCCTTTTCGAAAGA
221      G P V L L P D N H Y L S T Q S A L S K D

          730      740      750      760      770      780
721  TCCCAACGAAAAGAGAGACCACATGGTCTTCTTGAGTTTGTAACAGCTGCTGGGATTAC
241      P N E K R D H M V L L E F V T A A G I T

          790      800      810      820      830      840
781  ACATGGCATGGATGAACTATACAAAATCCGGCGGCTCCATGGCGCTCGAGCACCACCACCA
261      H G M D E L Y K S G G S M A L E H H H H

          850      860      870      880      890      900
841  CCACCCTAATAATGACTAGTCAGCTGATCCGGCTGCTAACAAAAGCCCGAAAGGAAGCTG
281      H H * *

          910      920
901  AGTTGGCTGCTGCCACCGCTG
301

```

W58XeGFP/wt eGFP in the context of the pcDNA3.1 vector (Invitrogen):

The gene was cloned using the BamH1 and Xba1 (underlined) restriction sites. The codon 58 (TAG) that is edited is highlighted in yellow. In the wildtype sequence = positive control plasmid this codon is TGG.

```

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

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

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

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

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

```

```

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

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

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

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

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

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

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

          730      740      750      760      770      780
721  GAACTATACAAATCCGGCTCTAGAGGGCCCTATTCTATAGTGTACACCTAAATGCTAGAGC
241  E L Y K S G S R G P Y S I V S P K C *

```

SNAPf-ADAR1 in the context of the pcDNA3.1 vector

```

          10      20      30      40      50      60
1  GGATCCACCATGGACAAAGACTGCGAAATGAAGCGCACCACCCTGGATAGCCCTCTGGGC
   BamH1  M D K D C E M K R T T L D S P L G

          70      80      90      100     110     120
61  AAGCTGGAAGTGTCTGGGTGCGAACAGGGCCTGCACCGTATCATCTTCCCTGGGCAAAGGA
21  K L E L S G C E Q G L H R I I F L G K G

          130     140     150     160     170     180
121 ACATCTGCCCGCCGACGCCGTGGAAGTGCCCTGCCCCAGCCCGCTGCTGGGCGGACCAGAG
41  T S A A D A V E V P A P A A V L G G P E

          190     200     210     220     230     240
181 CCACTGATGCAGGCCACCGCCTGGCTCAACGCCTACTTTCACCAGCCTGAGGCCATCGAG
61  P L M Q A T A W L N A Y F H Q P E A I E

          250     260     270     280     290     300
241 GAGTTCCCTGTGCCAGCCCTGCACCACCCAGTGTTCAGCAGGAGAGCTTTACCCGCCAG
81  E F P V P A L H H P V F Q Q E S F T R Q

          310     320     330     340     350     360
301 GTGCTGTGGAAACTGCTGAAAGTGGTGAAGTTCGGAGAGGTCATCAGCTACAGCCACCTG

```

101 V L W K L L K V V K F G E V I S Y S H L
 370 380 390 400 410 420
 361 GCCGCCCTGGCCGGCAATCCCGCCGCCACCGCCCGCTGAAAACCGCCCTGAGCGGAAAT
 121 A A L A G N P A A T A A V K T A L S G N
 430 440 450 460 470 480
 421 CCCGTGCCCATTCTGATCCCCTGCCACCGGGTGGTGCAGGGCGACCTGGACGTGGGGGGC
 141 P V P I L I P C H R V V Q G D L D V G G
 490 500 510 520 530 540
 481 TACGAGGGCGGGCTCGCCGTGAAAGAGTGGCTGCTGGCCCACGAGGGCCACAGACTGGGC
 161 Y E G G L A V K E W L L A H E G H R L G
 550 560 570 580 590 600
 541 AAGCCTGGGCTGGGTCTGCAGGCGGAGGCGCGCCAGGGTCTGGCGGCGGCAGTAAGGCA
 181 K P G L G P A G G G A P G S G G G S K A
 610 620 630 640 650 660
 601 GAACGCATGGGTTTTCACAGAGGTAACCCAGTGACAGGGGCCAGTCTCAGAAGAACTATG
 201 E R M G F T E V T P V T G A S L R R T M
 670 680 690 700 710 720
 661 CTCCTCCTCTCAAGGTCCCCAGAAGCACAGCCAAAAGACACTCCCCTCTCACTGGCAGCACC
 221 L L L S R S P E A Q P K T L P L T G S T
 730 740 750 760 770 780
 721 TTCCATGACCAGATAGCCATGCTGAGCCACCGGTGCTTCAACACTCTGACTAACAGCTTC
 241 F H D Q I A M L S H R C F N T L T N S F
 790 800 810 820 830 840
 781 CAGCCCTCCTTGCTCGGCCGCAAGATTCTGGCCGCCATCATTATGAAAAAGACTCTGAG
 261 Q P S L L G R K I L A A I I M K K D S E
 850 860 870 880 890 900
 841 GACATGGGTGTCGTCGTCAGCTTGGGAACAGGGAATCGCTGTGTAAAAGGAGATTCTCTC
 281 D M G V V V S L G T G N R C V K G D S L
 910 920 930 940 950 960
 901 AGCCTAAAAGGAGAAACTGTCAATGACTGCCATGCAGAAAATAATCTCCCGGAGAGGCTTC
 301 S L K G E T V N D C H A E I I S R R G F
 970 980 990 1000 1010 1020
 961 ATCAGGTTTCTCTACAGTGAGTTAATGAAATACAACCTCCAGACTGCGAAGGATAGTATA
 321 I R F L Y S E L M K Y N S Q T A K D S I
 1030 1040 1050 1060 1070 1080
 1021 TTTGAACCTGCTAAGGGAGGAGAAAAAGCTCCAAAATAAAAAAGACTGTGTTCATTCCATCTG
 341 F E P A K G G E K L Q I K K T V S F H L
 1090 1100 1110 1120 1130 1140
 1081 TATATCAGCACTGCTCCGTGTGGAGATGGCGCCCTCTTTGACAAGTCTGCAGCGACCGT
 361 Y I S T A P C G D G A L F D K S C S D R
 1150 1160 1170 1180 1190 1200
 1141 GCTATGGAAAGCACAGAATCCCGCCACTACCTGTCTTCGAGAATCCCAAACAAGGAAAG
 381 A M E S T E S R H Y P V F E N P K Q G K
 1210 1220 1230 1240 1250 1260
 1201 CTCCGCACCAAGGTGGAGAACGGAGAAGGCACAATCCCTGTGGAATCCAGTGACATTGTG
 401 L R T K V E N G E G T I P V E S S D I V

1261 1270 1280 1290 1300 1310 1320
 CCTACGTGGGATGGCATTTCGGCTCGGGGAGAGACTCCGTACCATGTCCTGTAGTGACAAA
 421 P T W D G I R L G E R L R T M S C S D K

1321 1330 1340 1350 1360 1370 1380
 ATCCTACGCTGGAACGTGCTGGGCCTGCAAGGGGCACTGTTGACCCACTTCCTGCAGCCC
 441 I L R W N V L G L Q G A L L T H F L Q P

1381 1390 1400 1410 1420 1430 1440
 ATTTATCTCAAATCTGTCACATTGGGTACCTTTTCAGCCAAGGGCATCTGACCCGTGCT
 461 I Y L K S V T L G Y L F S Q G H L T R A

1441 1450 1460 1470 1480 1490 1500
 ATTTGCTGTCGTGTGACAAGAGATGGGAGTGCATTTGAGGATGGACTACGACATCCCTTT
 481 I C C R V T R D G S A F E D G L R H P F

1501 1510 1520 1530 1540 1550 1560
 ATTTGCAACCACCCCAAGGTTGGCAGAGTCAGCATATATGATTCCAAAAGGCAATCCGGG
 501 I V N H P K V G R V S I Y D S K R Q S G

1561 1570 1580 1590 1600 1610 1620
 AAGACTAAGGAGACAAGCGTCAACTGGTGTCTGGCTGATGGCTATGACCTGGAGATCCTG
 521 K T K E T S V N W C L A D G Y D L E I L

1621 1630 1640 1650 1660 1670 1680
 GACGGTACCAGAGGCACTGTGGATGGGCCACGGAATGAATTGTCCCGGGTCTCCAAAAAG
 541 D G T R G T V D G P R N E L S R V S K K

1681 1690 1700 1710 1720 1730 1740
 AACATTTTTTCTTCTATTTAAGAAGCTCTGCTCCTTCCGTTACCGCAGGGATCTACTGAGA
 561 N I F L L F K K L C S F R Y R R D L L R

1741 1750 1760 1770 1780 1790 1800
 CTCTCCTATGGTGAGGCCAAGAAAGCTGCCCGTACTACGAGACGGCCAAGAATACTTTC
 581 L S Y G E A K K A A R D Y E T A K N Y F

1801 1810 1820 1830 1840 1850 1860
 AAAAAAGGCCTGAAGGATATGGGCTATGGGAACTGGATTAGCAAACCCAGGAGGAAAAG
 601 K K G L K D M G Y G N W I S K P Q E E K

1861 1870 1880 1890 1900 1910 1920
 AACTTTTATCTCTGCCAGTATCTAGAGGGCCCTTCGAACAAAACTCATCTCAGAAGAG
 621 N F Y L C P V S R G P F E Q K L I S E E

1921 1930 1940 1950 1960 1970 1980
 GATCTGAATATGCATACCGGTCATCATCACCATCACCATTGAGTTTAAACCCGCTGATCA
 641 D L N M H T G H H H H H H H *

SNAPf-ADAR2 in the context of the pcDNA3.1 vector

1 GGATCCACCATGGACAAAGACTGCGAAATGAAGCGCACCCCTGGATAGCCCTCTGGGC
 BamH1 M D K D C E M K R T T L D S P L G

61 70 80 90 100 110 120
 AAGCTGGAAGTGTCTGGGTGCGAACAGGGCCTGCACCGTATCATCTTCTGGGCAAAGGA
 21 K L E L S G C E Q G L H R I I F L G K G

130 140 150 160 170 180

121 ACATCTGCCGCCGACGCCGTGGAAGTGCCTGCCCCAGCCGCCGTGCTGGGCGGACCAGAG
41 T S A A D A V E V P A P A A V L G G P E

190 200 210 220 230 240
181 CCACTGATGCAGGCCACCGCCTGGCTCAACGCCTACTTTTCACCAGCCTGAGGCCATCGAG
61 P L M Q A T A W L N A Y F H Q P E A I E

250 260 270 280 290 300
241 GAGTTCCCTGTGCCAGCCCTGCACCACCCAGTGTTCAGCAGGAGAGCTTTACCCGCCAG
81 E F P V P A L H H P V F Q Q E S F T R Q

310 320 330 340 350 360
301 GTGCTGTGGAAACTGCTGAAAGTGGTGAAGTTCGGAGAGGTCATCAGCTACAGCCACCTG
101 V L W K L L K V V K F G E V I S Y S H L

370 380 390 400 410 420
361 GCCGCCCTGGCCGGCAATCCCGCCGCCACCGCCCGCTGAAAACCGCCCTGAGCGGAAAT
121 A A L A G N P A A T A A V K T A L S G N

430 440 450 460 470 480
421 CCCGTGCCCATTTCTGATCCCCTGCCACCGGGTGGTGCAGGGCGACCTGGACGTGGGGGGC
141 P V P I L I P C H R V V Q G D L D V G G

490 500 510 520 530 540
481 TACGAGGGCGGGCTCGCCGTGAAAGAGTGGCTGCTGGCCCACGAGGGCCACAGACTGGGC
161 Y E G G L A V K E W L L A H E G H R L G

550 560 570 580 590 600
541 AAGCCTGGGCTGGGTCTCAGGCGGAGGCGCCAGGGTCTGGCGGCGGCAGTAAGAAG
181 K P G L G P A G G G A P G S G G G S K K

610 620 630 640 650 660
601 CTTGCCAAGGCCCGGGCTGCGCAGTCTGCCCTGGCCGCCATTTTTAACTTGCACCTGGAT
201 L A K A R A A Q S A L A A I F N L H L D

670 680 690 700 710 720
661 CAGACGCCATCTCGCCAGCCTATTCCCAGTGAGGGTCTTCAGCTGCATTTACCGCAGGTT
221 Q T P S R Q P I P S E G L Q L H L P Q V

730 740 750 760 770 780
721 TTAGCTGACGCTGTCTCACGCCTGGTCTGGGTAAGTTTGGTGACCTGACCGACAACCTTC
241 L A D A V S R L V L G K F G D L T D N F

790 800 810 820 830 840
781 TCCTCCCCTCACGCTCGCAGAAAAGTGTGGCTGGAGTCGTCATGACAACAGGCACAGAT
261 S S P H A R R K V L A G V V M T T G T D

850 860 870 880 890 900
841 GTTAAAGATGCCAAGGTGATAAGTGTTCCTACAGGAACAAAATGTATTAATGGTGAATAC
281 V K D A K V I S V S T G T K C I N G E Y

910 920 930 940 950 960
901 ATGAGTGATCGTGGCCTTGCATTAATGACTGCCATGCAGAAAATAATATCTCGGAGATCC
301 M S D R G L A L N D C H A E I I S R R S

970 980 990 1000 1010 1020
961 TTGCTCAGATTTCTTTATACACAACCTTGAGCTTTACTTAAATAACAAAGATGATCAAAAA
321 L L R F L Y T Q L E L Y L N N K D D Q K

1030 1040 1050 1060 1070 1080
1021 AGATCCATCTTTTCAGAAATCAGAGCGAGGGGGTTTAGGCTGAAGGAGAATGTCCAGTTT

341 R S I F Q K S E R G G F R L K E N V Q F
 1090 1100 1110 1120 1130 1140
 1081 CATCTGTACATCAGCACCTCTCCCTGTGGAGATGCCAGAATCTTCTCACCACATGAGCCA
 361 H L Y I S T S P C G D A R I F S P H E P
 1150 1160 1170 1180 1190 1200
 1141 ATCCTGGAAGAACCAGCAGATAGACACCCAAATCGTAAAGCAAGAGGACAGCTACGGACC
 381 I L E E P A D R H P N R K A R G Q L R T
 1210 1220 1230 1240 1250 1260
 1201 AAAATAGAGTCTGGTGAGGGGACGATTCCAGTGCGCTCCAATGCGAGCATCCAAACGTGG
 401 K I E S G E G T I P V R S N A S I Q T W
 1270 1280 1290 1300 1310 1320
 1261 GACGGGGTGCTGCAAGGGGAGCGGCTGCTCACCATGTCCTGCAGTGACAAGATTGCACGC
 421 D G V L Q G E R L L T M S C S D K I A R
 1330 1340 1350 1360 1370 1380
 1321 TGGAACGTGGTGGGCATCCAGGGATCCCTGCTCAGCATTTTCGTGGAGCCCATTACTTC
 441 W N V V G I Q G S L L S I F V E P I Y F
 1390 1400 1410 1420 1430 1440
 1381 TCGAGCATCATCCTGGGCAGCCTTTACCACGGGGACCACCTTTCCAGGGCCATGTACCAG
 461 S S I I L G S L Y H G D H L S R A M Y Q
 1450 1460 1470 1480 1490 1500
 1441 CGGATCTCCAACATAGAGGACCTGCCACCTCTCTACACCCTCAACAAGCCTTTGCTCAGT
 481 R I S N I E D L P P L Y T L N K P L L S
 1510 1520 1530 1540 1550 1560
 1501 GGCATCAGCAATGCAGAAGCACGGCAGCCAGGGAAGGCCCCCAACTTCAGTGTCAACTGG
 501 G I S N A E A R Q P G K A P N F S V N W
 1570 1580 1590 1600 1610 1620
 1561 ACGGTAGGCGACTCCGCTATTGAGGTATCAACGCCACGACTGGGAAGGATGAGCTGGGC
 521 T V G D S A I E V I N A T T G K D E L G
 1630 1640 1650 1660 1670 1680
 1621 CGCGCGTCCCGCCTGTGTAAGCACGCGTTGTAAGTGTGCTGGATGCGTGTGCACGGCAAG
 541 R A S R L C K H A L Y C R W M R V H G K
 1690 1700 1710 1720 1730 1740
 1681 GTTCCCTCCCCTTACTACGCTCCAAGATTACCAAGCCCAACGTGTACCATGAGTCCAAG
 561 V P S H L L R S K I T K P N V Y H E S K
 1750 1760 1770 1780 1790 1800
 1741 CTGGCGGCAAAGGAGTACCAGGCCGCAAGGCGCTCTGTTTACAGCCTTTCATCAAGGCG
 581 L A A K E Y Q A A K A R L F T A F I K A
 1810 1820 1830 1840 1850 1860
 1801 GGGCTGGGGGCTGGGTGGAGAAGCCACCGAGCAGGACCAGTTCTCACTCACGCCCTCT
 601 G L G A W V E K P T E Q D Q F S L T P S
 1870 1880 1890 1900 1910 1920
 1861 AGAGGGCCCTTCGAACAAAACTCATCTCAGAAGAGGATCTGAATATGCATACCGGTCAT
 621 R G P F E Q K L I S E E D L N M H T G H
 1930 1940 1950 1960 1970 1980
 1921 CATCACCATCACCATTGAGTTTAAACCCGCTGATCAGCCTCGACTGTGCCCTTCTAGTTGC
 641 H H H H H *

In vitro editing

The general procedure of *in vitro* editing comprises the actual editing of the mRNA, reverse transcription and amplification of the cDNA by PCR. Subsequent sequencing of the PCR product served as a read-out for the editing yield.^[S4]

CFP Stop66 mRNA synthesis

The mRNA substrate for the *in vitro* editing experiments was generated by *in vitro* transcription of a PCR fragment containing the Stop66 gene under the control of a T7 promotor. Using a PCR fragment as template instead of a plasmid decreases the risk of RNase contamination. The *in vitro* transcription reaction mixture contained 100 ng template, transcription buffer (7.5 µl), DTT (100 mM, 7.5 µl), BSA (1 mg/ml, 3 µl), rNTPs (mix 25 mM, 12 µl) and T7 RNA polymerase (4.5 µl). The volume was adjusted to 50 µl with RNase free water. The mRNA was purified using the RNeasy mini kit (Quiagen). To remove traces of the template DNaseI (Qiagen) digest was performed according to the manufactures protocol. Afterwards, the mRNA was again purified using the RNeasy mini kit (Quiagen). Absence of DNA was controlled by PCR using 2 µl of diluted mRNA (OD = 0.2) as template and the primers Stop66 fw and Stop66 rv. If necessary the DNaseI digest was repeated until the PCR control reaction was negative.

Editing

All *in vitro* editings were performed on a 25 µl scale in PCR-tubes as previously described.^[S4] The editing buffer was composed of 25 mM Tris·HCl, 75 mM KCl, 10 mM DTT (pH 8.3) and 0.75 mM MgCl₂. The substrate mRNA (eCFP Stop66) was used in a concentration of 10 nM and NH₂⁻, BG- or Npom-guideRNA was added in 5-fold excess. The reaction was started by addition of 3.4 equivalents of editing enzyme (SNAP-ADAR1 or SNAP-ADAR2) relative to guideRNA resulting in a final enzyme concentration of 170 nM. Expression and purification of the SNAP-ADAR enzymes were previously described.^[S4,S5] Furthermore, the reaction mixture always contained heparin (2 µM, assuming an average molar mass of 20 kDa) and murine RNase inhibitor (0.5 u/µl, NEB) to prevent degradation of the guideRNAs and the mRNA substrate. Irradiation with 365 nm light was performed on a UV transilluminator (UVP TFL-40V, 25 W, intensity high) for the indicated amount of time at room temperature. The reaction was carried out by cycling between 30°C and 37°C [3x(30 min at 30°C and 30 min at 37°C) for SNAP-ADAR2, 3x(20 min at 30°C and 20 min at 37°C) for SNAP-ADAR1].

Reverse Transcription

Editing was stopped by adding an excess of an antisense DNA oligo (5'-TGACGGCTGGCTGCACCATT, final concentration 20 μ M), $MgCl_2$ (2.6 mM), dNTPs (0.27 mM each), a primer for reverse transcription (Stop66 rv, 0.5 μ M), DTT (5.3 mM), and M-MuLV RT buffer (1.25 μ l, NEB), and heating to 70°C for 3 min. After heating M-MuLV-reverse transcriptase (50 units, 0.5 μ l NEB) was added and the reaction mixture was incubated at 42°C for 2 h. Afterwards the cDNA was purified and concentrated using the NucleoSpin gel and PCR clean-up (Machery-Nagel).

Taq-PCR and agarose gel electrophoresis

The cDNA was amplified by PCR on a 50 μ l scale (denaturation: 15 s at 95°C, annealing: 30 s at 56°C, elongation: 60 s at 68°C, 27 cycles). The PCR reaction mix contained 5 μ l ThermoPol buffer (10 x, NEB), 2.5 μ l of Stop66 fw primer (10 μ M), 2.5 μ l Stop66 rv primer (10 μ M), 1.25 μ l dNTPs (10 mM each, NEB) and 0.5 μ l *Taq* DNA polymerase (5 units/ μ l). 5 μ l of cDNA (typical concentration ca. 10 ng/ μ l) were used as template. Elution buffer from the NucleoSpin kit was used as negative control to preclude contamination of the elution buffer or PCR components. PCR products of previous editings were used as positive control. All PCR reactions were applied to agarose gel electrophoresis (1 x TAE, 1.4%, staining with Rotisafe by Carl Roth GmbH). The eCFP band was cut out under UV light and the PCR product was recovered using the NucleoSpin gel and PCR clean-up kit by Machery-Nagel and sent for sequencing.

Sequencing and processing of sequencing traces

The PCR products (120 ng) were sequenced with Stop66 fw primer () at Eurofins MWG, Ebersberg. The editing yield was determined by comparison of the peak height of adenosine and guanosine in the sequencing traces according to the following formula:

$$\text{Editing yield} = \frac{h(\text{guanosine})}{h(\text{guanosine}) + h(\text{adenosine})} \times 100 \quad h = \text{peak height [cm]}$$

Sequencing traces were processed with DNAMAN 7. The Figure S15 on the next page shows the complete sequencing trace of the CFP s top66 mRNA edited *in vitro*.

Additional experimental data

The light-dependent RNA editing experiment shown in the manuscript in Figure 2c for SNAP-ADAR1 was also carried out with SNAP-ADAR2 as shown below (Figure S16). Again, the full dynamic range of editing is obtained from no editing prior to irradiation to the editing yield of the positive editing control after >60 sec irradiation. The maximum yield obtained for SNAP-ADAR2 is a little bit reduced compared to SNAP-ADAR1 which is caused by the stronger sensitivity of the first towards heparin.

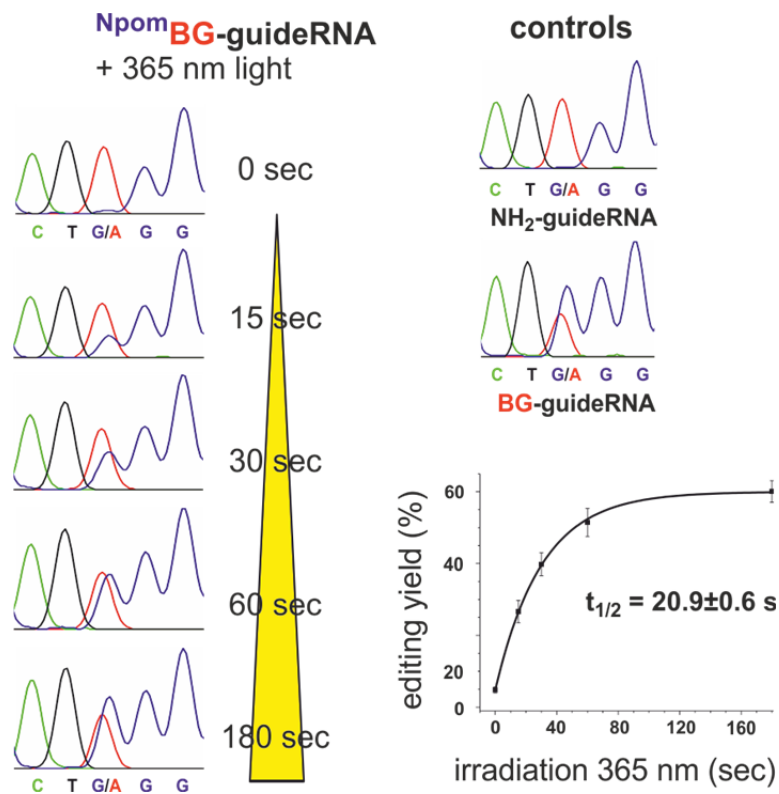


Figure S16 Light-dependent in-vitro site-directed RNA editing of the amber Stop codon at position 66 in the CFP gene by SNAP-ADAR2.

Editing in cell culture

Cell culture techniques

HEK293T cells (DSMZ Braunschweig, Germany, ACC-635) were grown in Dulbecco's Modified Eagle Medium (DMEM, Life Technologies) supplemented with 10% fetal bovine serum (FBS, Life Technologies) and 1% penicillin/streptomycin (P/S, Life Technologies) under standard conditions (37°C and 5% CO² in a water saturated steam atmosphere).

Plasmid transfection

HEK 293T cells (DSMZ code ACC-635, passage number < 15) were grown in a 25 cm² cell culture flask to a confluency of 70-90%. The medium was removed; the cells were washed with 5 ml PBS and trypsinated with 500 µl trypsin/EDTA for 3 min at 37°C. Then the protease was blocked by adding 4.5 ml DMEM+FBS+P/S. Cells were incubated with trypan blue for 5 min and the cell number was determined in a hemacytometer. Cells were seeded onto a 24 well plate (200 000 cells/well) in 500 µl DMEM+FBS+P/S grown for 24 h. 60-120 min before plasmid transfection the medium was replaced with 450 µl DMEM+FBS without antibiotics. Lipofectamine 2000 (Life Technologies, 4 µl per 1 µg DNA) was used as transfection reagent. For transfection, plasmid DNA and Lipofectamine were diluted in 50 µl OptiMem (Life Technologies) independently and incubated for 5 min at r.t. allowing micelle formation of the transfection agent. Then both solutions were mixed, incubated for another 20 min at r.t. for complex formation and applied to the cells. The transfected cells were incubated for 24 h.

In a typical experiment 500 ng of W58X coding plasmid (pcDNA3.1, Invitrogen) was co-transfected with 25 to 200 ng SNAPf-ADAR1 coding plasmid (pcDNA3.1, Invitrogen). Furthermore, W58X was co-transfected with empty pcDNA3.1 as negative control. The plasmid coding for functional EGFP was transfected together with SNAPf-ADAR1 as positive control.

guideRNA transfection

After 24 h incubation, the medium was removed; cells were washed with 500 µl PBS, trypsinated (60µl trypsin/EDTA) and taken up in 500 µl DMEM+FBS+P/S. The cell suspension was centrifuged (1600 rpm, 5 min, r.t.), the supernatant was removed, the cell pellet re-suspended in DMEM+FBS without antibiotics and the cell number was determined in a hemacytometer. The cells were reverse transfected onto a 96 well plate (60 000 cells/well, in 100 µl DMEM+FBS) containing the respective guideRNAs (typically 10 pmol/well) pre-treated with Lipofectamine 2000 (0.5 µL/well). guideRNA and Lipofectamine were diluted in 25 µl OptiMem independently, incubated for 5 min, mixed and incubated another 20 min before applying the mixture to the 96 well plate.

All experiments involving Npom-caged guideRNAs were performed in the absence of direct light or under red light (590-660 nm, LED Spot Luexon Red, Conrad Electronic). 96 well plates were wrapped in aluminum foil to protect Npom-guideRNAs from light.

Npom-decaging

After 4 h of guideRNA transfection, the medium was replaced by 100 μ l DMEM+HEPES (25 mM) without phenol red. Irradiation (365 nm) was performed in a fluorescence microscope (Zeiss CellObserverZ.1, equipped with a 365 nm Colibri.2 LED) at 100% LED power for the indicated amount of time.

Imaging & data processing

All cells were imaged 24 h after irradiation using a Zeiss AXIO Observer.Z1 with a Colibri.2 light source. For each sample, a phase contrast image acquired and EGFP fluorescence signal was recorded (Excitation: 465 nm, 50% LED power, 460-480 nm band pass filter, Emission: 500-557 nm band pass filter). All samples of one experiment were recorded with the same settings (exposure time, etc.). Images were processed in ImageJ 1.47h and the adjustment of contrast and brightness was carried out for all fluorescence images identically.

Cell harvesting and RNA isolation

After imaging, the cells were harvested and the RNA was isolated and reverse transcribed. The obtained cDNA was amplified by PCR, worked-up and sequenced.

The medium was removed from the 96 well-plates and the cells were trypsinated with 20 μ l Trypsin/EDTA for 1 min. Then, DMEM+FBS+P/S (80 μ l) was added, the cell suspension was transferred to a 1.5 ml reaction tube and centrifuged (1600 rpm, 5 min, r.t.). The supernatant was removed and the cell pellets were washed with PBS. After another centrifugation step (1600 rpm, 5 min, r.t.) and removal of the PBS, the pellets were frozen in liquid nitrogen and stored for up to 3 days at -80°C.

For RNA isolation, the pellets were thawed on ice and suspended in lysis buffer (RLT buffer from Qiagen RNeasy mini kit + 1% mercaptoethanol (v/v)). To enhance cell disruption, the cells were exposed to shearing stress by pipetting the suspension through a needle (\varnothing = 0.25 mm) several times. After lysis, RNA purification was performed according to the RNeasy mini kit manual (Qiagen). In a typical experiment 600-1800 ng RNA were recovered per well.

DNase digest and reverse transcription

For DNase digest 2.9 μ l 10 x RDD buffer and 1 μ l DNase I (Qiagen) were added to 25 μ l RNA sample and incubated for 30 min at 37°C. The DNase was inactivated by adding 1 μ l EDTA (25 mM) and heating to 65°C for 10 min. 14.75 μ l of the DNase digest were mixed with 1 μ l W58X RT (10 μ M) primer and 1 μ l dNTPs (10mM each) and incubated at 70°C for 3 min. Then, 2 μ l 10 x M-MuLV-RT buffer, 0.25 μ l murine RNase inhibitor (40 units/ μ l, NEB) and 1 μ l M-MuLV reverse transcriptase (200

units/ μl , NEB) were added and the solution was incubated at 42°C for 1 h. M-MuLV reverse transcriptase was inactivated by heating to 90°C for 10 min. The cDNA was purified using the NucleoSpin gel and PCR clean-up kit (Machery-Nagel). In a typical experiment, reverse transcription yielded 150-600 ng cDNA.

Taq-PCR and agarose gel electrophoresis

The cDNA was amplified by PCR on a 50 μl scale. Furthermore, a control PCR was performed using the DNase digest as template to verify the absence of interfering plasmid DNA. The PCR reaction mix contained 5 μl ThermoPol buffer (10 x), 2.5 μl of W58X fw primer (10 μM), 2.5 μl W58X rv primer (10 μM), 1.25 μl dNTPs (10 mM each) and 0.5 μl Taq DNA polymerase (5 U/ μl , NEB). 5 μl of cDNA sample or 2.5 μl RNA after DNase digest were used as template. Furthermore, RNase free water and elution buffer from the NucleoSpin kit were used as negative control to preclude contamination of the PCR components. W58X plasmid DNA was used as positive control. 10 μl of 6 x blue loading dye (NEB) were added and all PCR reactions were applied to agarose gel electrophoresis (1 x TAE, 1.4%, staining with Rotisafe by Carl Roth GmbH). The W58X/EGFP band was cut out and the PCR product was recovered using the NucleoSpin gel and PCR clean-up kit by Machery-Nagel. If any bands showed up in the DNase control PCR, the samples were processed again starting from DNase digest.

Sequencing

For sequencing 120 ng of the PCR products were send to Eurofins MWG. W58X fw (1.5 μl , 10 μM) was used as primer for all samples. The editing yield was determined as described for Stop66 editing *in vitro*. Figure S18 shows the complete sequencing trace of the W58X editing in cell culture. The amplified fragment corresponds to base pairs 44-668 of the EGFP ORF.

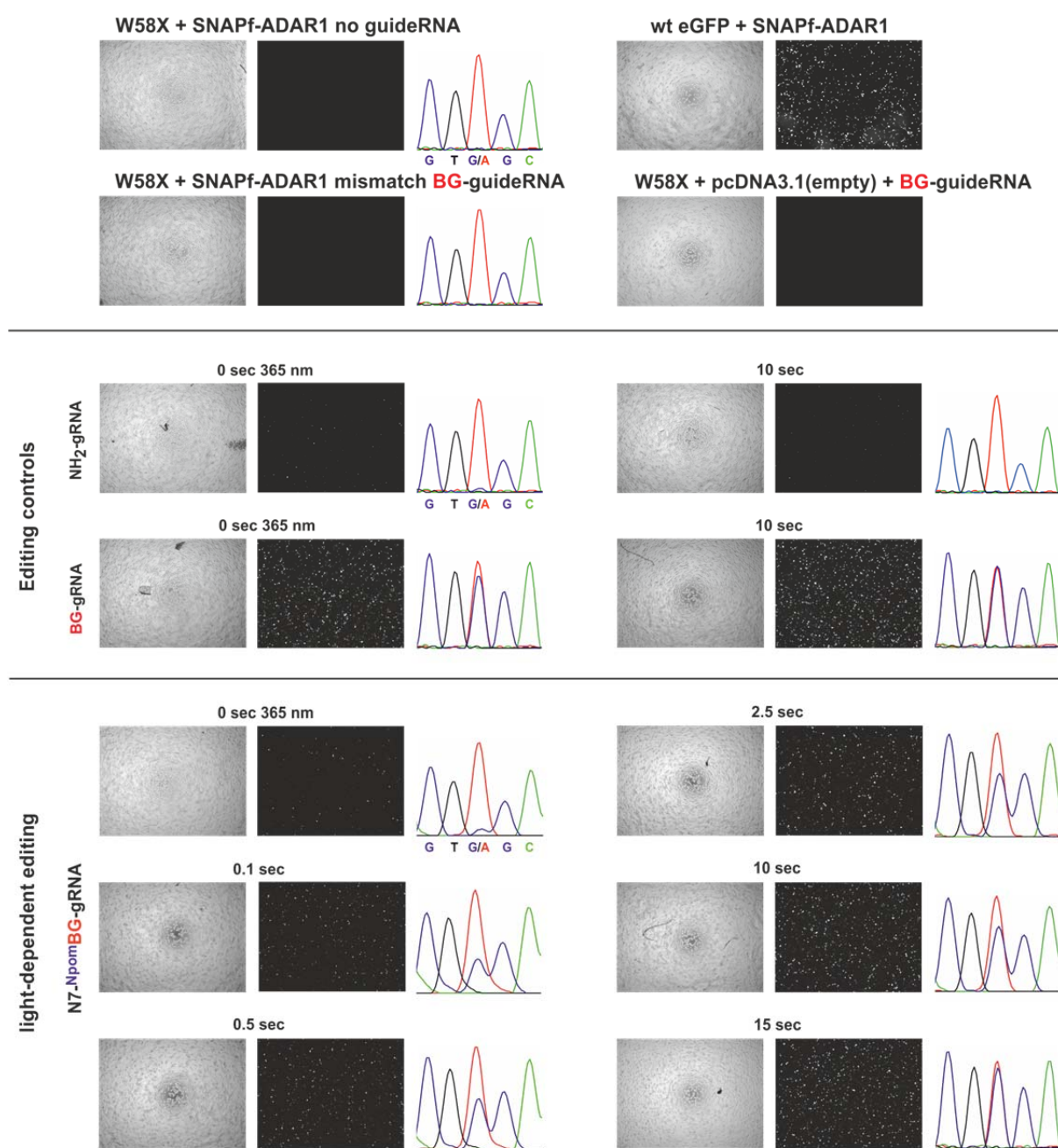


Figure S17. Light-controlled RNA editing in living 293T cells (corresponding to experiments shown in Figure 3 manuscript). For all samples phase contrast imaging (left), eGFP fluorescence imaging (middle) and Sanger sequencing traces (right) 24 h post transfection is shown.

Additional experimental data

Figure S19 and S20 demonstrate the dependence of the editing yield and Npom-guideRNA residual activity on the amount of transfected guideRNA and plasmid. Remarkably, the system yields reasonable editing (20-50%) even when the amount of plasmid or BG-guideRNA is varied by about 1 order of magnitude.

With a constant amount of BG-guideRNA (10 pmol/well) the editing yield decreases from ~50% with 200 ng SNAPf-ADAR1 plasmid to ~20% with 25 ng SNAPf-ADAR1. At the same time the residual activity of Npom-guideRNA in the dark is decreases from ~15% to <5%.

With a constant amount of SNAPf-ADAR1 enzyme (100 ng) the editing yield decreases from ~50% with 50 pmol BG-guideRNA to ~20% with 2 pmol BG-guideRNA. Residual activity of Npom-guideRNA can be efficiently suppressed by lowering the amount of guideRNA.

The presented results indicate that SNAPf-ADAR1 expression level and the amount guideRNA can be used to adjust both editing yield and residual activity for future applications.

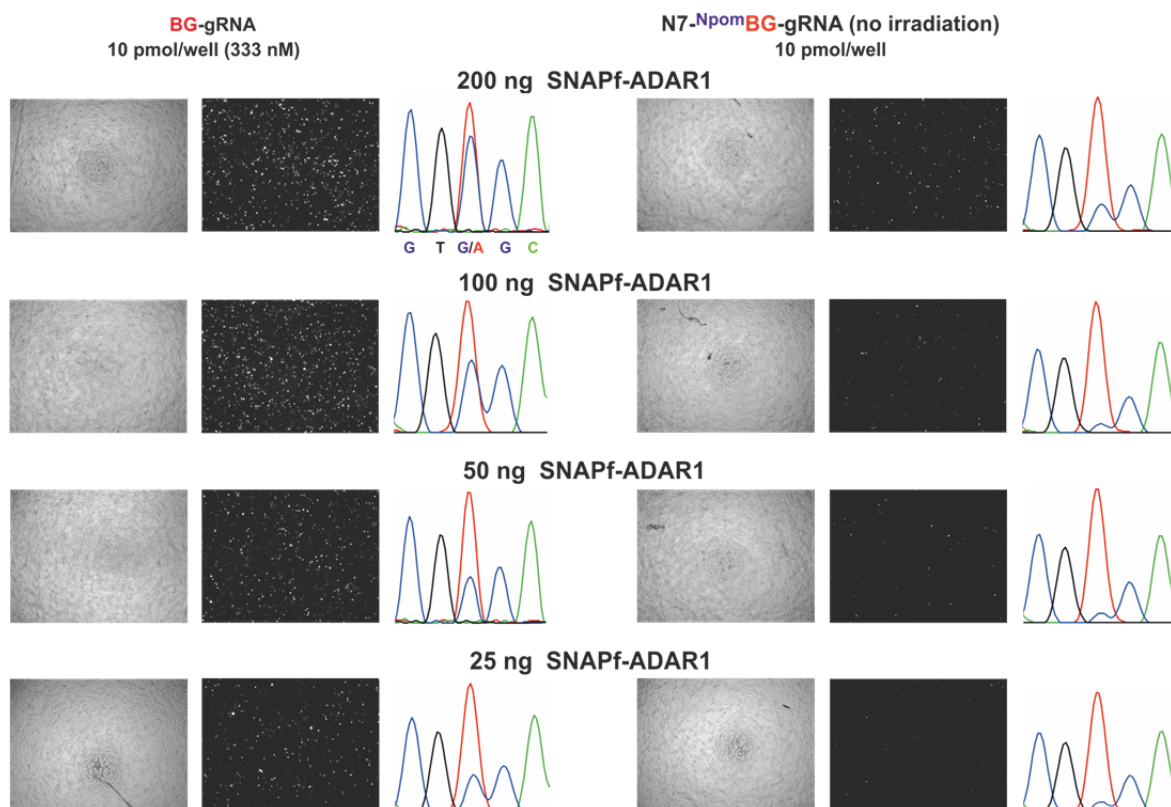


Figure S19. RNA editing in living 293T cells with different amounts of SNAPf-ADAR1 plasmid transfected. All cells were transfected with a constant amount of guideRNA (10 pmol/well).

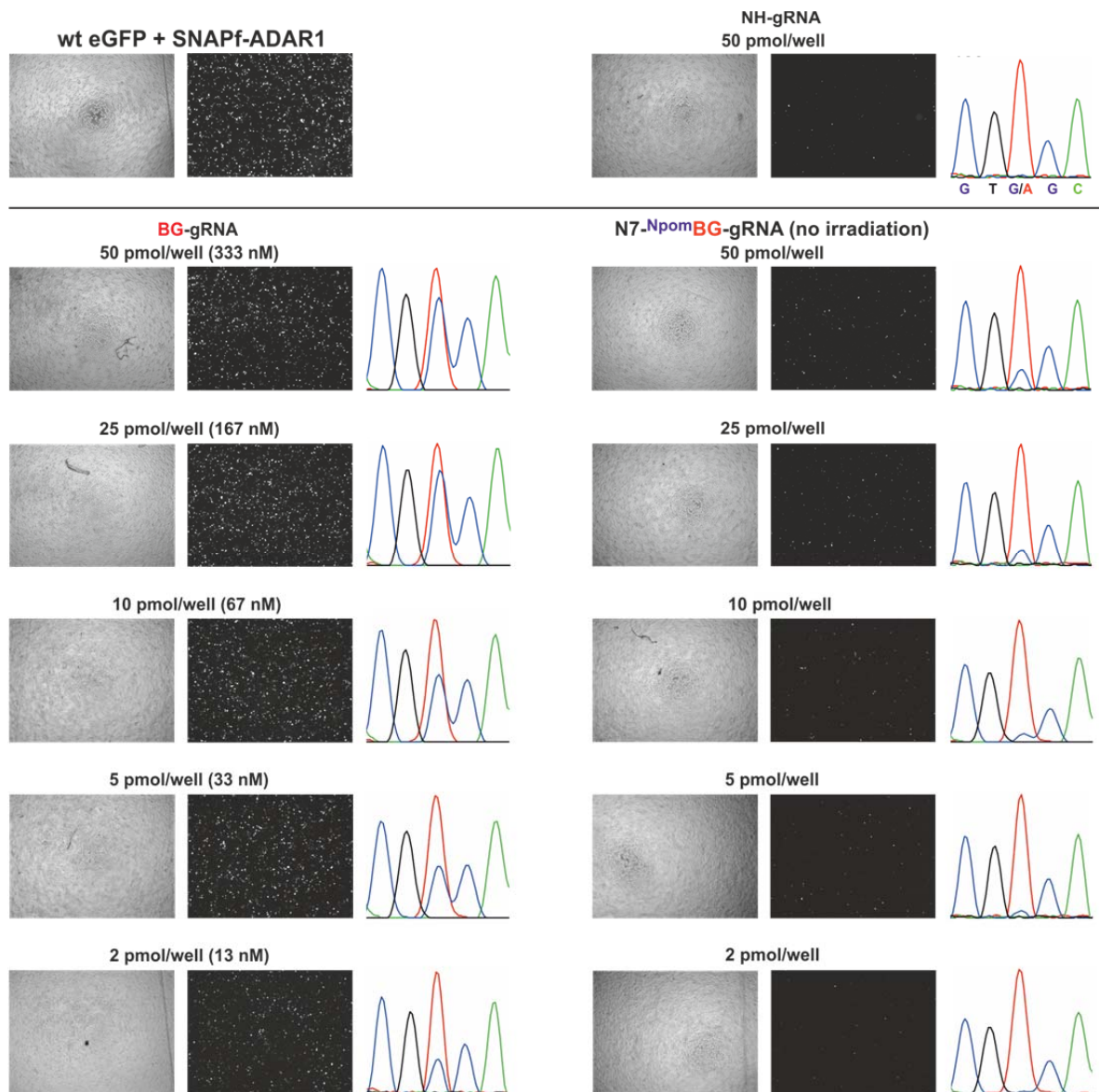


Figure S20. RNA editing in living 293T cells with different amounts of guideRNA transfected. All cells were transfected with a constant amount of SNAPf-ADAR1 plasmid (100ng).

Editing in Platynereis Dumerilii

Materials and Methods

Gene (transcript) sequences of eGFP and SNAP-ADAR1

W58X eGFP in the pUC57-T7-RPP2 context

```

          10          20          30          40          50          60
1  TAATACGACTCACTATAGGGAGATTTGATGTTTACAGGGCTATTTATAAAACAAATTGTTA
   T7 promoter

          70          80          90          100         110         120
61  ATAATTTAGGTGGAAATTATTTTGGTGTCTCAAACCCACATGTTTCAGGAAGCTGTGGCCC

          130         140         150         160         170         180
121  CCAAATCTTTCTTTCTGTTTGAGAAATTTTCTGGTGTGCACACGTTTTTCGTGTCTCTTGG

          190         200         210         220         230         240
181  AAGACTTAAAAAATGGCGCGCCTAGCTAGCAAAGGAGAAGAACTCTTCACTGGAGTTGTC
      M A R L A S K G E E L F T G V V

          250         260         270         280         290         300
241  CCAATTCTTGTGAATTAGATGGTGTGTTAACGGCCACAAGTTCTCTGTCACTGGAGAG
81  P I L V E L D G D V N G H K F S V S G E

          310         320         330         340         350         360
301  GGTGAAGGTGATGCAACATACGGAAAACCTTACCCTGAAGTTCATCTGCACTACTGGCAAA
101  G E G D A T Y G K L T L K F I C T T G K

          370         380         390         400         410         420
361  CTGCCTGTTCCGTAGCCGACACTAGTGACGACGCTCTGCTATGGCGTCCAGTGCTTTTCA
121  L P V P * P T L V T T L C Y G V Q C F S

          430         440         450         460         470         480
421  AGATACCCGGATCACATGAAACGGCATGACTTTTTTCAAGAGTGCCATGCCCGAAGTTAT
141  R Y P D H M K R H D F F K S A M P E G Y

          490         500         510         520         530         540
481  GTACAGGAAAGGACCATCTTCTTCAAAGATGACGGCAACTACAAGACACGTGCTGAAGTC
161  V Q E R T I F F K D D G N Y K T R A E V

          550         560         570         580         590         600
541  AAGTTTGAAGGTGATACCCCTTGTAAATAGAATCGAGTTAAAAGGTATTGACTTCAAGGAA
181  K F E G D T L V N R I E L K G I D F K E

          610         620         630         640         650         660
601  GATGGCAACATTCTGGGACACAAATTGGAATACAACCTATAACTCACACAATGTATACATC
201  D G N I L G H K L E Y N Y N S H N V Y I

          670         680         690         700         710         720
661  ATGGCAGACAAAACAAAAGAATGGAATCAAAGTGAACCTTCAAGACCCGCCACAACATTGAA
221  M A D K Q K N G I K V N F K T R H N I E

          730         740         750         760         770         780
721  GATGGAAGCGTTCAACTAGCAGACCATTATCAACAAAATACTCCAATTGGCGATGGCCCT
241  D G S V Q L A D H Y Q Q N T P I G D G P
```

```

781          790          800          810          820          830          840
261  GTCCCTTTTACCAGACAACCATTTACCTGTCCACACAATCTGCCCTTTTCGAAAGATCCCAAC
    V L L P D N H Y L S T Q S A L S K D P N

841          850          860          870          880          890          900
281  GAAAAGAGAGACCACATGGTCCTTCTTGAGTTTGTAACAGCTGCTGGGATTACACATGGC
    E K R D H M V L L E F V T A A G I T H G

901          910          920          930          940          950          960
301  ATGGATGAACTATACAAAACCGGTTAAAAGAAACACTTTTACAAACATCAGTCTGTAACA
    M D E L Y K T G *

961          970          980          990          1000          1010          1020
    TCTTTCCAATAAAAAAAAAAACAATGTAACCTTACTGGTCTGGAGTTGTTTAAAGAGAAAATTG

1021          1030          1040          1050
    GATCTAGATGCATTCGCGAGGTACCGAG(CTC)→ Polyadenylation
    Eco53kI

```

In yellow the transcription start site and the premature W58X site are highlighted.
The ORF is marked by the single letter amino acid code.

SNAP-ADAR1 in the pUC57-T7-RPP2 context

```

1          10          20          30          40          50          60
    TAATACGACTCACTATAAGGGAGATTTGATGTTTACAGGGCTATTTATAAACAATTTGTTA
    T7 promoter

61          70          80          90          100          110          120
    ATAATTTAGGTGGAATTATTTTGGTGTCTCAAACCCACATGTTTCAGGAAGCTGTGGCCC

121          130          140          150          160          170          180
    CCAAATCTTTCTTTCTGTTTGGAGAAATTTTCTGGTGTGCACACGTTTTTCGTGTCTCTTGG

181          190          200          210          220          230          240
    AAGACTTAAAAAATGGCGCTCTAGACAAAAGACTGCGAAATGAAGCGCACCACCCTGGAT
    M A R L D K D C E M K R T T L D

241          250          260          270          280          290          300
81  AGCCCTCTGGGCAAGCTGGAACCTGTCTGGGTGCGAACAGGGCCTGCACCGTATCATCTTC
    S P L G K L E L S G C E Q G L H R I I F

301          310          320          330          340          350          360
101 CTGGGCAAAGGAACATCTGCCGCCGACGCCGTGGAAGTGCCTGCCCCAGCCGCCGTGCTG
    L G K G T S A A D A V E V P A P A A V L

361          370          380          390          400          410          420
121 GGCGGACCAGAGCCACTGATGCAGGCCACCGCCTGGCTCAACGCCTACTTTCACCAGCCT
    G G P E P L M Q A T A W L N A Y F H Q P

421          430          440          450          460          470          480
141 GAGGCCATCGAGGAGTTCCCTGTGCCAGCCCTGCACCACCCAGTGTTCAGCAGGAGAGC
    E A I E E F P V P A L H H P V F Q Q E S

481          490          500          510          520          530          540
161 TTTACCCGCCAGGTGCTGTGGAAACTGCTGAAAGTGGTGAAGTTCGGAGAGGTCATCAGC
    F T R Q V L W K L L K V V K F G E V I S

```

541 550 560 570 580 590 600
181 TACAGCCACCTGGCCGCCCTGGCCGGCAATCCCGCCGCCACCGCCCGCTGAAAACCGCC
Y S H L A A L A G N P A A T A A V K T A

601 610 620 630 640 650 660
201 CTGAGCGGAAATCCCGTGCCATTCTGATCCCCTGCCACCGGGTGGTGCAGGGCGACCTG
L S G N P V P I L I P C H R V V Q G D L

661 670 680 690 700 710 720
221 GACGTGGGGGGCTACGAGGGCGGGCTCGCCGTGAAAGAGTGGCTGCTGGCCCACGAGGGC
D V G G Y E G G L A V K E W L L A H E G

721 730 740 750 760 770 780
241 CACAGACTGGGCAAGCCTGGGCTGGGTCTGCAGGCGGAGGCGCGCCAGGGTCTGGCGGC
H R L G K P G L G P A G G G A P G S G G

781 790 800 810 820 830 840
261 GGCAGTAAGGCAGAACGCATGGGTTTTCACAGAGGTAACCCAGTGACAGGGGCCAGTCTC
G S K A E R M G F T E V T P V T G A S L

841 850 860 870 880 890 900
281 AGAAGAACTATGCTCCTCCTCTCAAGGTCCCCAGAAGCACAGCCAAAGACACTCCCTCTC
R R T M L L L S R S P E A Q P K T L P L

901 910 920 930 940 950 960
301 ACTGGCAGCACCTTCCATGACCAGATAGCCATGCTGAGCCACCGGTGCTTCAACACTCTG
T G S T F H D Q I A M L S H R C F N T L

961 970 980 990 1000 1010 1020
321 ACTAACAGCTTCCAGCCCTCCTTGCTCGGCCGCAAGATTCTGGCCGCCATCATTATGAAA
T N S F Q P S L L G R K I L A A I I M K

1021 1030 1040 1050 1060 1070 1080
341 AAAGACTCTGAGGACATGGGTGTCGTCGTCAGCTTGGGAACAGGGAATCGCTGTGTA
K D S E D M G V V V S L G T G N R C V K

1081 1090 1100 1110 1120 1130 1140
361 GGAGATTCTCTCAGCCTAAAAGGAGAAACTGTCAATGACTGCCATGCAGAAATAATCTCC
G D S L S L K G E T V N D C H A E I I S

1141 1150 1160 1170 1180 1190 1200
381 CGGAGAGGCTTCATCAGGTTTCTCTACAGTGAGTTAATGAAATACAACCTCCAGACTGCG
R R G F I R F L Y S E L M K Y N S Q T A

1201 1210 1220 1230 1240 1250 1260
401 AAGGATAGTATATTTGAACCTGCTAAGGGAGGAGAAAAGCTCCAAATAAAAAAGACTGTG
K D S I F E P A K G G E K L Q I K K T V

1261 1270 1280 1290 1300 1310 1320
421 TCATTCCATCTGTATATCAGCACTGCTCCGTGTGGAGATGGCGCCCTCTTTGACAAGTCC
S F H L Y I S T A P C G D G A L F D K S

1321 1330 1340 1350 1360 1370 1380
441 TGCAGCGACCGTGTCTATGGAAAGCACAGAATCCCGCCACTACCTGTCTTCGAGAATCCC
C S D R A M E S T E S R H Y P V F E N P

1381 1390 1400 1410 1420 1430 1440
461 AAACAAGGAAAGCTCCGCACCAAGGTGGAGAACGGGAGAAAGGCACAATCCCCTGTGGAATCC
K Q G K L R T K V E N G E G T I P V E S

1450 1460 1470 1480 1490 1500

```

1441 AGTGACATTGTGCCTACGTGGGATGGCATTTCGGCTCGGGGAGAGACTCCGTACCATGTCC
481 S D I V P T W D G I R L G E R L R T M S

          1510          1520          1530          1540          1550          1560
1501 TGTAGTGACAAAATCCTACGCTGGAACGTGCTGGGCCTGCAAGGGGCACTGTTGACCCAC
501 C S D K I L R W N V L G L Q G A L L T H

          1570          1580          1590          1600          1610          1620
1561 TTCCTGCAGCCCATTTATCTCAAATCTGTCACATTGGGTTACCTTTTCAGCCAAGGGCAT
521 F L Q P I Y L K S V T L G Y L F S Q G H

          1630          1640          1650          1660          1670          1680
1621 CTGACCCGTGCTATTTGCTGTCGTGTGACAAGAGATGGGAGTGCATTTGAGGATGGACTA
541 L T R A I C C R V T R D G S A F E D G L

          1690          1700          1710          1720          1730          1740
1681 CGACATCCCTTTTATTGTCAACCACCCCAAGGTTGGCAGAGTCAGCATATATGATTCAAA
561 R H P F I V N H P K V G R V S I Y D S K

          1750          1760          1770          1780          1790          1800
1741 AGGCAATCCGGGAAGACTAAGGAGACAAGCGTCAACTGGTGTCTGGCTGATGGCTATGAC
581 R Q S G K T K E T S V N W C L A D G Y D

          1810          1820          1830          1840          1850          1860
1801 CTGGAGATCCTGGACGGTACCAGAGGCACTGTGGATGGGCCACGGAATGAATTGTCCCGG
601 L E I L D G T R G T V D G P R N E L S R

          1870          1880          1890          1900          1910          1920
1861 GTCTCCAAAAAAGAACATTTTTCTTCTATTTAAGAAGCTCTGCTCCTTCCGTTACCGCAGG
621 V S K K N I F L L F K K L C S F R Y R R

          1930          1940          1950          1960          1970          1980
1921 GATCTACTGAGACTCTCCTATGGTGAGGCCAAGAAAAGCTGCCCGTGACTACGAGACGGCC
641 D L L R L S Y G E A K K A A R D Y E T A

          1990          2000          2010          2020          2030          2040
1981 AAGAACTACTTCAAAAAAGGCCTGAAGGATATGGGCTATGGGAACTGGATTAGCAAACCC
661 K N Y F K K G L K D M G Y G N W I S K P

          2050          2060          2070          2080          2090          2100
2041 CAGGAGGAAAAGAACTTTTATCTCTGCCAGTATCCGGTTAAAAAGAAACACTTTTACAAA
681 Q E E K N F Y L C P V S G *

          2110          2120          2130          2140          2150          2160
2101 CATCAGTCTGTAACATCTTTCCAATAAAAAAAAAAACATGTAACCTACTGGTCTGGAGTTG

          2170          2180          2190          2200
2161 TTTAAGAGAAAATTGGATCTAGATGCATTTCGCGAGGTACCGAG (CTC) → Polyadenylation
Eco53kI

```

In yellow the transcription start site is highlighted. The ORF is marked by the single letter amino acid code.

SNAP-ADAR1 and reporter mRNA synthesis

For the in-vitro transcription of correctly m⁷G capped and poly(A)-tailed mRNAs of SNAP-ADAR1 and eGFP variants (wt, W58X) the mMESSAGING mMACHINE T7 Ultra Kit (Life Technologies) was used. For additional stabilization, the respective genes were subcloned into the pUC57-T7-RPP2 vector containing a 169 bp 5'-UTR from the Platynereis 60S acidic ribosomal protein P2 downstream of the T7 promoter. To avoid the formation of long heterogeneous RNA transcripts and to facilitate “run-off” transcription, the expression plasmids were linearized prior to in-vitro transcription. The linearization reaction mix contained 5 µL CutSmart buffer (10 x), 15 µg plasmid DNA and 2.5 µL Eco53kl (10 units/µL). The total reaction volume was adjusted to 50 µL with nuclease-free water. To terminate the linearization reaction 2.5 µL 0.5 M EDTA, 0.5 µL 3 M NaOAc and 100µl ethanol (RNase free) were added. The mixture was cooled for 1h at - 20°C, pelleted (15 min, 15000 rpm) and re-suspended in nuclease-free water at a concentration of 1 µg/ µL. 2 µg of the respective linearized plasmid template was in-vitro-transcribed with the mMESSAGING mMACHINE T7 Ultra Kit (Life Technologies). All steps were performed according to the manufacturer’s protocol. To verify the correct polyadenylation state, the synthesized transcripts were mixed with 2 µL of 5 x RNA loading buffer (95% formamide, 0.025 % bromophenol blue) denatured (70°C, 5 min) and applied to agarose gel electrophoresis (1 x TBE, 1.4%, staining with Rotisafe by Carl Roth GmbH, see Fig S21).

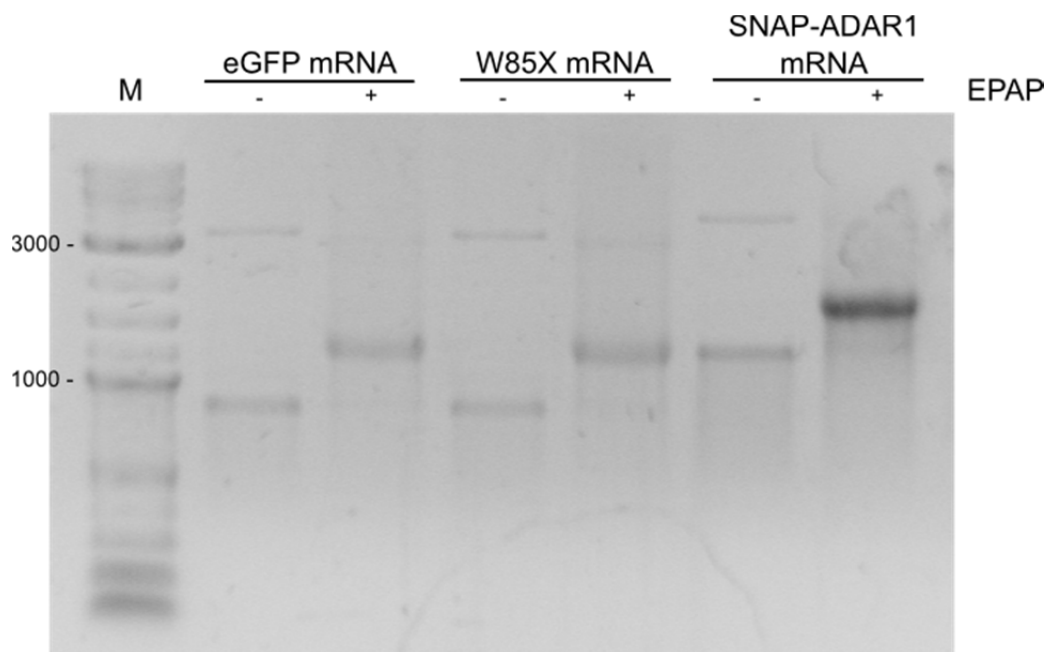


Figure S21. Analytical TBE agarose gel (1.4 %) of iv-transcribed mRNAs for microinjection prior to and after polyadenylation.

Ethanol precipitation of guideRNAs

Prior to the in vivo application of the editing system in *Platynereis Dumerilii* the synthesized guideRNAs had to be purified. Ethanol precipitation was used to get rid of impurities and detergents and to exchange the counter ion of the guideRNAs to potassium. Therefore, the volume of the guideRNA samples was adjusted to 300 μL with nuclease-free water. Subsequently 33 μL 3 M KOAc and 1 ml 100 % ethanol were added. The mixture was stored at -80°C overnight and centrifuged again (30 min, -4°C , 15000 rpm). The supernatant was removed carefully and the pellet was washed with 500 μL 70% ethanol (10 min, 4°C , 15000 rpm) and re-suspended in 15 μL nuclease-free water.

***Platynereis dumerilii* zygotes**

Full-grown and mature male and female *Platynereis* were obtained from the in-house breeding culture of the laboratory of Dr. Gáspár Jékely (Max Planck Institute for Developmental Biology, 72076 Tübingen, Germany). Reproduction of the worm was initiated by collecting two males and females in one beaker containing 75 mL natural seawater (NSW). The release of sperm and oocytes was monitored by eye and eventually the animals were removed carefully with a plastic pasteur pipette. The fertilized eggs were washed one time by pouring off half of the NSW and refilling the beaker. To facilitate microinjection, the zygotes were incubated for 55 min at 14.8°C and the egg jelly was removed by rinsing the zygotes with 500 mL NSW in a 100 μm sieve. The vitelline envelope was softened by a minute-long proteinase K treatment (final concentration: 70 $\mu\text{g}/\text{mL}$) and an additional washing step with 500 μL NSW. Approximately 100 zygotes were embedded into the channel of the injection stage (2% agarose in NSW). During the whole procedure of washing and the following microinjection, care has been taken to always keep the eggs covered with NSW to avoid draining.

Microinjection of *Platynereis* zygotes

Injection samples contained 1.5 $\mu\text{g}/\mu\text{L}$ Rhodamine-dextran (10 kDa MW, Sigma) for injection control, 250 ng/ μL of the respective reporter mRNA, 450 ng/ μL SNAP-ADAR1 mRNA, and 25 μM of the respective guideRNA. Prior to microinjection the samples were centrifuged and 3.5 μL of the supernatant were loaded into a Femtotips II microcapillarie (Eppendorf). Microinjection was performed with a Femtojet express microinjector (Eppendorf) on a Zeiss Axiovert 40 CL inverted microscope combined with a Luigs and Neumann micromanipulator. The temperature was controlled using a Luigs and Neumann Badcontroller V cooling system and a Cyclo 2 water pump (Roth). The injection session was carried out at 14.8°C . The start parameter for microinjection (injection pressure: 700 hPa; injection time: 0.1 s; compensation pressure: 35 hPa) were adjusted accordingly to the condition of the microcapillarie. Injection was started 1 hour post fertilization (hpf) and stopped when the first cleavage of the *Platynereis* zygotes could be detected

(ca. 2 hpf). Microinjected zygotes were bred at 19°C in Nunclon six-well plates containing 6 mL NSW until the desired stage of development was reached. For decaging of Npom-protected guideRNAs the injected zygotes were collected in a Nunclon petri dish containing 6 ml NSW, put on a UV-table (UVP high performance UV transilluminator, 25 W, 365 nm, power=high) and irradiated for the indicated amount of time.

Immobilization and imaging of Platynereis larvae

Platynereis larvae were imaged 24 hpf. Healthy animals had to be immobilized to allow proper imaging. Therefore, glass slides with 3 layers of adhesive tape on both sides were prepared. 10 µl of NSW containing 3-5 larvae were mounted in the middle of the slide and trapped with a coverslip, forming a small chamber for the animals.

Imaging was performed with an AxioCamHRc microscope camera connected to a Axio Imager Z1 widefield fluorescence microscope (Zeiss). The eGFP and the rhodamine-dextran signal were recorded with an exposure time of 750 ms and 350 ms respectively for 40x magnification. For 10x magnification an exposure time of 2000 ms (eGFP) and 1063 ms (rhodamine-dextran) were chosen.

Platynereis harvesting and RNA isolation

Larvae were harvested 25 hpf. Healthy ones were separated from unhealthy ones and collected in a 1.5 ml reaction tube. As much NSW as possible was removed from the tube and the larvae were frozen in liquid nitrogen and stored at - 80°C till further use.

Larvae were thawed on ice and 80–100 animals were used for RNA isolation. If needed, larvae of 2 injection sessions with the same injection sample were pooled. Lysis and RNA purification was performed with the RNeasy MinElute Cleanup Kit (Qiagen). All steps were performed according to the manufacturers protocol. To facilitate lysis of Platynereis, the RLT lysis buffer + 1% mercaptoethanol (v/v), shear forces (passing through 0.6 mm needle) and vortexing (10 s) was used. The purified RNA was eluted from the spin columns in 30 µl RNase free water and the RNA concentration was determined.

Reverse transcription

500 ng of extracted total RNA were reverse transcribed into cDNA. The sample was filled up to 14.75 µl with RNase free water, mixed with 1 µl W58X PD bw primer (10 µM) and 1 µl dNTPs (10 mM each) and incubated at 70 °C for 3 min. The mixture was immediately spun down and put on ice. Subsequently 2 µl 10 x M-MuLV-RT buffer, 0.25 µl murine RNase inhibitor (40 units/µl) and 1 µl M-MuLV reverse transcriptase (200 units/µl) were added and the solution was incubated at 42°C for 2 h. M-MuLV reverse transcriptase was inactivated by heating to 90°C for 10 min.

Taq-PCR and agarose gel electrophoresis

The cDNA was amplified by PCR on a 50 µl scale. The PCR reaction mix contained 5 µl ThermoPol buffer (10 x), 2.5 µl of W58X PD fw primer (10 µM), 2.5 µl W58X PD bw primer (10 µM), 1.25 µl dNTPs (10 mM each), 1 µl *Taq* DNA polymerase (5 units/µl) and 7 µl of the unpurified cDNA reaction mixture. 10 µl of 6 x blue loading dye (NEB) were added and all PCR reactions were applied to agarose gel electrophoresis (1 x TAE, 1.4%, staining with Rotisafe by Carl Roth GmbH). The W58X/eGFP band was cut out and the PCR product was recovered using the NucleoSpin gel and PCR clean-up kit by Macherey-Nagel.

Sequencing

For sequencing 120 ng of the PCR products were send to Eurofins MWG (Germany). W58X fw (1.5 µl, 10 µM) and W58Xpos327 fw (1.5 µl, 10µM) were used as primers for all samples respectively. The editing yield was determined as described for Stop66 editing *in vitro*. Figure S24-S26 show complete sequencing traces of W58X edited in *Platynereis Dumerilii*.

Additional experimental data

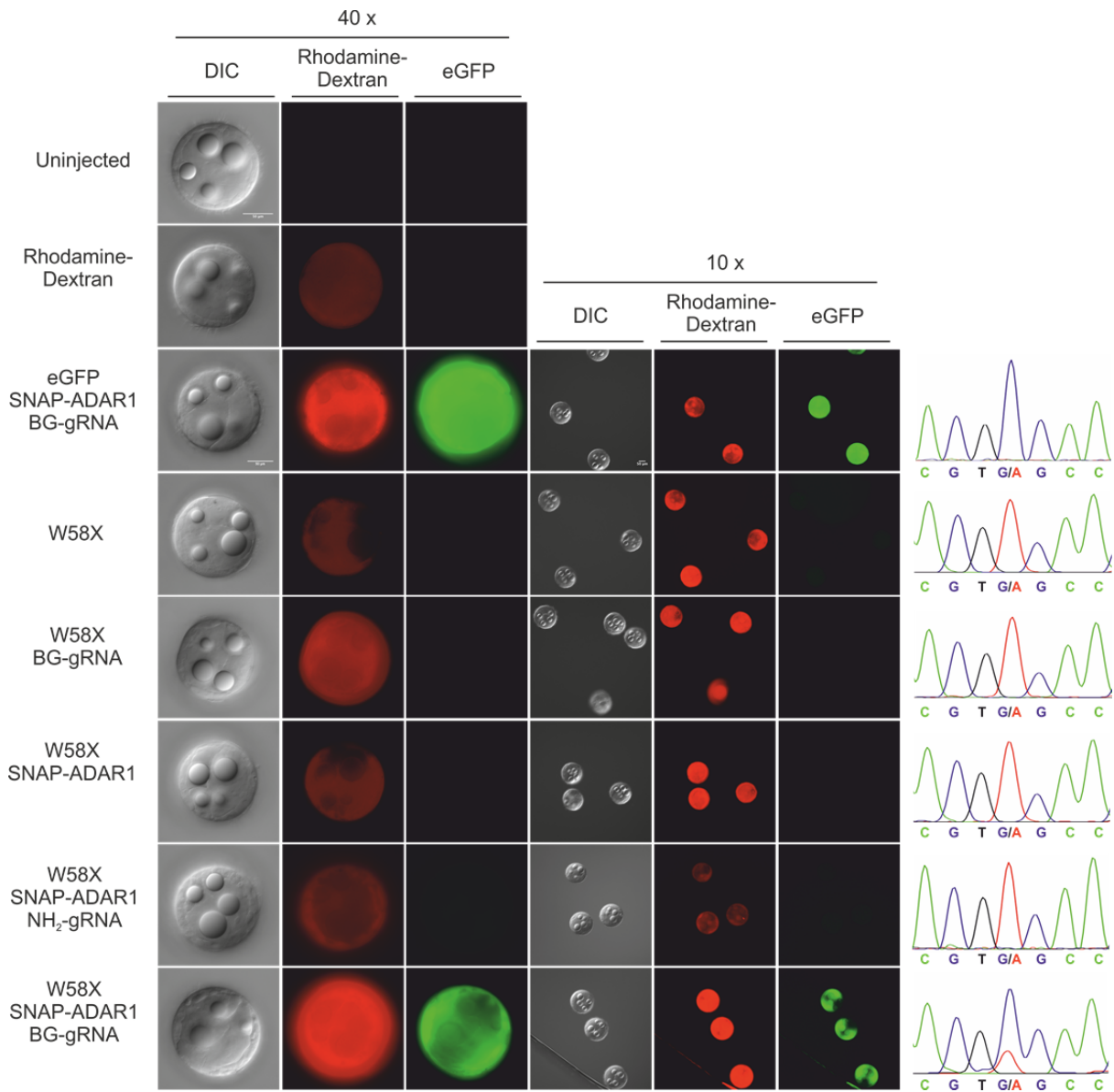


Figure S22. Editing in *Platynereis Dumerilii*. Showing the full set of controls and additional fluorescence and DIC images at 10x magnification corresponding to the experiment shown in Figure 4, manuscript.

Table S23. Statistical overview of detected GFP signal after microinjection. To score the phenotypes, the injected larvae (24hpf) were analyzed by an Imager Z1 widefield fluorescence microscope.

Injection Sample	No GFP signal [%]	Low GFP signal [%]	High GFP signal [%]	total number healthy larvae analyzed
eGFP SNAP-ADAR1 BG-guideRNA Figure 4a	0	4.6	95.4	195
W58X Figure 4b	100	0	0	200
W58X BG-guideRNA Figure 4c	100	0	0	110
W58X SNAP-ADAR1 Figure 4d	100	0	0	150
W58X SNAP-ADAR1 NH2-guideRNA Figure 4e	100	0	0	120
W58X SNAP-ADAR1 BG-guideRNA Figure 4f	29.1	8.2	62.7	110
W58X SNAP-ADAR1 NpomBG-guideRNA No light Figure 5a	81.2	15.2	3.6	250
W58X SNAP-ADAR1 NpomBG-guideRNA 5 min 365 nm Figure 5b	48	10	42	150

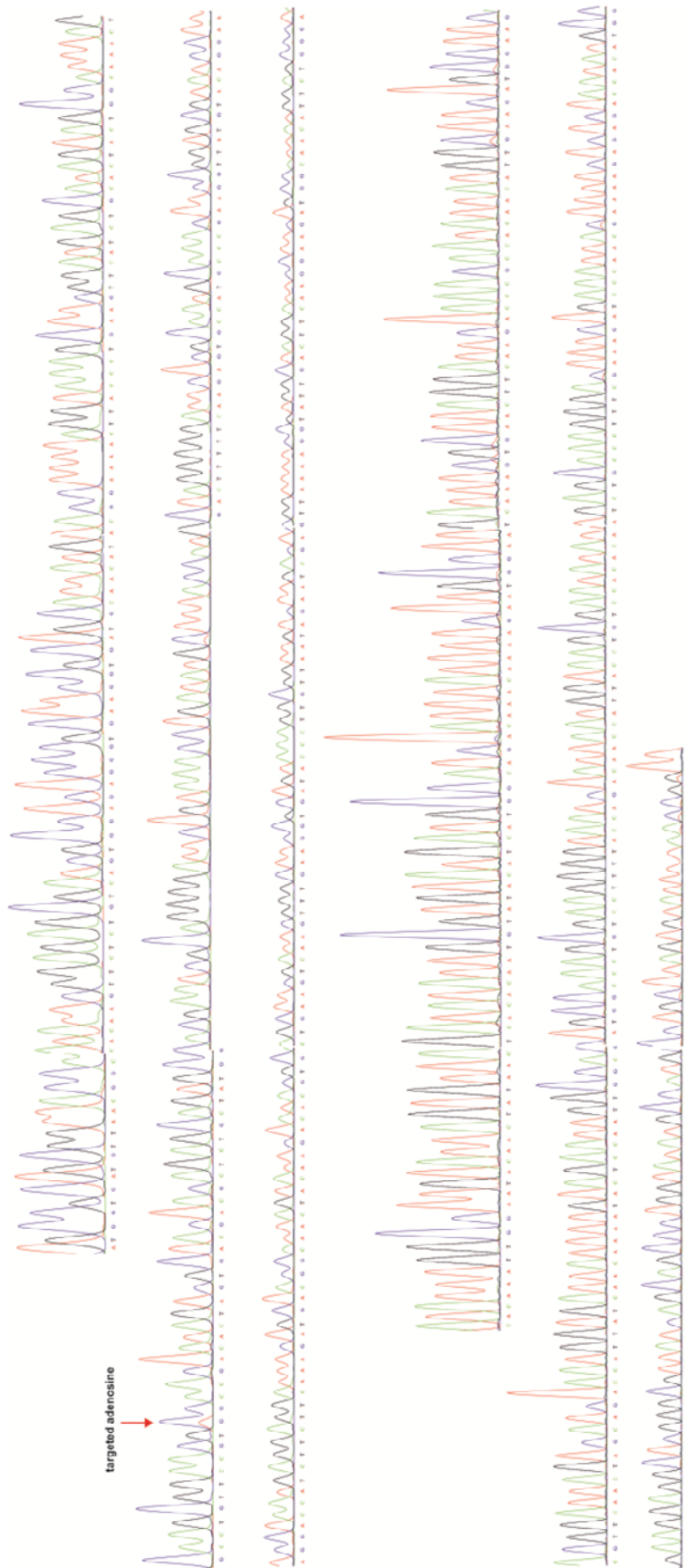


Figure S24. Sequencing obtained from the edited RNA from the experiment shown in Figure 4f in the manuscript (editing with BG-guideRNA & SNAP-ADAR1).

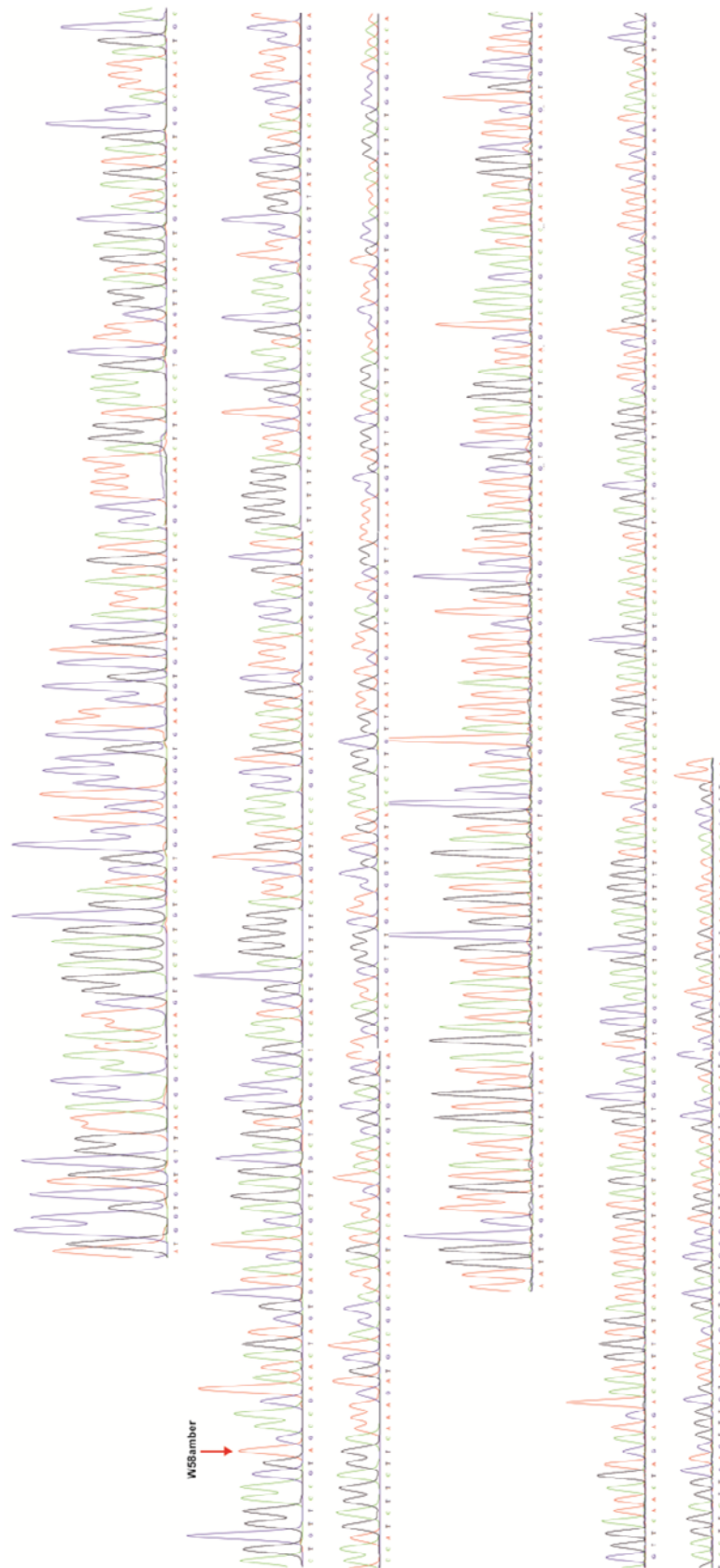


Figure S25. Sequencing obtained from the unedited RNA from the control experiment shown in Figure 4b in the manuscript (W58X eGFP & no SNAP-ADAR1).

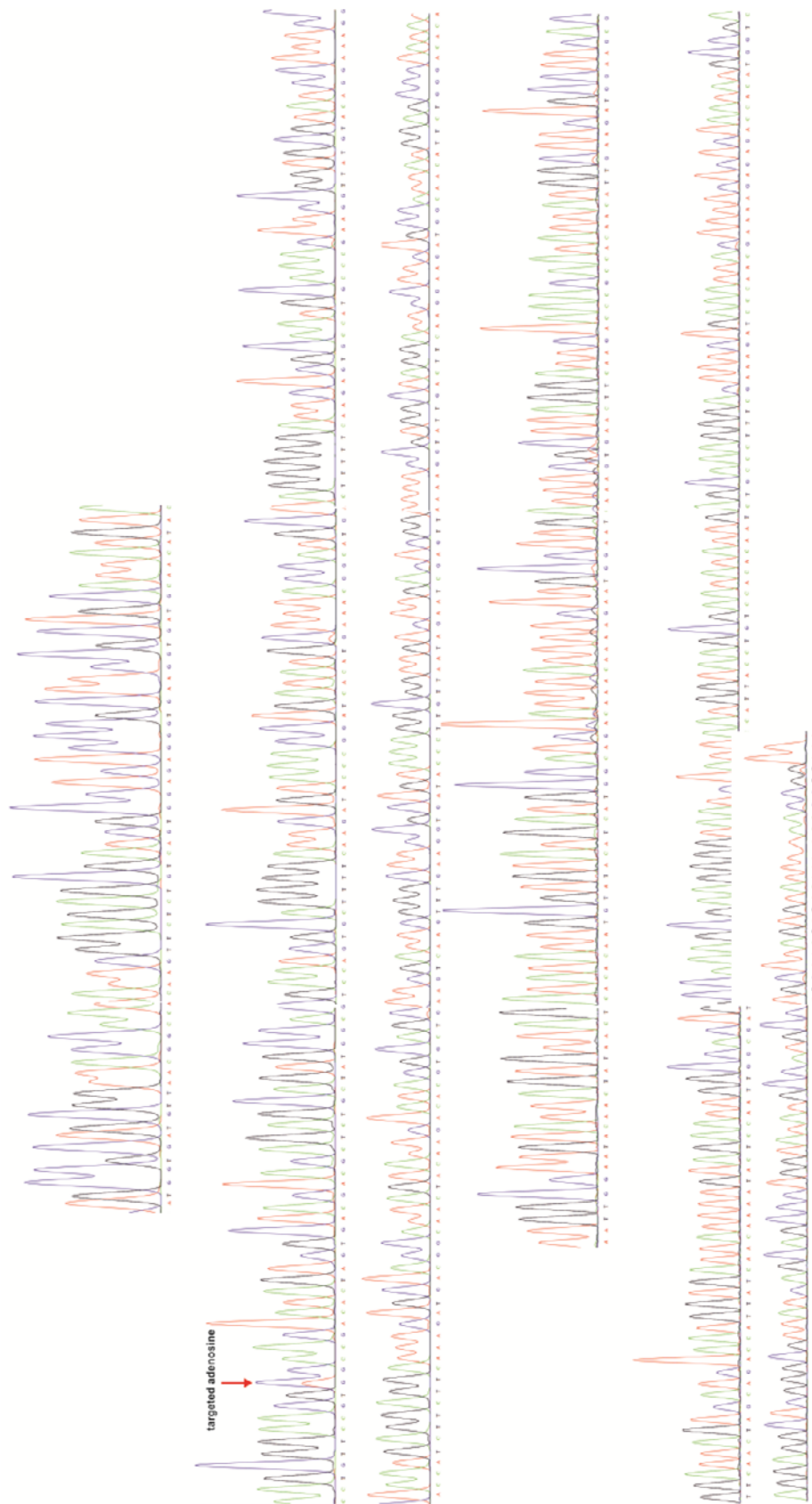


Figure S26. Sequencing obtained from the edited RNA from the experiment shown in Figure 5b in the manuscript (editing with NpomBG-guideRNA & SNAP-ADAR1 & 5 min UV light).

Literature

S1. Keppler, A., Gendreizig, S., Gronemeyer, T., Pick, H., Vogel, H., and Johnsson, K. A general method for the covalent labeling of fusion proteins with small molecules in vivo. *Nat. Biotechnol.* **21**, 86–89 (2002).

S2. Lusic, H., Deiters, A. A new photocaging group for aromatic N-heterocycles. *Synthesis* **13**, 2147-2150 (2006).

S3. Visintin, C., Aliev, A.E., Riddall, D., Baker, D., Okuyama, M., Hoi, P.M., Hiley, R., and Selwood, D.L. Membrane Receptor Probes: Solid-Phase Synthesis of Biotin-Asp-PEG-arvanil Derivatives. *Organic Letters* **7 (9)**, 1699-1702 (2005).

S4. Stafforst, T., Schneider, M.F. An RNA–Deaminase Conjugate Selectively Repairs Point Mutations. *Angew. Chem. Int. Ed.* **51**, 11166-11169 (2012).

S5. Schneider, M.F., Wettengel, J., Hoffmann, P.C., Stafforst, T. Optimal guideRNAs for re-directing deaminase activity of hADAR1 and hADAR2 in trans. *Nucl. Acids Res.* **42**, e87 (2014).

S6. Goeldner, M., Givens, R.S., Dynamic Studies in Biology: Phototriggers, Photoswitches and Caged Biomolecules, *Wiley-VCH Verlag GmbH & Co. KGaA* (2005).

Man. 3: accepted

Vogel, P., Hanswillemenke, A. & Stafforst, T. Switching protein localization by site-directed RNA editing under control of light. *ACS Synth. Biol.* **6**, 1642-1649 (2017).

Switching Protein Localization by Site-Directed RNA Editing under Control of Light

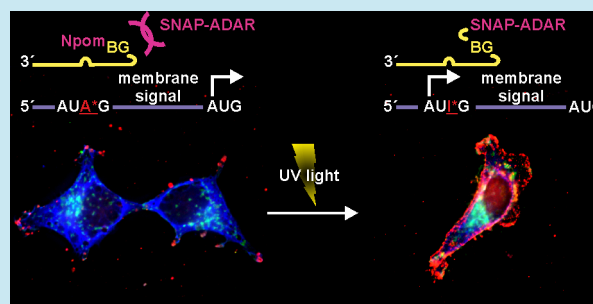
Paul Vogel, Alfred Hanswillemenke, and Thorsten Stafforst*^{1b}

University of Tübingen, Interfaculty Institute of Biochemistry, Auf der Morgenstelle 15, 72076 Tübingen, Germany

Supporting Information

ABSTRACT: Site directed RNA editing is an engineered tool for the posttranscriptional manipulation of RNA and proteins. Here, we demonstrate the inclusion of additional N- and C-terminal protein domains in an RNA editing-dependent manner to switch between protein isoforms in mammalian cell culture. By inclusion of localization signals, a switch of the subcellular protein localization was achieved. This included the shift from the cytoplasm to the outer-membrane, which typically is inaccessible at the protein-level. Furthermore, the strategy allows to implement photocaging to achieve spatiotemporal control of isoform switching. The strategy does not require substantial genetic engineering, and might well complement current optogenetic and optochemical approaches.

KEYWORDS: RNA editing, photocontrol, photocaging, protein localization, optogenetics, epitranscriptomics, gene regulation



During expression genetic information is diversified by various mechanisms. Even when encoded in a single genetic locus, many proteins occur in several isoforms, which result from alternative promoter usage or alternative splicing. Another way of diversification is a process called RNA editing.^{1,2} This refers to the insertion or deletion of nucleotides and to the enzymatic deamination of cytosine and adenosine resulting in the formation of uridine and inosine, respectively. Upon editing in the open reading frame (ORF) single amino acids can be recoded. Furthermore, RNA editing can interfere with RNA splicing, microRNA activity, and RNA stability. Such diversifications can affect almost any property of a protein including substrate specificity, catalytic efficiency, protein localization, stability, and others. As correct subcellular localization is essential for proper functioning, mislocalization can act as a strategy to control a protein's function. One example is the cytosolic sequestering of transcription factors like NF- κ B or the glucocorticoid receptors, which translocate to the nucleus in response to specific signaling cues. Synthetic biology has exploited the induction of translocation as a strategy to control genetic networks. One example for the latter is the engineered Cre-ER(T2) system for the conditional switch of gene function *in vivo*.^{3,4}

The information about the subcellular localization is typically encoded in short peptide-segments, so-called localization signals.⁵ Some signals, like the nuclear localization signal (NLS), can be found anywhere in a protein. Whereas others, like the ER-targeting sequence, are typically found in the N-terminus and are proteolytically cleaved off during translation.⁶ If protein isoforms differ in their localization, such signal peptides are typically in- or excluded by alternative promoter or

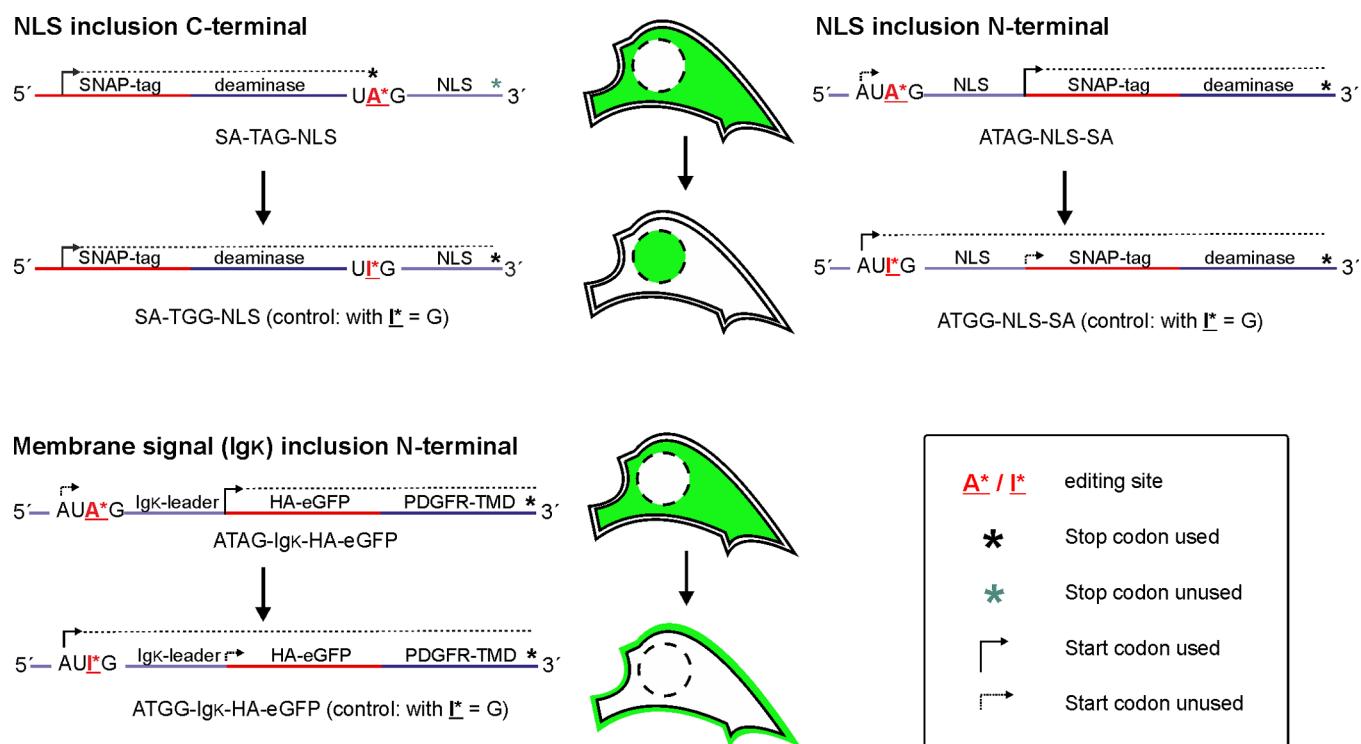
splice sites usage. We were wondering if site-directed RNA editing could be harnessed for that purpose.⁷

We and others have recently engineered artificial guideRNA-dependent editing machines that allow for the introduction of single A-to-I substitutions at targeted sites in selected transcripts inside living cells, a process called site-directed RNA editing.^{8,9} To achieve this, we have fused the catalytic domain of a human adenosine deaminase acting on RNA (hADAR) with a SNAP-tag that allows for the formation of highly defined 1-to-1 covalent conjugates between the guide-RNA and the deaminase.⁸ The approximately 20 nt long guideRNA steers the deaminase to any arbitrary transcript in a readily programmable way. As the deaminase acts only on double-stranded RNA the guideRNA component provides the basis for substrate specificity. By chemical modification and sequence refinement, the selectivity and efficiency of the editing reaction can be further fine-tuned.^{10,11} So far, we and others have applied site-directed RNA editing strategies in human cell culture,^{7,9} in living organisms,¹² and in *Xenopus* eggs,⁹ to manipulate reporter genes and to repair disease-related mutations in CFTR⁹ and PINK1¹³ mRNAs. Furthermore, we recently demonstrated the possibility of controlling the guideRNA–deaminase assembly by light, which enabled us to extent RNA editing by photocontrol *in vivo*.¹²

Here, we now demonstrate a simple strategy to apply RNA editing for triggering the inclusion of an additional peptide signal into both, the N- or the C-terminus of a protein. We apply the strategy for the inclusion of a nuclear localization signal (N- or C-terminal) and for the switching between a

Received: April 5, 2017

Published: May 31, 2017

Scheme 1. Three Different Constructs for Editing-Dependent Isoform Switching^a

^aThe NLS has been included either N- or C-terminally into the SNAP-deaminase protein. The Igk-leader sequence, which signals plasma membrane localization, has been included N-terminally into an HA-tagged eGFP. The C-terminal platelet-derived growth factor receptor transmembrane domain (PDGFR-TMD) is a single transmembrane α -helix that anchors the protein to the plasma membrane by pointing the N-terminus outside. The expected localization phenotype (cytoplasm, nucleoplasm or outer membrane) is indicated.

cytoplasmic and a membrane-bound isoform in human cell culture. Furthermore, we demonstrate the light control of the isoform switch.

RESULTS AND DISCUSSION

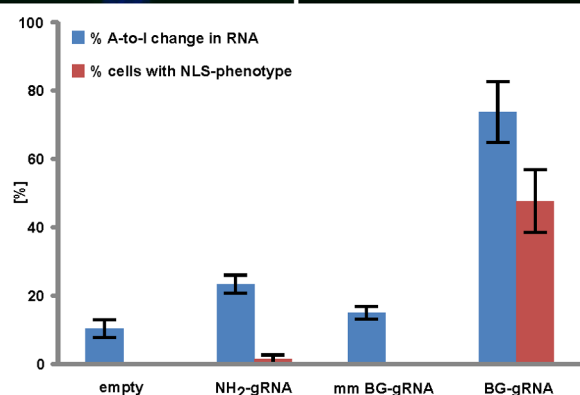
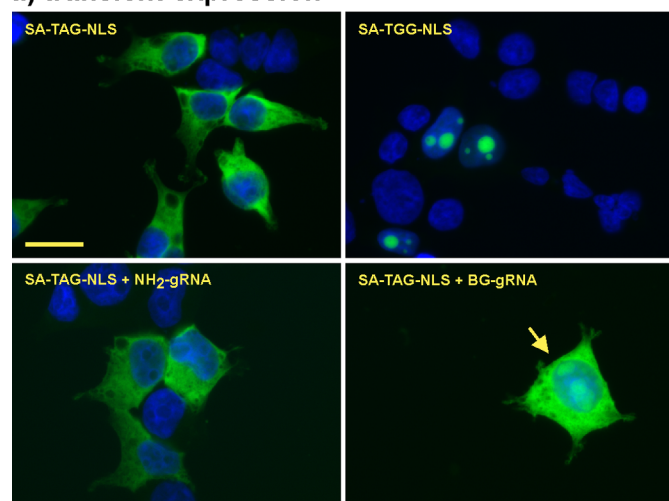
General Considerations. The C-terminal inclusion of an additional peptide signal appears particularly straightforward by putting the signal into the 3'-UTR directly behind an amber Stop codon (UAG), **Scheme 1**. Upon editing the Stop codon to Trp (UIG) the additional signal is inserted C-terminally. The analogous strategy at the N-terminus appears more challenging. We explore here the activation of an alternative Start codon in the 5'-UTR, as it is conceivable to edit an isoleucine codon (AUA) into a methionine/Start codon (AUI), **Scheme 1**. Prior to editing, the downstream Start codon would be used only. However, after editing the upstream Start should dominate, as cap-dependent translation typically applies the first Start codon after the cap.^{14–16} Nevertheless, site-directed RNA editing inside the 5'-UTR has not yet been reported. Also within natural editing sites, editing in the 5'-UTR is strongly underrepresented.¹⁷ Thus, it was unclear if the preinitiation complex of translation and the editing machinery will interfere. To assess both strategies in a comparable manner, we decided to start with the inclusion of a nuclear localization signal (NLS) derived from the SV 40 Large T-antigen (PKKKRKV)₃, which can be put to both, the N- or C-terminus.¹⁸ To visualize the localization phenotype, we chose the transcript of the editing enzyme (SNAP-ADAR2) as the editing target. On one hand, the enzyme is strictly localized in the cytoplasm when lacking an NLS. On the other hand, the enzyme is readily stained with fluoresceine-O6-benzylguanine (BG-FITC) to assess its local-

ization by fluorescence microcopy.¹² Furthermore, this procedure allowed us to stay with the ectopic expression of a single construct which simplified transfection and phenotypic analysis.

Editing-Dependent Inclusion of the NLS into the C-Terminus under Transient Expression. According to **Scheme 1**, two plasmids were constructed that contain SNAP-ADAR2 under control of the CMV promoter. In one construct, the NLS was put in frame at the C-terminus (SA-TAG-NLS). When transfected into 293T cells and BG-FITC-stained 48 h later, a clear nuclear localization was visible (**Figure 1a**). The other construct contained a single G-to-A mutation between the SNAP-ADAR and the NLS which inserts a premature Stop codon and thus shortens the open reading frame (SA-TAG-NLS). When expressed and stained comparably, a clear cytoplasmic phenotype was visible. The latter construct was the substrate to study the editing-dependent phenotype switch.

For editing, 293T cells were first transfected with SA-TAG-NLS (or SA-TGG-NLS in the control) and were then reverse transfected with a guideRNA. When the matching guideRNA was used, BG-FITC staining revealed a clear appearance of nuclear SNAP-ADAR2 protein (**Figure 1a**) that resembles the phenotype of the positive control. We found this new, mixed cyto-/nucleoplasmic phenotype in $48 \pm 9\%$ of the transfected cells. Sanger sequencing revealed an editing yield of $74 \pm 9\%$. We assume two reasons for the mixed (cytoplasmic/nuclear) phenotype after editing. First, editing was incomplete, and second, some of the stained SNAP-ADAR2 protein was old protein from the SNAP-ADAR expression prior to induction of the editing event by transfecting the guideRNA. The isoform

a) transient expression



b) genomic expression

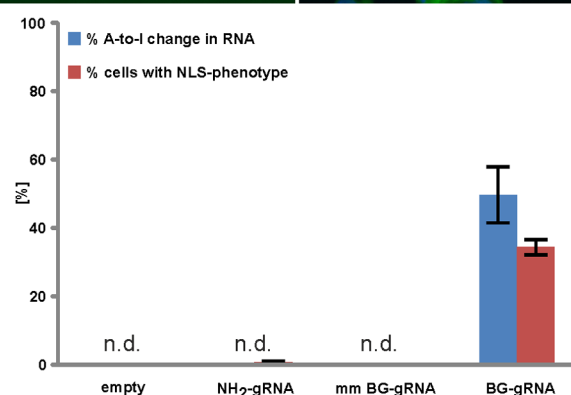
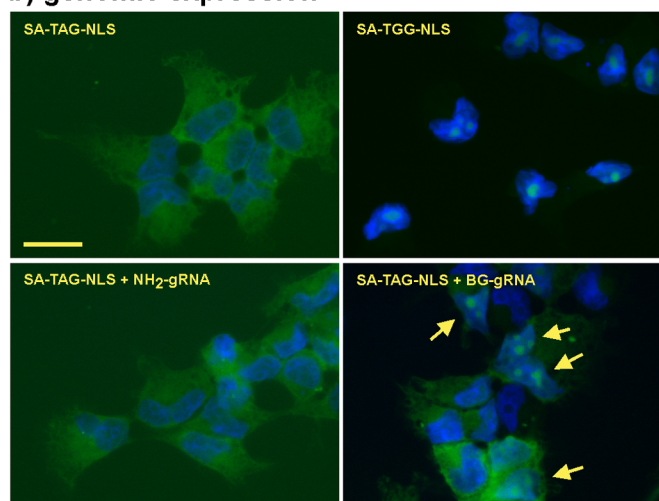


Figure 1. Editing-dependent switch from SNAP-ADAR2 to SNAP-ADAR2-NLS under transient (a) and genomic (b) expression. (a) Fluorescence imaging of FITC-stained SNAP-ADAR (green) and Hoechst 33342-stained nuclei (blue). SA-TGG-NLS is the positive control for the nuclear localization phenotype after editing. Quantitative analysis of the editing experiment: Blue shows the editing yield from Sanger sequencing. Red shows the amount of cells that are positive for SNAP-ADAR expression and show nuclear localization. mm BG-gRNA: mismatching BG-guideRNA. Black bars show the standard deviation from $N = 3$ independent experiments. The scale bars represent $20 \mu\text{m}$. (b) Analogous experiment as in panel (a), but under genomic expression of the SNAP-ADAR constructs. n.d. = neither RNA editing nor nuclear localization was detectable. Further data and controls are shown in the [Supporting Information](#), Figure S1–S3 for transient and S4–S6 for genomic expression.

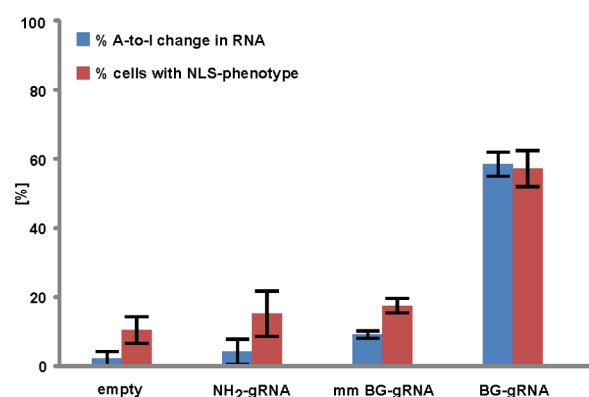
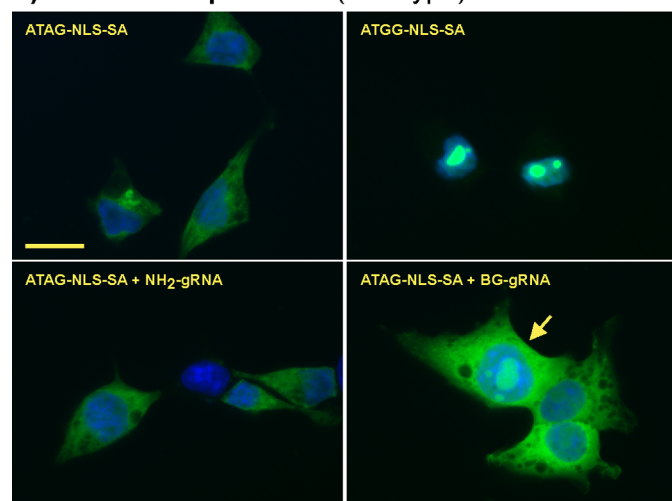
switch was strongly dependent on editing. It did neither occur in the presence of an NH₂-guideRNA incapable of conjugation,¹² nor in the presence of a BG-guideRNA with a mismatching (mm) sequence (Figure 1a). However, due to the high levels of SNAP-ADAR2 protein and its transcript under transient expression, low levels of guideRNA-independent editing were detectable (Figure 1a, graph). Even though this low-level editing did not result in a visible nuclear localization phenotype, we aimed to further improve the performance of the system by genomic integration of the SNAP-ADAR construct.

C-Terminal NLS-Inclusion Works Also under Genomic Expression. To obtain a weaker and more homogeneous expression, the respective constructs were integrated as a single copy into the genome of 293 Flip-In cells under control of the Tet-on CMV promoter (inducible genomic expression). Fluorescence microscopy confirmed the homogeneous, inducible and much weaker expression of the editase under genomic control (Figure 1b). Again, the cytoplasmic (SA-TAG-NLS) and nucleoplasmic (SA-TGG-NLS) phenotypes in the controls were clearly visible (Figure 1b). As expected, and in contrast to the conditions before, the editing was now fully dependent on the presence of the matching BG-guideRNA. Lacking the

guideRNA or applying a mismatching or an NH₂-guideRNA gave no detectable editing yield. The editing yield with the matching BG-guideRNA was $50 \pm 8\%$ and thus stayed a bit below that under transient expression. The same trend holds also true for the isoform switch. About $34 \pm 2\%$ of the cells showed the switch from pure cytoplasmic to a mixture of cytoplasmic and nuclear localization, demonstrating the C-terminal NLS inclusion in an editing-dependent manner under genomic expression of the construct.

Editing-Dependent Inclusion of the NLS into the N-Terminus (Transient Expression). As depicted in Scheme 1, two plasmids were constructed that contain two Start codons each embedded in a strong Kozak sequence ($5'$ -CCACC-AUG-G).¹⁹ One of the Start codons was located in front and one behind the NLS. In the construct ATGG-NLS-SA, both Start codons are appropriate to start translation. According to the scanning model of cap-dependent translation one expects this construct to predominantly use the Start codon prior to the NLS and thus to express the full NLS-SNAP-ADAR2 protein.^{14–16} Accordingly, transient expression of this construct in 293T cells showed exclusive nuclear localization of SNAP-ADAR (Figure 2a). The construct ATAG-NLS-SA differs from the latter by a single G-to-A mutation in the upstream Start

a) transient expression (wild-type)



b) genomic expression (E488Q)

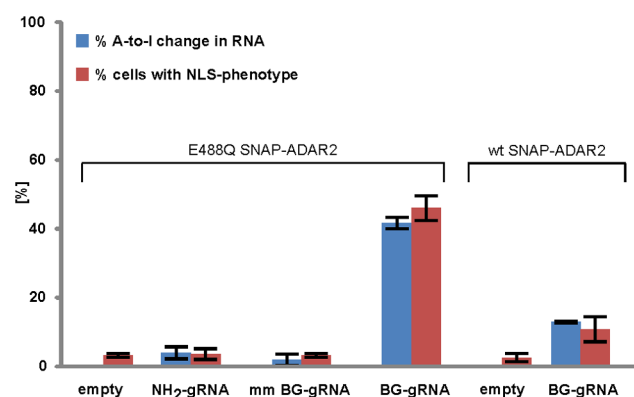
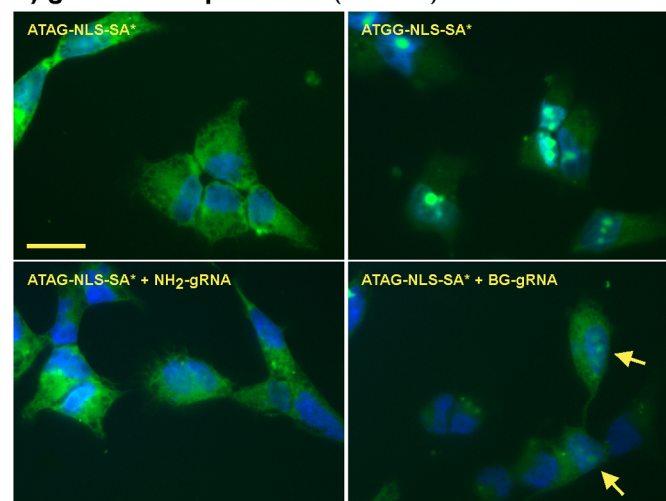


Figure 2. Editing-dependent switch of SNAP-ADAR2 to NLS-SNAP-ADAR2 under transient (a) and genomic (b) expression. Experiment and analysis follows the description given in Figure 1. SA* marks the construct with the activated E488Q deaminase. The scale bars represent 20 μ m. Further data and controls are shown in Figures S7–S9 for transient and S11–S13 for genomic expression.

codon, thereby creating a 5'-CCACCAUA*G sequence that is supposed to be inappropriate to start translation prior to editing (AUA*) but to turn into a strong initiation signal after editing (AUI*). Transient expression of this construct gave almost exclusive cytoplasmic localization of SNAP-ADAR. Only a small number of cells ($10 \pm 4\%$) showed a faint nuclear staining (Figure S10), which might result from a minor translation initiation from the unedited AUA Start codon, as it is embedded in a very strong sequence context. However, in a similar setting it was reported that the plasmid-borne sequence 5'-CCACCAUAG is unable to initiate translation when transfected into COS cells.²⁰ Clearly, the faint nuclear staining was not due to (guideRNA-independent) editing, as the editing yield in absence of the guideRNA was below the detection limit ($\leq 2\%$).

For editing, 293T cells were transfected with either of the two constructs and reverse transfected with a guideRNA. Protein localization was analyzed by fluorescence microscopy after BG-FITC staining. After transfection of the editing substrate (ATAG-NLS-SA) and the matching BG-guideRNA, we found a clearly visible nuclear staining that resembled that of the positive editing control (ATGG-NLS-SA), Figure 2a. Similar to the results at the 3'-UTR, we found a mixed nucleo-/cytoplasmic phenotype in $57 \pm 5\%$ of all cells. After editing the nuclear staining of the protein was much stronger compared to the occasional faint nuclear staining observed

prior to editing (Figure S10). Sequencing of the mRNA revealed an editing yield of $58 \pm 4\%$, in good agreement with the mixed phenotype. Again, the isoform switch was dependent on the editing event and did not happen in the presence of a mismatching or conjugation-incompetent NH₂-guideRNA.

Editing in the 5'-UTR under Genomic Expression Requires an Activated Deaminase. Again, we tested editing under genomic expression of the 5'-UTR constructs. Upon induction, both constructs behaved as expected showing either the strict nuclear (ATGG-NLS-SA) or cytoplasmic localization (ATAG-NLS-SA) with strongly reduced but homogeneous expression over the entire culture. Compared to the expression under transient conditions, the occasional appearance of faint nuclear staining in the ATAG construct was almost abolished (below 3%). However, the editing-dependent isoform switch was disappointing. The nuclear phenotype was visible in no more than $11 \pm 4\%$ of the cells. However, this matched the low editing levels of $13 \pm 1\%$ (Figure 2b). The editing reaction might suffer from the comparably low concentration of editase and substrate, which might slow down the editing reaction. To test if a faster enzyme would help to improve the performance, we engineered two new cell lines that contain again either the ATAG or ATGG construct, but now with a SNAP-ADAR2* protein that contains a well-described, single point mutation in the deaminase domain (E488Q) that is reported to speed up deamination by at least 1 order of magnitude.²¹

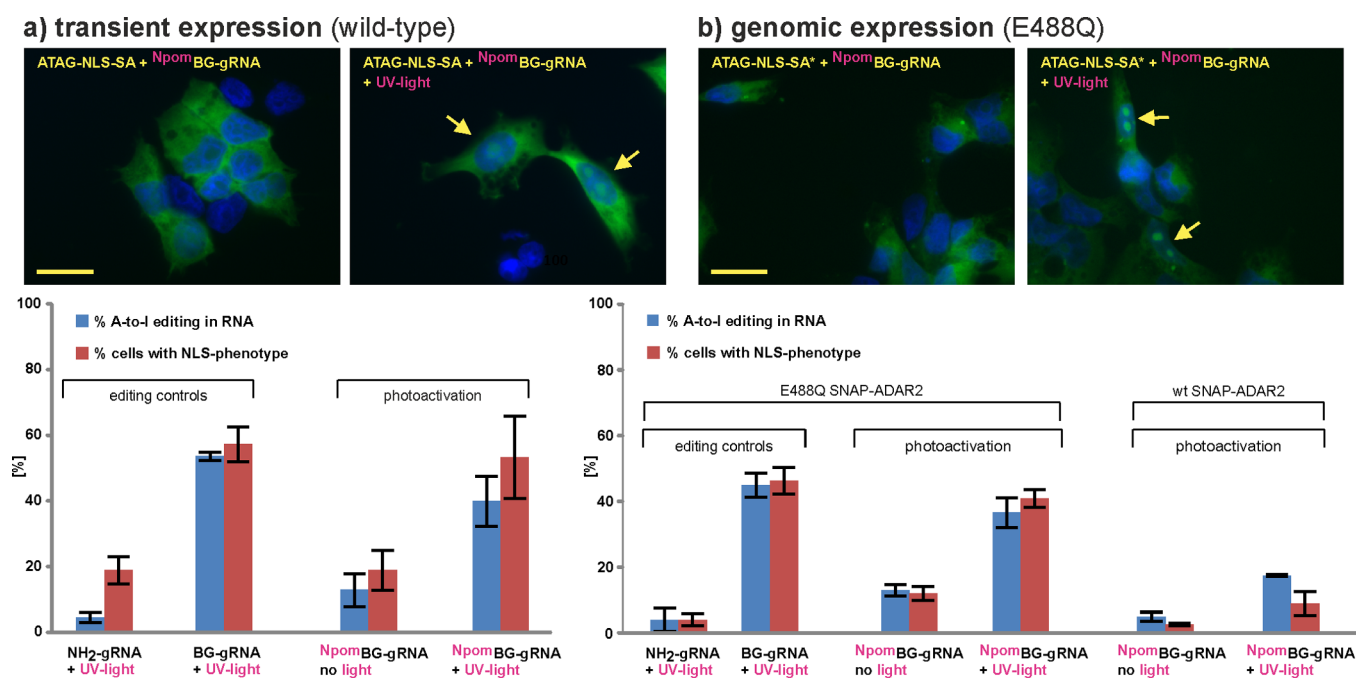


Figure 3. Photoinduced switch of SNAP-ADAR2 to NLS-SNAP-ADAR2 under transient (a) and genomic (b) expression. SA* marks the construct with the activated E488Q deaminase. The scale bars represent 20 μm . Further data is shown in Figures S7–S9 for transient and Figures S11–S16 for genomic expression.

The two new constructs behaved indistinguishable from their less active counterparts in terms of expression level, homogeneity and localization phenotype. However, in the editing experiment the new constructs showed a robust isoform switch. A clear change to the mixed nuclear/cytoplasmic phenotype was found in $46 \pm 4\%$ of the cells, resulting from an improved editing yield of $42 \pm 2\%$ (Figure 2b). As seen before, the switch was fully dependent on editing. Consequently, the nuclear phenotype was seen in $\leq 4\%$, when no guideRNA, a mismatching guideRNA, or an NH_2 -guideRNA was used, reflecting the low editing yields obtained under these conditions ($\leq 4\%$). Substitution of the deaminase by a more active variant boosted the performance of the system by 3-fold in terms of editing yields and number of cells with a phenotypic switch and brought the editing at the 5'-UTR to a level comparable to that at the 3'-UTR.

Isoform Switching Can Be Controlled by Light. Light is an attractive trigger to manipulate biological systems.²² We tested if a recently introduced strategy to control the editing process by controlling the assembly reaction could be applied for the light-control of 5'-UTR editing. As described earlier, guideRNAs have been made that mask the SNAP-reactive BG moiety with the Npom photocage to render it inactive for the assembly reaction.¹² Then, the editing reaction can be started by treating the cells under the microscope with a short UV-light pulse (365 nm, 5 s). First, we studied the system with the wildtype ATAG-NLS-SA construct under transient expression. The light flash had no effect on the negative (NH_2 -guideRNA) and the positive editing control (BG-guideRNA) in terms of editing yield ($4 \pm 2\%$ and $53 \pm 1\%$) and localization phenotype ($19 \pm 4\%$ and $57 \pm 5\%$), Figure 3a. However, when applying the Npom-protected BG-guideRNA, a clear photoinduction of editing yield and isoform switch was detectable. Without irradiation, $19 \pm 6\%$ of cells showed a faint nuclear staining, whereas $53 \pm 12\%$ of the cells showed the switch to a clear nuclear staining after irradiation (Figure 3a). This was in

accordance with the photoinduced change of editing levels from $13 \pm 5\%$ before to $40 \pm 8\%$ after irradiation. As before, the ATAG-NLS-SA construct suffered from the occasional formation of faint nuclear staining under transient expression.

Thus, we also tested photocontrol under genomic expression. As before, editing yields and phenotype switching were dissatisfying with wildtype enzyme and stayed below 20% (Figure 3b, graph). However, the E488Q variant of the deaminase was helpful again, and the positive editing control (BG-guideRNA) gave robust nuclear staining in $46 \pm 4\%$ of the cells, matching the respective editing yields of $45 \pm 4\%$ (Figure 3b). The negative editing control (NH_2 -guideRNA) showed virtually no editing ($\leq 4\%$) and also the occasional faint nuclear staining was strongly reduced ($\leq 4\%$). When applying the Npom-protected guideRNA, a clear photoactivation was visible. Before irradiation $12 \pm 2\%$ of the cells showed the nuclear staining, whereas $41 \pm 3\%$ showed nuclear staining after irradiation. Again the effect was clearly depending on the editing yields which changed from $13 \pm 2\%$ before to $37 \pm 5\%$ after irradiation, see Figure 3b. Thus, protein isoforms can be switched by light simply by photocontrolling the assembly reaction of editase and guideRNA, both under transient and genomic expression.

5'-UTR Editing Enables to Switch Localization from the Cytoplasm to the Outer Membrane. Induction of protein translocation from the cytoplasm to the nucleoplasm under control of small molecules and/or light has been achieved earlier, either by engineering fusion proteins to become controllable by small molecules (f.i. the Cre-ER(T2) system)²³ or by the ectopic expression of proteins with site-specifically photocaged amino acids.^{24–26} The latter strategies are feasible because trafficking into the nucleus is a posttranslational mechanism applied to fully folded proteins. As RNA editing happens before translation, isoform switches become feasible that are decided cotranslationally and thus are impossible to control at the protein level. A conceivable

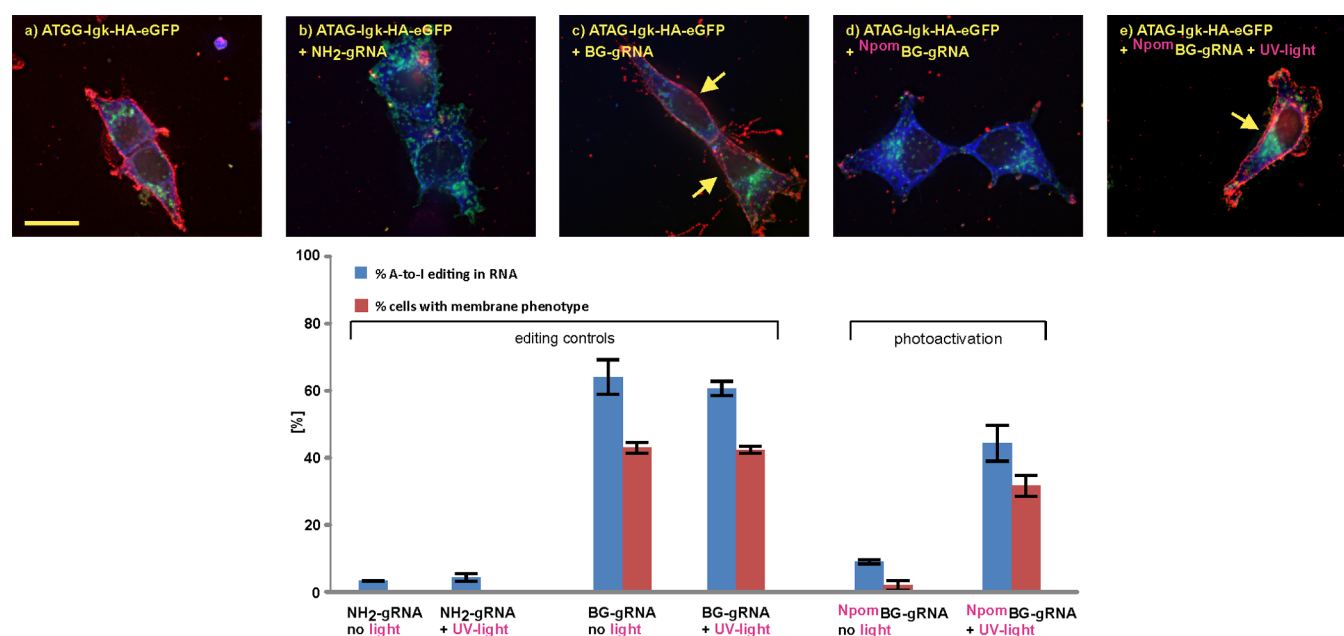


Figure 4. Editing-dependent switch of HA-GFP-PDGFR-TMD localization from the cytoplasm to the plasma membrane under transient coexpression with the BFP-tagged editase (SNAP-ADAR2). Imaging was carried out after fixation, without permeabilization: HA-immunostaining with AlexaFluor-594 (red), GFP (green), and BFP (blue). (a) positive control for plasma membrane localization; (b) negative editing control; (c) editing; (d,e) light-dependent editing experiment. The scale bar represents 20 μ m. Further data and controls are shown in Figures S17–S21.

example is plasma membrane localization. The respective signal peptides are found in the very N-terminus of a protein.⁵ Once the nascent signal peptide leaves the exit at the ribosome, it is recognized by the signal recognition particle that recruits the translating ribosome to the ER. At the ER, translation continues, the signal peptide is cleaved off during translation inside the ER and the protein is inserted into the membrane cotranslationally.⁶

We explored how RNA editing can be used to switch protein isoforms from a cytoplasmic to a membrane-anchored localization. For this a construct was made that contains an editable Start codon (AUA*) followed by the 22 amino acid Igk chain leader sequence, an alternative Start codon (AUG), and an HA-tagged GFP protein (Scheme 1). At the very C-terminus, the construct contains the transmembrane domain (TMD) of the PDGF receptor that anchors the protein to the plasma membrane displaying the GFP and the HA-tag to the extracellular side of the cell. The analogous ATGG construct served as the positive editing control. To assess the phenotype, immunofluorescence microscopy was applied.

Under transient expression (293T cells) of the positive control (ATGG), the HA-GFP is clearly localized to the outer membrane, as visualized by a rim-like anti-HA-immunostaining in fixed but not permeabilized cells (Figure 4a). In contrast, the negative editing control (ATAG) gave no rim-like anti-HA-staining (f.i. Figure 4b). However, when cells were permeabilized prior to immunostaining (Figure S21), the cytoplasmic expression of the construct was clearly detectable. When cotransfecting the ATAG construct with SNAP-ADAR2-BFP and reverse transfecting the matching BG-guideRNA, the HA-immunofluorescence showed again the rim-like staining of the outer membrane in $43 \pm 2\%$ of the cells that have been positive for GFP and BFP fluorescence (Figure 4c). This phenotypic switch was again clearly depending on the editing yield ($64 \pm 5\%$). It did not occur in the absence of a guideRNA or in the presence of a mismatching or NH₂-guideRNA (Figure 4b).

Translocation to the Outer Membrane Can Be Controlled by Light.

Finally, we tested to switch the isoforms under control of light. As before, we put the Npom photocage on the guideRNA. When applying the Npom-BG-guideRNA, a modest residual editing activity was detected ($9 \pm 1\%$), however, no outer-membrane staining was detectable ($<2\%$, Figure 4d). After irradiation with 365 nm light a clear membrane staining became visible in $31 \pm 3\%$ of the cotransfected cells (Figure 4e). Accordingly, the editing yield increased from $9 \pm 1\%$ prior to $44 \pm 5\%$ after irradiation. UV-irradiation had no influence on the editing yield or localization phenotypes of the positive (BG-guideRNA) or negative (NH₂-guideRNA) editing controls (Figures S17–S20). Overall, isoform switch from cytoplasmic to the outer membrane can be controlled at the posttranscriptional level, and photocontrol is readily included.

CONCLUSION AND OUTLOOK

RNA editing can be applied to switch protein isoforms. This is not restricted to the recoding of amino acids or splice sites, but can be harnessed for the inclusion of additional N- or C-terminal peptide signals by editing of Start and Stop codons. UTRs in mammals are typically around 100 nt long, but can extend to 1000 nt or longer. Thus, even the N- or C-terminal inclusion of large protein domains is conceivable.¹⁹ Our artificial editing strategy that relies on the RNA-guided SNAP-tagged deaminases enables this without detectable interference with translation and translation initiation. It can be accomplished either under transient or genomic expression. The usage of the SNAP-deaminases further allows for a ready inclusion of light-control. The method might well complement current methods in synthetic biology, including optogenetics²⁷ and other optochemical approaches.²² On one hand it enables light-controlled isoform switches that are impossible at the protein-level. This holds particularly true for phenotypes that separate already during translation and thus are inaccessible

with caged or otherwise engineered proteins,^{23–26} as demonstrated by the switch to an outer-membrane anchored isoform. On the other hand, the method might complement approaches that depend on the light-dependent (in)activation of genes,^{28,29} which typically require massive genetic engineering. To our knowledge, this is the first report about redirecting protein localization from the cytoplasm to the membrane. In combination with light-control, our tool could provide new opportunities to address biological questions in basic research. In the future, proteins might be steered to the cell surface by using light-activated RNA editing to manipulate intracellular signaling but also extracellular events like cell–cell and cell–matrix interactions in a spatiotemporal manner.

METHODS

Editing under Transient Expression. 293T cells were grown in DMEM + 10% FBS + 1× P/S, 5% CO₂. Plasmid transfection was done with 300 ng of the respective plasmid/well with Lipofectamine 2000 in DMEM + 10% FBS. The respective guideRNA (2.5–10 pmol/well) was reverse transfected with Lipofectamine 2000 in DMEM + 1% FBS. Cells were seeded on coverslips (DMEM + 1% FBS + HEPES). After 24 h cells were harvested for RNA sequencing or stained with BG-FITC for fluorescence microscopy as described before.^{10,12}

Editing under Genomic Expression. 293-Flp-In T-REx cells were induced in DMEM + 10% FBS + 15 μg/mL blasticidinS + 100 μg/mL hygromycinB + 10 ng/mL doxycycline, 5% CO₂. The respective guideRNA (5–20 pmol/well) was reverse transfected with Lipofectamine 2000 in DMEM + 10% FBS + 10 ng/mL doxycycline. Cells were seeded on coverslips (DMEM + 10% FBS + HEPES + doxycycline). After 24 h cells were harvested for RNA sequencing or stained with BG-FITC for fluorescence microscopy.

Light-Induced RNA Editing. Experiments were carried out as described above with an additional irradiation step 4 h after guideRNA transfection. Cells were washed and the entire well was irradiated with 365 nm light on the microscope (Zeiss CellObserverZ1, 365 nm LED light source) for 5 s under full power at 5× magnification. Then the protocol was continued as described above.

BG-FITC Staining. To visualize the localization of SNAP-ADAR2, acetylated BG-FITC (final concentration 2 μM) was applied to the cells together with a blue Hoechst stain (Thermo Fisher, R37605) for 30 min. Cells were fixed with formaldehyde and permeabilized with 0.1% Triton X-100. Cover glasses were mounted using Shandon Immu-Mount (Thermo Fisher, USA).

Immunofluorescence Microscopy. Cells were fixed with formaldehyde and blocked with PBS + 10% FBS at 4 °C overnight. Cells were stained with a primary mouse anti-HA-antibody (Sigma-Aldrich, H9658) diluted 1:1250 in PBS + 5% FBS for 1.5 h at room temperature, and a secondary antimouse antibody conjugated to AlexaFluor-594 (Thermo Fisher, A-11005) diluted 1:1500 in PBS + 10% FBS for 45 min at room temperature. Cover glasses were mounted using Dako mounting medium (Dako North America, USA). Microscopy was performed with a Zeiss CellObserverZ1 under 600× total magnification.

ASSOCIATED CONTENT

Supporting Information

The Supporting Information is available free of charge on the ACS Publications website at DOI: 10.1021/acssynbio.7b00113.

General methods, details of materials and all constructs, the full primary data and some additional experiments; Figures S1–S21; Tables S1–S7 (PDF)

AUTHOR INFORMATION

Corresponding Author

*E-mail: thorsten.stafforst@uni-tuebingen.de.

ORCID

Thorsten Stafforst: 0000-0001-9359-3439

Author Contributions

PV performed the experiments, AH synthesized Npom-caged BG-guideRNAs, all authors planned experiments, all authors analyzed the data, TS instructed the research and wrote the paper under assistance of all authors.

Notes

The authors declare no competing financial interest.

ACKNOWLEDGMENTS

We gratefully acknowledge support from the University of Tübingen and the Deutsche Forschungsgemeinschaft (STA1053/3-2; STA1053/4-1; STA1053/7-1). This work has received funding from the European Research Council (ERC) under the European Union's Horizon 2020 research and innovation program (grant agreement No 647328).

REFERENCES

- (1) Nishikura, K. (2010) Functions and regulation of RNA editing by ADAR deaminases. *Annu. Rev. Biochem.* 79, 321–349.
- (2) Bass, B. (2002) RNA editing by adenosine deaminases that act on RNA. *Annu. Rev. Biochem.* 71, 817–846.
- (3) Lewandoski, L. (2001) Conditional control of gene expression in the mouse. *Nat. Rev. Genet.* 2, 743–755.
- (4) Feil, S., Valtcheva, N., and Feil, R. (2009) Inducible Cre mice. *Methods Mol. Biol.* 530, 343–63.
- (5) Blobel, G. (2000) Protein targeting. *ChemBioChem* 1, 86–102.
- (6) Reid, D. W., and Nicchitta, C. V. (2015) Diversity and selectivity in mRNA translation on the endoplasmic reticulum. *Nat. Rev. Mol. Cell Biol.* 16, 221–231.
- (7) Vogel, P., and Stafforst, T. (2014) Site-directed RNA editing with antagomir deaminases — A tool to study protein and RNA function. *ChemMedChem* 9, 2021–2025.
- (8) Stafforst, T., and Schneider, M. F. (2012) An RNA–deaminase conjugate selectively repairs point mutations. *Angew. Chem., Int. Ed.* 51, 11166–11169.
- (9) Montiel-Gonzalez, M. F., Guillermo, I., Yudowski, A., and Rosenthal, J. J. C. (2013) Correction of mutations within the cystic fibrosis transmembrane conductance regulator by site-directed RNA editing. *Proc. Natl. Acad. Sci. U. S. A.* 110, 18285–290.
- (10) 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.
- (11) Schneider, M. F., Wettengel, J., Hoffmann, P. C., and Stafforst, T. (2014) Optimal guideRNAs for re-directing deaminase activity of hADAR1 and hADAR2 in trans. *Nucleic Acids Res.* 42, e87.
- (12) Hanswillemenke, A., Kuzdere, T., Vogel, P., Jékely, G., and Stafforst, T. (2015) Site-directed RNA editing *in vivo* can be triggered by the light-driven assembly of an artificial riboprotein. *J. Am. Chem. Soc.* 137, 15875–81.
- (13) Wettengel, J., Reautschnig, J., Geisler, S., Kahle, P. J., and Stafforst, T. (2017) Harnessing human ADAR2 for RNA repair — Recoding a PINK1 mutation rescues mitophagy. *Nucleic Acids Res.* 45, 2797–2808.
- (14) Kozak, M. (1999) Initiation of translation in prokaryotes and eukaryotes. *Gene* 234, 187–208.

- (15) Sonenberg, N., and Hinnebusch, A. G. (2009) Regulation of translation initiation in eukaryotes: mechanisms and biological targets. *Cell* 137, 731–745.
- (16) Van Der Kelen, K., Beyaert, R., Inzé, D., and De Veylder, L. (2009) Translational control of eukaryotic gene expression. *Crit. Rev. Biochem. Mol. Biol.* 44, 143–168.
- (17) Peng, Z., Cheng, Y., Tan, B. C.-M., Kang, L., Tian, Z., Zhu, Y., Zhang, W., Liang, Y., Hu, X., Tan, X., Guo, J., Dong, Z., Liang, Y., Bao, L., and Wang, J. (2012) Comprehensive analysis of RNA-Seq data reveals extensive RNA editing in a human transcriptome. *Nat. Biotechnol.* 30, 253–60.
- (18) Lange, A., Mills, R. E., Lange, C. L., Stewart, M., Devine, S. E., and Corbett, A. H. (2007) Classical nuclear localization signals: definition, function, and interaction with importin. *J. Biol. Chem.* 282, 5101–5105.
- (19) Kozak, M. (1987) An analysis of 5'-noncoding sequences from 699 vertebrate messenger RNAs. *Nucleic Acids Res.* 15, 8125–8148.
- (20) Kozak, M. (1989) Context effects and inefficient initiation at non-AUG codons in eukaryotic cell-free translation systems. *Mol. Cell. Biol.* 9, 5073–5080.
- (21) 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.
- (22) Brieke, C., Rohrbach, F., Gottschalk, A., Mayer, G., and Heckel, A. (2012) Light-controlled tools. *Angew. Chem., Int. Ed.* 51, 8446–8476.
- (23) Inlay, M. A., Choe, V., Bharathi, S., Fernhoff, N. B., Baker, J. R., Weissman, I. L., and Choi, S. K. (2013) Synthesis of a photocaged tamoxifen for light-dependent activation of Cre-ER recombinase-driven gene modification. *Chem. Commun.* 49, 4971–4973.
- (24) Gautier, A., Nguyen, D. P., Lusic, H., An, W., Deiters, A., and Chin, J. W. (2010) Genetically encoded photocontrol of protein localization in mammalian cells. *J. Am. Chem. Soc.* 132, 4086–4088.
- (25) Engelke, H., Chou, C., Uprety, R., Jess, P., and Deiters, A. (2014) Control of protein function through optochemical translocation. *ACS Synth. Biol.* 3, 731–736.
- (26) Edwards, W. F., Young, D. D., and Deiters, A. (2009) Light-activated Cre recombinase as a tool for the spatial and temporal control of gene function in mammalian cells. *ACS Chem. Biol.* 4, 441–445.
- (27) Fenno, L., Yizhar, O., and Deisseroth, K. (2011) The development and application of optogenetics. *Annu. Rev. Neurosci.* 34, 389–412.
- (28) Konermann, S., Brigham, M. D., Trevino, A. E., Hsu, P. D., Heidenreich, M., Cong, L., Platt, R. J., Scott, D. A., Church, G. M., and Zhang, F. (2013) Optical control of mammalian endogenous transcription and epigenetic states. *Nature* 500, 472–476.
- (29) Nihongaki, Y., Yamamoto, S., Kawano, F., Suzuki, H., and Sato, M. (2015) CRISPR-Cas9-based photoactivatable transcription system. *Chem. Biol.* 22, 169–74.

Electronic Supporting Information

Switching protein localization by site-directed RNA editing under control of light

Paul Vogel, Alfred Hanswillemenke and Thorsten Stafforst*

University of Tübingen, Interfaculty Institute of Biochemistry, Auf der Morgenstelle 15, 72076 Tübingen (Germany); *correspondence to thorsten.stafforst@uni-tuebingen.de

Methods

Editing under transient expression: 2×10^5 cells (293T) cells were grown in 24-well format in DMEM+10%FBS+1xP/S, 5% CO₂. After 24h, forward transfection was carried out in 24-wells with 300 ng of the respective plasmid/well with Lipofectamine 2000 (4 μ l per 1 μ g plasmid) in DMEM+10%FBS. As indicated, the respective guideRNA (2.5-10 pmol/well) was reverse transfected (4 h) in a 96-well format with Lipofectamine 2000 (0.5 μ L/well) in DMEM + 1% FBS. Cells were seeded on cover slips (DMEM+1%FBS+HEPES). After incubation (24h) cells were harvested for RNA sequencing or stained with BG-FITC for fluorescence microscopy as described before.

Editing under genomic expression: 4×10^4 293-Flp-In T-REx cells were induced in DMEM+10%FBS + 15 μ g/mL blasticidin S + 100 μ g/mL hygromycin B + 10 ng/mL doxycycline, 5%CO₂ in a 24 well-format. After 24 h, the respective guideRNA (5-20 pmol/well) was reverse transfected (4 h) in a 96-well format with Lipofectamine 2000 (0.75 μ L/well) in DMEM + 10% FBS +10 ng/mL doxycycline. Cells were seeded on cover slips (DMEM+10%FBS+HEPES+doxycycline). After incubation (24h) cells were harvested for RNA sequencing or stained with BG-FITC for fluorescence microscopy as described before.

Light-induced RNA editing: Experiments were carried out as described above with irradiation as an additional step. After reverse transfection of the guideRNA for 4hrs, cells were washed and the whole well was irradiated with 365 nm light on an inverted cell culture microscope (Zeiss Cell ObserverZ1, Colibri.2 with 365nm LED light source) for 5 sec under full power (DAPI filter) at 5x magnification. Then the protocol was continued as described above.

Quantification of A-to-I change at RNA: In case of changing the protein localization from the cytoplasm to the nucleus, cells were collected for RNA isolation with the RNeasy MinElute Kit (Qiagen, Germany) according to the manufacturer's protocol. After DNaseI digestion, RNA was converted in cDNA for the subsequent amplification by Taq DNA PCR. The DNA was delivered to Eurofins Genomics (Germany) for Sanger sequencing. The A-to-I change was calculated by the peak height of the resulting guanosine divided by the sum of the heights of the guanosine and adenosine. For RNA editing quantification in case of changing the protein localization from the cytoplasm to the membrane, RNA was isolation using 500 μ l TRI Reagent (Sigma Aldrich, Germany). The RNA was separated by 100 μ l chloroform and precipitated by 350 μ l isopropanol supplemented with 1.5 μ l linear acrylamide (5 mg/ml). The RNA pellet was washed twice with 500 μ l of 75% ethanol and dissolved in 25 μ l RNase-free water. Afterwards the steps were performed as described above.

BG-FITC staining: To determine the localization of SNAP-ADAR2, 303 μ l medium was removed and 1 μ l of 400 μ M acetylated BG-FITC (for final concentration of 2 μ M) together with 2 μ l of NucBlue Live ReadyProbes Reagent (Thermo Fisher Scientific, R37605) were added to the cells. After 30 min in the incubator, cells were fixed by 21.6 μ l aqueous deionized formaldehyde (37%) at room temperature for 3 min and washed threefold with PBS. 200 μ l of a 0.1% Triton X-100/PBS solution was used to permeabilized the cells at room temperature for 15 min. After washing the threefold with PBS, cover glasses were mounted onto microscope slides using Shandon Immu-Mount (Thermo Fisher Scientific, USA).

Immunofluorescence microscopy: From each well (24 well) 300 μ l medium was removed and 21.6 μ l aqueous deionized formaldehyde (37%) were added to obtain a final concentration of 4%. After 3 min incubation at room temperature, cells were washed threefold with PBS (5 min) and incubated with 500 μ l PBS+10% FBS at 4 °C overnight for blocking. After removing the blocking solution, cells were incubated with 200 μ l of a mouse anti-HA-antibody (Sigma Aldrich, H9658) diluted 1:1250 in PBS+5% FBS for 1.5h at

room temperature. Cells were washed threefold with PBS and incubated with 200 μ l of a secondary anti-mouse antibody conjugated to Alexa Fluor 594 (Thermo Fisher Scientific, A-11005) diluted 1:1500 in PBS+10% FBS for 45 min at room temperature. After threefold washing with PBS, cover glasses were fixed on microscope slides with Dako mounting medium (Dako North America, USA). Microscopy was performed with a Zeiss CellObserverZ1 under 600x total magnification.

Materials

Full DNA and protein sequence of the construct SNAP-ADAR2-TAG-NLS in the context of the pcDNA3.1 vector for transient expression (green = editing site, yellow = NLS).

```

      10      20      30      40      50      60
1      CTCGGATCCACCATGGACAAAGACTGCGAAATGAAGCGCACCACCCTGGATAGCCCTCTG
1      M D K D C E M K R T T L D S P L
      BamHI
      70      80      90      100     110     120
61     GGCAAGCTGGAAGTGTCTGGGTGCGAACAGGGCCTGCACCGTATCATCTTCCTGGGCAAA
21     G K L E L S G C E Q G L H R I I F L G K

      130     140     150     160     170     180
121    GGAACATCTGCCGCCGACGCCGTGGAAGTGCCTGCCCCAGCCGCCGTGCTGGGCGGACCA
41    G T S A A D A V E V P A P A A V L G G P

      190     200     210     220     230     240
181    GAGCCACTGATGCAGGCCACCGCCTGGCTCAACGCCTACTTTTACCAGCCTGAGGCCATC
61    E P L M Q A T A W L N A Y F H Q P E A I

      250     260     270     280     290     300
241    GAGGAGTTCCTGTGCCAGCCCTGCACCACCCAGTGTTCAGCAGGAGAGCTTTACCCGC
81    E E F P V P A L H H P V F Q Q E S F T R

      310     320     330     340     350     360
301    CAGGTGCTGTGGAAGTGTGAAAGTGGTGAAGTTCCGGAGAGGTCATCAGCTACAGCCAC
101   Q V L W K L L K V V K F G E V I S Y S H

      370     380     390     400     410     420
361    CTGGCCGCCCTGGCCGGCAATCCCGCCGCCACCGCCCGTGA AAAACCGCCCTGAGCGGA
121   L A A L A G N P A A T A A V K T A L S G

      430     440     450     460     470     480
421    AATCCCGTGCCATTCTGATCCCCTGCCACCGGGTGGTGCAGGGCGACCTGGACGTGGGG
141   N P V P I L I P C H R V V Q G D L D V G

      490     500     510     520     530     540
481    GGCTACGAGGGCGGGCTCGCCGTGAAAGAGTGGCTGCTGGCCCACGAGGGCCACAGACTG
161   G Y E G G L A V K E W L L A H E G H R L
```

541 550 560 570 580 590 600
181 GGCAAGCCTGGGCTGGGTCTGCAGGCGGAGGCGCGCCAGGGTCTGGCGGGCGGCAGTAAG
G K P G L G P A G G G A P G S G G G S K

601 610 620 630 640 650 660
201 AAGCTTGCCAAGGCCCGGGCTGCGCAGTCTGCCCTGGCCGCCATTTTTAACTTGCACTTG
K L A K A R A A Q S A L A A I F N L H L

661 670 680 690 700 710 720
221 GATCAGACGCCATCTCGCCAGCCTATTCCCAGTGAGGGTCTTCAGCTGCATTTACCGCAG
D Q T P S R Q P I P S E G L Q L H L P Q

721 730 740 750 760 770 780
241 GTTTTAGCTGACGCTGTCTCACGCCTGGTCCTGGGTAAGTTTGGTGACCTGACCGACAAC
V L A D A V S R L V L G K F G D L T D N

781 790 800 810 820 830 840
261 TTCTCCTCCCCTCACGCTCGCAGAAAAGTGCTGGCTGGAGTCGTCATGACAACAGGCACA
F S S P H A R R K V L A G V V M T T G T

841 850 860 870 880 890 900
281 GATGTTAAAGATGCCAAGGTGATAAGTGTCTTACAGGAACAAAATGTATTAATGGTGAA
D V K D A K V I S V S T G T K C I N G E

901 910 920 930 940 950 960
301 TACATGAGTGATCGTGGCCTTGCAATTAATGACTGCCATGCAGAAATAATATCTCGGAGA
Y M S D R G L A L N D C H A E I I S R R

961 970 980 990 1000 1010 1020
321 TCCTTGCTCAGATTTCTTTATACACAACCTTGAGCTTTACTTAAATAACAAAGATGATCAA
S L L R F L Y T Q L E L Y L N N K D D Q

1021 1030 1040 1050 1060 1070 1080
341 AAAAGATCCATCTTTTCAGAAATCAGAGCGAGGGGGTTTAGGCTGAAGGAGAATGTCCAG
K R S I F Q K S E R G G F R L K E N V Q

1081 1090 1100 1110 1120 1130 1140
361 TTTTCATCTGTACATCAGCACCTCTCCCTGTGGAGATGCCAGAATCTTCTCACCACATGAG
F H L Y I S T S P C G D A R I F S P H E

1150 1160 1170 1180 1190 1200

1141 CCAATCCTGGAAGAACCAGCAGATAGACACCCAAATCGTAAAGCAAGAGGACAGCTACGG
381 P I L E E P A D R H P N R K A R G Q L R

1201 ACCAAAATAGAGTCTGGTGAGGGGACGATTCCAGTGCCTCCAATGCGAGCATCCAAACG
401 T K I E S G E G T I P V R S N A S I Q T

1261 TGGGACGGGGTGCTGCAAGGGGAGCGGCTGCTCACCATGTCCTGCAGTGACAAGATTGCA
421 W D G V L Q G E R L L T M S C S D K I A

1321 CGCTGGAACGTGGTGGGCATCCAGGGATCCCTGCTCAGCATTTTTCGTGAGCCCATTAC
441 R W N V V G I Q G S L L S I F V E P I Y

1381 TTCTCGAGCATCATCCTGGGCAGCCTTTACCACGGGGACCACCTTTCCAGGGCCATGTAC
461 F S S I I L G S L Y H G D H L S R A M Y

1441 CAGCGGATCTCCAACATAGAGGACCTGCCACCTCTCTACACCCTCAACAAGCCTTTGCTC
481 Q R I S N I E D L P P L Y T L N K P L L

1501 AGTGGCATCAGCAATGCAGAAGCACGGCAGCCAGGGAAGGCCCCCAACTTCAGTGTCAAC
501 S G I S N A E A R Q P G K A P N F S V N

1561 TGGACGGTAGGCGACTCCGCTATTGAGGTCATCAACGCCACGACTGGGAAGGATGAGCTG
521 W T V G D S A I E V I N A T T G K D E L

1621 GGCCGCGCGTCCCGCCTGTGTAAGCACGCGTTGTACTGTGCTGGATGCGTGTGCACGGC
541 G R A S R L C K H A L Y C R W M R V H G

1681 AAGGTTCCCTCCCACTTACTACGCTCCAAGATTACCAAGCCCAACGTGTACCATGAGTCC
561 K V P S H L L R S K I T K P N V Y H E S

1741 AAGCTGGCGGCAAAGGAGTACCAGGCCGCAAGGCGCGTCTGTTTCACAGCCTTCATCAAG


```

581      K L A A K E Y Q A A K A R L F T A F I K
      1810      1820      1830      1840      1850      1860
1801    GCGGGGCTGGGGCCTGGGTGGAGAAGCCCACCGAGCAGGACCAGTTCTCACTCACGCC
601      A G L G A W V E K P T E Q D Q F S L T P
      1870      1880      1890      1900      1910      1920
1861    TCCACCGGCGGCATGGACGAGCTGTACAAGGCTAGCCCCGGGCCCCCAAAGTGCCTGTT
621      S T G G M D E L Y K A S P G P P K L P V
      1930      1940      1950      1960      1970      1980
1921    CCGTAGCCGACACTAGGTACCCCAAAAAAGAAGAAAGGTGCCGAAGAAGAAGGAAG
641      P * P T L G T P K K K R K V P K K K R K
      1990      2000      2010      2020      2030      2040
1981    GTGGATCCTAAGAAAAAAGGAAAGTTTCTAGAGGGCCCTTCGAACAAAAACTCATCTCA
661      V D P K K K R K V S R G P F E Q K L I S
      XbaI
      2050      2060      2070      2080      2090
2041    GAAGAGGATCTGAATATGCATACCGGTCATCATCACCATCACCATTGA
681      E E D L N M H T G H H H H H H *

```

Full DNA and protein sequence of the construct SNAP-ADAR2-TAG-NLS in the context of the pcDNA5 vector for genomic expression (green = editing site, yellow = NLS).

```

      10      20      30      40      50      60
1    CTCGGATCCACCATGGACAAAGACTGCGAAATGAAGCGCACCACCCTGGATAGCCCTCTG
1      M D K D C E M K R T T L D S P L
      BamHI
      70      80      90      100      110      120
61    GGCAAGCTGGAAGTGTCTGGGTGCGAACAGGGCCTGCACCGTATCATCTTCCTGGGCAAA
21      G K L E L S G C E Q G L H R I I F L G K
      130      140      150      160      170      180
121    GGAACATCTGCCCGCAGCCGTGGAAGTGCCTGCCCCAGCCCGCTGCTGGGCGGACCA
41      G T S A A D A V E V P A P A A V L G G P
      190      200      210      220      230      240
181    GAGCCACTGATGCAGGCCACCGCCTGGCTCAACGCCTACTTTTACCAGCCTGAGGCCATC
61      E P L M Q A T A W L N A Y F H Q P E A I

```

241 250 260 270 280 290 300
81 GAGGAGTTCCTGTGCCAGCCCTGCACCACCCAGTGTTCAGCAGGAGAGCTTTACCCGC
E E F P V P A L H H P V F Q Q E S F T R

301 310 320 330 340 350 360
101 CAGGTGCTGTGGAAACTGCTGAAAGTGGTGAAGTTCGGAGAGGTCATCAGCTACAGCCAC
Q V L W K L L K V V K F G E V I S Y S H

361 370 380 390 400 410 420
121 CTGGCCGCCCTGGCCGGCAATCCCGCCGCCACCGCCCGGTGAAAACCGCCCTGAGCGGA
L A A L A G N P A A T A A V K T A L S G

421 430 440 450 460 470 480
141 AATCCCGTGCCCATTTCTGATCCCCTGCCACCGGGTGGTGCAGGGCGACCTGGACGTGGGG
N P V P I L I P C H R V V Q G D L D V G

481 490 500 510 520 530 540
161 GGCTACGAGGGCGGGCTCGCCGTGAAAGAGTGGCTGCTGGCCCACGAGGGCCACAGACTG
G Y E G G L A V K E W L L A H E G H R L

541 550 560 570 580 590 600
181 GGCAAGCCTGGGCTGGGTCTGCAGGCGGAGGCGCGCCAGGGTCTGGCGGCGGCAGTAAG
G K P G L G P A G G G A P G S G G G S K

601 610 620 630 640 650 660
201 AAGCTTGCCAAGGCCCGGGCTGCGCAGTCTGCCCTGGCCGCCATTTTTAACTTGCACTTG
K L A K A R A A Q S A L A A I F N L H L

661 670 680 690 700 710 720
221 GATCAGACGCCATCTCGCCAGCCTATTCCCAGTGAGGGTCTTCAGCTGCATTTACCGCAG
D Q T P S R Q P I P S E G L Q L H L P Q

721 730 740 750 760 770 780
241 GTTTTAGCTGACGCTGTCTCACGCCTGGTCCTGGGTAAGTTTGGTGACCTGACCGACAAC
V L A D A V S R L V L G K F G D L T D N

781 790 800 810 820 830 840
261 TTCTCCTCCCCTCACGCTCGCAGAAAAGTGTGGCTGGAGTCGTCATGACAACAGGCACA
F S S P H A R R K V L A G V V M T T G T

841 850 860 870 880 890 900
281 GATGTTAAAGATGCCAAGGTGATAAGTGTCTTCTACAGGAACAAAATGTATTAATGGTGAA
D V K D A K V I S V S T G T K C I N G E

901 910 920 930 940 950 960
301 TACATGAGTGATCGTGGCCTTGCATTAATGACTGCCATGCAGAAATAATATCTCGGAGA
Y M S D R G L A L N D C H A E I I S R R

961 970 980 990 1000 1010 1020
321 TCCTTGCTCAGATTTCTTTATACACAACCTTGAGCTTTACTTAAATAACAAAGATGATCAA
S L L R F L Y T Q L E L Y L N N K D D Q

1021 1030 1040 1050 1060 1070 1080
341 AAAAGATCCATCTTTTCAGAAATCAGAGCGAGGGGGGTTTAGGCTGAAGGAGAATGTCCAG
K R S I F Q K S E R G G F R L K E N V Q

1081 1090 1100 1110 1120 1130 1140
361 TTTTCATCTGTACATCAGCACCTCTCCCTGTGGAGATGCCAGAATCTTCTCACCACATGAG
F H L Y I S T S P C G D A R I F S P H E

1141 1150 1160 1170 1180 1190 1200
381 CCAATCCTGGAAGAACCAGCAGATAGACACCCAAATCGTAAAGCAAGAGGACAGCTACGG
P I L E E P A D R H P N R K A R G Q L R

1201 1210 1220 1230 1240 1250 1260
401 ACCAAAATAGAGTCTGGTGAGGGGACGATTCCAGTGCCTCCAATGCGAGCATCCAAACG
T K I E S G E G T I P V R S N A S I Q T

1261 1270 1280 1290 1300 1310 1320
421 TGGGACGGGGTGCTGCAAGGGGAGCGGCTGCTCACCATGTCCTGCAGTGACAAGATTGCA
W D G V L Q G E R L L T M S C S D K I A

1321 1330 1340 1350 1360 1370 1380
441 CGCTGGAACGTGGTGGGCATCCAGGGATCCCTGCTCAGCATTTCGTTGGAGCCCATTAC
R W N V V G I Q G S L L S I F V E P I Y

1381 1390 1400 1410 1420 1430 1440
461 TTCTCGAGCATCATCCTGGGCAGCCTTTACCACGGGGACCACCTTTCCAGGGCCATGTAC
F S S I I L G S L Y H G D H L S R A M Y

1450 1460 1470 1480 1490 1500

1441 CAGCGGATCTCCAACATAGAGGACCTGCCACCTCTCTACACCCTCAACAAGCCTTTGCTC
481 Q R I S N I E D L P P L Y T L N K P L L

1501 1510 1520 1530 1540 1550 1560
AGTGGCATCAGCAATGCAGAAGCACGGCAGCCAGGGAAGGCCCAACTTCAGTGTCAAC
501 S G I S N A E A R Q P G K A P N F S V N

1561 1570 1580 1590 1600 1610 1620
TGGACGGTAGGCGACTCCGCTATTGAGGTCATCAACGCCACGACTGGGAAGGATGAGCTG
521 W T V G D S A I E V I N A T T G K D E L

1621 1630 1640 1650 1660 1670 1680
GGCCGCGCGTCCCGCCTGTGTAAGCACGCGTTGTACTGTCGCTGGATGCGTGTGCACGGC
541 G R A S R L C K H A L Y C R W M R V H G

1681 1690 1700 1710 1720 1730 1740
AAGGTTCCCTCCCACTTACTACGCTCCAAGATTACCAAGCCCAACGTGTACCATGAGTCC
561 K V P S H L L R S K I T K P N V Y H E S

1741 1750 1760 1770 1780 1790 1800
AAGCTGGCGGCAAAGGAGTACCAGGCCCAAGGCGCGTCTGTTCACAGCCTTCATCAAG
581 K L A A K E Y Q A A K A R L F T A F I K

1801 1810 1820 1830 1840 1850 1860
GCGGGGCTGGGGGCCTGGGTGGAGAAGCCCACCGAGCAGGACCAGTTCTCACTCACGCCC
601 A G L G A W V E K P T E Q D Q F S L T P

1861 1870 1880 1890 1900 1910 1920
GCTAGCCCCGGGCCCCCAAAGTGCCTGTTCCGTAGCCGACACTAGGTACCCAAAAAAG
621 A S P G P P K L P V P * P T L G T P K K

1921 1930 1940 1950 1960 1970 1980
AAGAGAAAGGTGCCGAAGAAGAAGAGAAAGGTAGATCCTAAGAAAAAAGGAAAGTTGCG
641 K R K V P K K K R K V D P K K K R K V A

1981 1990
661 GCCGCTCGAGTCTAG
A A R V *

NotI

Full DNA and protein sequence of the construct ATAG-NLS-SNAP-ADAR2 in the context of the pcDNA3.1 vector for transient expression (green = editing site, yellow = NLS).

```

      10      20      30      40      50      60
1  TTGGTACC GAGCTCCACCATAGCCCCAAAAAAGAAGAGAAAGGTGCCGAAGAAGAAGAGG
1  A P K K K R K V P K K K R
      KpnI
      70      80      90      100     110     120
61  AAGGTGGATCCTAAGAAAAAAGGAAAGTTGGATCCACCATGGACAAAGACTGCGAAATG
21  K V D P K K K R K V G S T M D K D C E M

      130     140     150     160     170     180
121 AAGCGCACCACCCTGGATAGCCCTCTGGGCAAGCTGGAAGTGTCTGGGTGCGAACAGGGC
41  K R T T L D S P L G K L E L S G C E Q G

      190     200     210     220     230     240
181 CTGCACCGTATCATCTTCTGGGCAAAGGAACATCTGCCGCCGACGCCGTGGAAGTGCCT
61  L H R I I F L G K G T S A A D A V E V P

      250     260     270     280     290     300
241 GCCCCAGCCGCCGTGCTGGGCGGACCAGAGCCACTGATGCAGGCCACCGCCTGGCTCAAC
81  A P A A V L G G P E P L M Q A T A W L N

      310     320     330     340     350     360
301 GCCTACTTTTACCAGCCTGAGGCCATCGAGGAGTTCCCTGTGCCAGCCCTGCACCACCCA
101 A Y F H Q P E A I E E F P V P A L H H P

      370     380     390     400     410     420
361 GTGTTCCAGCAGGAGAGCTTTACCCGCCAGGTGCTGTGGAAACTGCTGAAAGTGGTGAAG
121 V F Q Q E S F T R Q V L W K L L K V V K

      430     440     450     460     470     480
421 TTCGGAGAGGTCATCAGCTACAGCCACCTGGCCGCCCTGGCCGGCAATCCCGCCGCCACC
141 F G E V I S Y S H L A A L A G N P A A T

      490     500     510     520     530     540
481 GCCGCCGTGAAAACCGCCCTGAGCGGAAATCCCGTGCCCATCTGATCCCCTGCCACCGG
161 A A V K T A L S G N P V P I L I P C H R

```

541 181 550 560 570 580 590 600
GTGGTGCAGGGCGACCTGGACGTGGGGGGCTACGAGGGCGGGCTCGCCGTGAAAGAGTGG
V V Q G D L D V G G Y E G G L A V K E W

601 201 610 620 630 640 650 660
CTGCTGGCCCACGAGGGCCACAGACTGGGCAAGCCTGGGCTGGGTCTGCAGGCGGAGGC
L L A H E G H R L G K P G L G P A G G G

661 221 670 680 690 700 710 720
GCGCCAGGGTCTGGCGGGCAGTAAGAAGCTTGCCAAGGCCCGGGCTGCGCAGTCTGCC
A P G S G G G S K K L A K A R A A Q S A

721 241 730 740 750 760 770 780
CTGGCCGCCATTTTTAACTTGCACCTGGATCAGACGCCATCTCGCCAGCCTATTCCCAGT
L A A I F N L H L D Q T P S R Q P I P S

781 261 790 800 810 820 830 840
GAGGGTCTTCAGCTGCATTTACCGCAGGTTTTAGCTGACGCTGTCTCACGCCTGGTCTG
E G L Q L H L P Q V L A D A V S R L V L

841 281 850 860 870 880 890 900
GGTAAGTTTGGTGACCTGACCGACAATTCTCCTCCCCTCACGCTCGCAGAAAAGTGCTG
G K F G D L T D N F S S P H A R R K V L

901 301 910 920 930 940 950 960
GCTGGAGTCGTCATGACAACAGGCACAGATGTTAAAGATGCCAAGGTGATAAGTGTCTTCT
A G V V M T T G T D V K D A K V I S V S

961 321 970 980 990 1000 1010 1020
ACAGGAACAAAATGTATTAATGGTGAATACATGAGTGATCGTGGCCTTGCATTAATGAC
T G T K C I N G E Y M S D R G L A L N D

1021 341 1030 1040 1050 1060 1070 1080
TGCCATGCAGAAATAATATCTCGGAGATCCTTGCTCAGATTTCTTTATACACAACCTTGAG
C H A E I I S R R S L L R F L Y T Q L E

1081 361 1090 1100 1110 1120 1130 1140
CTTTACTTAAATAACAAAGATGATCAAAAAAGATCCATCTTTTCAGAAATCAGAGCGAGGG
L Y L N N K D D Q K R S I F Q K S E R G

1150 1160 1170 1180 1190 1200

1141 GGGTTTAGGCTGAAGGAGAATGTCCAGTTTTCATCTGTACATCAGCACCTCTCCCTGTGGA
381 G F R L K E N V Q F H L Y I S T S P C G

1201 1210 1220 1230 1240 1250 1260
GATGCCAGAATCTTCTCACCACATGAGCCAATCCTGGAAGAACCAGCAGATAGACACCCA
401 D A R I F S P H E P I L E E P A D R H P

1261 1270 1280 1290 1300 1310 1320
AATCGTAAAGCAAGAGGACAGCTACGGACAAAATAGAGTCTGGTGAGGGGACGATTCCA
421 N R K A R G Q L R T K I E S G E G T I P

1321 1330 1340 1350 1360 1370 1380
GTGCGCTCCAATGCGAGCATCCAAACGTGGGACGGGGTCTGCAAGGGGAGCGGCTGCTC
441 V R S N A S I Q T W D G V L Q G E R L L

1381 1390 1400 1410 1420 1430 1440
ACCATGTCCTGCAGTGACAAGATTGCACGCTGGAACGTGGTGGGCATCCAGGGATCCCTG
461 T M S C S D K I A R W N V V G I Q G S L

1441 1450 1460 1470 1480 1490 1500
CTCAGCATTTCGTGGAGCCATTTACTTCTCGAGCATCATCCTGGGCAGCCTTTACCAC
481 L S I F V E P I Y F S S I I L G S L Y H

1501 1510 1520 1530 1540 1550 1560
GGGGACCACCTTTCCAGGGCCATGTACCAGCGGATCTCCAACATAGAGGACCTGCCACCT
501 G D H L S R A M Y Q R I S N I E D L P P

1561 1570 1580 1590 1600 1610 1620
CTCTACACCCTCAACAAGCCTTTGCTCAGTGGCATCAGCAATGCAGAAGCACGGCAGCCA
521 L Y T L N K P L L S G I S N A E A R Q P

1621 1630 1640 1650 1660 1670 1680
GGGAAGGCCCCCAACTTCACTGACTGCAACTGGACGGTAGGCGACTCCGCTATTGAGGTCATC
541 G K A P N F S V N W T V G D S A I E V I

1681 1690 1700 1710 1720 1730 1740
AACGCCACGACTGGGAAGGATGAGCTGGGCCGCGCTCCCGCCTGTGTAAGCACGCGTTG
561 N A T T G K D E L G R A S R L C K H A L

1741 1750 1760 1770 1780 1790 1800
TACTGTCGCTGGATGCGTGTGCACGGCAAGGTTCCCTCCCACTTACTACGCTCCAAGATT

```

581      Y C R W M R V H G K V P S H L L R S K I
      1810      1820      1830      1840      1850      1860
1801  ACCAAGCCCAACGTGTACCATGAGTCCAAGCTGGCGGCAAAGGAGTACCAGGCCGCCAAG
601      T K P N V Y H E S K L A A K E Y Q A A K
      1870      1880      1890      1900      1910      1920
1861  GCGCGTCTGTTCACAGCCTTCATCAAGGCGGGGCTGGGGGCCTGGGTGGAGAAGCCCACC
621      A R L F T A F I K A G L G A W V E K P T
      1930      1940      1950      1960      1970      1980
1921  GAGCAGGACCAGTTCTCACTCACGCCCTCTAGAGGGCCCTTCGAACAAAACTCATCTCA
641      E Q D Q F S L T P S R G P F E Q K L I S
      xbaI
      1990      2000      2010      2020      2030
1981  GAAGAGGATCTGAATATGCATACCGGTCATCATCACCATCACCATTGA
661      E E D L N M H T G H H H H H H *

```

Full DNA and protein sequence of the construct ATAG-NLS-SNAP-ADAR2* in the context of the pcDNA5 vector for genomic expression (green = editing site, yellow = NLS, activating E/Q site = grey).

```

1      TTGGTACC GAGCTCCACCATAG CCCCCAAAAAAGAAGAGAAAGGTGCCGAAGAAGAAGAGA
1      A P K K K R K V P K K K R
      KpnI
      70      80      90      100      110      120
61      AAGGTAGATCCTAAGAAAAAAGGAAAGTTGGATCCACCATGGACAAAGACTGCGAAATG
21      K V D P K K K R K V G S T M D K D C E M
      130      140      150      160      170      180
121  AAGCGCACCACCCTGGATAGCCCTCTGGGCAAGCTGGAAGTGTCTGGGTGCGAACAGGGC
41      K R T T L D S P L G K L E L S G C E Q G
      190      200      210      220      230      240
181  CTGCACCGTATCATCTTCTGGGCAAAGGAACATCTGCCGCCGACGCCGTGGAAGTGCCT
61      L H R I I F L G K G T S A A D A V E V P
      250      260      270      280      290      300
241  GCCCCAGCCCGCGTGTCTGGGCGGACCAGAGCCACTGATGCAGGCCACCGCCTGGCTCAAC
81      A P A A V L G G P E P L M Q A T A W L N

```


301 310 320 330 340 350 360
GCCTACTTTTACCAGCCTGAGGCCATCGAGGAGTTCCCTGTGCCAGCCCTGCACCACCCA
101 A Y F H Q P E A I E E F P V P A L H H P

361 370 380 390 400 410 420
GTGTTCCAGCAGGAGAGCTTTACCCGCCAGGTGCTGTGGAAACTGCTGAAAGTGGTGAAG
121 V F Q Q E S F T R Q V L W K L L K V V K

421 430 440 450 460 470 480
TTCGGAGAGGTCATCAGCTACAGCCACCTGGCCGCCCTGGCCGGCAATCCCGCCGCCACC
141 F G E V I S Y S H L A A L A G N P A A T

481 490 500 510 520 530 540
GCCGCCGTGAAAACCGCCCTGAGCGGAAATCCCGTGCCATTCTGATCCCCTGCCACCGG
161 A A V K T A L S G N P V P I L I P C H R

541 550 560 570 580 590 600
GTGGTGCAGGGCGACCTGGACGTGGGGGGCTACGAGGGCGGGCTCGCCGTGAAAGAGTGG
181 V V Q G D L D V G G Y E G G L A V K E W

601 610 620 630 640 650 660
CTGCTGGCCCACGAGGGCCACAGACTGGGCAAGCCTGGGCTGGGTCTGCAGGCGGAGGC
201 L L A H E G H R L G K P G L G P A G G G

661 670 680 690 700 710 720
GCGCCAGGGTCTGGCGGCGGCAGTAAGAAGCTTGCCAAGGCCCGGGCTGCGCAGTCTGCC
221 A P G S G G G S K K L A K A R A A Q S A

721 730 740 750 760 770 780
CTGGCCGCCATTTTTAACTTGCACCTGGATCAGACGCCATCTCGCCAGCCTATTCCCAGT
241 L A A I F N L H L D Q T P S R Q P I P S

781 790 800 810 820 830 840
GAGGGTCTTCAGCTGCATTTACCGCAGGTTTTAGCTGACGCTGTCTCACGCCTGGTCCTG
261 E G L Q L H L P Q V L A D A V S R L V L

841 850 860 870 880 890 900
GGTAAGTTTGGTGACCTGACCGACAATTCTCCTCCCCTCACGCTCGCAGAAAAGTGCTG
281 G K F G D L T D N F S S P H A R R K V L

910 920 930 940 950 960

901 GCTGGAGTCGTCATGACAACAGGCACAGATGTTAAAGATGCCAAGGTGATAAGTGTTTCT
301 A G V V M T T G T D V K D A K V I S V S

970 980 990 1000 1010 1020
961 ACAGGAACAAAATGTATTAATGGTGAATACATGAGTGATCGTGGCCTTGCATTAAATGAC
321 T G T K C I N G E Y M S D R G L A L N D

1030 1040 1050 1060 1070 1080
1021 TGCCATGCAGAAATAATATCTCGGAGATCCTTGCTCAGATTTCTTTATAACAACCTTGAG
341 C H A E I I S R R S L L R F L Y T Q L E

1090 1100 1110 1120 1130 1140
1081 CTTTACTTAAATAACAAAGATGATCAAAAAAGATCCATCTTTTCAGAAATCAGAGCGAGGG
361 L Y L N N K D D Q K R S I F Q K S E R G

1150 1160 1170 1180 1190 1200
1141 GGGTTTAGGCTGAAGGAGAATGTCCAGTTTTCATCTGTACATCAGCACCTCTCCCTGTGGA
381 G F R L K E N V Q F H L Y I S T S P C G

1210 1220 1230 1240 1250 1260
1201 GATGCCAGAATCTTCTCACCACATGAGCCAATCCTGGAAGAACCAGCAGATAGACACCCA
401 D A R I F S P H E P I L E E P A D R H P

1270 1280 1290 1300 1310 1320
1261 AATCGTAAAGCAAGAGGACAGCTACGGACCAAAATAGAGTCTGGTCAGGGGACGATTCCA
421 N R K A R G Q L R T K I E S G Q G T I P

1330 1340 1350 1360 1370 1380
1321 GTGCGCTCCAATGCGAGCATCCAAACGTGGGACGGGGTGCTGCAAGGGGAGCGGCTGCTC
441 V R S N A S I Q T W D G V L Q G E R L L

1390 1400 1410 1420 1430 1440
1381 ACCATGTCCTGCAGTGACAAGATTGCACGCTGGAACGTGGTGGGCATCCAGGGATCCCTG
461 T M S C S D K I A R W N V V G I Q G S L

1450 1460 1470 1480 1490 1500
1441 CTCAGCATTTTCGTGGAGCCATTTACTTCTCGAGCATCATCCTGGGCAGCCTTTACCAC
481 L S I F V E P I Y F S S I I L G S L Y H

1510 1520 1530 1540 1550 1560
1501 GGGGACCACCTTTCCAGGGCCATGTACCAGCGGATCTCCAACATAGAGGACCTGCCACCT

501 G D H L S R A M Y Q R I S N I E D L P P
 1570 1580 1590 1600 1610 1620
 1561 CTCTACACCCTCAACAAGCCTTTGCTCAGTGGCATCAGCAATGCAGAAGCACGGCAGCCA
 521 L Y T L N K P L L S G I S N A E A R Q P
 1630 1640 1650 1660 1670 1680
 1621 GGGAAGGCCCCCAACTTCAGTGTCAACTGGACGGTAGGCGACTCCGCTATTGAGGTCATC
 541 G K A P N F S V N W T V G D S A I E V I
 1690 1700 1710 1720 1730 1740
 1681 AACGCCACGACTGGGAAGGATGAGCTGGGCCGCGCTCCCGCCTGTGTAAGCACGCGTTG
 561 N A T T G K D E L G R A S R L C K H A L
 1750 1760 1770 1780 1790 1800
 1741 TACTGTCGCTGGATGCGTGTGCACGGCAAGGTTCCCTCCCCTTACTACGCTCCAAGATT
 581 Y C R W M R V H G K V P S H L L R S K I
 1810 1820 1830 1840 1850 1860
 1801 ACCAAGCCCAACGTGTACCATGAGTCCAAGCTGGCGGCAAAGGAGTACCAGGCCGCCAAG
 601 T K P N V Y H E S K L A A K E Y Q A A K
 1870 1880 1890 1900 1910 1920
 1861 GCGCGTCTGTTACAGCCTTCATCAAGGCGGGGCTGGGGGCCTGGGTGGAGAAGCCCACC
 621 A R L F T A F I K A G L G A W V E K P T
 1930 1940 1950 1960
 1921 GAGCAGGACCAGTTCTCACTCACGCCCTCTAGAGGGCCCGTTTAA
 641 E Q D Q F S L T P S R G P V *

ApaI

Full DNA sequence of the construct ATAG-IgK-HA-eGFP-PDGFR-TMD. Full DNA and protein sequence of the construct ATAG-IgK-HA-eGFP-PDGFR-TMD in the context of the pcDNA3.1 vector for transient expression (green = editing site, yellow = IgK chain leader sequence).

```

          10      20      30      40      50      60
1      CTCGGATCCACCATAGAGACAGACACACTCCTGCTCTGGGTACTGCTGCTCTGGGTTCCA
1      E T D T L L L W V L L L W V P
      BamHI
          70      80      90      100     110     120
61      GGTTCCTACTGGTGACTCCACCATGTATCCATATGATGTTCCAGATTATGCTGGGGCCCAG
21      G S T G D S T M Y P Y D V P D Y A G A Q

          130     140     150     160     170     180
121     CCGGCTAGCAAAGGAGAAGAAGTCTTCACTGGAGTTGTCCCAATTCTTGTTGAATTAGAT
41     P A S K G E E L F T G V V P I L V E L D

          190     200     210     220     230     240
181     GGTGATGTTAACGGCCACAAGTTCTCTGTCACTGGAGAGGGTGAAGGTGATGCAACATAC
61     G D V N G H K F S V S G E G E G D A T Y

          250     260     270     280     290     300
241     GGAAACTTACCCTGAAGTTCATCTGCACTACTGGCAAAGTGCCTGTTCCGTGGCCGACA
81     G K L T L K F I C T T G K L P V P W P T

          310     320     330     340     350     360
301     CTAGTGACGACGCTCTGCTATGGCGTCCAGTGCTTTTTCAAGATACCCGGATCACATGAAA
101    L V T T L C Y G V Q C F S R Y P D H M K

          370     380     390     400     410     420
361     CGGCATGACTTTTTCAAGAGTGCCATGCCCCGAAGTTATGTACAGGAAAGGACCATCTTC
121    R H D F F K S A M P E G Y V Q E R T I F

          430     440     450     460     470     480
421     TTCAAAGATGACGGCAACTACAAGACACGTGCTGAAGTCAAGTTTGAAGGTGATACCCTT
141    F K D D G N Y K T R A E V K F E G D T L

          490     500     510     520     530     540
481     GTTAATAGAATCGAGTTAAAAGGTATTGACTTCAAGGAAGATGGCAACATTCTGGGACAC
161    V N R I E L K G I D F K E D G N I L G H

          550     560     570     580     590     600

```

541 AAATTGGAATACAACACTATAACTCACACAATGTATACATCATGGCAGACAAAACAAAAGAAT
181 K L E Y N Y N S H N V Y I M A D K Q K N

610 620 630 640 650 660
601 GGAATCAAAGTGAACCTTCAAGACCCGCCACAACATTGAAGATGGAAGCGTTCAACTAGCA
201 G I K V N F K T R H N I E D G S V Q L A

670 680 690 700 710 720
661 GACCATTATCAACAAAATACTCCAATTGGCGATGGCCCTGTCCTTTTACCAGACAACCAT
221 D H Y Q Q N T P I G D G P V L L P D N H

730 740 750 760 770 780
721 TACCTGTCCACACAATCTGCCCTTTTCGAAAGATCCCAACGAAAAGAGAGACCACATGGTC
241 Y L S T Q S A L S K D P N E K R D H M V

790 800 810 820 830 840
781 CTTCTTGAGTTTTGTAACAGCTGCTGGGATTACACATGGCATGGATGAACTATACAAATCC
261 L L E F V T A A G I T H G M D E L Y K S

850 860 870 880 890 900
841 GGCGGTACCGAACAAAAACTCATCTCAGAAGAGGATCTGAATGCTGTGGGCCAGGACACG
281 G G T E Q K L I S E E D L N A V G Q D T

910 920 930 940 950 960
901 CAGGAGGTCATCGTGGTGCCACACTCCTTGCCCTTTAAGGTGGTGGTGATCTCAGCCATC
301 Q E V I V V P H S L P F K V V V I S A I

970 980 990 1000 1010 1020
961 CTGGCCCTGGTGGTGCTCACCATCATCTCCCTTATCATCCTCATCATGCTTTGGCAGAAG
321 L A L V V L T I I S L I I L I M L W Q K

1030
1021 AAGCCACGTTAGTCTAGAGGG
341 K P R *

XbaI

Table S1. Sequences and extinction coefficients ($\epsilon_{260\text{nm}}$) of guideRNAs applied in this study. *NpomBG-* and *BG-*conjugated guideRNAs were synthesized and PAGE-purified from commercially acquired oligonucleotides coupled with an 5'-amino-C6 linker (BioSpring, Germany) as described in Hannswillemecke et al. (*J. Am. Chem. Soc.* 2015, 137, 15875-15881). The nucleotides highlighted in bold are unmodified ribonucleotides and face the nucleotide triplet with the target adenosine in the middle. Nucleotides highlighted in italic are modified with 2'-*O*-methylation. Terminal phosphorothioate linkages are indicated by "s". The first three nucleotides at the 5'-end do not bind to the mRNA substrate, but function as linker between guideRNA and SNAP-tag. The BG and the *Npom* group add 2.5 mM⁻¹ cm⁻¹ and 6.5 mM⁻¹ cm⁻¹, respectively, to the 260 nm extinction coefficient of the NH₂-guideRNAs.

C-terminal NLS inclusion	Sequence	$\epsilon_{260\text{nm}}$ [mM⁻¹ cm⁻¹]
NH ₂ /BG-guideRNA	<i>UsAsUGUGUCGG</i> CCA <i>CGGAAsCsAsGsG</i>	226/228.5
mm BG-guideRNA	<i>UsCsGGAACACC</i> CCA <i>GCAAsCsAsGsA</i>	232.5
N-terminal NLS inclusion	Sequence	$\epsilon_{260\text{nm}}$ [mM⁻¹ cm⁻¹]
NH ₂ /BG/ <i>Npom</i> BG-guideRNA	<i>AsCsAUUUGGGG</i> CCA <i>UGGUGsGsAsGsC</i>	226/228.5/232.5
mm BG-guideRNA	<i>UsAsUGUGUCGG</i> CCA <i>CGGAAsCsAsGsG</i>	228.5
N-terminal Igk inclusion	Sequence	$\epsilon_{260\text{nm}}$ [mM⁻¹ cm⁻¹]
NH ₂ /BG/ <i>Npom</i> BG-guideRNA	<i>AsCsAUCUGUCU</i> CCA <i>UGGUGsGsAsUsC</i>	218/220.5/224.5
mm BG-guideRNA	<i>UsAsUGUGUCGG</i> CCA <i>CGGAAsCsAsGsG</i>	228.5

Primary data

3'UTR Editing to switch protein localization

Transient Expression of SNAP-ADAR2-TAG-NLS

Figure S1. Sanger sequencing of SNAP-ADAR2-TAG-NLS in 293T cells, with 10 pmol guideRNA/well, mm = mismatched.

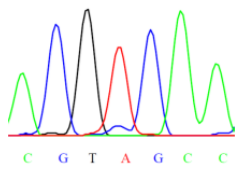
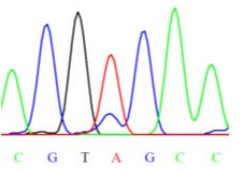
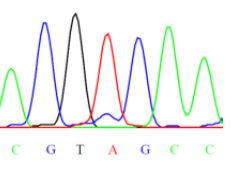
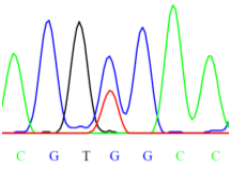
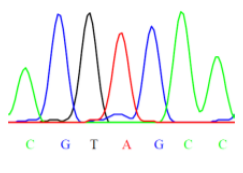
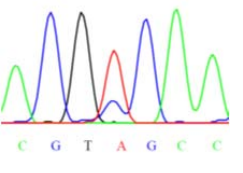
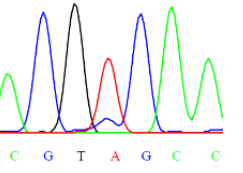
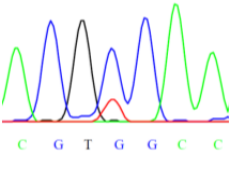
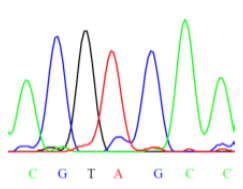
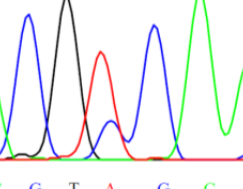
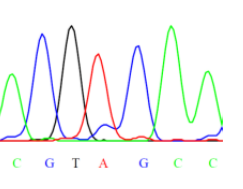
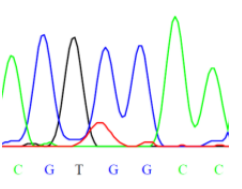
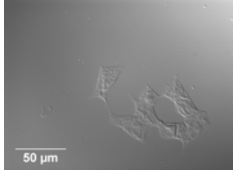

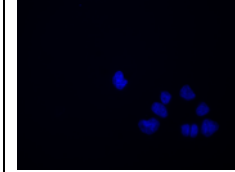
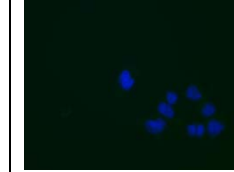
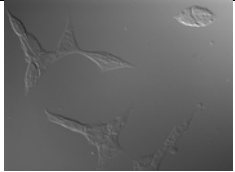

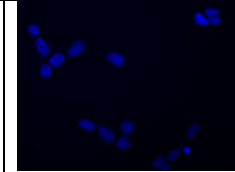
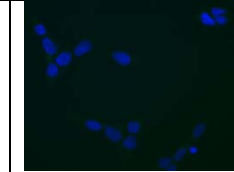
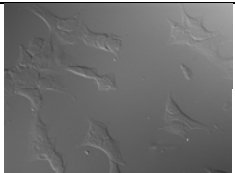
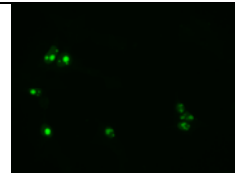
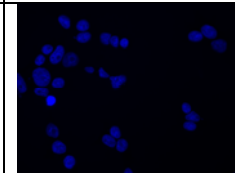
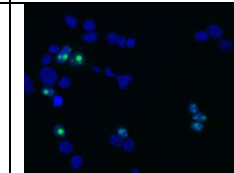
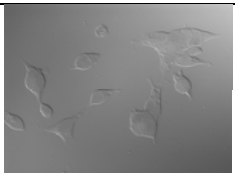
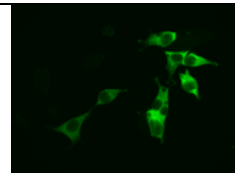
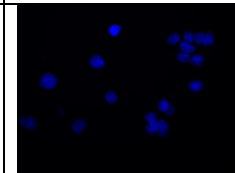
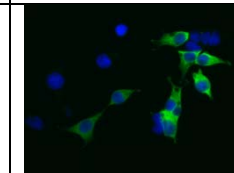
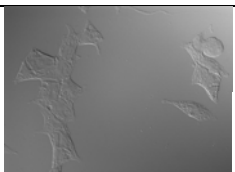
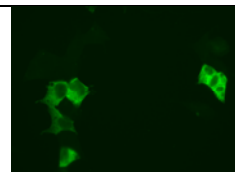
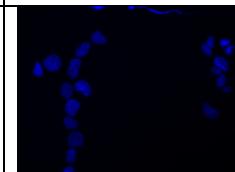
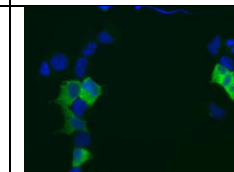
Exp. No.	Lipofectamine	NH2-gRNA	mm BG-gRNA	BG-gRNA
1				
2				
3				

Figure S2. Localization phenotype via FITC staining in transfected 293T cells, mm = mismatched.

Sample	DIC (63×)	FITC (F)	Hoechst (H)	F + H
293T + pcDNA3.1, only Hoechst added				
293T + pcDNA3.1				
293T + SNAP-ADAR2- TGG-NLS				
293T + SNAP-ADAR2- TAG-NLS				
293T + SNAP-ADAR2- TAG-NLS + NH2-gRNA				

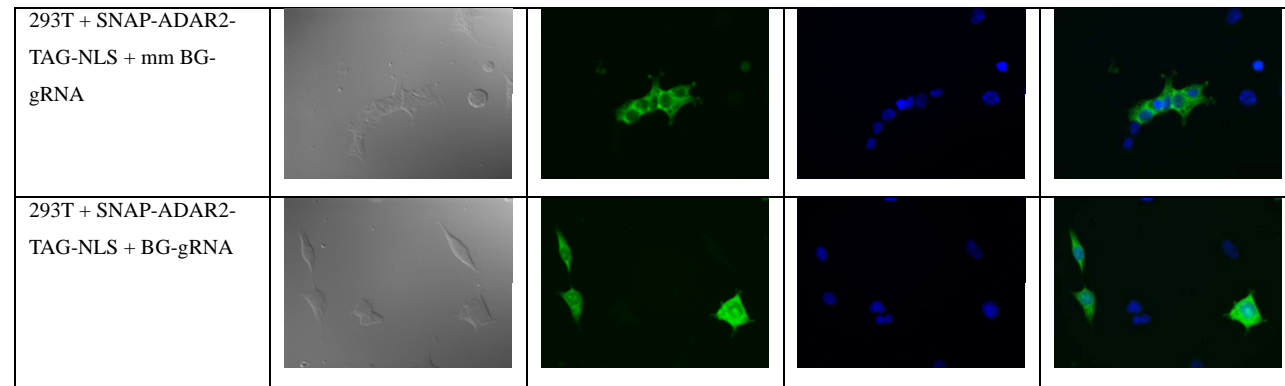
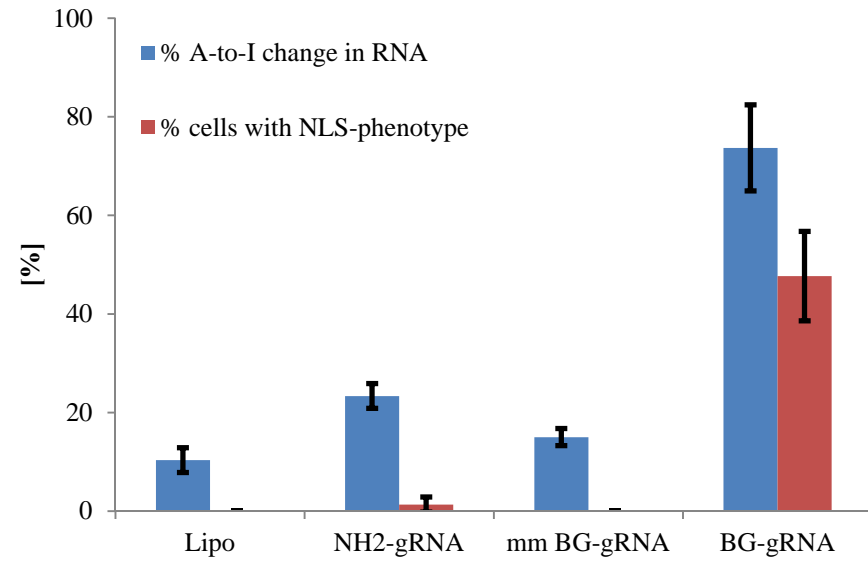


Table S2. Quantification of phenotype switch by fluorescence microscopy, transfected cells with NLS-phenotype [%], 50-110 cells were counted for each sample per experiment, mm = mismatched.

Exp. No.	Lipo	NH2-gRNA	mm BG-gRNA	BG-gRNA
1	0	1	0	38
2	0	3	0	49
3	0	0	0	56
Average	0.00	1.33	0.00	47.67
Standard Deviation	0.00	1.53	0.00	9.07

Figure S3. Comparison of change in RNA (A-to-I change in RNA) with localization switch (cells with NLS-phenotype) in 293T cells, mm = mismatched.



Editing of genomically integrated SNAP-ADAR2-TAG-NLS

Figure S4. Sanger sequencing of 293-SNAP-ADAR2-TAG-NLS cells with 20 pmol guideRNA/well, mm = mismatched.

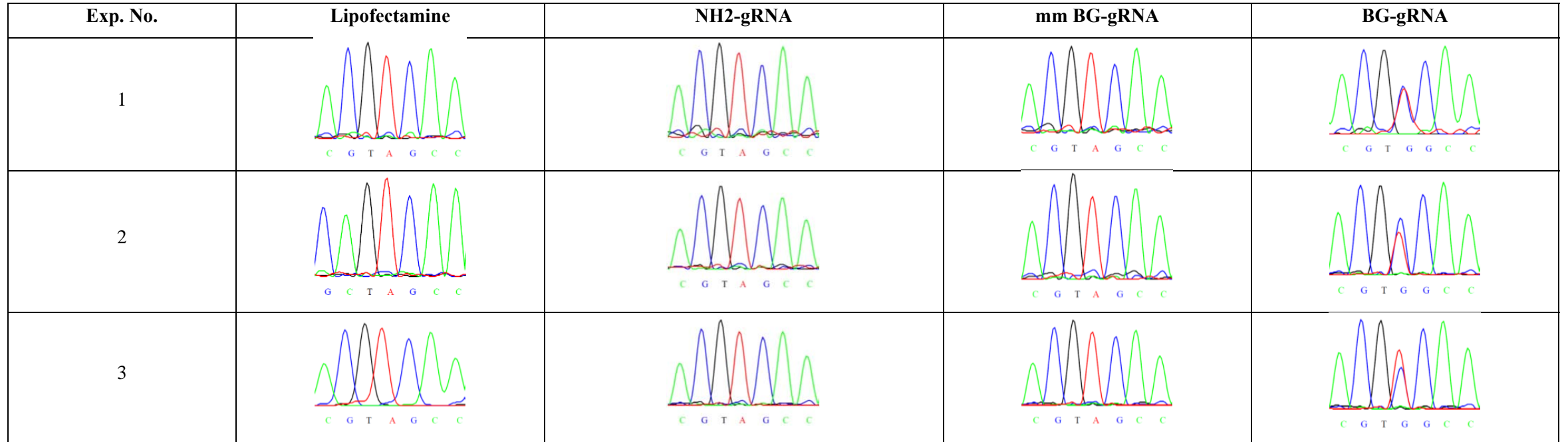
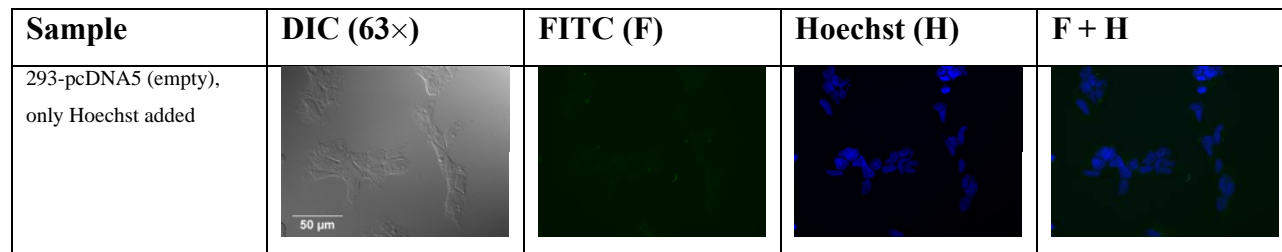


Figure S5. Localization phenotype via FITC staining in 293-Flip-In T-REx cells, mm = mismatched.




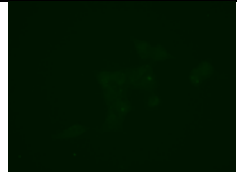
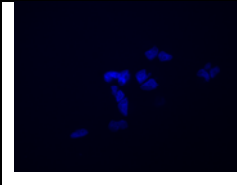
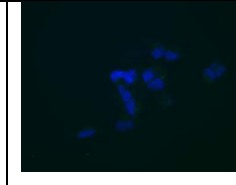

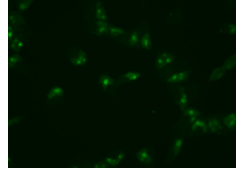
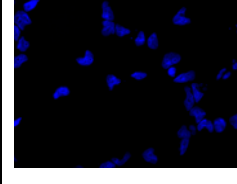
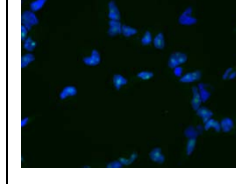
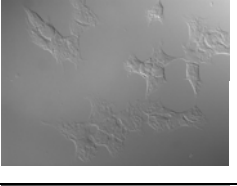
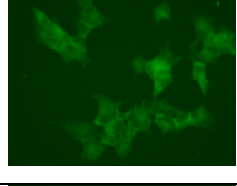
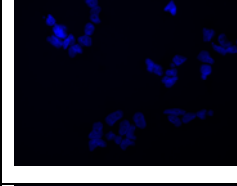
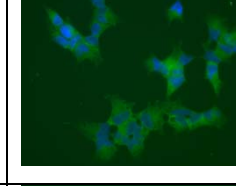

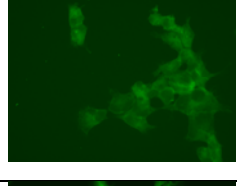
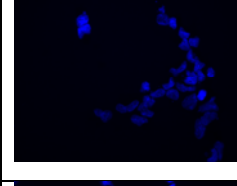
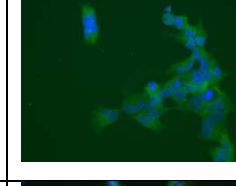
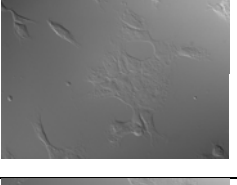
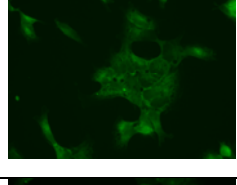
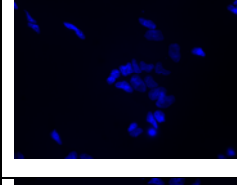
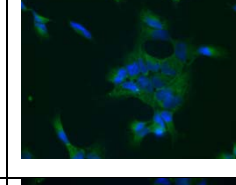
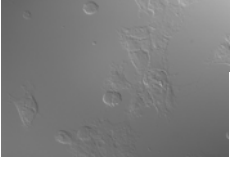
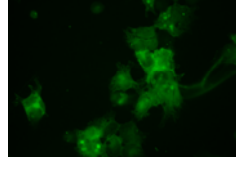
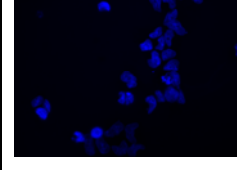
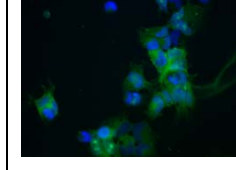
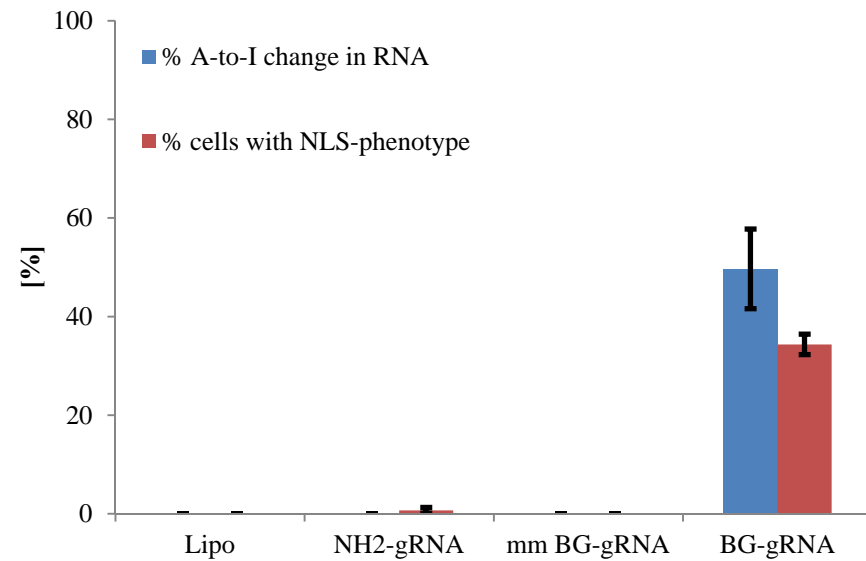
293-pcDNA5 (empty)				
293-SNAP-ADAR2-TGG-NLS				
293-SNAP-ADAR2-TAG-NLS				
293-SNAP-ADAR2-TAG-NLS + NH2-gRNA				
293-SNAP-ADAR2-TAG-NLS + mm BG-gRNA				
293-SNAP-ADAR2-TAG-NLS + BG-gRNA				

Table S3. Quantification of phenotype switch by fluorescence microscopy, cells with NLS-phenotype [%], 140-280 cells were counted for each sample per experiment, mm = mismatched.

Exp. No.	Lipo	NH2-gRNA	mm BG-gRNA	BG-gRNA
1	0	1	0	32
2	0	1	0	36
3	0	0	0	35
Average	0.00	0.67	0.00	34.33
Standard Deviation	0.00	0.58	0.00	2.08

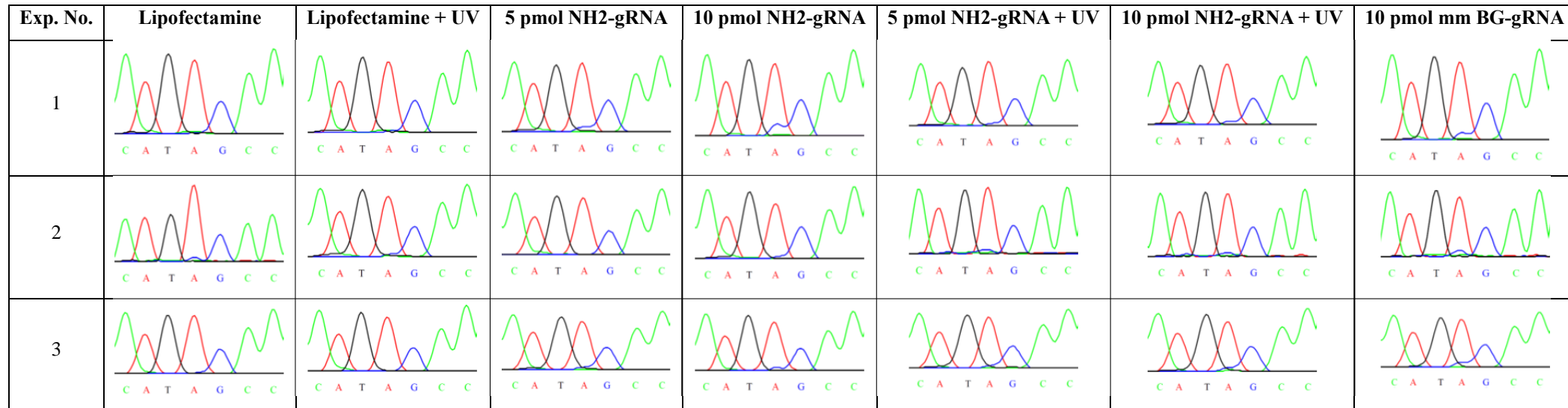
Figure S6. Comparison of change in RNA (A-to-I change in RNA) with localization switch (cells with NLS-phenotype) in 293-SNAP-ADAR-TAG-NLS cells, mm = mismatched.



Editing in the 5'UTR to switch protein localization

Transient Expression of ATAG-NLS-SNAP-ADAR2, including light control

Figure S7. Sanger sequencing of ATAG-NLS-SNAP-ADAR2 in 293T cells, guideRNA: 5 and 10 pmol/well, mm = mismatched.



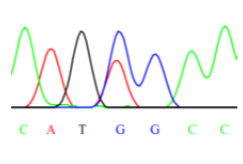
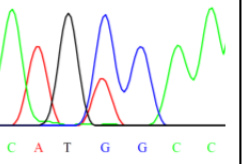
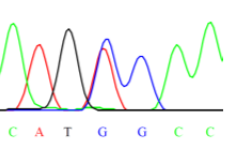
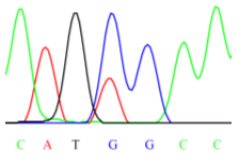
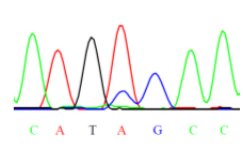
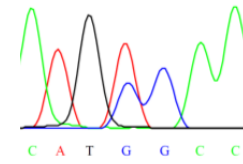
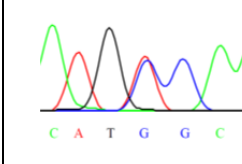
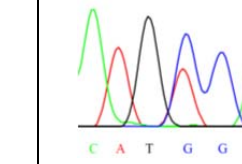
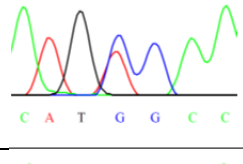
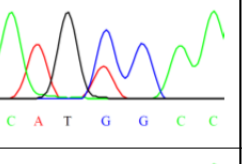
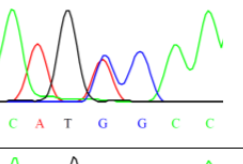
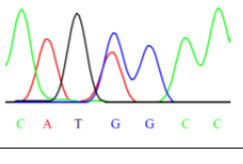
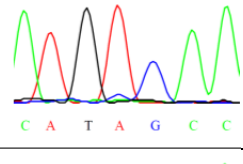
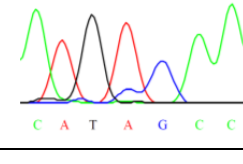
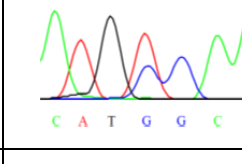
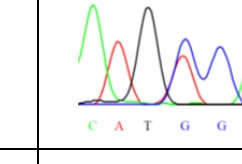
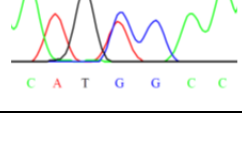
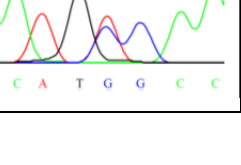
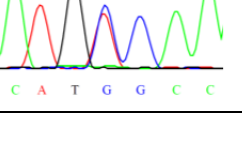
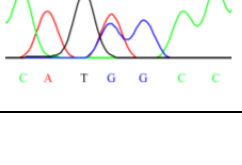
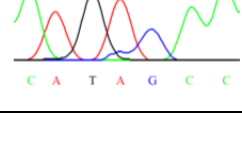
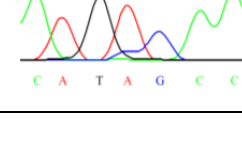
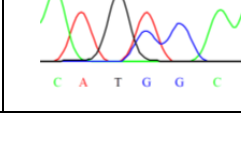
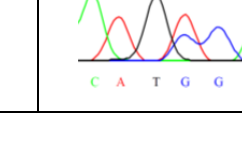
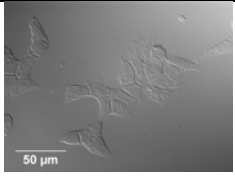
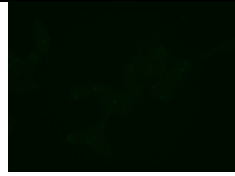
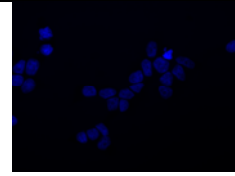

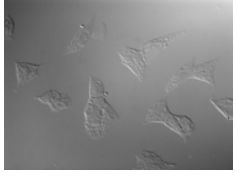

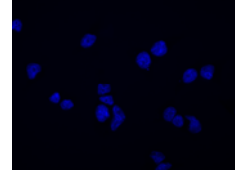
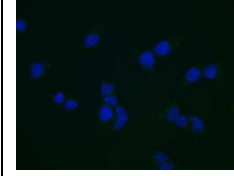

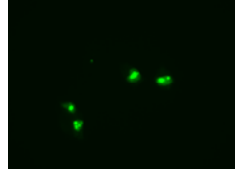


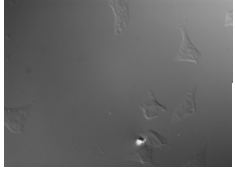
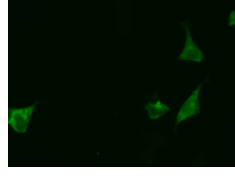

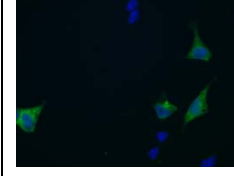
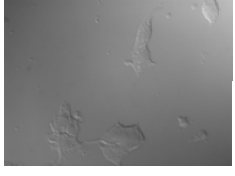
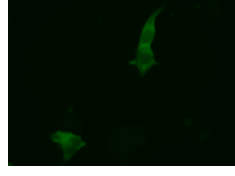
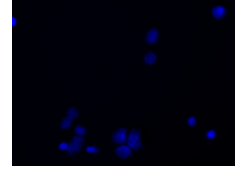
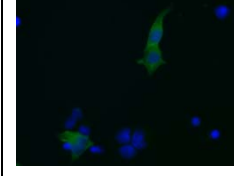
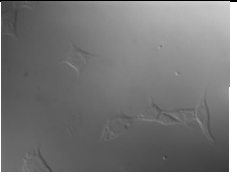
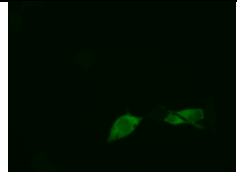
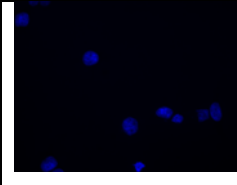
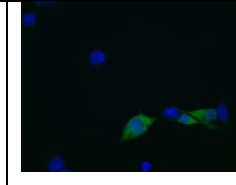
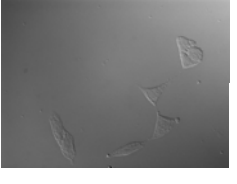
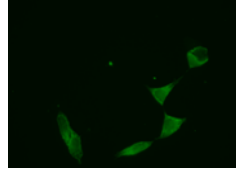
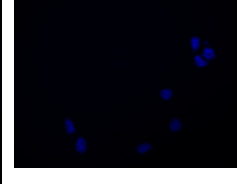
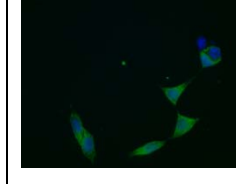
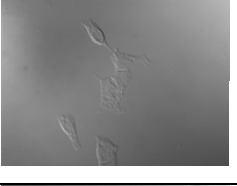
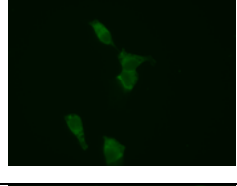
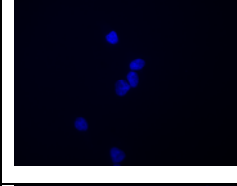
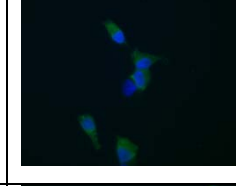
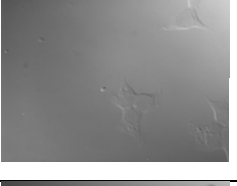
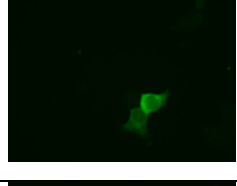
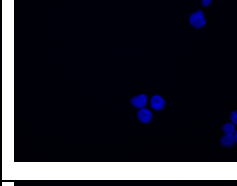
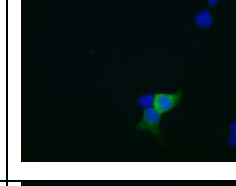
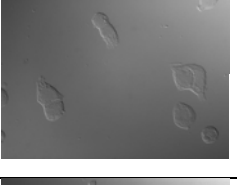
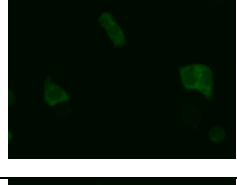
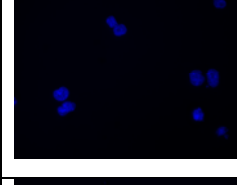
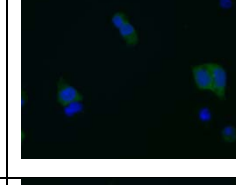
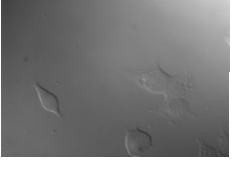
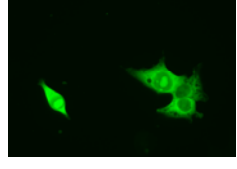
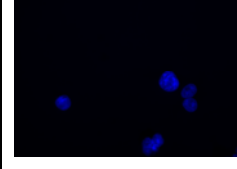
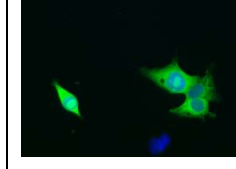
Exp. No.	5 pmol BG-gRNA	10 pmol BG-gRNA	5 pmol BG-gRNA + UV	10 pmol BG-gRNA + UV	5 pmol NpomBG-gRNA	10 pmol NpomBG-RNA	5 pmol NpomBG-gRNA + UV	10 pmol NpomBG-gRNA + UV
1								
2								
3								

Figure S8. Localization phenotype by FITC staining in transfected 293T cells, mm = mismatched.

Sample	DIC (63×)	FITC (F)	Hoechst (H)	F + H
293T + pcDNA3.1 (empty), only Hoechst added				
293T + pcDNA3.1 (empty)				
293T + ATGG-NLS- SNAP-ADAR2				
293T + ATAG-NLS- SNAP-ADAR2				
293T + ATAG-NLS- SNAP-ADAR2 + UV				

<p>293T + ATAG-NLS- SNAP-ADAR2 + 5 pmol NH2-gRNA</p>				
<p>293T + ATAG-NLS- SNAP-ADAR2 + 10 pmol NH2-gRNA</p>				
<p>293T + ATAG-NLS- SNAP-ADAR2 + 5 pmol NH2-gRNA + UV</p>				
<p>293T + ATAG-NLS- SNAP-ADAR2 + 10 pmol NH2-gRNA + UV</p>				
<p>293T + ATAG-NLS- SNAP-ADAR2 + 10 pmol mm BG-gRNA</p>				
<p>293T + ATAG-NLS- SNAP-ADAR2 + 5 pmol BG-gRNA</p>				

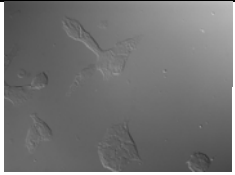
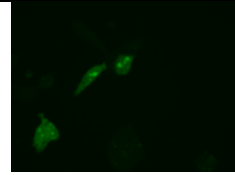
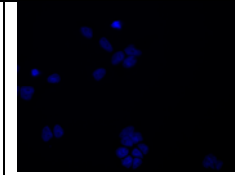
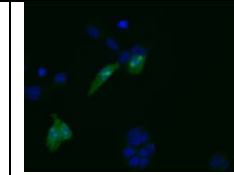
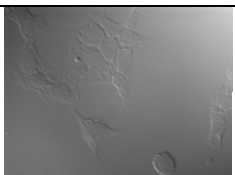
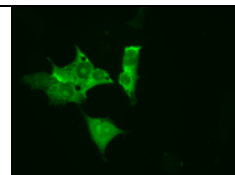
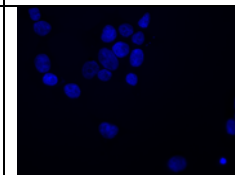
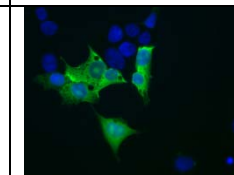
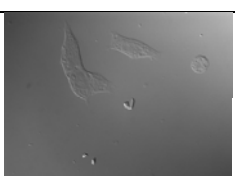
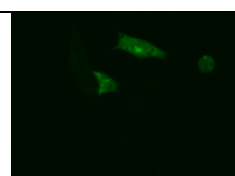
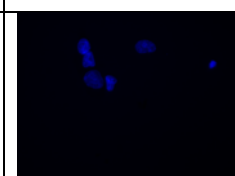

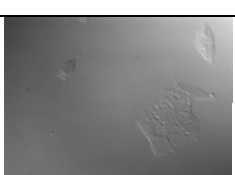
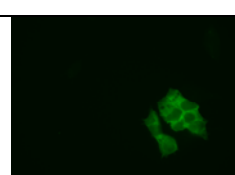
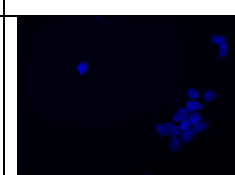
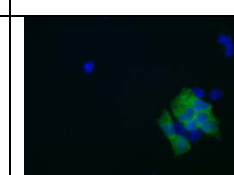
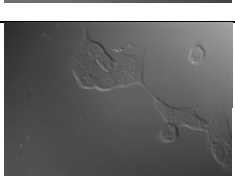
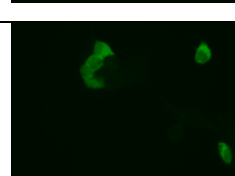
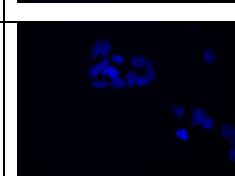
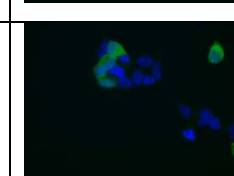
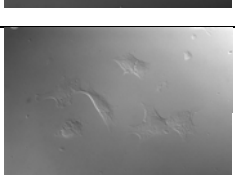
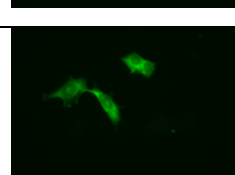
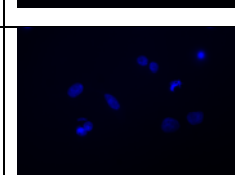
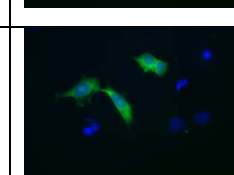
293T + ATAG-NLS- SNAP-ADAR2 + 10 pmol BG-gRNA				
293T + ATAG-NLS- SNAP-ADAR2 + 5 pmol BG-gRNA+ UV				
293T + ATAG-NLS- SNAP-ADAR2 + 10 pmol BG-gRNA + UV				
293T + ATAG-NLS- SNAP-ADAR2 + 5 pmol NpomBG-gRNA				
293T + ATAG-NLS- SNAP-ADAR2 + 10 pmol NpomBG-gRNA				
293T + ATAG-NLS- SNAP-ADAR2 + 5 pmol NpomBG-gRNA + UV				



Table S4. Quantification of phenotype switch by fluorescence microscopy, transfected cells with NLS-phenotype [%], 50-200 cells were counted for each sample per experiment, mm = mismatched.

Exp. No.	Lipo	Lipo + UV	5 pmol NH2-gRNA	10 pmol NH2-gRNA	5 pmol NH2-gRNA + UV	10 pmol NH2-gRNA + UV	10 pmol mm BG-gRNA	5 pmol BG-gRNA	10 pmol BG-gRNA	5 pmol BG-gRNA + UV	10 pmol BG-gRNA + UV	5 pmol NpomBG-gRNA	10 pmol NpomBG-gRNA	5 pmol NpomBG-gRNA + UV	10 pmol NpomBG-gRNA + UV
1	12	13	8	13	14	19	15	54	70	51	65	18	26	63	67
2	13	12	16	19	20	23	18	54	67	59	63	13	19	39	63
3	6	7	21	13	22	12	19	63	65	61	60	25	22	57	52
Average	10.33	10.67	15.00	15.00	18.67	18.00	17.33	57.00	67.33	57.00	62.67	18.67	22.33	53.00	60.67
Standard Deviation	3.79	3.21	6.56	3.46	4.16	5.57	2.08	5.20	2.52	5.29	2.52	6.03	3.51	12.49	7.77

Figure S9. Comparison of change on RNA (A-to-I change at RNA) with localization switch (cells with NLS-phenotype) in 293T cells, mm = mismatched.

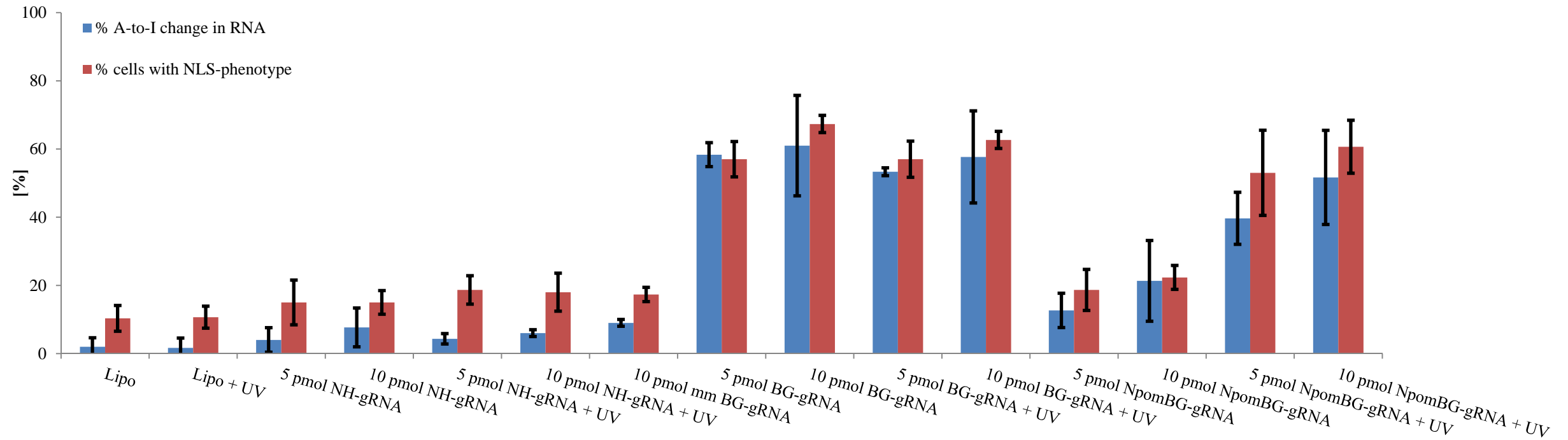
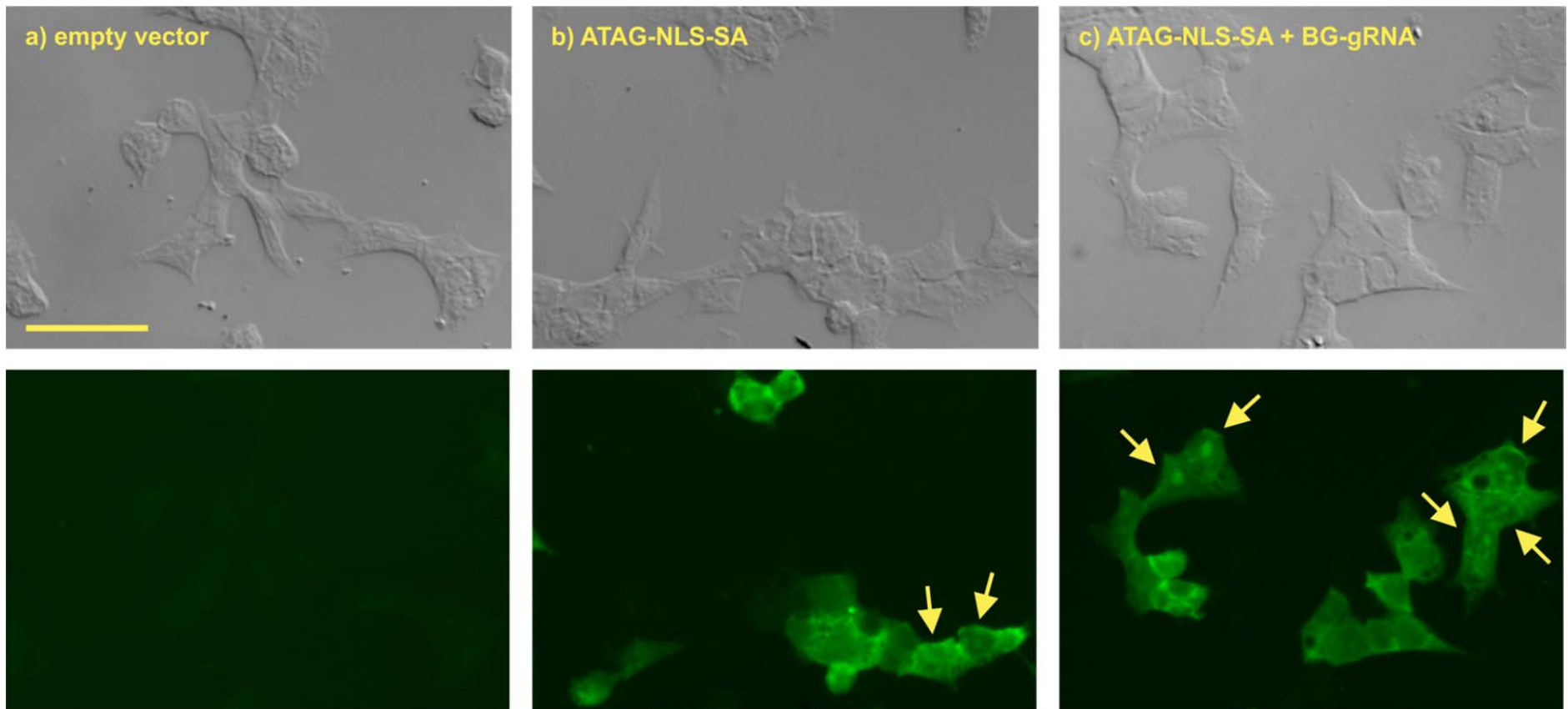


Figure S10. Example of faint nuclear staining with the ATAG-NLS-SNAP-ADAR2 construct prior to editing compared to clear staining after editing. Shown are DIC and FITC staining at 40× magnification of 293T cells transfected with a) empty vector, b) negative editing control: ATAG-NLS-SNAP-ADAR2, or c) editing: ATAG-NLS-SNAP-ADAR2 + BG-guideRNA. The arrows indicate the nuclear phenotype of the SNAP-ADAR protein. 10-15% of cells transfected with ATAG-NLS-SNAP-ADAR2 (negative editing control) show some faint nuclear staining, compared to the cells which are additionally transfected with BG-guideRNA, which show a clearly visible nuclear staining. Nevertheless, when analyzing the phenotype in the negative editing control, cells showing faint staining have already been counted as positive. The scale bar represents 50 μm.



Editing of genomically integrated ATAG-NLS-SNAP-ADAR2, including light control

Figure S11. Sanger sequencing of editing in 293-ATAG-NLS-SNAP-ADAR2 cells, 20 pmol guideRNA/well, mm = mismatched.

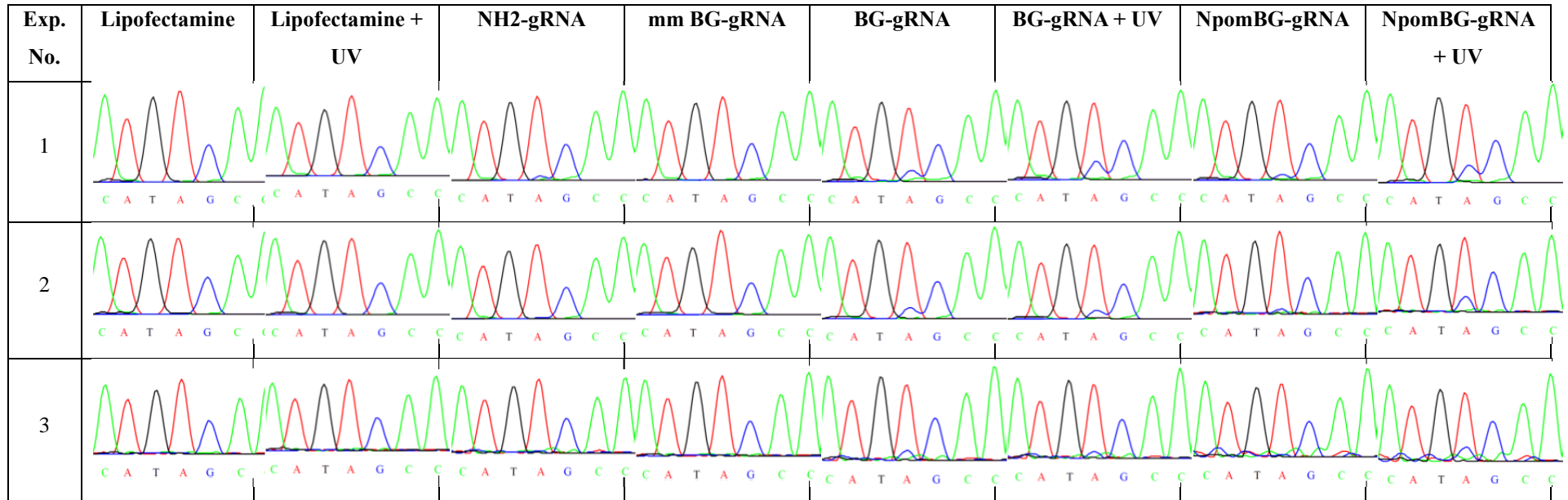
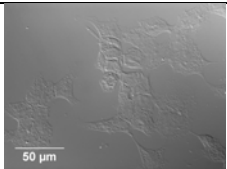
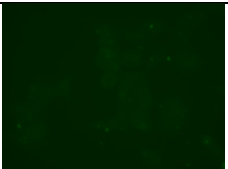
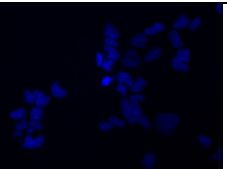
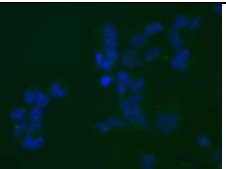
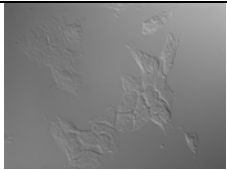
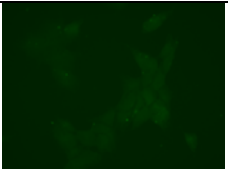
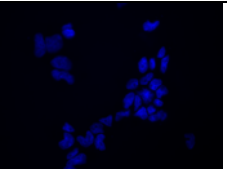
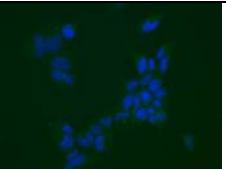
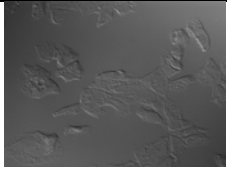
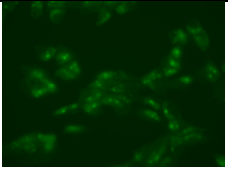
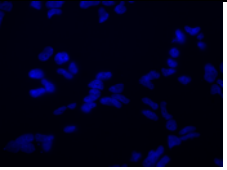
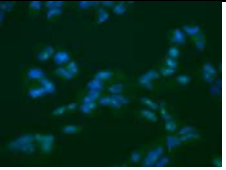
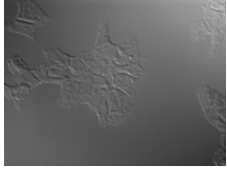
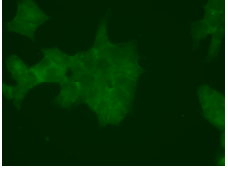
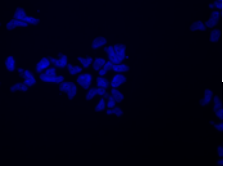
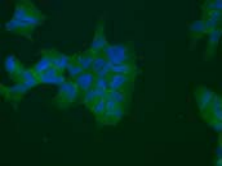
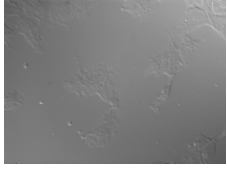
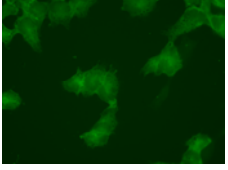
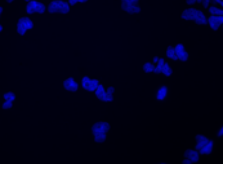
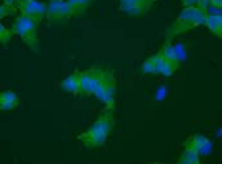
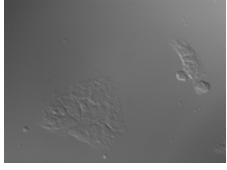
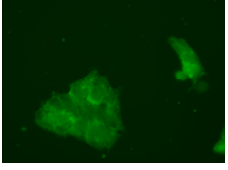
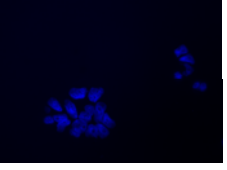
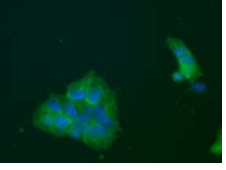


Figure S12. Localization phenotype by FITC staining of transfected 293-Flip-In T-REx cells, mm = mismatched.

Sample	DIC (63×)	FITC (F)	Hoechst (H)	F + H
293-pcDNA5 (empty), only Hoechst added				
293-pcDNA5 (empty)				
293-ATGG-NLS-SNAP- ADAR2				
293-ATAG-NLS-SNAP- ADAR2				
293-ATAG-NLS-SNAP- ADAR2 + UV				
293-ATAG-NLS-SNAP- ADAR2 + NH2-gRNA				

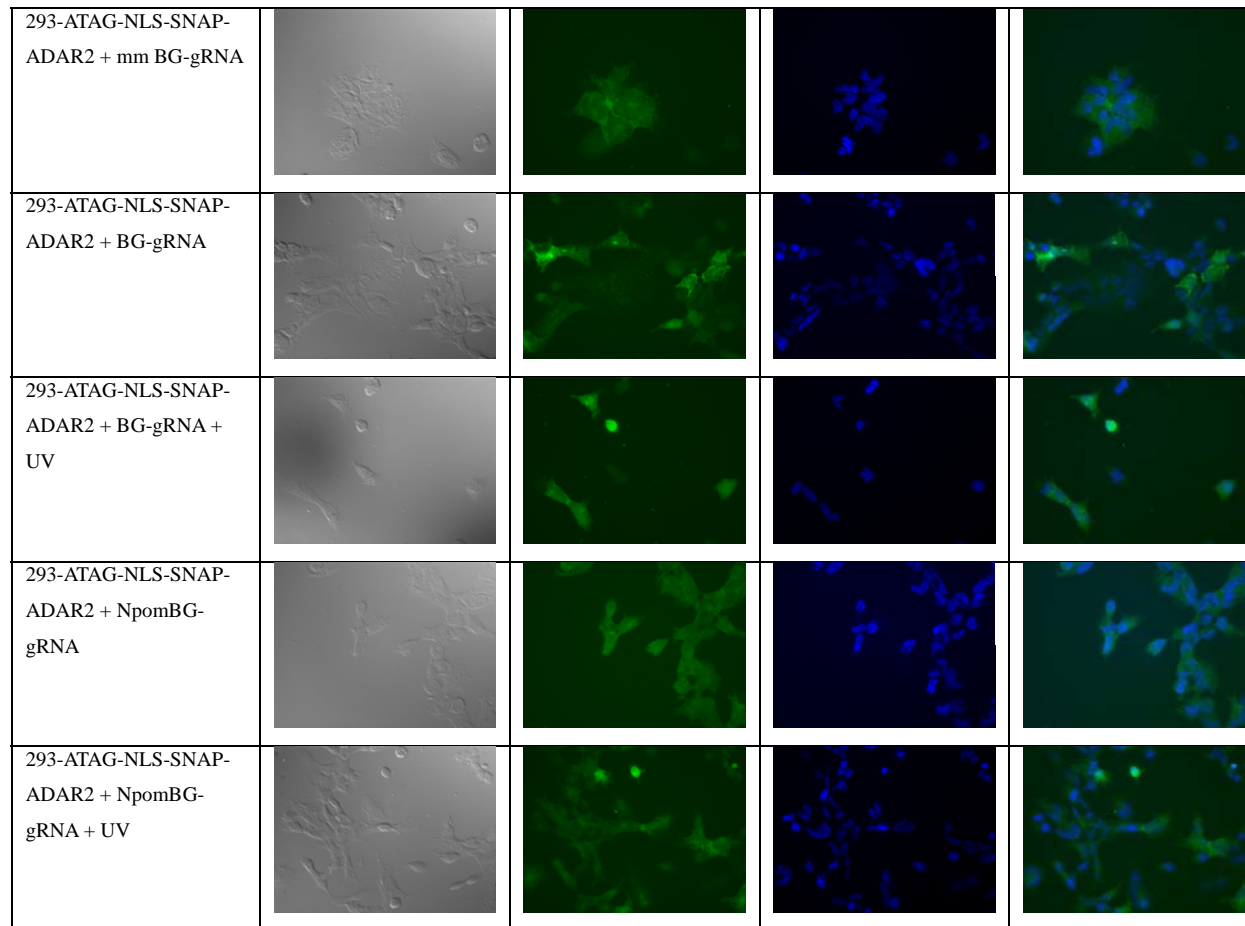
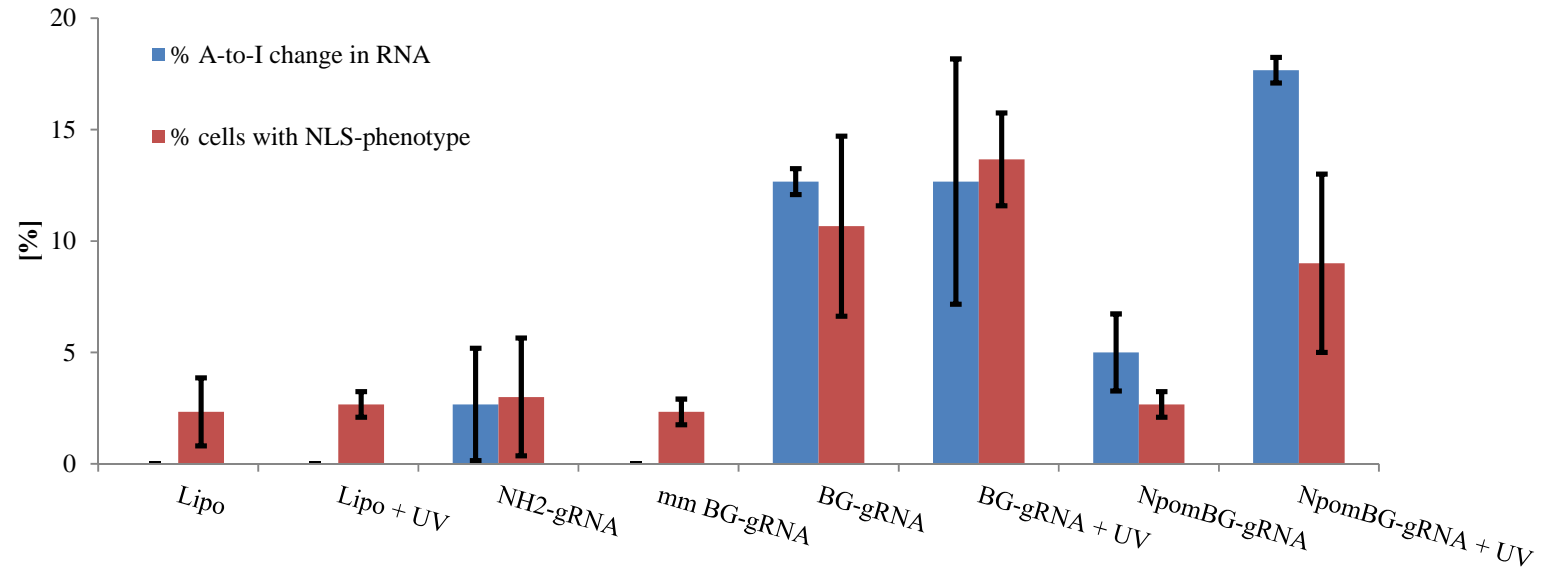


Table S5. Quantification of phenotype switch by fluorescence microscopy, cells with NLS-phenotype [%], 110-310 cells were counted for each sample per experiment, mm = mismatched.

Exp. No.	Lipo	Lipo + UV	NH2-gRNA	mm BG-gRNA	BG-gRNA	BG-gRNA + UV	NpomBG-gRNA	NpomBG-gRNA+ UV
1	1	3	6	3	13	12	3	9
2	2	2	2	2	6	13	3	5
3	4	3	1	2	13	16	2	13
Average	2.33	2.67	3.00	2.33	10.67	13.67	2.67	9.00
Standard Deviation	1.53	0.58	2.65	0.58	4.04	2.08	0.58	4.00

Figure S13. Comparison of change in RNA (A-to-I change in RNA) with localization switch (cells with NLS-phenotype) in 293-ATAG-NLS-SNAP-ADAR2 cells, mm = mismatched.



Editing under genomic expression of ATAG-NLS-SNAP-ADAR2*(E488Q), including light control

Figure S14. Sanger sequencing of editing in 293-ATAG-NLS-SNAP-ADAR2* cells, 10 pmol guideRNA/well, mm = mismatching.

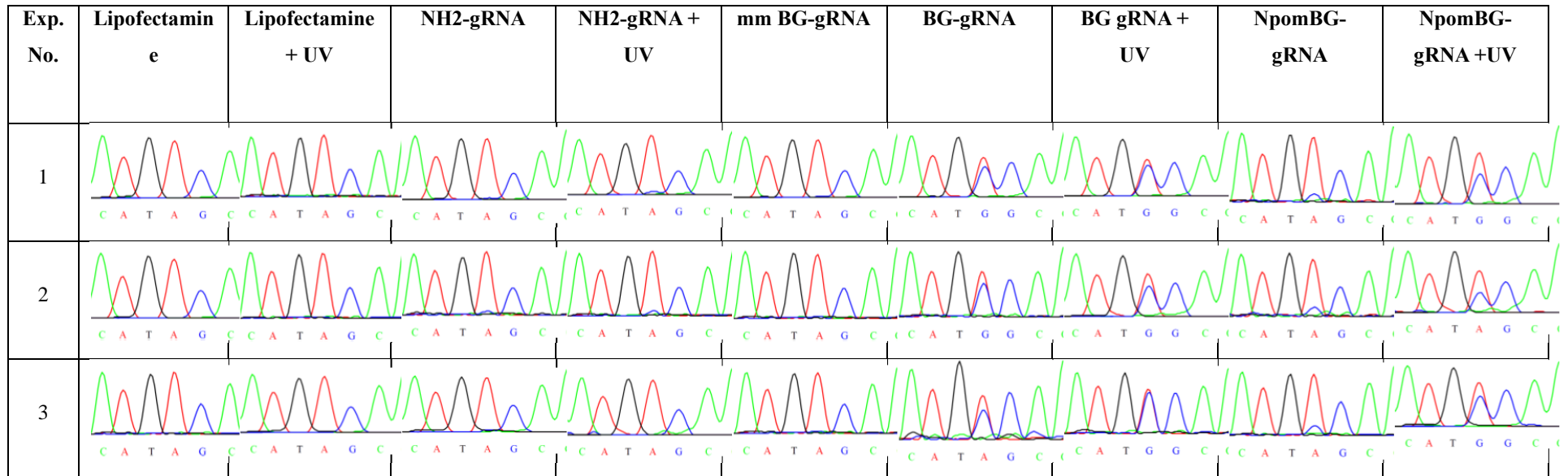
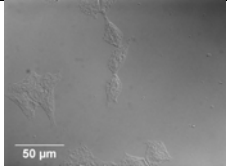
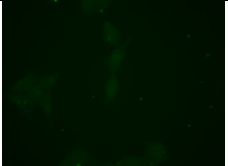
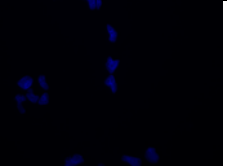
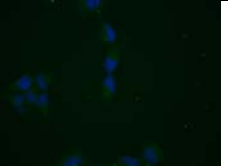
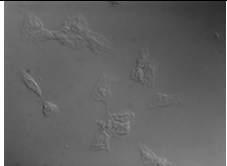
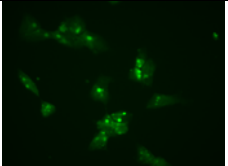
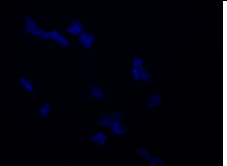
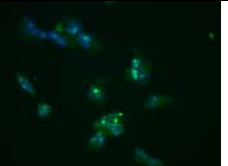

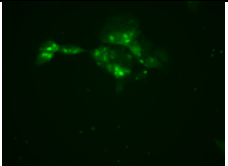

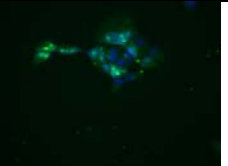
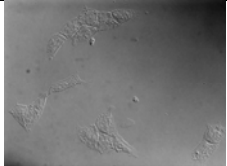
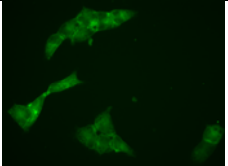
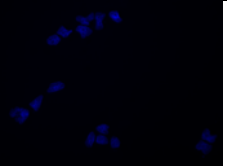
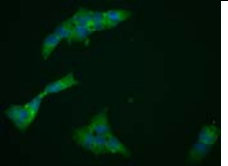

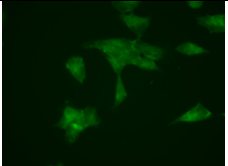
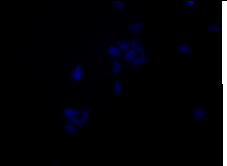
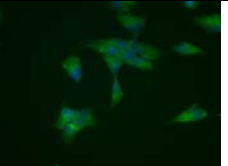
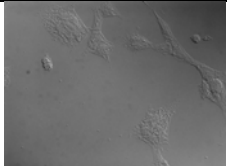
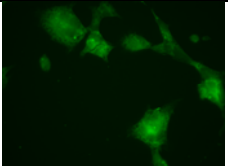
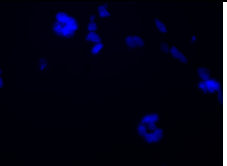
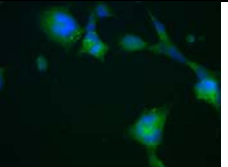


Figure S15. Localizaiton phenotype by FITC staining of transfected 293-Flip-In T-REx cells, mm = mismatched.

Sample	DIC (63×)	FITC (F)	Hoechst (H)	F + H
293-pcDNA5 (empty)				
293-ATGG-NLS-SNAP-ADAR2*				
293-ATGG-NLS-SNAP-ADAR2* + BG-guideRNA				
293-ATAG-NLS-SNAP-ADAR2*				
293-ATAG-NLS-SNAP-ADAR2* + UV				
293-ATAG-NLS-SNAP-ADAR2* + NH2-gRNA				


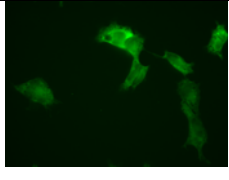
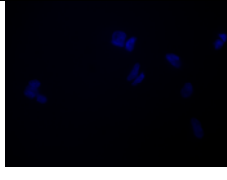
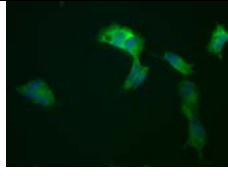

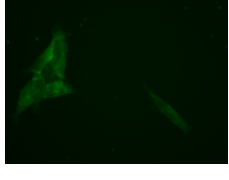
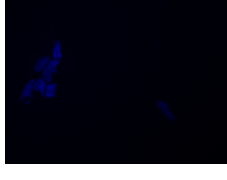
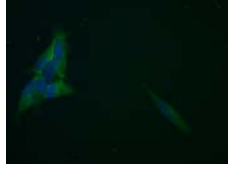

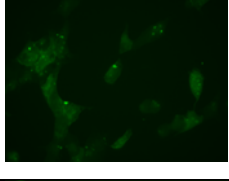
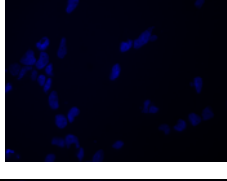
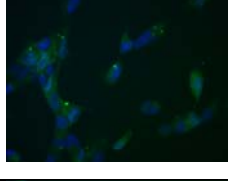

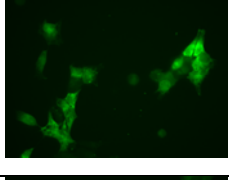
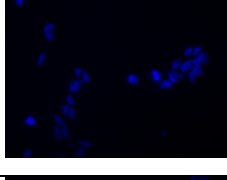
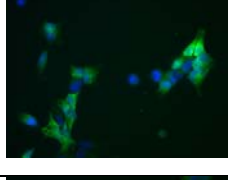

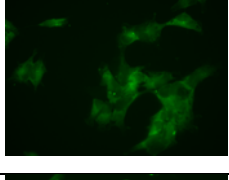
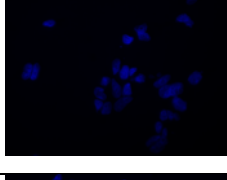
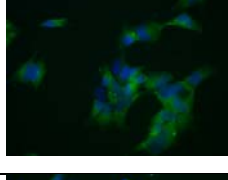
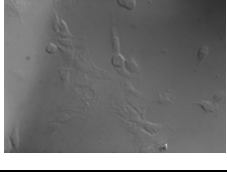
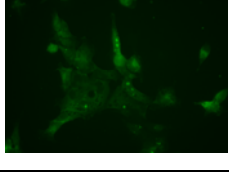
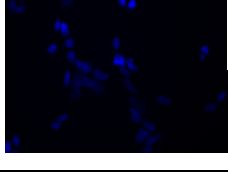
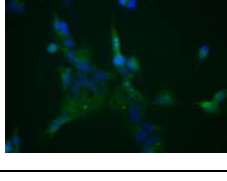
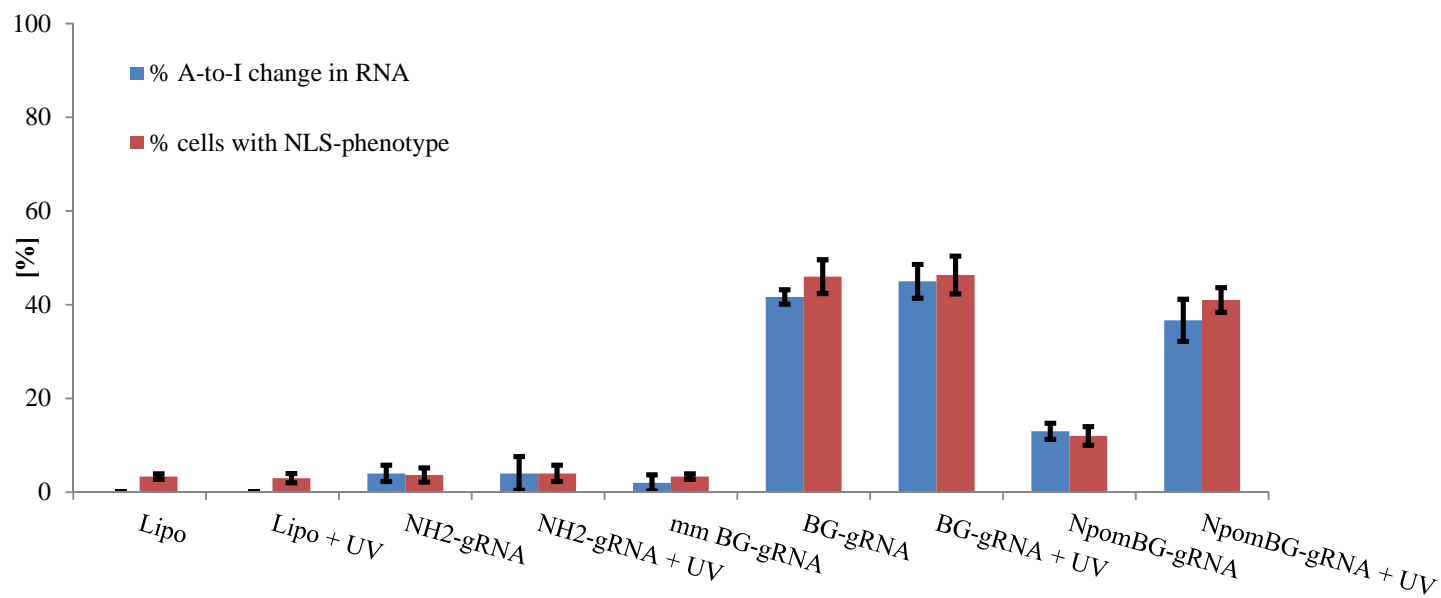
293-ATAG-NLS-SNAP-ADAR2* + NH2-gRNA + UV				
293-ATAG-NLS-SNAP-ADAR2* + mm BG-gRNA				
293-ATAG-NLS-SNAP-ADAR2* + BG-gRNA				
293-ATAG-NLS-SNAP-ADAR2* + BG-gRNA + UV				
293-ATAG-NLS-SNAP-ADAR2* + NpomBG-gRNA				
293-ATAG-NLS-SNAP-ADAR2* + NpomBG-gRNA + UV				

Table S6. Quantification of phenotype switch by fluorescence microscopy, cells with NLS-phenotype [%], 120-350 cells were counted for each sample per experiment, mm = mismatched.

Exp. No.	Lipo	Lipo + UV	NH2-gRNA	NH2-gRNA + UV	mm BG-gRNA	BG-gRNA	BG-gRNA+ UV	NpomBG-gRNA	NpomBG-gRNA + UV
1	3	3	2	2	3	47	51	10	42
2	4	2	5	5	3	49	44	14	38
3	3	4	4	5	4	42	44	12	43
Average	3.33	3.00	3.67	4.00	3.33	46.00	46.33	12.00	41.00
Standard Deviation	0.58	1.00	1.53	1.73	0.58	3.61	4.04	2.00	2.65

Figure S16. Comparison of change in RNA (A-to-I change in RNA) with localization switch (cells with NLS-phenotype) in 293-ATAG-NLS-SNAP-ADAR2*, mm = mismatched.



Switch to outer-membrane localization by 5'-UTR editing

Transient expression of ATAG-IgK-HA-eGFP, including light control

Figure S17. Sanger sequencing of ATAG-IgK-HA-eGFP in 293T cells co-transfected with SNAP-ADAR2-BFP, and guideRNA 2.5 pmol/well, mm = mismatched.

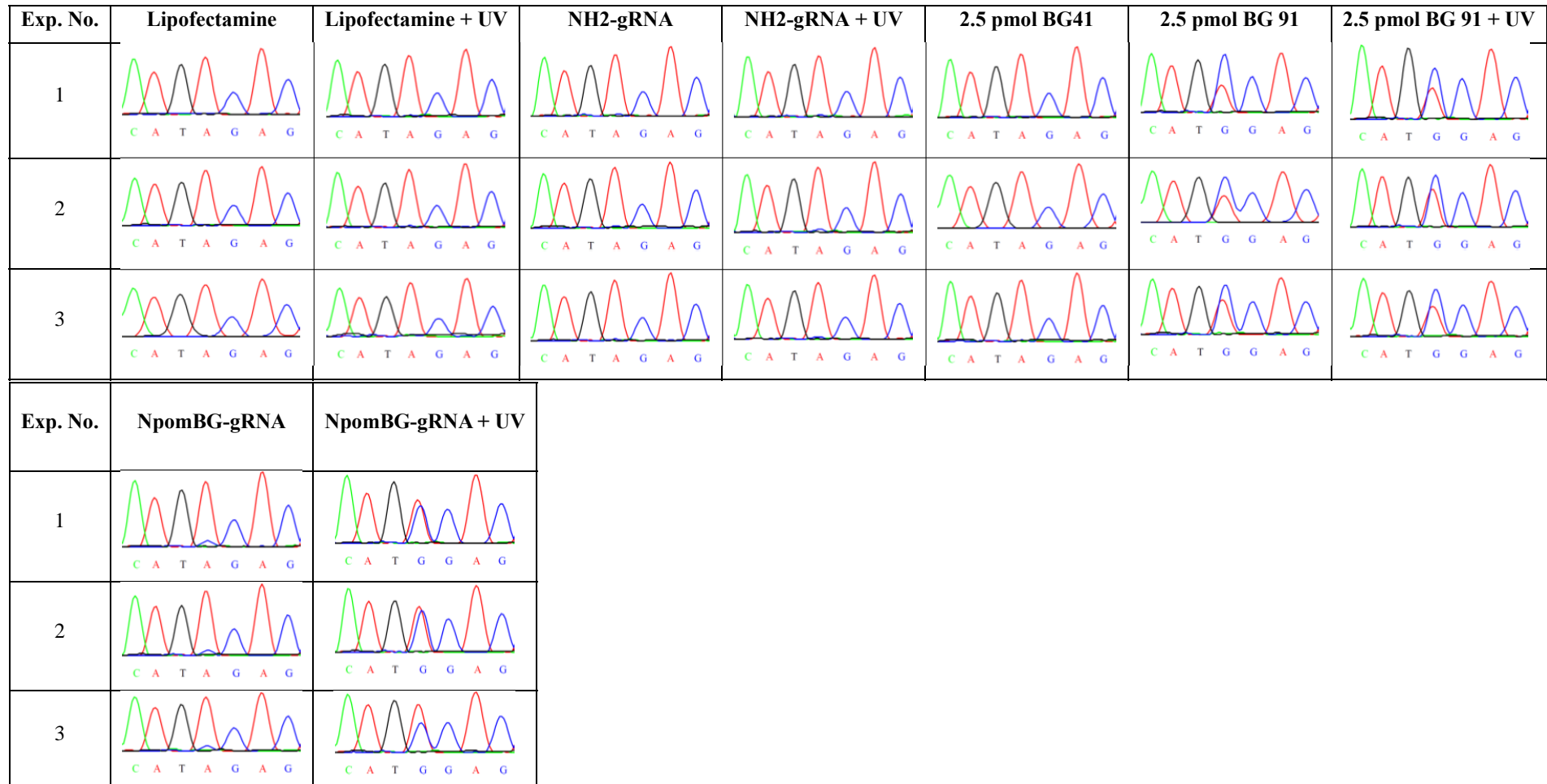


Figure S18. Additional control: Editing depends on co-transfection of SNAP-ADAR2. Sanger sequencing of ATAG-IgK-HA-eGFP in 293T cells co-transfected with empty vector, and guideRNA 2.5 pmol/well.

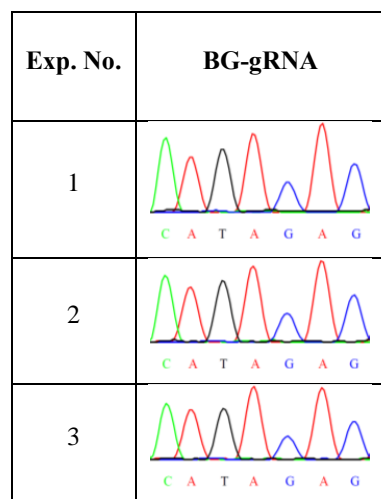
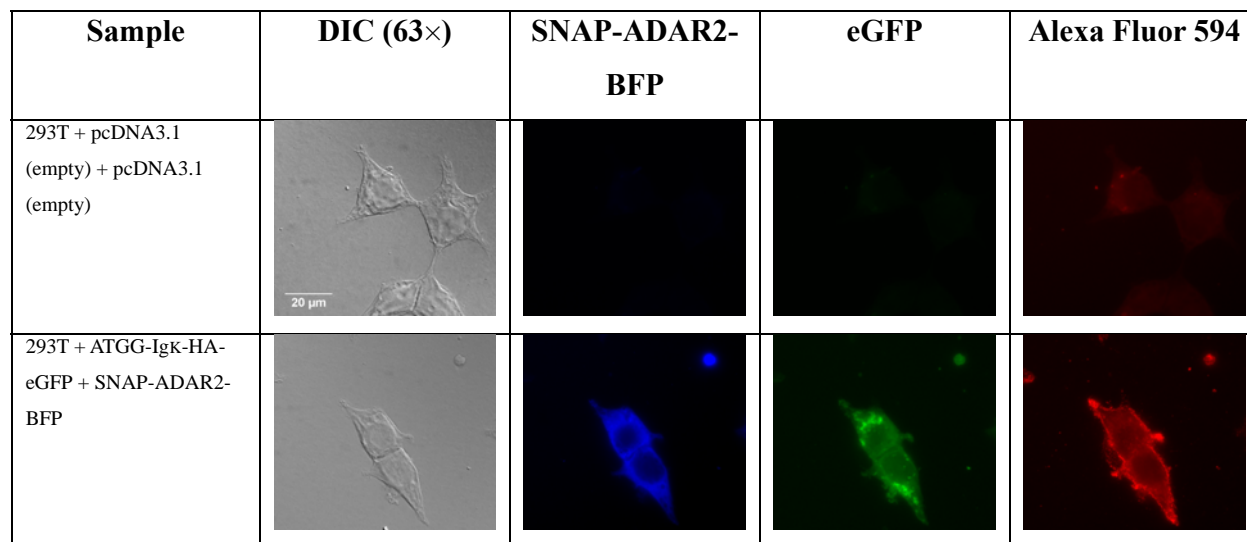
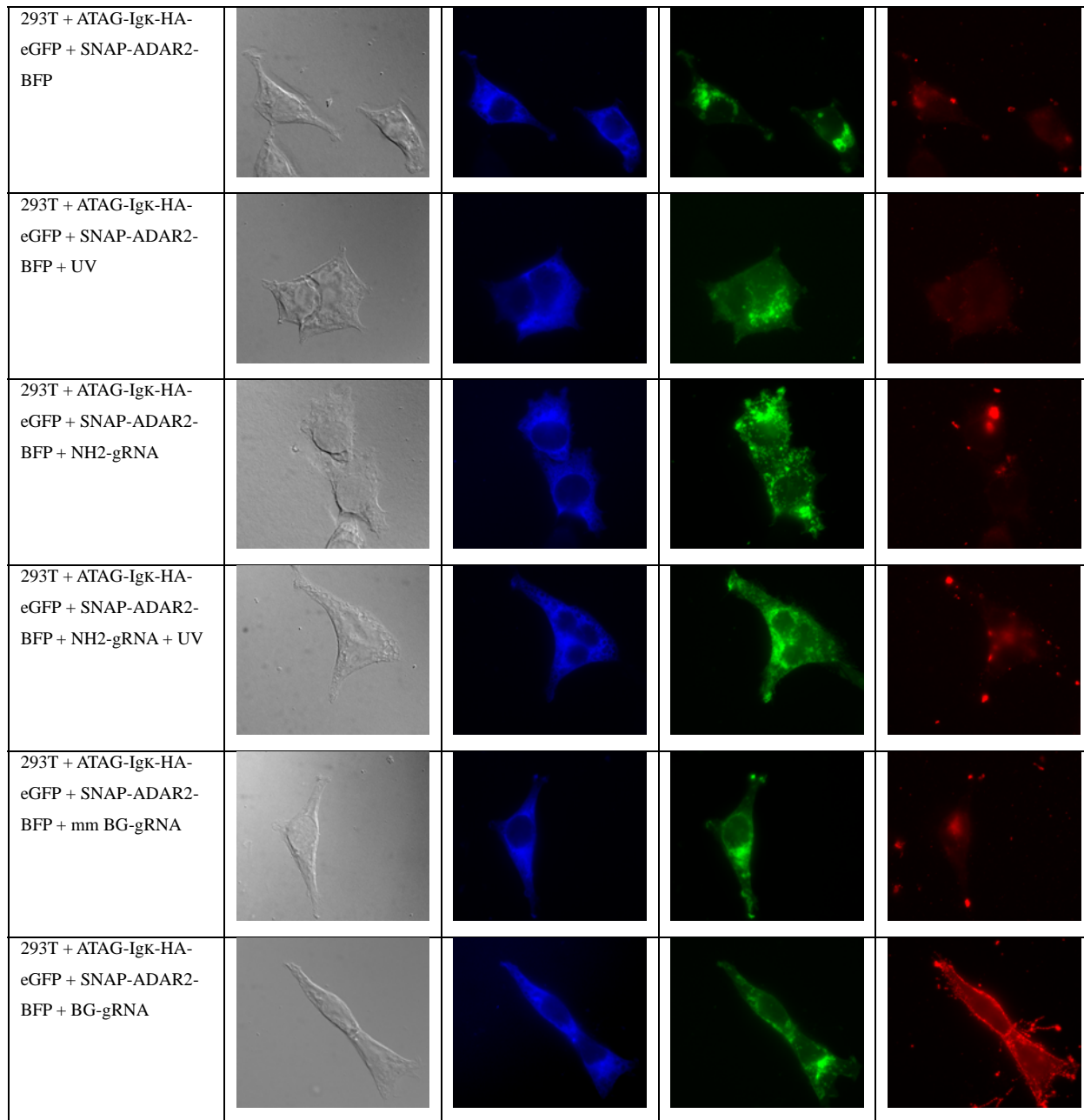


Figure S19. Localization phenotype via immunostaining in transfected 293T cells, mm = mismatched.





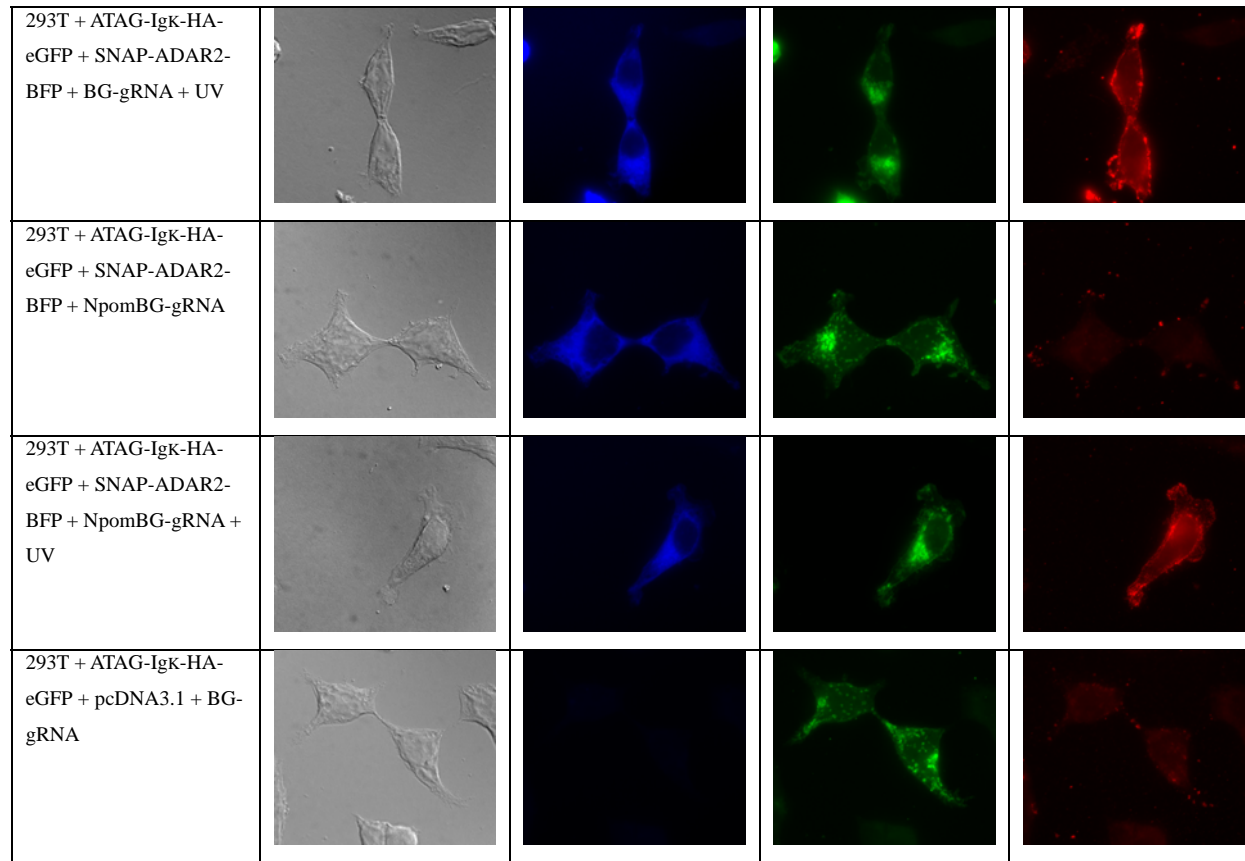


Table S7. Quantification of phenotype switch by fluorescence microscopy, transfected cells with membrane phenotype [%], 50-90 cells were counted for each sample per experiment, mm = mismatched.

Changed phenotype [%]	Lipo	Lipo + UV	NH2-gRNA	NH2-gRNA +UV	mm BG-gRNA	BG-gRNA	BG-gRNA + UV	NpomBG-gRNA	NpomBG-gRNA + UV
1	0	0	0	0	0	44	43	3	32
2	0	0	0	0	0	41	41	2	34
3	0	0	0	0	0	43	42	0	28
AVERAGE	0.00	0.00	0.00	0.00	0.00	42.67	42.00	1.67	31.33
Standard Deviation	0.00	0.00	0.00	0.00	0.00	1.53	1.00	1.53	3.06

Figure S20. Comparison of change in RNA (A-to-I change in RNA) with localization switch (cells with membrane phenotype) in 293T cells, mm = mismatched.

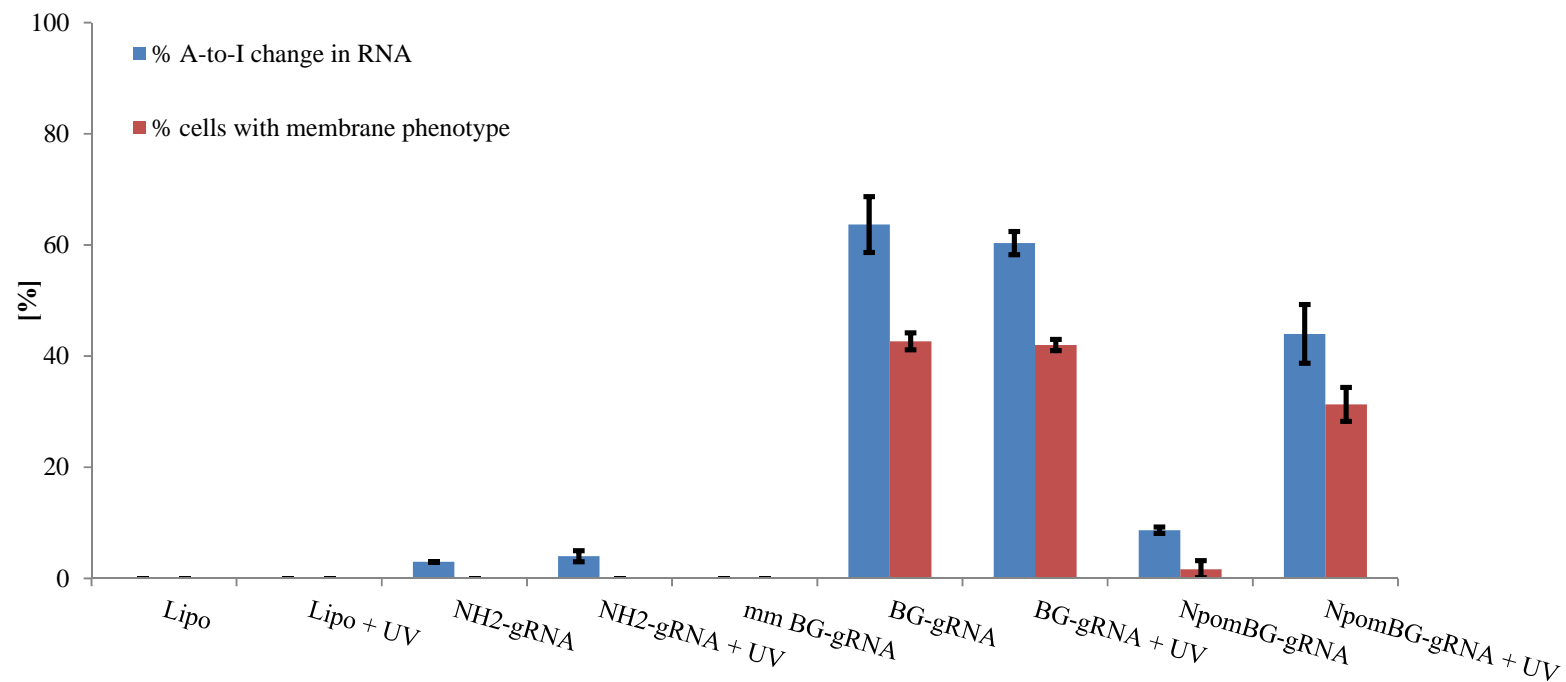
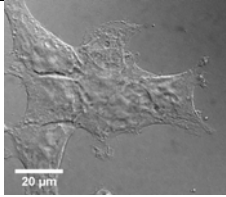
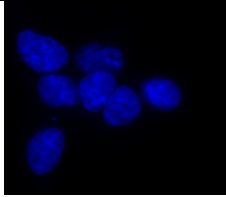
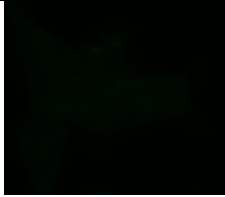

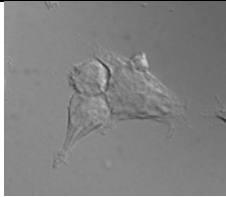
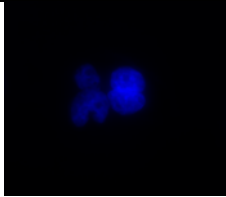
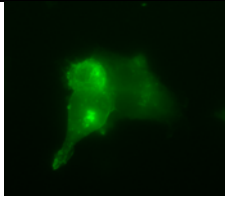
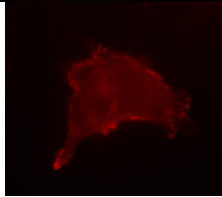

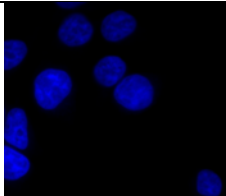
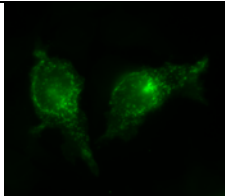
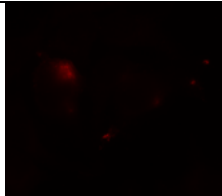
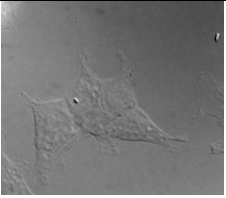
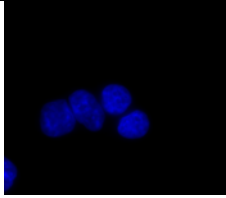
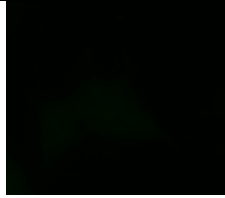

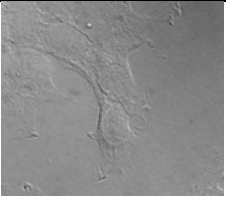
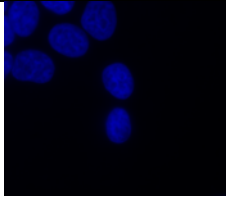
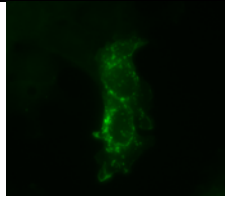
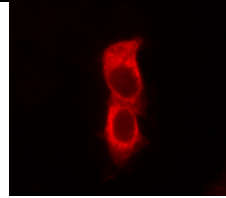
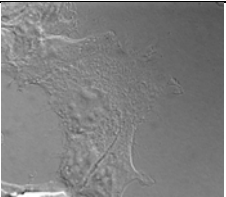
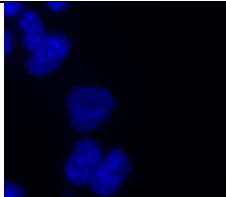
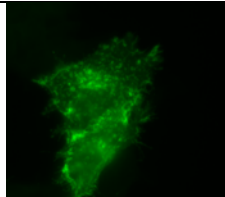
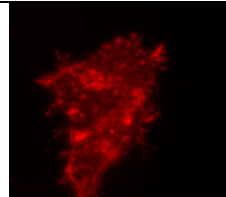


Figure S21. Effect of permeabilization on the detection of HA-IgK-HA-eGFP via GFP fluorescence and anti-HA immunostain. Cells were transfected with empty vector, ATGG-Igk-HA-eGFP, or ATAG-Igk-HA-eGFP. 24h after transfection, cells were plated on cover slips and incubated for 1 d. Immunostaining was performed as described above. In case of permeabilization, cells were incubated in 500 μ l of 1% Triton X-100/PBS for 5 min before blocking. Additionally, cells were incubated in 200 μ l Hoechst solution (1:100 in PBS) before mounting. Shown are DIC images (63 \times magnification), nuclear staining by Hoechst 33342 (blue), GFP fluorescence (green) and HA-immunostain with Alexa Fluor 594 (red). Without permeabilization, the anti-HA antibody stains only the outer-membrane-bound form, not the cytosolic one. However, after permeabilization also the Ha-tagged protein in the cytoplasm is stained. This is supported by the GFP channel.

	Sample	DIC (63 \times)	Hoechst	eGFP	Alexa Fluor 594
Without permeabilization	293T + pcDNA3.1 (empty)				
	293T + ATGG-Igk-HA-eGFP				
	293T + ATAG-Igk-HA-eGFP				

With permeabilization	293T + pcDNA3.1 (empty)				
	293T + ATGG-Igk-HA-eGFP				
	293T + ATAG-Igk-HA-eGFP				

Man. 4: accepted

Vogel, P., Moschref, M., Li, Q., Merkle, T., Selvasaravanan, K. D., Li, J. B. & Stafforst, T. Efficient and precise editing of endogenous transcripts with SNAP-tagged ADARs. *Nat. Methods* **15**, 535-538 (2018).

Efficient and precise editing of endogenous transcripts with SNAP-tagged ADARs

Paul Vogel¹, Matin Moschref¹, Qin Li², Tobias Merkle¹, Karthika D. Selvasaravanan¹, Jin Billy Li² and Thorsten Stafforst^{1*}

Molecular tools that target RNA at specific sites allow recoding of RNA information and processing. SNAP-tagged deaminases guided by a chemically stabilized guide RNA can edit targeted adenosine to inosine in several endogenous transcripts simultaneously, with high efficiency (up to 90%), high potency, sufficient editing duration, and high precision. We used adenosine deaminases acting on RNA (ADARs) fused to SNAP-tag for the efficient and concurrent editing of two disease-relevant signaling transcripts, KRAS and STAT1. We also demonstrate improved performance compared with that of the recently described Cas13b-ADAR.

Tools for efficient and precise RNA manipulation are highly desired¹. We recently introduced SNAP-tagged ADARs, which can be used to substitute adenosine by inosine in RNA in a rational and programmable way with a guide RNA (gRNA)^{2,3} (Supplementary Fig. 1). Because inosine is interpreted as guanosine, RNA editing can alter splicing, start and stop codons, and microRNA action, and can reprogram the protein product⁴. Manipulation at the RNA level is tunable in yield and reversible in time. This might be particularly useful for substitutions that are either lethal or compensated when introduced at the DNA level⁵, for example, in signaling proteins⁶. A further advantage is safety, as off-site RNA editing can be considered reversible. Current methods^{7–9} typically apply overexpression of (engineered) deaminases, which may result in massive global off-target editing. In contrast, the deaminase and gRNA are covalently linked in our SNAP-ADAR approach, which allows for efficient RNA-targeting after single-copy, genomic integration of the editase.

We validated four editases: SNAP-ADAR1 (SA1) and SNAP-ADAR2 (SA2)², and their respective hyperactive E>Q variants¹⁰ SA1Q and SA2Q. We initiated editing by transfection of a short, chemically stabilized benzylguanine-modified gRNA (BG-gRNA) (Supplementary Fig. 1), and analyzed the results for formal A-to-G conversion in cDNA at specific 5'-UAG triplets in the 3' untranslated regions (UTRs) of the four targeted endogenous mRNAs: *ACTB*, *GAPDH*, *GUSB*, and *SAI/2*. For both wild-type enzymes (SA1 and SA2), editing yields of 40–80% were achieved (Fig. 1a), depending on the target. Application of the hyperactive mutants (SA1Q and SA2Q) raised the yields to 65–90%; in particular, the weaker edited transcripts *GUSB* and *SAI/2* profited from this application. The maximum editing yield (80–90%) was nearly achieved 3 h after transfection (Fig. 1b), remained constant for 3 d, and then declined slowly, probably as a result of dilution of the gRNA-enzyme conjugate by cell division. The activated enzymes (SA1Q and SA2Q) were up to 12-fold more potent than the wild-type enzymes (SA1 and SA2), achieving the half-maximum editing yield at concentrations of 0.15 pmol per well, compared with 1–2 pmol per well for the

wild type (Fig. 1c). We tested the concurrent editing of all four transcripts by cotransfection of four gRNAs. Notably, the yields stayed unchanged (Fig. 1a). We obtained similar results for the concurrent editing of three sites in the *GAPDH* mRNA (Supplementary Fig. 2). Editing yields were higher in the 3'-UTR than in the open reading frame (ORF) and 5'-UTR (Fig. 1d), probably because of interference with translation. Accordingly, the faster enzymes (SA1Q and SA2Q) boosted the yields from 25–50% to 60–75% in the 5'-UTR and from 15–60% to 50–85% in the ORF (Fig. 1d). Furthermore, translation inhibition with puromycin increased ORF editing in SA1/2⁺ cells to the level of 3'-UTR editing (Supplementary Fig. 3). To assess the codon scope, we targeted all 16 conceivable 5'-NAN triplets in the ORF of endogenous *GAPDH* for SA1Q and SA2Q. We obtained yields ranging from very little to almost quantitative, reflecting the well-known preferences of ADARs^{10,11} (Fig. 1e). Although editing was generally difficult for 5'-GAN triplets (<30%), we obtained significant yields (>50%) for 10/16 triplets. For 7/16 triplets, we obtained excellent editing yields (>70%) for at least one enzyme.

An important aspect is specificity. A major advantage of our strategy² (compared with others^{7–9,12–14}) is the suppression of off-site editing within the gRNA-mRNA duplex by chemical modification of our gRNA. Only for adenosine-rich triplets (AAC, AAA, UAA, and CAA) did we detect some off-target editing, mainly with SA2Q (5–75%) and mainly for the CAA triplet (Fig. 2a, left). Off-target editing was due to three natural nucleotides in the gRNA opposite the targeted adenosine² (Supplementary Fig. 4). Careful inclusion of further chemical modifications (2'-methoxy, 2'-fluoro; Fig. 2a, right) restricted off-target editing at the CAA triplet to 20% and limited off-target editing at all other sites to <10% without reducing on-target editing. Notably, for AAA, the additional modification even elevated the on-target yield from 40% to 50%. Global off-target editing is the main obstacle for RNA editing, in particular with overexpression of editases^{9,12,13,15}. To test this for SNAP-ADARs under genomic expression, we conducted deep RNA-seq when editing the *ACTB* transcript. We also assessed the role of gRNA-dependent misguiding. The wild-type enzymes (SA1/2) were extremely precise. Among the 50,000 editing sites called (Supplementary Data), only very few were significantly differently edited compared with the negative control (6 for SA1, 30 for SA2; Fig. 2b). Most of these sites are known¹⁶ sites in the 3'-UTRs (Table 1) and were edited less than 25% (Supplementary Fig. 5a). For SA1, there was a single nonsynonymous edit (*TMX3*; 10%) that was gRNA dependent (Supplementary Table 1). For SA2, there were two nonsynonymous edits (*AAGAB*, 42%; *CHFR*, 32%), with the former being gRNA dependent. Off-targets were much more frequent with the hyperactive enzymes (835/1,310 sites for SA1Q/SA2Q; Table 1,

¹Interfaculty Institute of Biochemistry, University of Tübingen, Tübingen, Germany. ²Department of Genetics, Stanford University, Stanford, CA, USA. *e-mail: thorsten.stafforst@uni-tuebingen.de

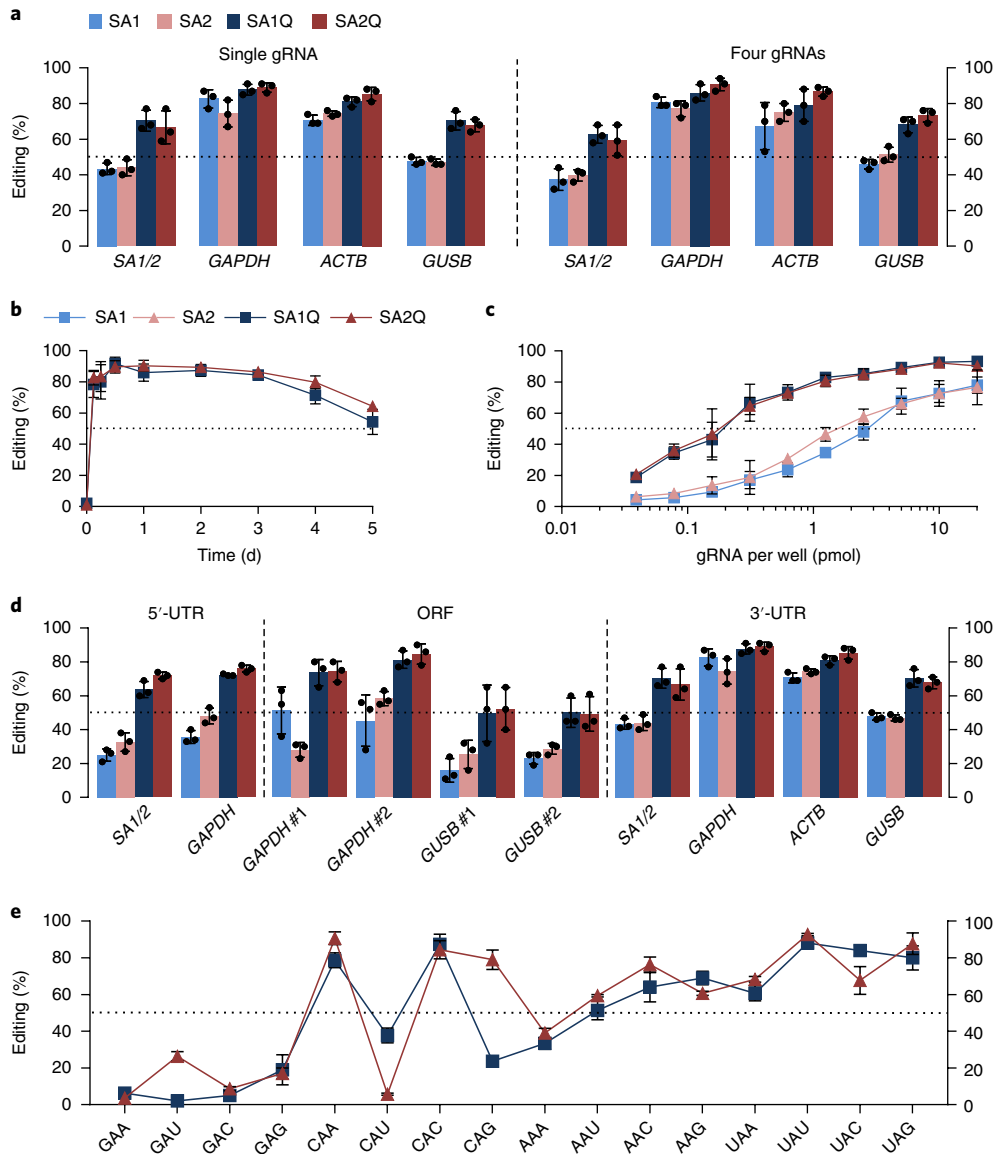


Fig. 1 | Editing performance of four SNAP-ADARs. **a**, Engineered 293 cell lines expressing one of four editase enzymes (see key) were transfected with either a single gRNA or four gRNAs targeting 5'-UAG triplets in the indicated endogenous transcripts. **b,c**, Time (**b**) and dose (**c**) dependency of editing in the *GAPDH* transcript. **d**, Editing of 5'-UAG sites in various transcripts; the plot shows 5'-UTR editing versus ORF and 3'-UTR editing. **e**, Comparative editing of all 16 triplets (5'-NAN) in the ORF of the endogenous *GAPDH* transcript. All data are shown as the mean \pm s.d.; $n=3$ independent experiments; black dots in **a,d** represent individual data points.

Fig. 2b), were caused by the free-floating enzyme, and comprised mainly novel sites (74–85%). Only a small number of sites were edited in a gRNA-dependent manner (~ 30 sites for each editase; Fig. 2c). A vast amount of sites were located in the ORF (347–496 sites) and gave rise to nonsynonymous editing (230–347 sites). However, none of the nonsynonymous editing exceeded that at the target site, and the majority of these edits occurred at a low level. This was particularly true for SA1Q, where only 4 of 227 sites were edited more than 50%, and 167 of 227 sites were edited less than 25% (Fig. 2d). For SA2Q, however, the average editing level was higher, with 20/344 sites above 50% and 240/344 below 25% editing yield. We found SA1Q and SA2Q to share only 414 of their off-target sites. SA1Q and SA2Q differ in their off-target codon preferences, with SA2Q accepting 5'-CAN triplets better (Supplementary Fig. 5b). All SNAP-ADAR cell lines behaved indistinguishably from normal 293 cells with respect to doubling times and morphology, and analysis

of the number of fragments per kilobase of transcript per million mapped reads (FPKM) revealed no difference in gene expression due to the presence of (off-target) editing activity (Supplementary Fig. 6). Because SA1(Q) showed the best balance of efficiency and specificity, we continued with that editase.

RNA editing would be particularly attractive for the manipulation of signaling networks. Also, the editable codons (5'-UAG, 5'-UAC, 5'-UAU, 5'-UAA, and 5'-AAG) indicate amino acid substitutions (Thr-to-Ala, Tyr-to-Cys, Ser-to-Gly, and Lys-to-Arg; Supplementary Fig. 5c) suitable for the manipulation of signaling proteins. For illustration, we edited two 5'-UAG sites in *KRAS* mRNA (sites 1 and 2) and the Tyr701 site (5'-UAU) in *STAT1* mRNA, its most relevant phosphorylation site¹⁷ for signaling. With SA1Q, we achieved editing levels of $55\% \pm 8\%$ (*KRAS* site 1), $46\% \pm 2\%$ (*KRAS* site 2), and $76\% \pm 6\%$ (*STAT1*) (Fig. 2e). We found no detectable off-target editing in the gRNA-mRNA duplex (Supplementary Fig. 7).

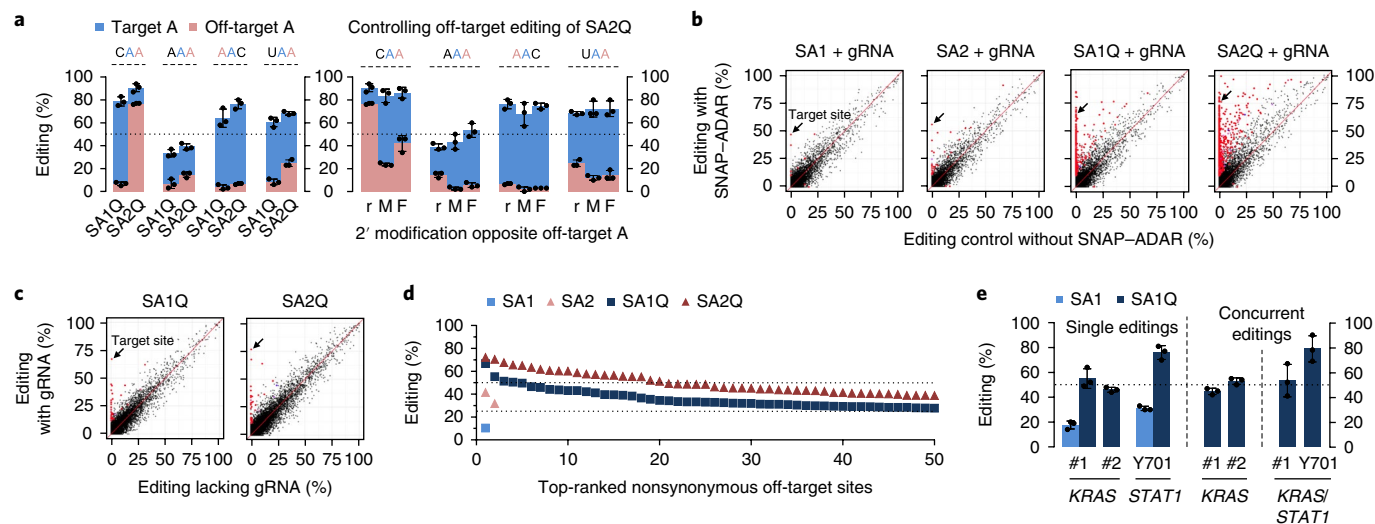


Fig. 2 | Editing specificity and application. **a**, Off-target editing of adjacent adenines (A) in A-rich triplets. r, M, and F refer to the chemical modification opposite the off-target A; r, natural ribonucleotide; M, 2'-methoxy; F, 2'-fluoro. **b, c**, Scatter plots of differential editing at ~50,000 sites per experiment. The target site (*ACTB*) is indicated by an arrow in each plot. Significantly differently edited sites ($P < 0.01$) are indicated by red dots. In **b**, editing is compared with that in a control cell line that did not express any SA protein. **c**, Editing in the presence versus in the absence of gRNA. **d**, Nonsynonymous off-target sites ranked by editing yield. Experiments were carried out in duplicate. **e**, Editing of signaling transcripts. Two 5'-UAG sites in the ORF of the *KRAS* transcript (sites 1 and 2) and a 5'-UUA site in the *STAT1* transcript (Tyr701) were targeted. For concurrent editing, two respective gRNAs were cotransfected into SA1Q⁺ cells. Data in **a, e** are shown as the mean \pm s.d.; $n = 3$ independent experiments; black dots represent individual data points. Significance in **b, c** was tested by Fisher's exact test (two-sided); $n = 2$ independent experiments.

Table 1 | Global off-target editing

Enzyme ^a	Total	Location in mRNA								
		Known		Novel		Coding region				
		Alu	Non-Alu	Alu	Non-Alu	5'-UTR	Syn.	Nonsyn.	3'-UTR	Others ^b
SA1	6	2	1	0	3	0	0	1	3	2
SA2	30	15	8	1	6	0	0	2	22	6
SA1Q	835	70	59	7	699	11	117	230	402	75
SA2Q	1,310	267	71	24	948	13	149	347	637	164

Numbers represent the number of sites that were significantly differently edited compared with sites in a related control cell line that did not express the respective SA editase. Syn., synonymous; nonsyn., nonsynonymous. ^aEditing was carried out in cells expressing the given SNAP-ADAR in the presence of a BG-gRNA targeting the *ACTB* transcript. ^bOthers refers to editing in introns, intergenic regions, and noncoding RNA.

Again, concurrent editing of either two sites on the *KRAS* transcript or sites on two transcripts (*KRAS* and *STAT1*) was possible without a loss of editing efficiency (Fig. 2e). The highly precise editase SA1 was less active, but was still able to obtain yields of $18\% \pm 3\%$ (*KRAS* site 1) and $31\% \pm 2\%$ (*STAT1*).

The chemical modification of our gRNA restricted off-target editing in the mRNA-gRNA duplex. This is in contrast to two competing approaches (one based on Cas13b)^{9,12,13} that were shown to induce massive global off-target editing caused by the overexpressed editases^{9,15} (Supplementary Tables 2 and 3). For SNAP-ADARs, global off-target editing was restricted by genomic integration. It was almost eliminated with the precise editases SA1 and SA2, and editing of endogenous targets was still sufficient for some codons (UAG and UAU). The performance of SA1 was also better than that of the 'high-specificity variant' of Cas13b-ADAR⁹ (Supplementary Note 1). Notably, our integrated hyperactive SA1Q and SA2Q showed off-target editing that was orders of magnitude less than that observed with overexpressed Cas13b-ADAR version 1⁹ or λ N-deaminases¹⁵ (Supplementary Fig. 8). We found that further reduction of SA1Q/SA2Q expression (up to 25-fold) is possible to

further reduce off-target editing (Supplementary Fig. 9). One could further improve on the gRNA chemistry¹⁸ or the editase used in our approach^{9,10}. Notably, we tested the reported high-specificity variant of Cas13b-ADAR (T375G), but in the context of SNAP-ADAR (Supplementary Fig. 10). In contrast to previous claims⁹, we found this mutant to be much less efficient than SA1Q/SA2Q, and even inferior to SA1/2. Compared with those used in other approaches, our gRNAs are extremely short (22 nt). Thus editing clearly depends on the targeting mechanism and will not interfere with endogenous ADARs⁸. However, we found that the long Cas13b gRNAs (85 nt) recruited overexpressed human ADAR2, as well as SA2Q, to elicit editing of a cotransfected reporter at levels similar to those observed with Cas13b-ADAR (Supplementary Fig. 11). This observation raises the question of the extent to which previously reported edits⁹ were affected by overexpression artifacts (Methods, Supplementary Note 2). Finally, the small size (20 kDa) and human origin of the SNAP-tag provide further advantages over Cas13-ADAR. Together, our results set a new benchmark for site-directed RNA editing and provide a tool ready for use in concurrent editing of endogenous transcripts.

Methods

Methods, including statements of data availability and any associated accession codes and references, are available at <https://doi.org/10.1038/s41592-018-0017-z>.

Received: 8 November 2017; Accepted: 9 April 2018;

Published online: 2 July 2018

References

1. Frye, M., Jaffrey, S. R., Pan, T., Rechavi, G. & Suzuki, T. *Nat. Rev. Genet.* **17**, 365–372 (2016).
2. Vogel, P., Schneider, M. F., Wettengel, J. & Stafforst, T. *Angew. Chem. Int. Ed. Engl.* **53**, 6267–6271 (2014).
3. Stafforst, T. & Schneider, M. F. *Angew. Chem. Int. Ed. Engl.* **51**, 11166–11169 (2012).
4. Nishikura, K. *Nat. Rev. Mol. Cell Biol.* **17**, 83–96 (2016).
5. Rossi, A. et al. *Nature* **524**, 230–233 (2015).
6. Vogel, P. & Stafforst, T. *ChemMedChem* **9**, 2021–2025 (2014).
7. Montiel-Gonzalez, M. F., Vallecillo-Viejo, I., Yudowski, G. A. & Rosenthal, J. J. C. *Proc. Natl. Acad. Sci. USA* **110**, 18285–18290 (2013).
8. Wettengel, J., Reautschnig, P., Geisler, S., Kahle, P. J. & Stafforst, T. *Nucleic Acids Res.* **45**, 2797–2808 (2017).
9. Cox, D. B. T. et al. *Science* **358**, 1019–1027 (2017).
10. Kuttan, A. & Bass, B. L. *Proc. Natl. Acad. Sci. USA* **109**, E3295–E3304 (2012).
11. Eggington, J. M., Greene, T. & Bass, B. L. *Nat. Commun.* **2**, 319 (2011).
12. Sinnamon, J. R. et al. *Proc. Natl. Acad. Sci. USA* **114**, E9395–E9402 (2017).
13. Montiel-González, M. F., Vallecillo-Viejo, I. C. & Rosenthal, J. J. C. *Nucleic Acids Res.* **44**, e157 (2016).
14. Fukuda, M. et al. *Sci. Rep.* **7**, 41478 (2017).
15. Vallecillo-Viejo, I. C., Liscovitch-Brauer, N., Montiel-Gonzalez, M. F., Eisenberg, E. & Rosenthal, J. J. C. *RNA Biol.* **15**, 104–114 (2018).
16. Ramaswami, G. & Li, J. B. *Nucleic Acids Res.* **42**, D109–D113 (2014).
17. Bromberg, J. J. *Clin. Invest.* **109**, 1139–1142 (2002).
18. Bennett, C. F., Baker, B. F., Pham, N., Swayze, E. & Geary, R. S. *Annu. Rev. Pharmacol. Toxicol.* **57**, 81–105 (2017).

Acknowledgements

We gratefully acknowledge support from the Deutsche Forschungsgemeinschaft (STA 1053/3-2 and STA 1053/7-1 to T.S.), the European Research Council (ERC) under the European Union's Horizon 2020 research and innovation program (grant agreement no. 647328 to T.S.), and the US National Institutes of Health (NIH) (grants R01GM102484 and R01GM124215 to J.B.L.).

Author contributions

P.V., M.M., T.M., K.D.S., and T.S. conceived, performed, and analyzed the wet lab experiments; Q.L. and J.B.L. analyzed and all authors interpreted next-generation sequencing data; and all authors contributed to writing of the manuscript.

Competing interests

The authors declare no competing interests.

Additional information

Supplementary information is available for this paper at <https://doi.org/10.1038/s41592-018-0017-z>.

Reprints and permissions information is available at www.nature.com/reprints.

Correspondence and requests for materials should be addressed to T.S.

Publisher's note: Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Methods

BG-gRNA synthesis. Synthesis of chemically modified BG-gRNAs does not require any chemistry equipment. All chemical modifications used in this study are commercially available. The benzylguanine (BG) modification can be achieved by application of a commercial amino or thiol reactive BG derivative such as BG-maleimide (New England Biolabs). The sequences and chemical modifications of all gRNAs are presented in Supplementary Table 4. For this study, all NH₂-gRNAs were purchased from Biospring (Germany) as HPLC-purified single-stranded RNAs with a 5'-C6 amino linker. As an alternative to commercial BG derivatives, our protocol can be used to introduce the BG moiety. BG connected to a carboxylic acid linker²³ (12 μl, 60 mM in DMSO) was activated in situ as an OSu-ester by incubation with EDCI-HCl (12 μl, 17.4 mg/ml in DMSO) and NHS (12 μl, 17.8 mg/ml in DMSO) for 1 h at 30 °C. Then, the NH₂-gRNA (25 μl, 6 μg/μl) and DIPEA (12 μl, 1:20 in DMSO) were added to the preactivation mix and incubated (90 min, 30 °C)¹⁹. The crude BG-gRNA was purified from unreacted NH₂-gRNA by 20% urea PAGE and then extracted with H₂O (700 μl; overnight at 4 °C). RNA precipitation was done with sodium acetate (0.1 volumes, 3.0 M) and ethanol (3 volumes, 100%, overnight at -80 °C). The BG-gRNA was washed with ethanol (75%) and dissolved in water (60 μl).

SNAP-ADAR-expressing cell lines. Each enzyme was integrated as a single copy under control of the doxycycline-inducible CMV promoter at the FRT site into the genome of Flp-In 293 cells (R78007; Thermo Fisher Scientific) as described⁸. The exact cDNAs are listed in Supplementary Note 4. Enzyme expression of all four enzymes was inducible by doxycycline (10 ng/ml) to roughly similar levels as validated by western blotting and fluorescence microscopy (Supplementary Fig. 12 and Supplementary Note 3). Also at the RNA level, the expression levels of SA1 (wild-type and Q) and SA2 (wild-type and Q) were roughly similar, with average FPKM values of 679 and 814 for SA1(Q) and SA2(Q), respectively. The E>Q mutation did not change the protein localization (Supplementary Note 3). SA1(Q) is localized to cytoplasm and nucleoplasm; SA2(Q) is mainly localized to cytoplasm. To determine the location of the different SNAP-ADAR proteins, we seeded 1 × 10⁵ cells in 500 μl of selection media with or without doxycycline (10 ng/ml) on poly-D-lysine-coated coverslips in a 24-well format. After 1 d, we carried out BG-FITC labeling of the SNAP-tag and nuclear staining. To validate SNAP-ADAR protein amounts, we performed western blotting analysis. For this, 3 × 10⁵ cells were seeded in 500 μl of selection media with or without doxycycline (10 ng/ml) in a 24-well format for 1 d. Then, cells were lysed with urea buffer (8 M urea in 10 mM Tris, 100 mM NaH₂PO₄, pH 8.0). Protein lysate (5 μg) was separated by SDS-PAGE and transferred onto a PVDF membrane (Bio-Rad Laboratories) for immunoblotting with primary antibodies to the SNAP-tag (1:1,000; P9310S; New England Biolabs) and β-actin (1:40,000; A5441; Sigma-Aldrich). Afterwards, the blot was incubated with HRP-conjugated secondary antibodies against rabbit (1:10,000; 111-035-003; Jackson ImmunoResearch Laboratories) and mouse (1:10,000; 115-035-003; Jackson ImmunoResearch Laboratories) and visualized by enhanced chemiluminescence.

RNA-editing experiments. *General.* Flp-In T-REx 293 cells stably transfected with the respective SNAP-ADAR-expressing pcDNA5 vector were grown in DMEM with 10% FBS, 100 μg/ml hygromycin B, and 15 μg/ml blasticidin S. For experiments, 3 × 10⁵ cells per well were seeded in 24-well plates, and gene expression was induced by doxycycline (10 ng/ml) for 1 d. Then, 8 × 10⁴ cells per well were resuspended in 100 μl of DMEM with 10% FBS and 15 ng/ml doxycycline and reverse-transfected in a 96-well format with the gRNA transfection mixture (39 fmol to 40 pmol of gRNA and 0.75 μl of Lipofectamine 2000 in 50 μl of OptiMEM; the exact amounts of gRNA used in this study are given in Supplementary Table 4). After 24 h, cells were collected for RNA isolation. When determining editing yields at later time points, we resuspended the cells in DMEM with 10% FBS and 10 ng/ml doxycycline and seeded them into 24-well plates. 48 h later, we added fresh medium containing 10% FBS and 10 ng/ml doxycycline to the cells. RNA was extracted with the RNeasy MinElute kit (Qiagen) and treated with DNase I. After DNA digestion, RNA was converted into cDNA for subsequent amplification by Taq DNA PCR. The DNA was analyzed by Sanger sequencing (Eurofins Genomics, Germany). We quantified A-to-I editing yields by measuring the height of the resulting guanosine peak divided by the sum of the peak heights of the guanosine and adenosine peaks at a respective site. In general, negative controls were run for all experiments and never showed detectable editing.

Potential editing at the DNA versus the RNA level. To check for potential A-to-I editing of the genomic DNA beside the targeted RNA, we used the innuPREP DNA/RNA mini kit (Analytik Jena, Germany) to extract genomic DNA and RNA from cells in parallel. We followed the manufacturer's protocol. Cellular RNA was further reverse-transcribed as described above, and the genomic DNA was immediately amplified by Taq DNA PCR and sequenced without reverse transcription. No A-to-G change in the DNA was detectable (Supplementary Fig. 13).

Potency and time dependency. For the potency and the time-dependence experiments, RNA was isolated with 500 μl of TRI reagent (Sigma-Aldrich).

Chloroform (100 μl) was added to extract the RNA for precipitation with isopropanol (350 μl) in the presence of linear acryl amide (1.5 μl; 5 mg/ml). The RNA pellet was washed twice (500 μl of 75% ethanol) and was then dissolved in RNase-free water (30 μl). Furthermore, we tested whether the editing efficiency and potency were dependent on the formation of the covalent bond between gRNA and SNAP-ADAR. gRNAs that lacked the BG moiety could elicit substantial editing only with the hyperactive enzymes (up to 70% editing yield), and required ~50-fold higher amounts of gRNA (ED₅₀ (effector dose for a half-maximum response) ~ 6–7 pmol per well; Supplementary Fig. 14). With the wild-type enzymes, no substantial editing was obtained even at the highest gRNA concentration (20 pmol per well). The target site in the potency screen was UAG site 2 in the ORF of endogenous *GAPDH* mRNA. The target in the time-dependency screen was a 5'-UAG site in the 3'-UTR of endogenous *GAPDH* mRNA.

Triplet scope. When studying the editing scope with all 16 5'-NAN triplets, we chose targets such that no amino acid change resulted. For four triplets, sites had to be chosen that elicited amino acid changes. Then, sites were selected that were expected not to interfere with *GAPDH* activity (Supplementary Note 4).

Applicability. In terms of maximum yield (up to 90%), potency (≥1 pmol per well), and duration (several days), site-directed RNA editing behaves similarly to RNA interference with transfected short interfering RNAs³⁰ in cell culture and may allow numerous applications. However, it is difficult to reliably predict the outcome of an editing reaction from the triplet preference alone (Fig. 1e). The accessibility of an arbitrary target might be limited by RNA secondary structure, RNA-binding proteins²¹, low mRNA copy numbers, and short half-lives.

Off-target editing. Accurate analyses uncovered an example of off-target editing at the targeted transcript but outside the gRNA-mRNA duplex. This was undetectable for SA1/2, but was found for SA1Q (50% editing of one AAG triplet in *GAPDH* mRNA) and for SA2Q (70% editing of a CAG site in *GAPDH* mRNA). These two strongly edited sites in *GAPDH* mRNA were predicted by mfold to be located in highly double-stranded regions of the transcript (Supplementary Fig. 15). In accordance, editing yields correlated with the proximity of the gRNA binding site, reminiscent of the recently described TRIBE method to elucidate binding sites of RNA-binding proteins²².

Next-generation RNA-sequencing experiments. The RNA editing was done by transfection of 5 pmol of gRNA targeting a 5'-UAG triplet in the 3'-UTR of *ACTB* mRNA into the respective Flp-In cell line as described above. Overall, seven settings were implemented, each with an independent duplicate: (1) empty lipofection into empty (i.e., not expressing SA proteins) Flp-In 293 cells, (2) gRNA lipofection into SA1⁺ cells, (3) gRNA transfection into SA2⁺ cells, (4) empty transfection into SA1Q⁺ cells, (5) empty transfection into SA2Q⁺ cells, (6) gRNA transfection into SA1Q⁺ cells, and (7) gRNA transfection into SA2Q⁺ cells. RNA was isolated with the RNeasy MinElute kit, treated with DNase I, and purified again with the RNeasy MinElute kit. Purified RNA (1.2 μg) was delivered to CeGaT (Germany) for poly(A)⁺ mRNA sequencing. The library was prepared from 100 ng of RNA with the TruSeq stranded mRNA library prep kit (Illumina) and sequenced with a HiSeq 4000 (50 million reads, 2 × 100 bp paired end; Illumina).

Mapping of RNA-seq reads. We adopted a previously published pipeline to accurately align RNA-seq reads onto the genome^{23,24}. We used BWA²⁵ to align the reads to a combination of the reference genome sequences and exonic sequences surrounding known splicing junctions from known gene models. Each of the paired-end reads was mapped separately using the commands “bwa aln fastqfile” and “bwa samse -n4.” We then chose a length of the splicing junction that was slightly shorter than the RNA-seq reads to prevent redundant alignment (i.e., 95 bp for reads 100 bp in length). The reference genome used was hg19, and the gene models were obtained through the UCSC Genome Browser for Gencode, RefSeq, Ensembl, and UCSC Genes. We considered only uniquely mapped reads with mapping quality $q > 10$ and used SAMtools rmdup²⁶ to remove clonal reads (PCR duplicates) mapped to the same location. Of these identical reads, only the read with the highest mapping quality was kept for downstream analysis. Unique and nonduplicate reads were subjected to local realignment and base-score recalibration using the IndelRealigner and TableRecalibration from the Genome Analysis Toolkit (GATK)²⁷. The above steps were applied separately to each of the RNA-seq samples.

Identification of editing sites from RNA-seq data. We used the UnifiedGenotyper from GATK²⁷ to call variants from the mapped RNA-seq reads. In contrast to the usual practice of variant calling, we identified the variants with relatively loose criteria by using the UnifiedGenotyper tool with options stand_call_conf 0, stand_emit_conf 0, and output mode EMIT_VARIANTS_ONLY. Variants from nonrepetitive and repetitive non-Alu regions were required to be supported by at least three reads containing mismatches between the reference genome sequences and RNA-seq data. Supporting of one mismatched read was required for variants in Alu regions. We subjected this set of variant candidates to several filtering steps to increase the accuracy of editing-site calling. We first removed all known

human single-nucleotide polymorphisms (SNPs) present in dbSNP (except SNPs of molecular type 'cDNA'; database version 135; <http://www.ncbi.nlm.nih.gov/SNP/>), the 1000 Genomes Project, and the University of Washington Exome Sequencing Project (<http://evs.gs.washington.edu/EVS/>). To remove false positive RNA-seq variant calls due to technical artifacts, we applied further filters as previously described^{23,24}. In brief, we required a variant call quality $Q > 20$ ^{23,24}, discarded variants if they occurred in the first six bases of a read²⁵, removed variants in simple repeats²⁶, removed intronic variants that were within 4 bp of splice junctions, and discarded variants in homopolymers²⁷. Moreover, we removed sites in highly similar regions of the genome by BLAT²⁸. Finally, variants were annotated with ANNOVAR²⁹ on the basis of gene models from Gencode, RefSeq, Ensembl, and UCSC. The resulting sets of sites identified from RNA-seq data were compared with all sites available in the RADAR database¹⁶ and were subsequently referred to as 'known' sites if also found in RADAR, or 'novel' sites if not found.

Identification of significantly differently edited sites. We quantified editing levels of edited sites with coverage of ≥ 50 reads (combined coverage of both replicates) and performed Fisher's exact tests (adjusted $P < 0.01$) to identify significantly differently edited sites across the samples (editing difference $> 10\%$). Additional next-generation sequencing (NGS) quality data are given in Supplementary Note 4.

Benchmarking with Cas13b-ADAR and λ N-deaminases. The SNAP-ADAR approach was benchmarked against the recently published Cas13b-ADAR approach (Supplementary Notes 1 and 2, Supplementary Table 2, and Supplementary Figs. 10 and 11). First, we repeated the editing of *KRAS* mRNA sites 1 and 2 with SA1 and SA1Q. We observed that SA1Q achieved better editing yields than Cas13b-ADAR version 1 (e.g., 50–65% compared with 15–25% for *KRAS* site 1), SA1 was better than Cas13b-ADAR version 2 (e.g., 18–20% versus ~12%), editing depended strictly on the targeting mechanism, and there was no off-target editing in the mRNA-gRNA duplex (Supplementary Note 1). ADARs are known to edit double-stranded RNA substrates of > 30 bp readily. We wondered whether large Cas13-gRNAs (85 nt, 50-bp duplex) are able to recruit human ADAR or any other ADAR fusion protein independently of a specific targeting mechanism. Indeed, we found that such 50-bp gRNAs recruited overexpressed ADAR2 but also engineered SA2Q to elicit editing of a cotransfected reporter transcript at levels similar to those achieved with Cas13-ADAR (~25–30%; Supplementary Fig. 11, Supplementary Note 2). This medium-level editing was apparently due to self-targeting of the deaminase (domain) alone and independent of a specific targeting mechanism. Most of the experiments reported by Cox et al.⁹ were done under such co-overexpression conditions, and it remains unclear to what extent their results rely on a true (Cas-dependent) targeting mechanism and which, if any, are overexpression artifacts (self-targeting). The lack of codon preference reported for repairV1 (with 10–35% editing yields) could be impaired by this. Cox et al.⁹ argue that Cas-ADAR has a weak codon preference due to tight binding of the Cas protein to the mRNA-gRNA complex, but in our opinion they do not report sufficient data or controls to support this. In the worst case, a very stable long RNA duplex wrapped by Cas-ADAR could inhibit translation, in particular when the start codon is close or even included, as this is given for the *KRAS* transcript they reported on (Supplementary Note 1). As we have shown here in the context of SNAP-ADARs, translation inhibition with puromycin can indeed increase editing

levels in the ORF (Supplementary Fig. 3). In this respect, it is notable that we have tested the mutation from their 'high-specificity' Cas-ADAR repair version 2 (T375G), but in the context of SNAP-ADAR. For this, we genomically integrated SA2QG (E488Q + T375G) and tested it side-by-side with SA1 and SA2 for the editing of five codons in the ORF of the *GAPDH* transcript (UAG, CAA, CAG, AAG, and GAU). SA2QG elicited only minor editing at the UAG codon (15%) and no significant yield with the other four codons (Supplementary Fig. 10). It was always less active than the two wild-type SA enzymes, which produced editing at some of the codons (~40% at UAG, 23–66% at CAA, 18% at CAG). In the ORF, SA2QG seemed unable to edit even the preferred UAG codon sufficiently. However, editing was successful when we targeted a UAG triplet in the 3'-UTR of *GAPDH* mRNA (80% SA2QG, 85–90% for wild-type SA enzyme). Unfortunately, Cox et al.⁹ do not comprehensively characterize repairV2 or show whether and how it promotes the editing reaction. Notably, our data predict that the wild-type deaminase would always be the better choice (compared with repairV2) to achieve decent editing at preferred codons with manageable off-target edits also in the context of Cas-ADAR. The true mechanism of Cas-ADAR-directed RNA editing and how it can be best applied remain partly unclear. We also provide a side-by-side comparison with the λ N-deaminase approach (Supplementary Table 3) and reanalyzed the NGS data from Vallecillo-Viejo et al.¹⁵ with our pipeline (Supplementary Fig. 8). In comparison, our wild-type SA1/SA2 enzymes were highly precise and provoked several-hundred-fold less off-target editing. Our hyperactive enzymes SA1Q and SA2Q were less prone to off-target editing than the wild-type versions of the λ N-deaminases and much less off-target prone than the hyperactive version of the λ N-deaminases.

Reporting Summary. Further information on experimental design is available in the Nature Research Reporting Summary linked to this article.

Data availability. All original NGS data have been deposited in the NCBI GEO database under accession [GSE112787](https://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE112787). Our NGS data analysis is available online as Supplementary Data. All programs used are publicly available. The gene sequences of all constructs are given in the Supplementary Information; plasmids can be obtained from the corresponding author upon request.

References

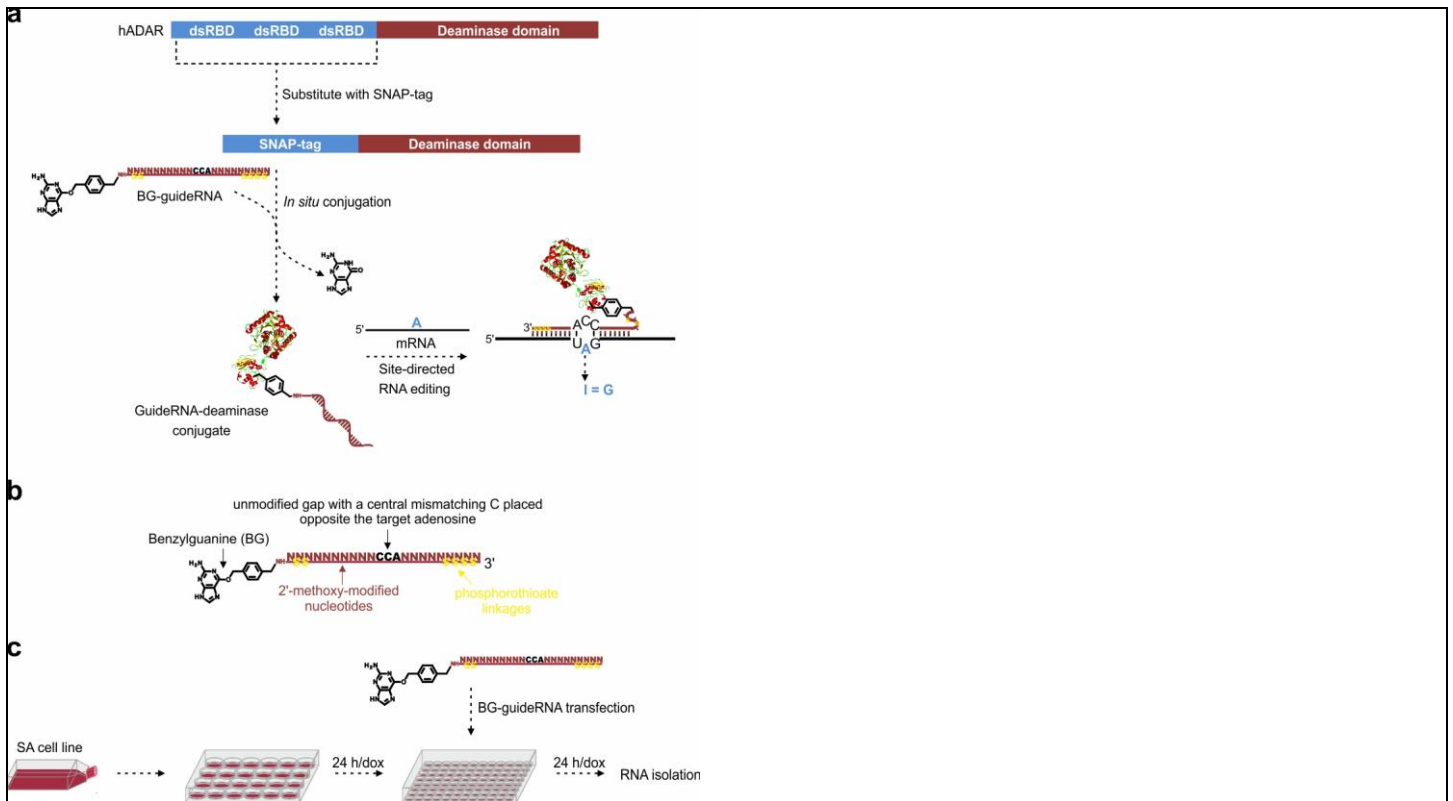
- Hanswillemenke, A., Kuzdere, T., Vogel, P., Jékely, G. & Stafforst, T. *J. Am. Chem. Soc.* **137**, 15875–15881 (2015).
- Kim, D.-H. et al. *Nat. Biotechnol.* **23**, 222–226 (2005).
- Deffit, S. N. & Hundley, H. A. *WIREs RNA* **7**, 113–127 (2016).
- McMahon, A. C. et al. *Cell* **165**, 742–753 (2016).
- Ramaswami, G. et al. *Nat. Methods* **9**, 579–581 (2012).
- Ramaswami, G. et al. *Nat. Methods* **10**, 128–132 (2013).
- Li, H. & Durbin, R. *Bioinformatics* **26**, 589–595 (2010).
- Li, H. et al. *Bioinformatics* **25**, 2078–2079 (2009).
- McKenna, A. et al. *Genome Res.* **20**, 1297–1303 (2010).
- Kent, W. J. *Genome Res.* **12**, 656–664 (2002).
- Wang, K., Li, M. & Hakonarson, H. *Nucleic Acids Res.* **38**, e164 (2010).

In the format provided by the authors and unedited.

Efficient and precise editing of endogenous transcripts with SNAP-tagged ADARs

Paul Vogel¹, Matin Moschref¹, Qin Li², Tobias Merkle¹, Karthika D. Selvasaravanan¹, Jin Billy Li² and Thorsten Stafforst^{1*}

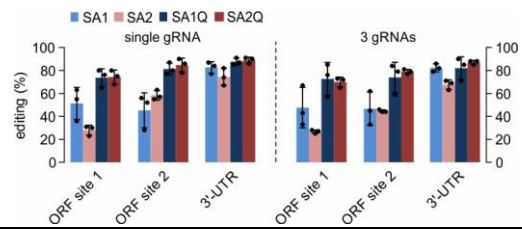
¹Interfaculty Institute of Biochemistry, University of Tübingen, Tübingen, Germany. ²Department of Genetics, Stanford University, Stanford, CA, USA.
*e-mail: thorsten.stafforst@uni-tuebingen.de



Supplementary Figure 1

Site-directed RNA editing by SNAP-tagged ADARs driven by short, chemically modified guide RNAs.

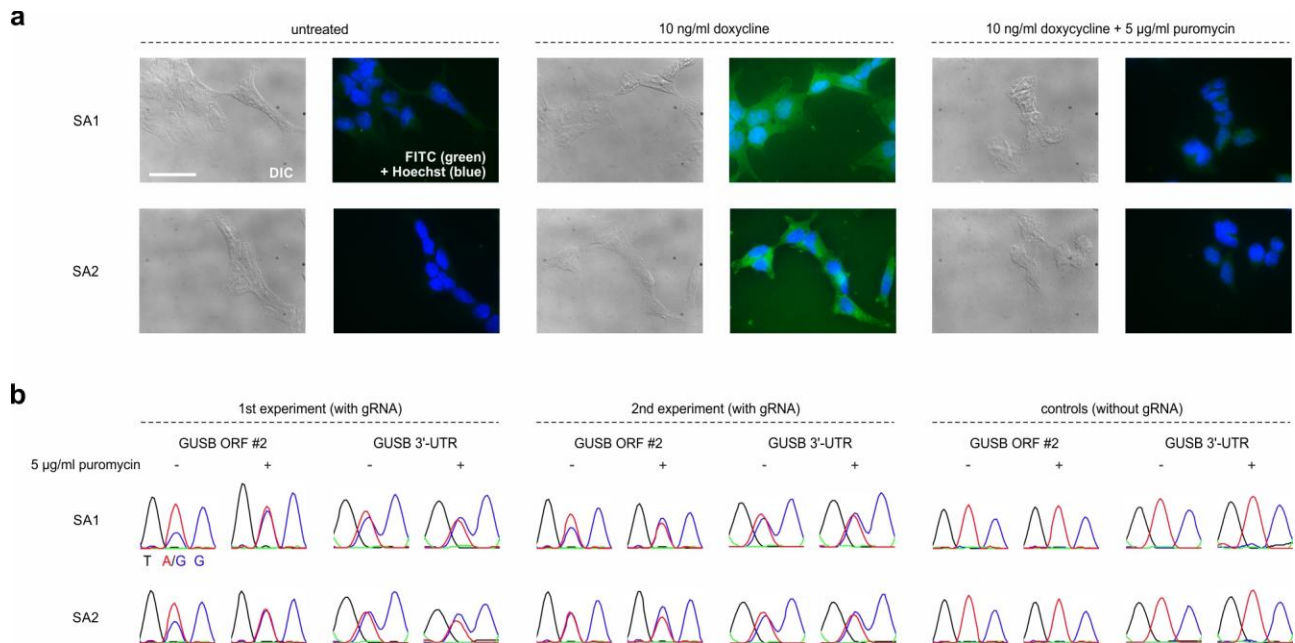
a) General concept: The double-stranded RNA-binding domains (dsRBDs) of hADAR have been substituted with the SNAP-tag. The latter is able to form a covalent bond to a guideRNA that is modified with benzylguanine (BG). When bound to the SNAP-ADAR, the guideRNA steers the attached SNAP-ADAR protein to the target RNA and forms the necessary secondary structure for A-to-I editing catalyzed by the deaminase domain. **b)** A typical BG-guideRNA that targets a UAG site with a 5'-CCA anticodon. The guideRNA is 22-nt long and is densely chemically stabilized by 2'-methoxylation and terminal phosphorothioate linkages (commercially available). The first three 5'-terminal nucleotides do not base pair with the target RNA, but serve as a linker. The sequence comprises an unmodified ribonucleotide gap (5'-CCA) which faces the target site and contains a central mismatching cytosine opposite the targeted adenosine for efficient deamination. A commercial C6-amino-linker is located at the 5'-end of the guideRNA to introduce the BG modification to the full length oligonucleotide. Modification of the guideRNA with OSu-activated BG can be performed in any reaction tube. **c)** Experimental setup. Cells with stably integrated SNAP-ADAR (SA) are seeded into 24-well plates with medium containing doxycycline (dox) to induce SA expression. 24 h later, the cells were reverse-transfected with the guideRNA (see online methods). After 24 h, the cells were lysed for RNA isolation to analyze RNA editing.



Supplementary Figure 2

Concurrent editing of three 5'-UAG sites in endogenous *GAPDH* transcript.

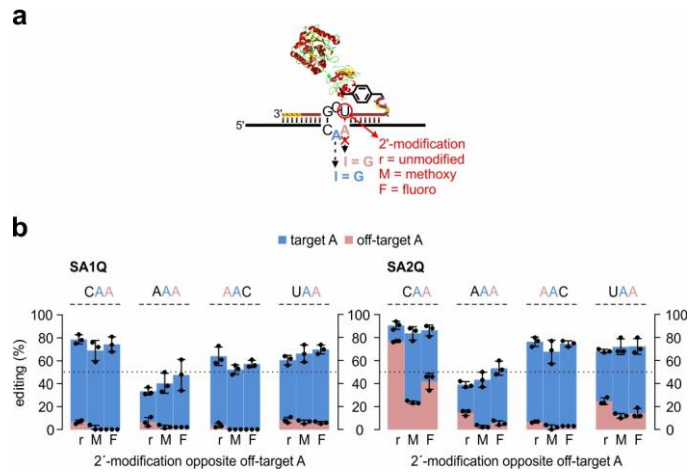
The respective SA-expressing cells were transfected with either a single gRNA or 3 gRNAs against distinct sites on the endogenous *GAPDH* transcript. Data are shown as the mean \pm SD, N=3 independent experiments, black dots represent individual data points.



Supplementary Figure 3

Effect of global translation inhibition (puromycin) on RNA editing.

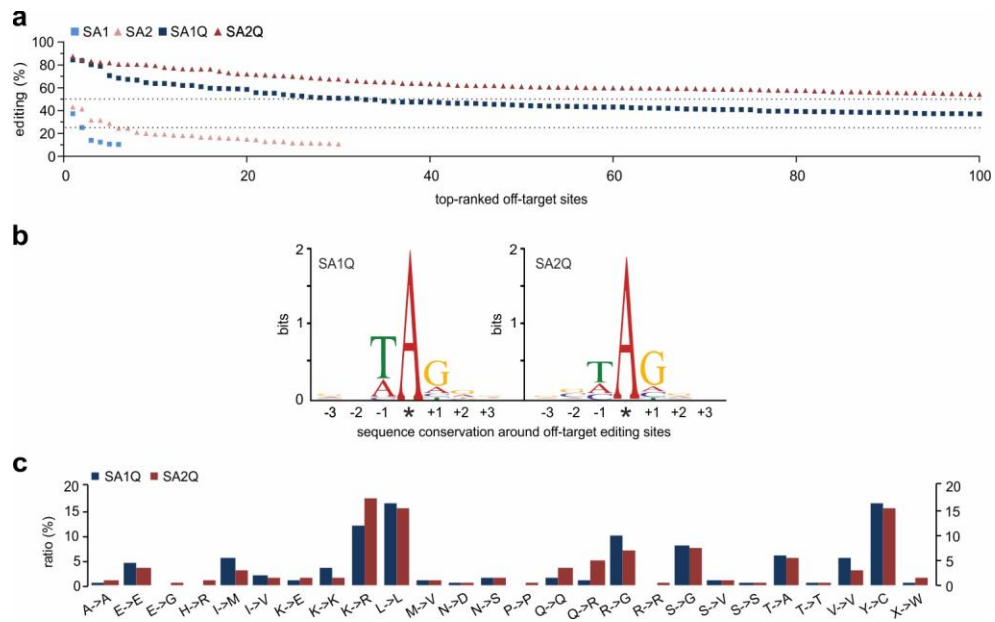
a) SA1 and SA2 cells were seeded on poly-D-lysine-coated glass slides. After one day, doxycycline (end concentration = 10 ng/ml) was added to induce SA enzyme expression. To inhibit translation, cells were additionally supplemented with 5 µg/ml puromycin, respectively. After 12h, cells were stained with BG-FITC and Hoechst. The staining shows that the applied amount of puromycin (5 µg/ml) is sufficient to block translation. The scale bar represents 40 µm. **b)** Cells were reverse transfected with BG-gRNA (5 pmol/96 well) targeting a UAG site either in the ORF (site #2) or in the 3'-UTR of GUSB. After 4 h, the cells were optionally incubated with 5 µg/ml puromycin for 12 h. Then, RNA was isolated and reverse transcribed for Sanger sequencing. As one can see, the editing levels differ between ORF and 3'-UTR in the absence of puromycin with less efficient editing in the ORF than in the 3'-UTR. After addition of puromycin translation is blocked (panel a) and the editing levels in the ORF increase to the levels obtained in the 3'-UTR (panel b), which don't change notably under puromycin treatment. This supports our assumption that editing in the ORF can be kinetically limited by the process of translation. a), b) Two independent experiments were performed with similar results.



Supplementary Figure 4

Controlling off-target editing in SA1Q/SA2Q⁺ cells.

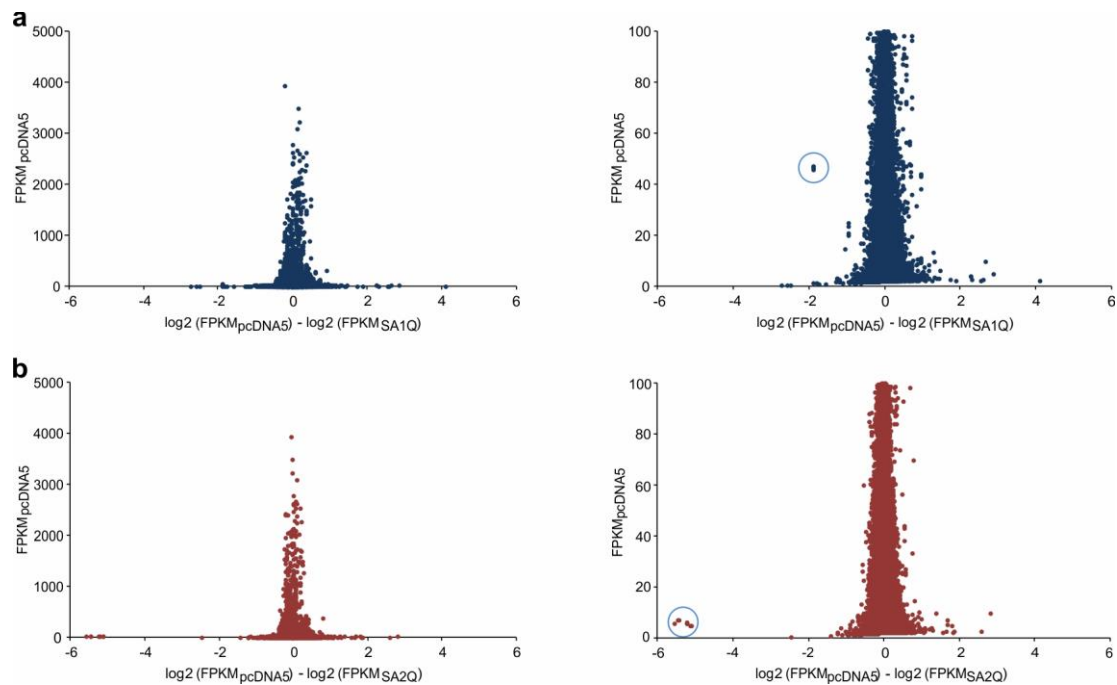
a) General strategy. To avoid unintended editing of an adjacent adenosine at the target site, the opposing base in the guideRNA can be modified by 2'-methoxylation (M) or 2'-fluorination (F). This is exemplary shown for the triplet CAA. **b)** In the study, off-target editing of an adjacent adenosine was detected in the triplets CAA, AAA, AAC and UAA when particularly using SA2Q cells. However, off-target editing was remarkably reduced when the strategy was applied. Data are shown as the mean±SD, N=3 independent experiments, black dots represent individual data points.



Supplementary Figure 5

Off-target editing and off-target codon preferences caused by SA enzymes.

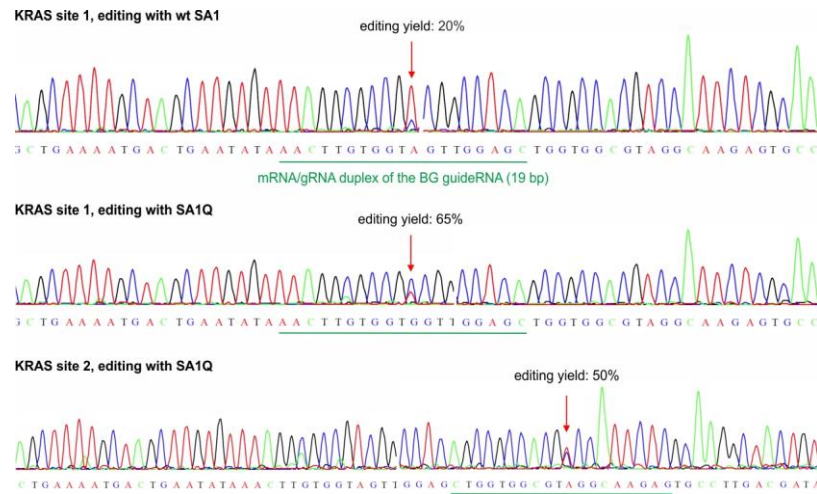
a) Overall off-target sites are ranked by editing yields. **b)** Logo represents the sequence conservation around all significant off-target sites for SA1Q and SA2Q. **c)** Analysis of the codon changes for all off-target editings that were found in SA1Q and SA2Q cells. The ratio was calculated in relation to the total number of editing events happened in the coding region of the transcripts.



Supplementary Figure 6

Gene expression analysis.

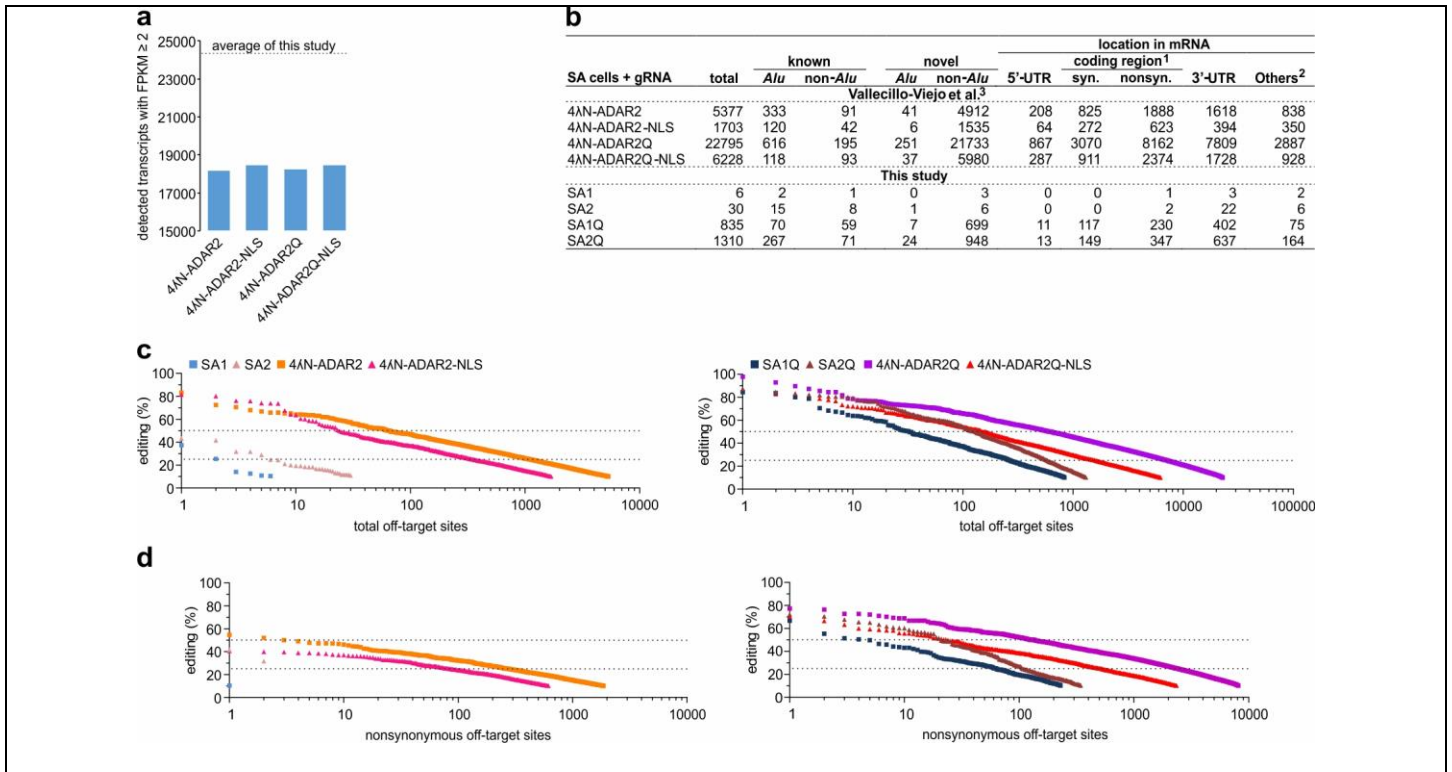
FPKM values of approximately 25,000 expressed transcripts are compared between cells containing the empty pcDNA5 vector with SA1Q cells + gRNA (a) or SA2Q cells + gRNA (b). Plotted is the \log_2 fold change in expression against the FPKM of the respective transcript in the control cell line (pcDNA5). The left plots show the data for all transcripts, the right plots for the low expressing transcripts only (FPKM < 100). Analysis was restricted to transcripts with FPKM values ≥ 2 in either the control or the SNAP-ADAR-expressing cell line. No strongly expressed transcripts (FPKM > 100) show \log_2 -fold changes > 1. \log_2 -fold changes of low expressing genes originate from transcripts with low FPKM and very low read coverage (typically non-coding RNAs, read-coverage below 50). The significance of such expression changes are difficult to assess. Clearly visible was the different expression of SNAP-ADAR in the engineered versus control cell line as highlighted by light blue circles.



Supplementary Figure 7

Targeting of *KRAS* mRNA by SA1 and SA1Q.

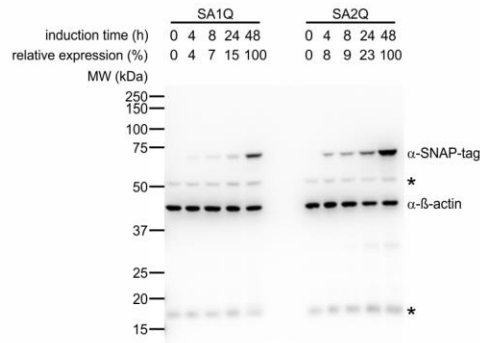
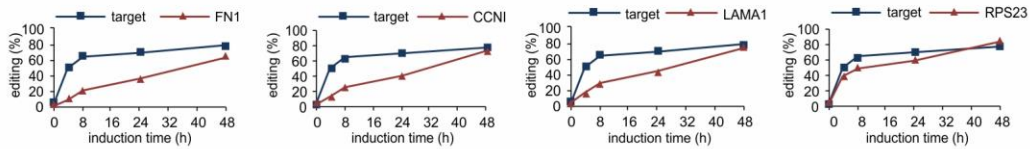
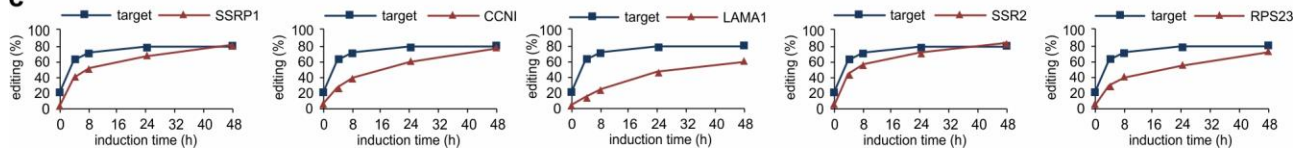
The applied BG-gRNAs form 19 bp duplex structures with the target transcript. No off-target editing was detected within these mRNA/gRNA duplexes in *KRAS* mRNA. For further discussion, see also Supplementary Note 1. N=3 independent experiments were performed with similar results.



Supplementary Figure 8

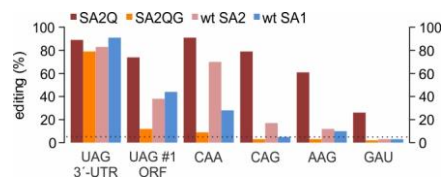
Reanalysis of NGS data produced by Vallecillo-Viejo et al.¹⁵ according to our pipeline (Methods).

a) Number of transcripts covered in RNA sequencing of the samples with 2boxB-driven 4λN-ADAR2 enzymes. Shown are numbers of detected transcripts with a FPKM value ≥ 2 . The dashed line shows the average of detected transcripts with FPKM value ≥ 2 produced in this study. **b**) Summary of off-target sites produced by 2boxB-driven 4λN-ADAR2 enzymes. Given are the numbers of off-target sites significantly differently edited compared to the related cells lacking editing enzyme and gRNA. NGS data were re-analyzed according to the protocol for detecting off-target editing by SNAP-ADAR enzymes (see online methods). ¹Nonsynonymous refers to editing that results in amino acid change (syn. = synonymous; nonsyn. = nonsynonymous); ²others refers to editing in introns, intergenic regions, and ncRNA. ³Edits were carried out in 293T cells transfected with 4λN-ADAR2 enzyme, CFTR Y122X reporter and 2boxB-gRNA by Vallecillo-Viejo et al., RNA Biol. (2018). **c**) Ranking of all off-target editing sites by the editing level. Left panel: wildtype SA versus wt λN-ADAR; right panel: hyperactive SA versus hyperactive λN-ADAR. **d**) Like c) but ranking of all nonsynonymous off-target edits.

a**b****c****Supplementary Figure 9**

Changes in editing efficiency and specificity upon variation of SNAP-ADAR induction time (0–48 h, as indicated).

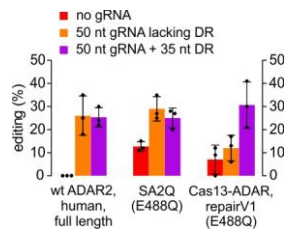
a) The expression of SA1Q or SA2Q was varied and quantified by western blot analysis (shown in relative expression, asterisks (*) indicate unspecific protein bands). We assessed the effect of reduced SA expression levels on editing the on-target (GAPDH, ORF site #2) versus several high-ranked off-targets in SA1Q (**b**) and SA2Q cells (**c**). For SA1Q (**b**), we tested three top-ranked nonsynonymous off-targets (FN1, CCNI, LAMA1) and one top-ranked, known 3'-UTR editing site (RPS23). For SA2Q (**c**), we tested three top-ranked nonsynonymous off-targets (SSRP1, CCNI, LAMA1) and two top-ranked 3'-UTR editing sites (SSR2, RPS23). On-target editing tolerated the reduction of SA expression much better compared to most off-targets. At 4h induction (4-8% SNAP-ADAR protein expression compared to full induction after 48h), most off-target editing yields were reduced by 2- to 3-fold while the on-target editing was only reduced by 35% (SA1Q) and 25% (SA2Q) compared to the editing level at full induction (48h). a)-c) The data presented are obtained from a single experiment.



Supplementary Figure 10

Studying the Cas13b-ADAR repairV2 mutant (E488Q + T375G) in the context of the SNAP-ADAR system (SA2QG).

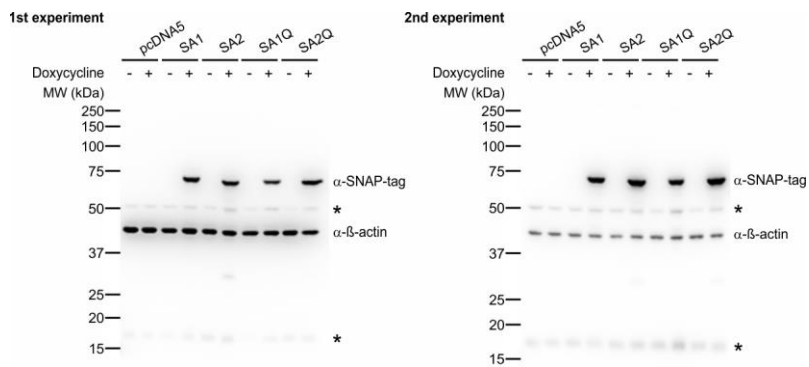
The respective double mutant (SA2QG) was genomically integrated into Flip-In cells analog as described for the other four enzymes (SA1, SA2, SA1Q, SA2Q). We then studied the editing of 6 different sites in the GAPDH transcript entirely analog as described. Interestingly, SA2QG was only active in the 3'-UTR. It was almost unable to edit targets in the ORF. SA2QG lagged behind the wildtype enzymes SA1 and SA2 in all studied codons. This is in contrast to Cox et al. (Ref. 10) who claim Cas13b-ADAR repairV2 to be a more specific mutant that still enables good editing yields. Editing levels between 5% and 10% are difficult to detect precisely by Sanger sequencing, editing levels below 5% (dotted line) cannot be detected. The data presented here is a single experiment (N = 1).



Supplementary Figure 11

Recruitment of various editases by Cas13-gRNAs (with and without DR domain).

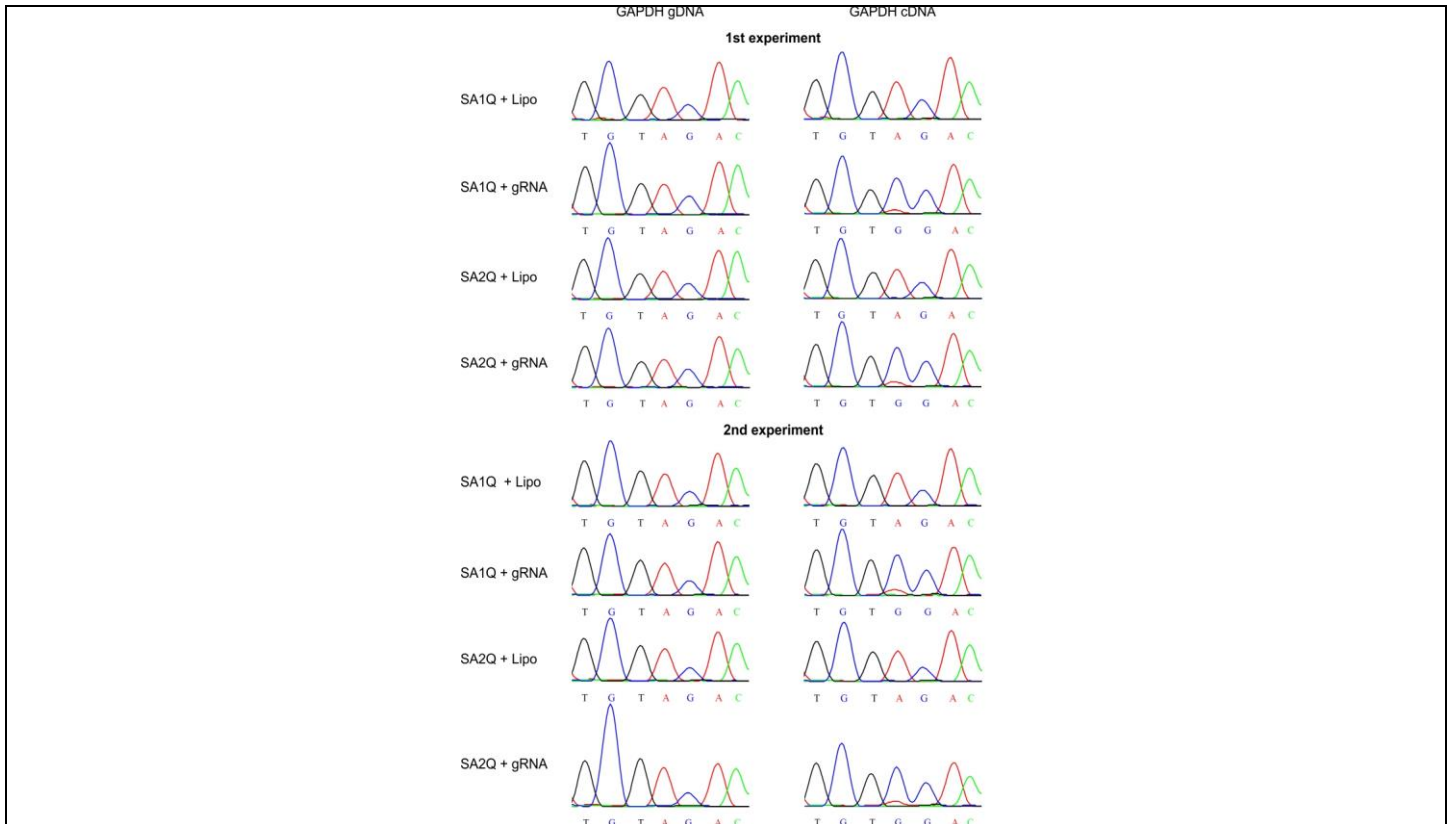
Overexpressed Cas13-guideRNAs can also recruit human ADAR2 or SA2Q to elicit editing in co-transfected reporter transcripts to yields similar as Cas13b-ADAR repairV1 does. For further details and discussion, see Supplementary Note 2. Data are shown as the mean±SD, N=3 independent experiments, black dots represent individual data points. DR = Cas13 directing domain



Supplementary Figure 12

Western blotting analysis of SNAP-ADAR expression.

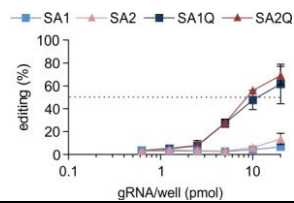
Five Flip-In T-REx cell lines were generated expressing either the empty vector (pcDNA5) or SNAP-ADAR genes (SA1, SA2, SA1Q, SA2Q) under the control of a doxycycline-inducible CMV promoter. Western blot analysis was done after the cells were incubated with and without doxycycline for 1 day. A SNAP-tag antibody was used to evaluate the protein levels of the SNAP-ADAR enzymes. The expression of β -actin served as reference. Asterisks (*) indicate unspecific protein bands. N = 2 independent experiments were performed with similar results.



Supplementary Figure 13

Editing control gDNA versus cDNA .

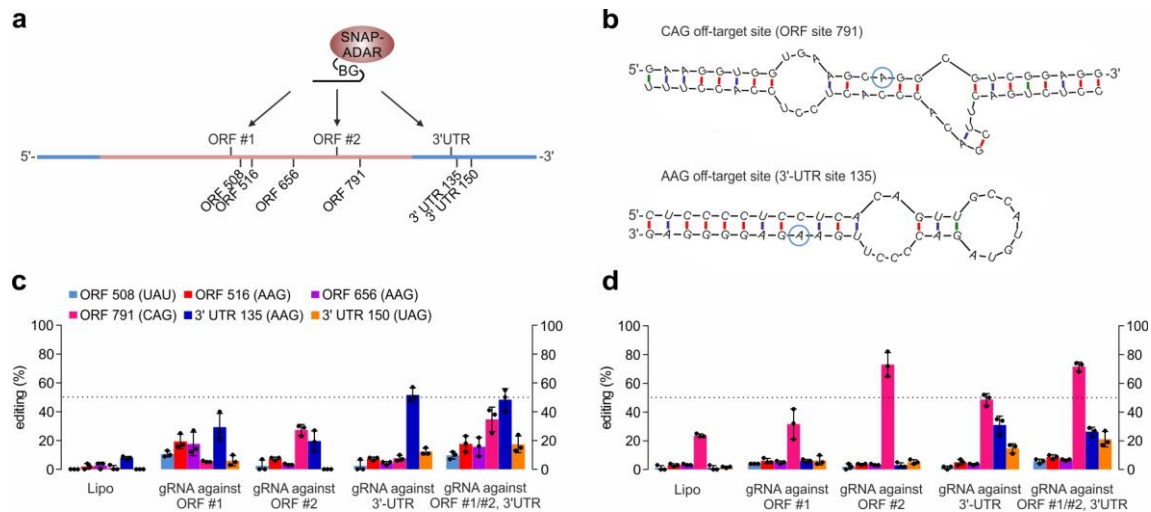
To ensure that editing occurred only at the transcript and not the genomic DNA (gDNA), sequencing traces of gDNA and cDNA derived from mRNA were compared after site-directed editing in the 3'-UTR of GAPDH. Only the cDNA traces showed an A-to-G change for SA1Q and SA2Q, indicating that both enzymes target only RNA but not DNA. N = 2 independent experiments were performed with similar results.



Supplementary Figure 14

Dose-dependence of editing when using a control gRNA lacking the benzylguanine (BG) moiety.

gRNA lacking the benzylguanine moiety was transfected in amounts of 0.625 pmol – 20 pmol, showing that efficient editing requires BG-dependent covalent bond formation via the SNAP-tag BG reaction. In particular, the wt enzymes SA1 and SA2 do not elicit editing. The hyper-active enzymes (SA1Q/SA2Q) require much higher doses to elicit editing when lacking the BG moiety. Data are shown as the mean \pm SD, N=3 independent experiments.



Supplementary Figure 15

gRNA-dependent off-target editing in *GAPDH* transcript outside the gRNA–mRNA duplex.

Different from the three other targets (*ACTB*, *GUSB*, *SA*), off-target editing was found in the *GAPDH* transcript when targeting the *GAPDH* transcript. These off-target sites are all outside the mRNA/guideRNA duplex. **a)** The *GAPDH* transcript was targeted with gRNAs at two sites in the ORF (#1, #2) and one site in the 3'UTR. Six off-target sites were observed (ORF 508/516/656/791 and 3'-UTR 135/150). **b)** Secondary structures of the off-target sites with strongest editing (ORF 791 and 3'-UTR 135) were predicted with mfold. 250 nt up- and downstream from the editing site were chosen for the analysis. The light blue circles highlight the off-target site. **c)** Editing of the respective six off-target sites in SA1Q cells transfected with the respective guideRNA(s) against the three target adenosines in the transcript. Off-target editing was promoted when the editase was directed into vicinity of the off-target site. **d)** The same observation was made for SA2Q cells, but with higher editing levels and different off-target preference. c), d) Data are shown as the mean \pm SD, N=3 independent, black dots represent individual data points.

Supplementary Table 1. Sequence similarity between top-ranked off-targets (TMX3 and AAGAB) and the target site in β -actin (ACTB) reveals sequence similarity as the cause of guideRNA-dependent off-target editing.

mRNA	sequence bound by gRNA ^a
ACTB	5'-GGGAGGUGAU AG CAUUGCU-3'
TMX3	5'- A GGAGGUGAU AG CAUU UUG -3'
AAGAB	5'- CC AGG U UGAU AG CAUU UG -3'

^a edited adenosines are highlighted in bold and not matching nucleotides in red.

Supplementary Table 2. Comparison SNAP-ADAR and dCas13b-ADAR system (Cox et al. Science 2017)

	SNAP-ADAR (SA) system	dCas13b-ADAR system
Targeting System	SNAP-tag – gRNA covalent bond SNAP-tag: human, < 200 aa gRNA: ca. 22 nt, chemically stabilized	guideRNA / dCAS13b RNP assembly ^{a)} Cas13: bacterial >1000 aa gRNA: ~85 nt, genetically encoded
Deaminase tested	4 enzymes fully tested: ADAR1 and ADAR2 each wildtype and E488Q	1 enzyme strongly tested: ADAR2 E488Q (REPAIRv1) 1 enzyme briefly tested: ADRA2 E488Q/T375G (REPAIRv2)
Delivery	SNAP-ADAR: single genomic copy, inducible gRNA: lipofection of chemically stabilized gRNA (22 nt)	dCas-ADAR: massive overexpression via plasmid lipofection guideRNA: massive overexpression via plasmid lipofection
Editing of endogenous targets	ACTB, GAPDH, GUSB, SA, KRAS, STAT1	KRAS and PPIB
Concurrent editing	3 sites or 4 endogenous house keeping transcripts, no loss in efficiency 2 sites or 2 endogenous signaling transcripts (KRAS, STAT1), no loss in efficiency	Nothing shown
Editing range for the best editable codon (UAG) on endogenous targets	wild-type SA: 15 - 90%, (12 sites on 6 targets, ORF & UTRs) SAQ variants: 46 - 90%, (13 sites on 6 targets, ORF & UTRs)	REPAIRv2: 7-25%, (5 sites on 2 targets, only ORF) REPAIRv1: 15-40%, (5 sites on 2 targets, only ORF)
Codon scope	all 16 codons tested on an endogenous target with SA1Q and SA2Q	all 16 codons tested, but on an overexpressed reporter transcript with overexpressed Cas-ADAR. The co-overexpression together with the low editing yields suggest that the shallow codon specificity observed could be an overexpression artefact. Codon scope was only tested for REPAIRv1, not for version 2
Applications in the manuscript	Manipulation of signaling transcripts, KRAS and STAT1, recoding of phosphorylation switch Tyr701 in STAT1	Manipulation of the signaling transcript KRAS, but not at a phosphorylation site. The claimed editing of 34 “ <i>release-relevant transcripts</i> ” (Figure 4) is somewhat misleading. ^{b)}
Editing duration	stable over several days	Nothing shown
Off-targets in gRNA/substrate duplex	the guideRNA/mRNA duplex is small (19 bp), chemical modification of guideRNA blocks off-target editing almost entirely even in A-rich codons	General: the guideRNA/mRNA duplex is large (50 bp) REPAIRv1: massive problem, several sites, high yields REPAIRv2: better, but present, too little data is shown yet
Global off-target editing	Wild-type SA: almost absent SAQ variants: moderate (~1000 sites, might be further decreased by lowering SAQ expression)	REPAIRv2: almost absent (but the 125x coverage/deep sequencing analysis (Figure 6D) was done with 15fold less Cas13-ADAR plasmid (10 ng instead of 150 ng) than used in the relevant editing reactions on KRAS and PPIB (Figure 6F & Figure 5). It is unclear if KRAS/PPIB editing would be effective with 15fold less CAS13-ADAR plasmid. ^{c)} REPAIRv1: extremely high (>18 000 sites, even though 15fold less Cas13-ADAR was transfected then in almost all other experiments)
Unique property	1) Chemically stabilized guideRNAs enable perfect specificity inside gRNA/mRNA duplex 2) low expression of editase enables high editing yields with reduced global off-target editing	1) Fashionable there are at least two other RNA editing systems that apply encodable guideRNAs which encounter the same specificity problems as Cas13-ADAR does (local off-target editing in the guideRNA/mRNA duplex, global off-target editing due to

	2) clearly proven, covalent RNA targeting 3) very short guideRNA/mRNA duplex, unlikely to interfere with endogenous ADARs or translation 4) simple co-transfection of guideRNAs enables concurrent editing	overexpression, in particular with hyperactive ADAR deaminases, low editing yields with wildtype or less active ADAR domains like version2)
--	--	---

a) It remains to be determined to which extent the RNA-targeting via the 35 nt DR-helix in the Cas13-guideRNAs and dCas13b interaction contributes to Cas13-ADAR editing, in particular under overexpression conditions on reporter constructs. From previous control experiments we know that under overexpression conditions editing can be obtained even in absence of any RNA targeting mechanism by self-targeting of the ADAR, in particular for long RNA duplexes (like >30 bp). When carefully reading the Cox et al. paper, the evidence is lacking that the dCAS13/guideRNA RNP assembly is strictly required for editing; the respective important control for this (Figure S8 in the Cox et al. paper) is flawed: it shows that overexpression of the ADAR2 deaminase lacking Cas13 doesn't give editing, but the guideRNA is missing too. There is also no proof that the ADAR deaminase domain they express is giving stable, catalytically functional protein. On one hand, they claim that the free-floating deaminase is giving rise to off-target editing. On the other hand, their control ADAR deaminase alone (ADAR2DD) gives much less off-targets compared to REPAIRv1 (Figure S8, C) indicating that the truncation is less functional per se. The proper control would have been to mutate the guideRNA (at the DR domain or leave the DR domain away). We tested the Cas13 guideRNAs and found them similarly active (editing yields around 25%) when overexpressing them together with either wildtype ADAR2 or SNAP-ADAR2Q, independent of the DR domain (see Supplementary Figure 11, and further Supplementary Notes 1 and 2 below). This shows that any overexpressed highly active ADAR fusion can edit 50 bp guideRNA/mRNA duplexes independent of a targeting mechanism to similar yields under the conditions reported by Cox et al. (their Figure 2-4).

b) Cox et al. suggest that 34 disease-relevant editings have been achieved (Figure 4E). This is somewhat misleading, in particular the suggestive Figure 4G that pretends that the data from the codon screen can be transferred to thousands of clinical variants. As the 34 disease-relevant transcripts are only small pieces of cDNA (ca. 200 bp) that have been overexpressed within a reporter cassette it is unlikely that one will be able to edit the respective real transcripts with the suggested editing yields in a relevant cell with the current Cas-ADAR versions (in particular version2) and the current delivery methods. It is also unclear if any of the mutations (all selected for simple-to-edit 5'-UAG codons) is really relevant for human disease (incidence, penetrance), and what editing yield might be required for therapy. Anyway, only hyperactive, off-target-prone REPAIRv1 has been used, the more precise REPAIRv2, which has a lower editing activity (similar or lower than wildtype ADAR2, see Supplementary Fig. 10), has not been characterized in this respect. Similar experiments with disease-relevant, and overexpressed cDNAs like CFTR, and PINK1 have anyway already been described before by others, however, additionally including a relevant phenotypic change.

c) Cox et al. use very high amounts of plasmids (150 ng/96 well Cas-ADAR, 300 ng/96 well guideRNA plasmid) for the editings. However, for the deep sequencing analysis they transfect only 10 ng/96 well Cas-ADAR plasmid (if understood correctly from their manuscript). One can expect that 15fold less plasmid will strongly reduce the transfection efficiency, thus the background of many untransfected cells will clearly reduce global off-target editing, while editing on a co-transfected reporter transcript (Cluc) is less affected by lowering Cas-ADAR (Cox et al. Fig S15). Nevertheless, one can expect that editing of an endogenous target (like KRAS, PPIB) will strongly suffer if less cells are transfected. If we understand the paper correctly, the editing on endogenous targets was not shown with low plasmid transfection. For the SNAP-ADAR system, however, we can much better and more homogeneously control the enzyme expression levels (by doxycycline induction) and we did show to what extent the reduction of SNAP-ADAR does change the editing at endogenous targets and at selected off-targets (see our Supplementary Figure 9).

Supplementary Table 3. Comparison SNAP-ADAR and 4λN-DD / BoxB system (Vallecillo-Viejo et al. RNA Biol 2018 & Sinnamon et al. PNAS 2017)^{a)}

	SNAP-ADAR (SA) system	4λN-DD / BoxB system
Targeting System	SNAP-tag – gRNA covalent bond SNAP-tag: human, < 200 aa gRNA: ca. 22 nt, chemically stabilized	λN / BoxB RNA peptide interaction λN (typically 4 copies): bacteriophage, ca. 100 aa optional 3x NLS: ca. 30 aa gRNA: ~84 nt, genetically encoded
Deaminase tested	4 enzymes fully tested: ADAR1 and ADAR2 each wildtype and E488Q	several versions, all based on ADAR2 deaminase domain, either wt or E488Q in combination with 1-4 copies λN peptide, with and without NLS 4 copies λN increase efficiency; 3xNLS can reduce off-target editing by ca. 50%
Delivery	SNAP-ADAR: single genomic copy, inducible gRNA: lipofection of chemically stabilized gRNA (22 nt)	Enzyme: currently massive overexpression via plasmid lipofection (or AAV) guideRNA: massive overexpression via plasmid lipofection (or AAV)
Editing of endogenous targets	ACTB, GAPDH, GUSB, SA, KRAS, STAT1	This system has mainly been characterized with reporter constructs, in particular GFP and CFTR; to my knowledge only a single example of an endogenous target has been described (MeCP2); the targeting of endogenous transcripts has not yet been tested systematically
Concurrent editing	3 sites or 4 endogenous housekeeping transcripts, no loss in efficiency 2 sites or 2 endogenous signaling transcripts (KRAS, STAT1), no loss in efficiency	Not shown; it is unclear if several different guideRNAs can ever be co-expressed as very high amounts of U6-guideRNA plasmids are currently used already for a single target (like 4- 15fold more than the editase plasmid)
Editing range for the best editable codon (UAG) on endogenous targets	wild-type SA: 15 - 90%, (12 sites on 6 targets, ORF & UTRs) SAQ variants: 46 - 90%, (13 sites on 6 targets, ORF & UTRs)	With the E488Q variant editing levels of 70-80% have been observed on reporter transcripts GFP and CFTR; with the wildtype enzyme editing levels typically stay below (more like 40-60%); so far only a few preferred codons have been targeted, mostly UAG and mostly in reporter transcripts
Codon scope	all 16 codons tested on an endogenous target with SA1Q and SA2Q	There is no systematic test on the full codon scope published
Applications in the manuscript	Manipulation of signaling transcripts, KRAS and STAT1, recoding of phosphorylation switch Tyr701 in STAT1	The system has been explored for the repair of CFTR (cDNA) and endogenous MeCP2
Editing duration	stable over several days	Nothing shown yet
Off-targets in gRNA/substrate duplex	the guideRNA/mRNA duplex is small (19 bp), chemical modification of guideRNA blocks off- target editing almost entirely even in A-rich codons	General: the guideRNA/mRNA duplex is large (50 bp, twice interrupted by the two 17 nt BoxB hairpins) The system suffers from major off-target editing inside the gRNA/mRNA duplex (e.g. PNAS 2017), even though endogenous MeCP2 was repaired in primary cells to ca. 75% yield, this came along with 5 off-target editings in the duplex (10-50% yield). The system also elicits strong guideRNA dependent off-target editing in the target transcript but outside the gRNA/mRNA duplex due to a proximity effect; e.g. RNA Biol 2018, depending on the enzyme 5-14 off-target editings (10-55%) have been found along the CFTR transcript

Global off-target editing	Wild-type SA: almost absent SAQ variants: moderate (≈ 1000 sites, decreased by lowering SAQ expression)	The E488Q version of Vallecillo-Viejo et al. was also tested by Cox et al. (Supporting Figure S9 in their paper) and showed massive global off-editing at rates very similar to Cas13-ADAR repairV1. We performed a re-analysis of Vallecillo-Viejo et al.'s NGS analysis with our pipeline (see Supplementary Figure 8). The wildtype enzymes elicit several hundred-fold more off-target edits compared to the wt SA. The wt Vallecillo-Viejo et al. enzymes are even more off-target-prone than our hyperactive SA1Q/SA2Q mutants. The hyperactive Vallecillo-Viejo et al. enzymes seem extremely off-target-prone.
Unique property	<ol style="list-style-type: none"> 1) Chemically stabilized guideRNAs enable proper specificity inside gRNA/mRNA duplex 2) low expression of editase enables high editing yields with reduced global off-target editing 2) clearly proven, covalent RNA targeting 3) very short guideRNA/mRNA duplex, unlikely to interfere with endogenous ADARs or translation 4) simple co-transfection of guideRNAs enables concurrent editing 	<ol style="list-style-type: none"> 1) the system is fully genetically encoded 2) the entire system (editase + 6 copies guideRNA) has been delivered as a single AAV

a) This system has already undergone several rounds of refinement. We focused on the results reported in the two most recent papers.

Supplementary Table 4. Sequences of gRNAs applied in this study. BG-conjugated gRNAs were synthesized and PAGE-purified from commercially acquired oligonucleotides containing a 5'-amino-C6 linker (BioSpring, Germany) as described by Hanswillemenke et al. (*J. Am. Chem. Soc.* **2015**, *137*, 15875-15881). Nucleotides highlighted in bold are unmodified and are placed opposite the triplet with the target adenosine in the middle. Nucleotides highlighted in italic are modified with 2'-O-methylation, those highlighted in red are 2'-fluorinated nucleotides. The backbone contains terminal phosphorothioate linkages as indicated by "s". The first three nucleotides at the 5'-end are not complementary to the mRNA substrate, but serve as linker sequence between gRNA and SNAP-tag.

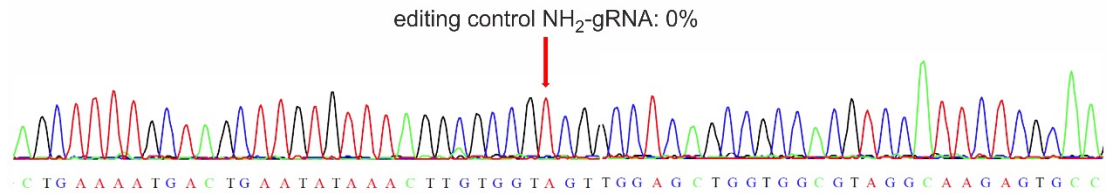
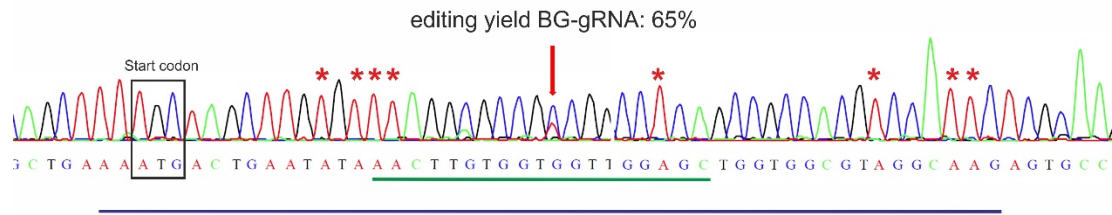
target	gRNA sequence	applied gRNA amount ^{a)}
editing of various endogenous transcripts		
5'-UTR SNAP-ADAR	5'-UsCsAUUAAACG CCA GAGUCsCsGsGsA-3'	5 pmol
5'-UTR GAPDH isoform 2	5'-UsCsUGAAUAAU CCA GGAAAsAsGsCsA-3'	5 pmol
ORF #1 GAPDH	5'-UsAsUAGGGGUG CCA AGCAGsUsUsGsG-3'	5 pmol
ORF #2 GAPDH ^{b)}	5'-UsAsUGGUUUUU CCA GACGGsCsAsGsG-3'	5 pmol
ORF #1 GUSB	5'-GsGsUGCAGAUU CCA GGUGGsGsAsCsG-3'	5 pmol
ORF #2 GUSB	5'-AsCsAGACUUGG CCA CUGAGsUsGsGsG-3'	5 pmol
3'-UTR SNAP-ADAR	5'-UsAsUGUGUCGG CCA CGGAAAsCsAsGsG-3'	5 pmol
3'-UTR GAPDH ^{c)}	5'-AsAsUAAGGGGU CCA CAUGGsCsAsA-3'	5 pmol
3'-UTR ACTB	5'-UsCsGAGCAAUG CCA UAACCsUsCsCsC-3'	5 pmol
3'-UTR GUSB	5'-UsAsUUUCCUG CCA GAAUAsGsAsUsG-3'	5 pmol
KRAS target A/1	5'-GsAsUGCUCCAA CCA CCACAsAsGsUsU-3'	SA1: 40 pmol , SA1Q: 10 pmol
KRAS target 2	5'-CsGsUCUCUUGC CCA CGCCAsCsCsAsG-3'	20 pmol
STAT1 Y701	5'-GsUsCUCUUGAU ACA UCCAGsUsUsCsC-3'	20 pmol
editing of all 16 adenosine-containing triplets in GAPDH isoform 1		
5'-GAA	5'-CsAsCAUGGGAU UCC CAUUGsAsUsGsA-3'	5 pmol
5'-GAU	5'-UsAsUCGACCAA ACC CGUUGsAsCsUsC-3'	5 pmol
5'-GAC	5'-CsAsCGUCAUGA GCC CUUCCsAsCsGsA-3'	5 pmol
5'-GAG	5'-AsAsCGAGGGAU CCC GCUCCsUsGsGsA-3'	5 pmol
5'-CAA	5'-GsAsAGAGGCUG UCG UCAUAsCsUsUsC-3'	5 pmol
5'-CAU	5'-CsAsAGAGGUCA ACG AAGGGsGsUsCsA-3'	5 pmol
5'-CAC	5'-AsAsCGCCAGGG GCG CUAAGsCsAsGsU-3'	5 pmol
5'-CAG	5'-UsAsCGCAUGGA CCG UGGUCsAsUsGsA-3'	5 pmol
5'-AAA	5'-UsAsCAUGACCC UCU UGGCUsCsCsCsC-3'	5 pmol
5'-AAU	5'-GsAsCUAGCCAA ACU CGUUGsUsCsAsU-3'	5 pmol
5'-AAC	5'-AsGsUCGCCACA GCU UCCCCGsGsAsGsG-3'	5 pmol
5'-AAG	5'-UsGsUAUAUCCA CCU UACCAsGsAsGsU-3'	5 pmol
5'-UAA	5'-AsGsGAGGGGUC UCA CUCCUsUsGsGsA-3'	5 pmol
5'-UAU	5'-CsUsAGGCAACA ACA UCCACsUsUsUsA-3'	5 pmol
5'-UAC	5'-CsCsGAGGCCA GCA GAGGCsAsGsGsG-3'	5 pmol
5'-UAG	5'-UsAsUGGUUUUU CCA GACGGsCsAsGsG-3'	5 pmol
avoiding off-target editing of neighboring adenosine		
5'-CAA methoxy	5'-GsAsAGAGGCUGU CG UCAUAsCsUsUsC-3'	5 pmol
5'-CAA fluoro	5'-GsAsAGAGGCUGU CG UCAUAsCsUsUsC-3'	5 pmol
5'-AAA methoxy	5'-UsAsCAUGACCCU CU UGGCUsCsCsCsC-3'	5 pmol
5'-AAA fluoro	5'-UsAsCAUGACCCU CU UGGCUsCsCsCsC-3'	5 pmol
5'-AAC methoxy	5'-AsGsUCGCCACA GC UCCCCGsGsAsGsG-3'	5 pmol
5'-AAC fluoro	5'-AsGsUCGCCACA GC UCCCCGsGsAsGsG-3'	5 pmol
5'-UAA methoxy	5'-AsGsGAGGGGUCU CA CUCCUsUsGsGsA-3'	5 pmol
5'-UAA fluoro	5'-AsGsGAGGGGUCU CA CUCCUsUsGsGsA-3'	5 pmol

- The indicated gRNA amounts were used for single and concurrent editings.
- This gRNA was additionally applied to test the dose dependency of RNA editing (Fig. 1c)
- This gRNA was additionally applied to test the time dependency of RNA editing (Fig. 1b)

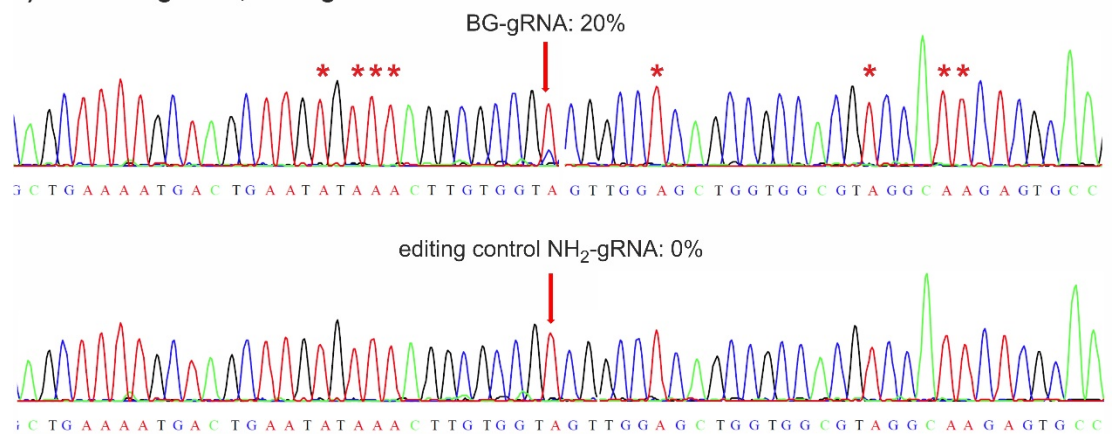
Editing of two sites in endogenous KRAS as previously reported by Cox et al. with Cas13b-ADAR

a) KRAS Target 1/A, editing with SA1Q

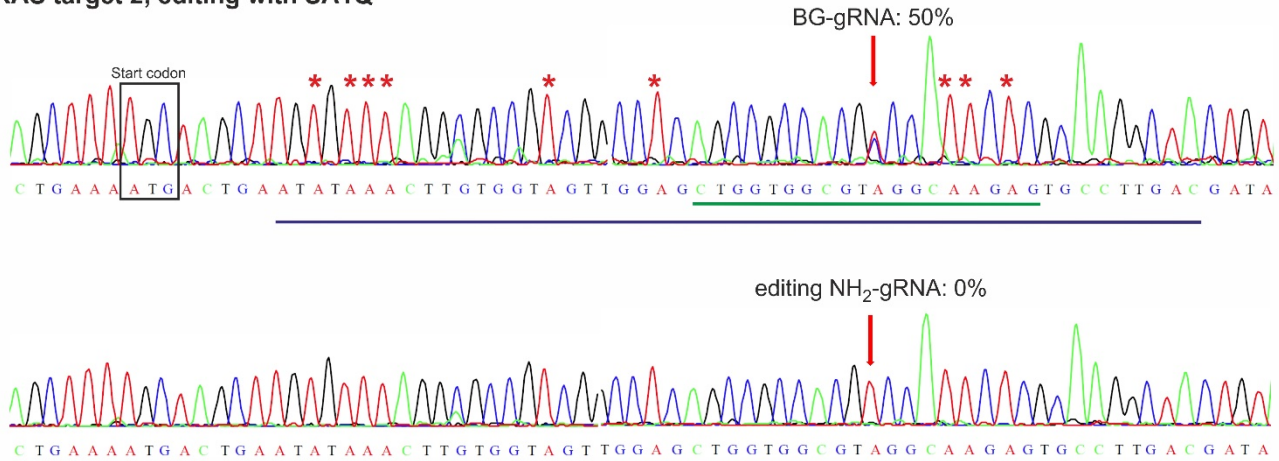
mRNA/gRNA duplex of the BG guideRNA (19 bp)
mRNA/gRNA duplex of the Cas13 guideRNA (50 bp)
* off-target sites for Cas13-ADAR



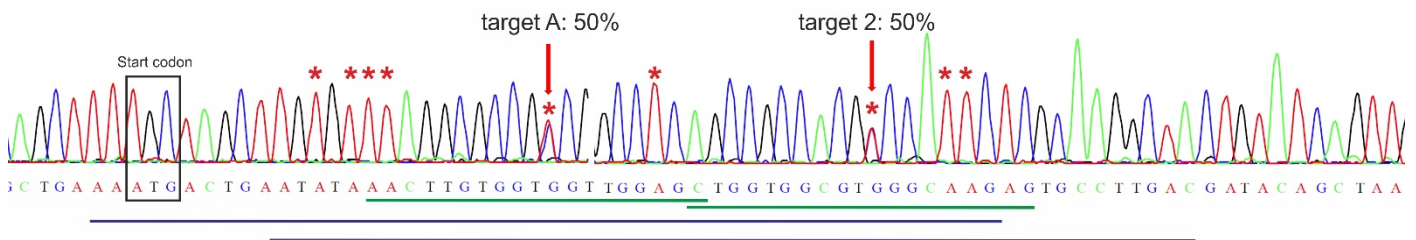
b) KRAS Target 1/A, editing with wt SA1



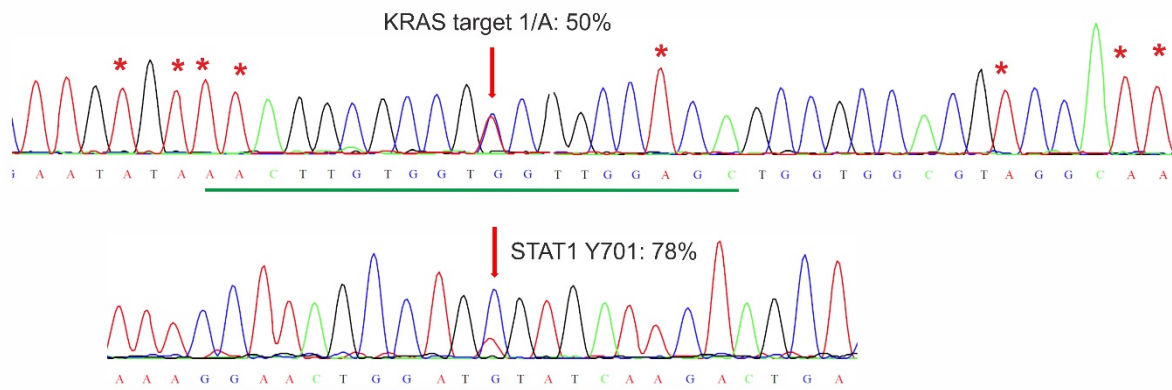
c) KRAS target 2, editing with SA1Q



d) KRAS concurrent editing target1/A and target 2 with SA1Q / BG-guideRNAs



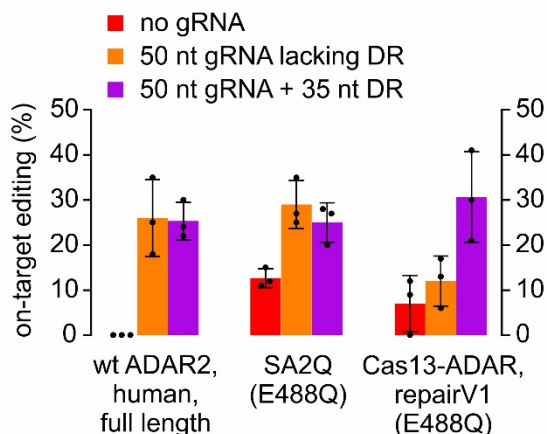
e) Concurrent editing KRAS target 1/A and STAT1 Tyr701 with SA1Q



Supplementary Note 1. Editing of KRAS target #1, #2, and STAT1 with SNAP-ADARs. Editing of KRAS target #1/A gives very high yields with SA1Q and absolutely no off-target editing at the sites reported for Cas13b-ADAR (*). Note also the large mRNA/gRNA duplexes applied for Cas13b-ADAR guideRNAs (50 bp, blue lines) versus the short ones applied for SNAP-ADAR (green lines). For target #1/A, the long Cas13 guideRNA even overlaps with the translation start site (boxed ATG) of the KRAS transcript (translation inhibition?). Also note the strong dependency of the SNAP-ADAR on the targeting mechanism. The same guideRNA lacking the BG modification (NH₂-guideRNA) cannot form the covalent bond with the deaminase and is incapable of editing the target at all (a-c). Panel a), the editing yield is significantly larger (50-65%) compared to off-target prone Cas13b-ADAR version 1 (ca. 25%). The precise wildtype SA1 edits target #1/A better than the precise Cas13-ADAR version 2 (20% versus ca. 12%). Target #2 (panel c) is also better edited by SA1Q than Cas13b version 1 (50% compared to 32%). Finally, we show efficient concurrent editing of KRAS site #1 + site #2, with yields of 50% both (d). And we show concurrent editing of KRAS site #1 with the most important phosphorylation site of STAT1 (Y701) with very good yields (50% and 78%, panel e). a-e) N=3 independent experiments were performed with similar results.

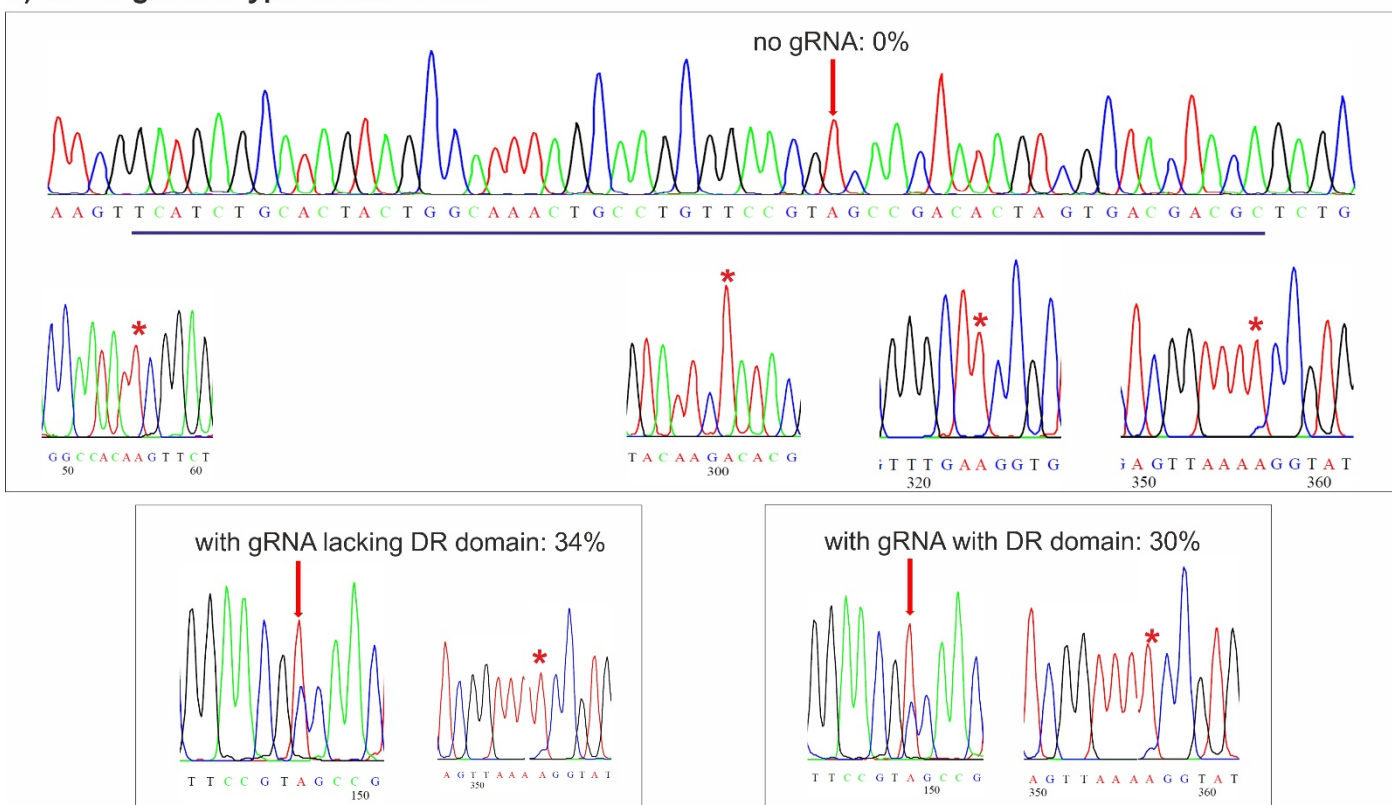
Editing of overexpressed GFP reporter W58X with co-overexpressed Cas13-guideRNAs and co-overexpressed, different ADAR fusions (SA2Q, human full length ADAR2, and Cas13-ADAR version1) following exactly the protocol given by Cox et al.

a) Overview of n=3 experiments

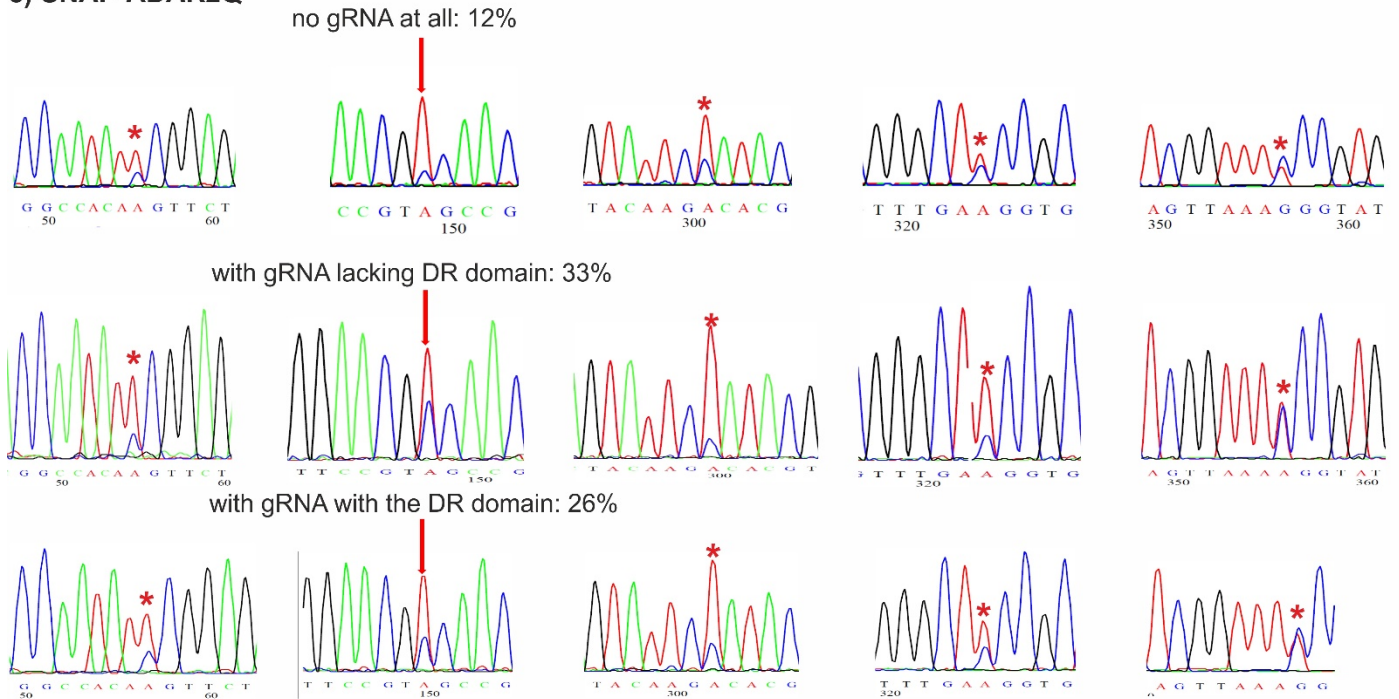


Selected Sanger sequencing traces (selected was always the trace with the highest on-target yield out of 3 experiments)

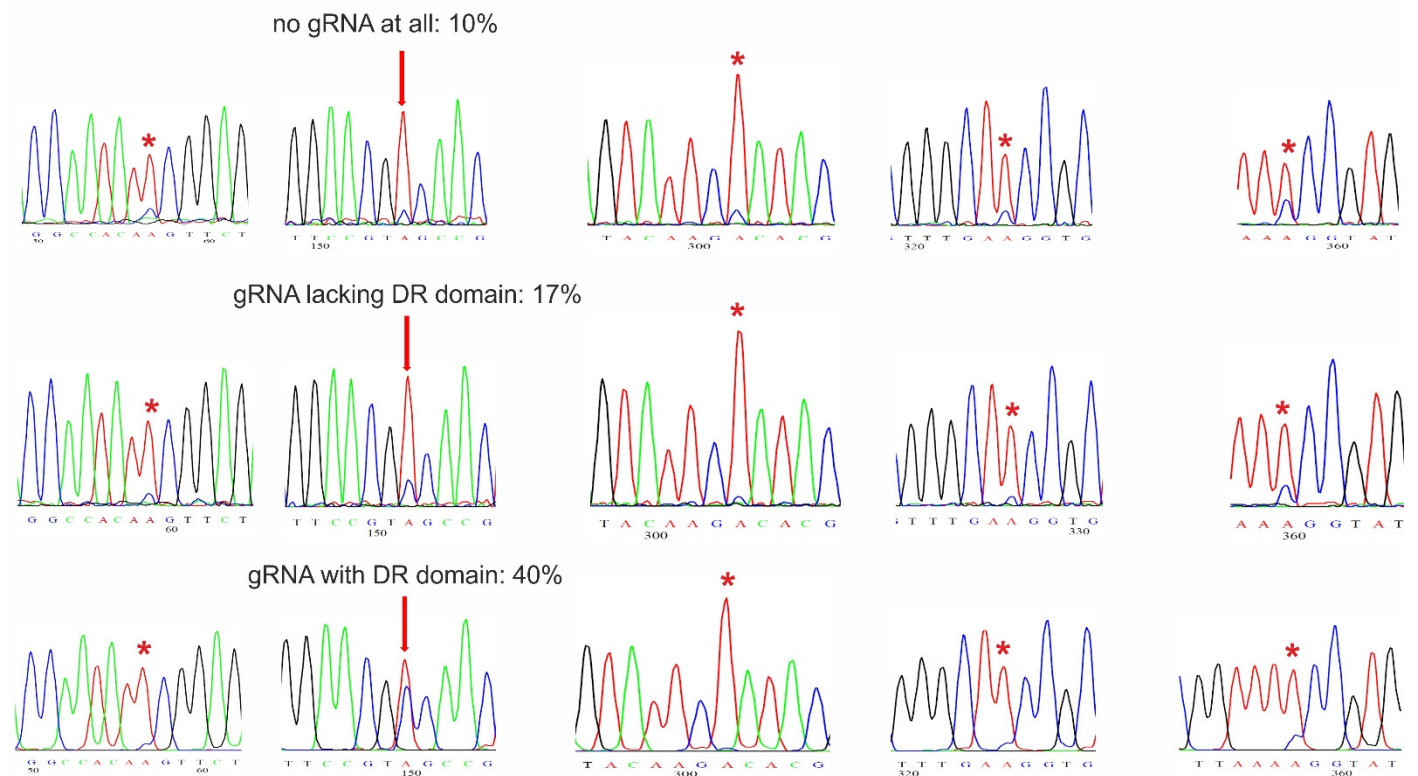
b) full length wildtype human ADAR2



c) SNAP-ADAR2Q



d) Cas-ADAR13, repairV1



Supplementary Note 2. Lacking specificity of overexpressed Cas13-guideRNAs. Cox et al. repeatedly claim a unique Cas-dependent targeting mechanism which is the reason for the claimed higher effectiveness of “repair” compared to other editing systems, the reason for the lacking codon preference they find, and the reason for the lack of a PFS dependency. However, all those claims are built on co-overexpression experiments of Cas-ADAR together with a guideRNA and reporter constructs. Here, we show that the Cas13-guideRNAs, they apply, are able to elicit editing with ADAR2 but also with SNAP-ADAR2Q in yields comparable to Cas-ADAR repair1, demonstrating that the applied guideRNAs under the applied conditions are not specific for Cas-ADAR and that many of the findings, in particular under overexpression / reporter conditions could be partly flawed by self-targeting of the deaminase (domain) itself. Unfortunately, Cox et al.

did not properly address this question in their paper (e.g. control experiments with guideRNAs lacking the DR domain are completely missing).

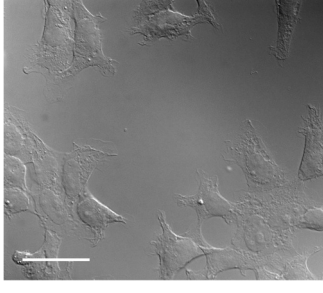
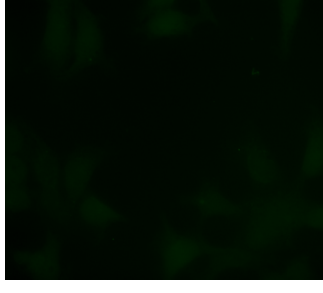
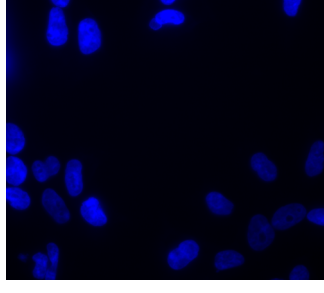
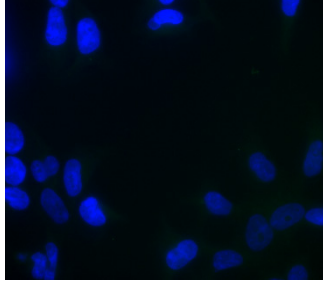
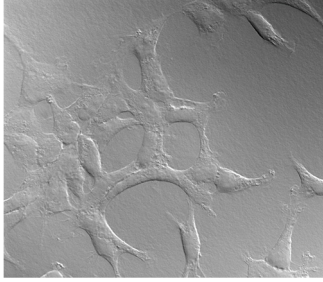
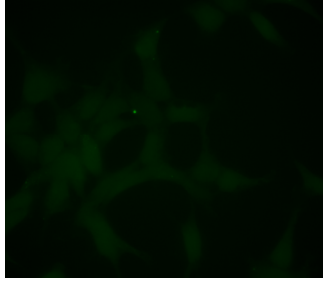
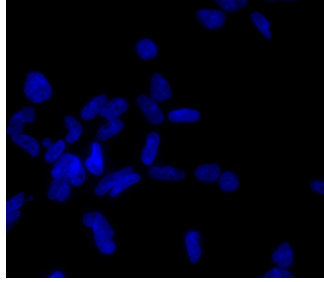
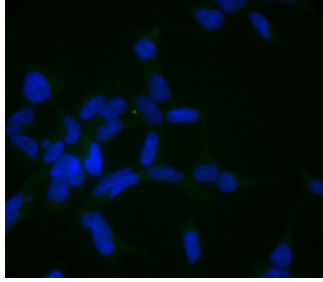
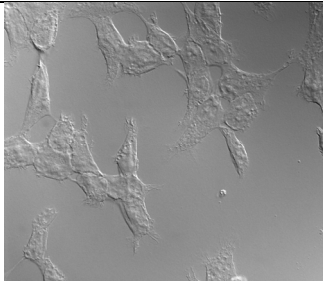

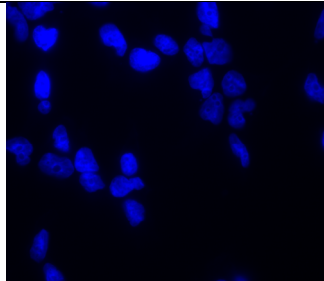
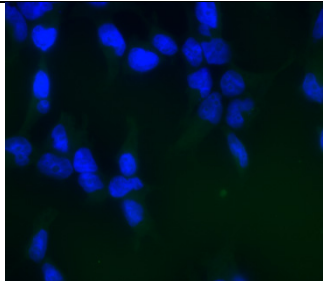
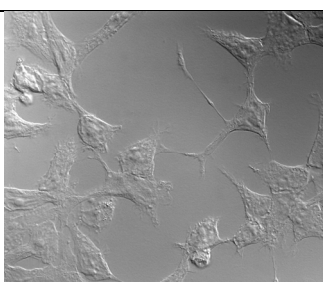
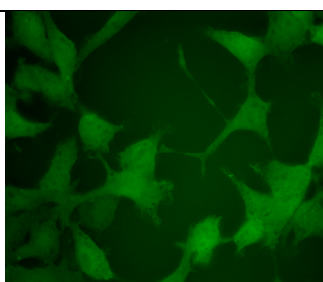
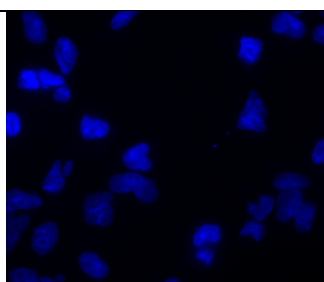
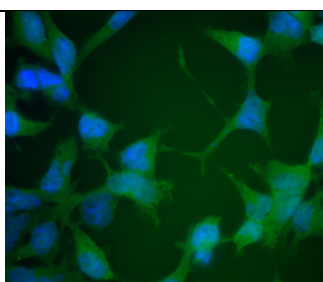
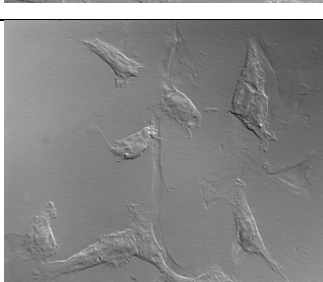
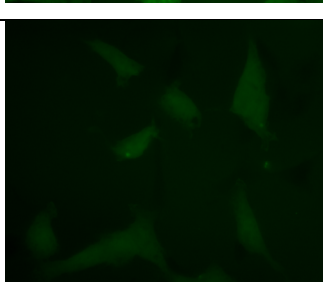
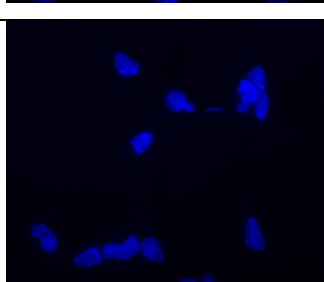
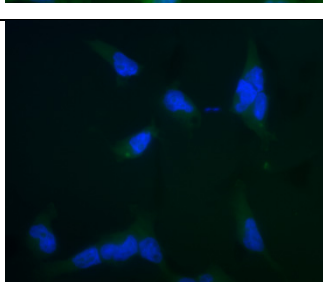
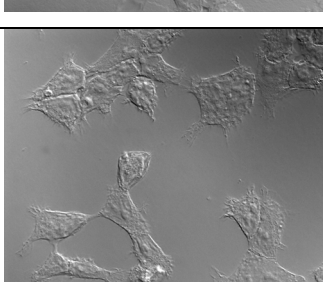
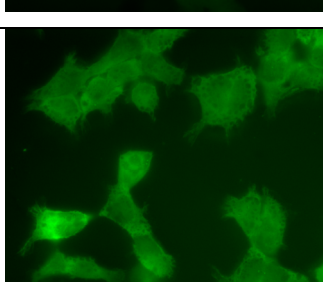
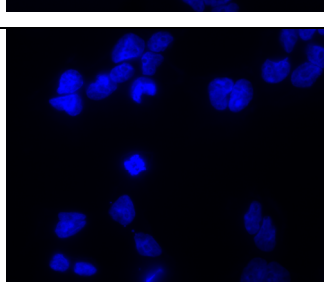
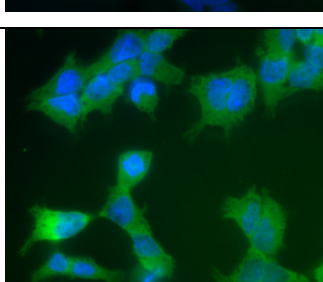
For this, we designed a Cas13 guideRNA according to Cox et al. containing a 50 nt part antisense to our GFP reporter (W58amber), putting the targeted A into mismatch with C. Mismatch position was 34. We constructed guideRNAs with the 3'-terminal DR hairpin for Cas-targeting but also lacking the DR motif (the DR motif is a 34 nt hairpin that has the function to recruit Cas13). The guideRNAs were expressed from a U6 promotor (pSilencer plasmid), as applied by Cox et al. Co-transfection was carried out as described by Cox et al.: 150 ng editing enzyme, 300 ng guideRNA vector, 40 ng GFP reporter plasmid in a coated 96 well into 293T cells. As enzymes, we co-transfected either full length human ADAR2 (wildtype), or the respective hyperactive SNAP-ADAR2Q, or Cas13-ADAR repairV1 (containing the same mutated deaminase domain of ADAR2 E488Q as SA2Q). guideRNA (antisense part: capital letters; DR domain: small letters):GCGTCACTAGTGTCGGCCACGGAACAGGCAGTTTGCCAGTAGTGAGtggtggaaggtccagtttgaggggctattacaac. In panel b), the position and length of the gRNA is indicated as a blue line under the sequence, the on-target site is marked by a red arrow, main off-target sites are marked by red asterisks.

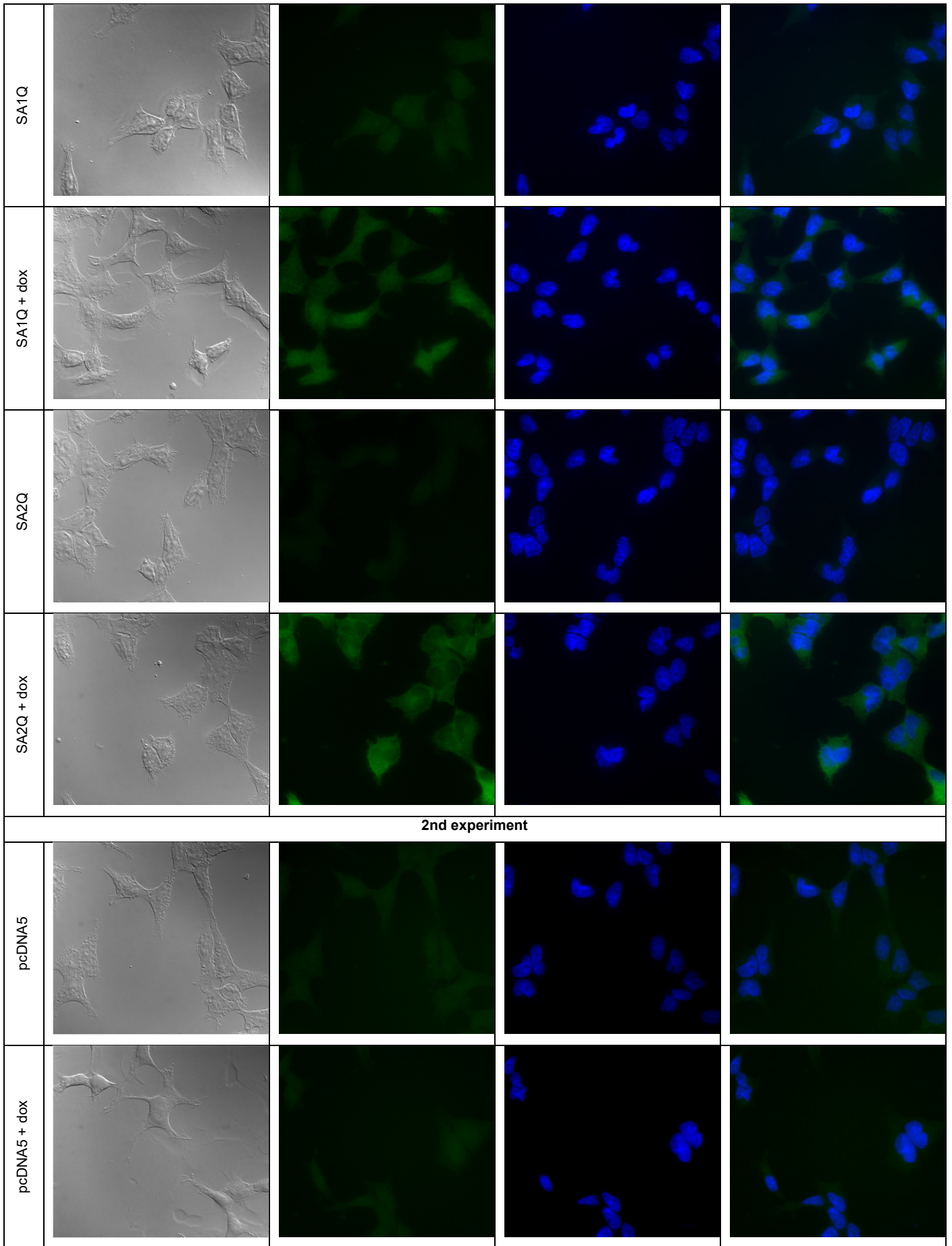
a) shows that the Cas13-guideRNA can also recruit human ADAR2 or SNAP-ADAR2Q to elicit editing yields similar to Cas13-ADAR. The average editing levels (25-30%) are very similar to those described by Cox et al. for various similar overexpression / reporter experiments in their Figures 2-4 (15-30%). As expected the recruitment of ADAR2 and SNAP-ADAR2Q is independent of the DR motif. In contrast, we have shown in the past that short chemically stabilized (BG)-guideRNAs (as we apply) are unable to recruit ADAR2 (see NAR 2016, gkw911, Figure S9A); and as we have shown repeatedly in our manuscript that SNAP-ADARs are only recruited by short chemically stabilized guideRNAs when the BG moiety is present, clearly demonstrating the SNAP-tag-dependent targeting mechanism. The editing control with Cas13-ADAR shows several interesting things. First, editing is to some extent depending on the DR motif, but second, editing also occurs without a guideRNA and also with a guideRNA lacking the DR motif, even though with reduced editing yields; this indicates that the editing yields reported by Cox et al. are composed of an unknown Cas-dependent and an unknown Cas-independent (self-targeting) part, probably differing for each respective target and condition; third, the editing yield with Cas13-ADAR with the ideal guideRNA (30%) was not notably better than that with other deaminases (25-30%); d) the off-target editing of Cas13-ADAR was higher than that of ADAR2 but lower than that of SA2Q. Finally, we want to mention that editing yields are strongly varying under co-overexpression conditions as seen in the error bars of N=3 independent experiments (Data are shown with the mean±SD, black dots represent individual data points). This is in agreement with our earlier experience.

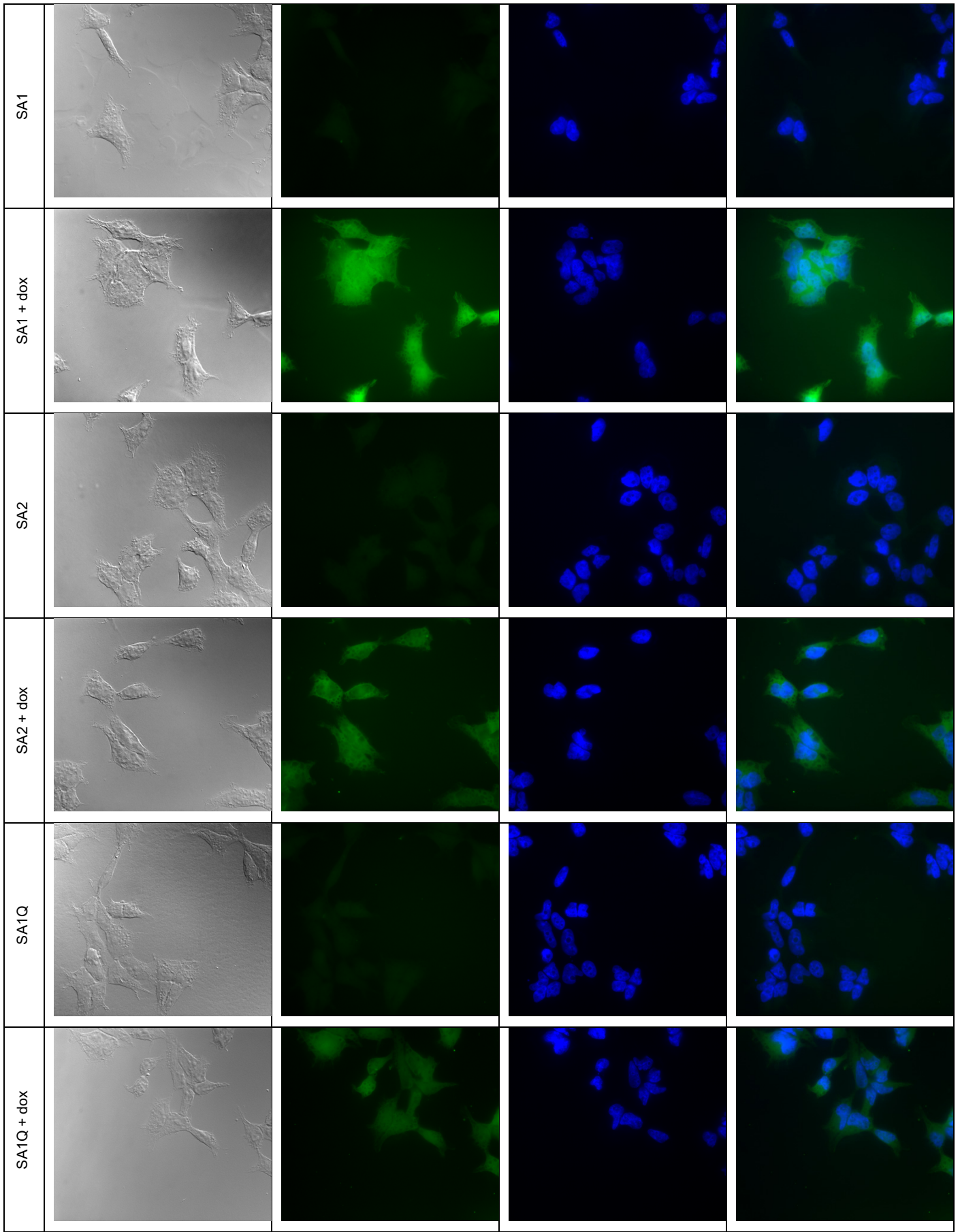
b-d) show selected Sanger sequencing traces (always the trace with the highest on-target editing yield was chosen) to give an idea of off-target editing. While ADAR2 (**b**) gives decent on-target editing (25%) there was only very little off-target editing seen and on-target editing was fully dependent on the presence of the guideRNA, even though not on the DR motif in the guideRNA. The respective single off-target editing site was described before by us (NAR 2017). Co-transfection with hyperactive SA2Q (**c**) largely shows the misery of overexpressing hyperactive deaminases (like Cas13-ADAR repairV1 too): even in absence of the guideRNA, there is massive off-target editing all over the transcript (only few sites are picked here). On-target editing was achieved with 10% yield if though no gRNA was transfected. With the Cas13-guideRNA, on-target editing increased to 25%, independent of the DR-motif. With respect to off-target editing, the experiment with Cas13-ADAR overexpression (**d**) shows results similar to the overexpression of SA2Q, which contains the same ADAR deaminase mutant (E488Q). Off-target editing is found all over the transcript, on-target editing is already found prior to the expression of the guideRNA. However, such off-target yields are roughly half that strong as found for SA1Q, which might be due to lower expression levels. After adding the guideRNA, editing levels increase and there is a targeting effect, however, there is also a notable increase in editing yield with the guideRNA lacking the DR domain. N=3 independent experiments were performed with similar results.

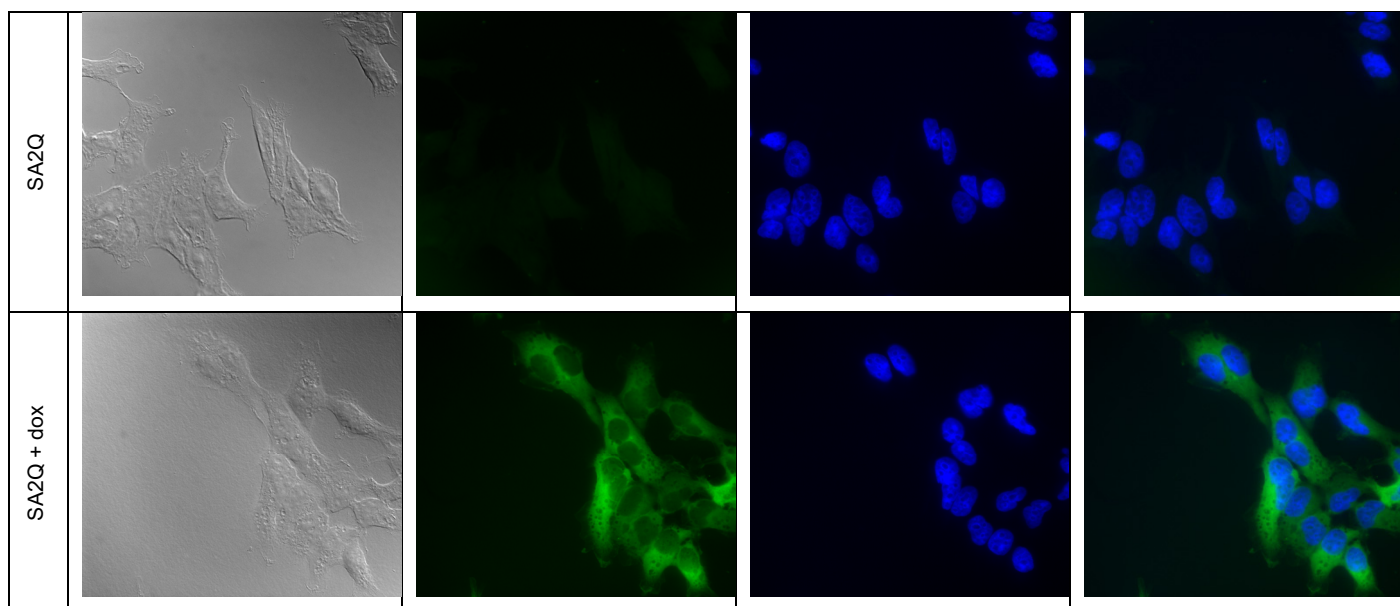
Together, panels **a-d**) suggest that the conditions (overexpression & reporters) under which Cas13-ADAR has mostly been characterized today are not sufficient to support the general claims made by Cox et al.

Determination of intracellular SNAP-ADAR localization by fluorescence microscopy

	DIC (63×)	FITC (F)	Hoechst (H)	F + H
1st experiment				
pcDNA5				
pcDNA5 + dox				
SA1				
SA1 + dox				
SA2				
SA2 + dox				



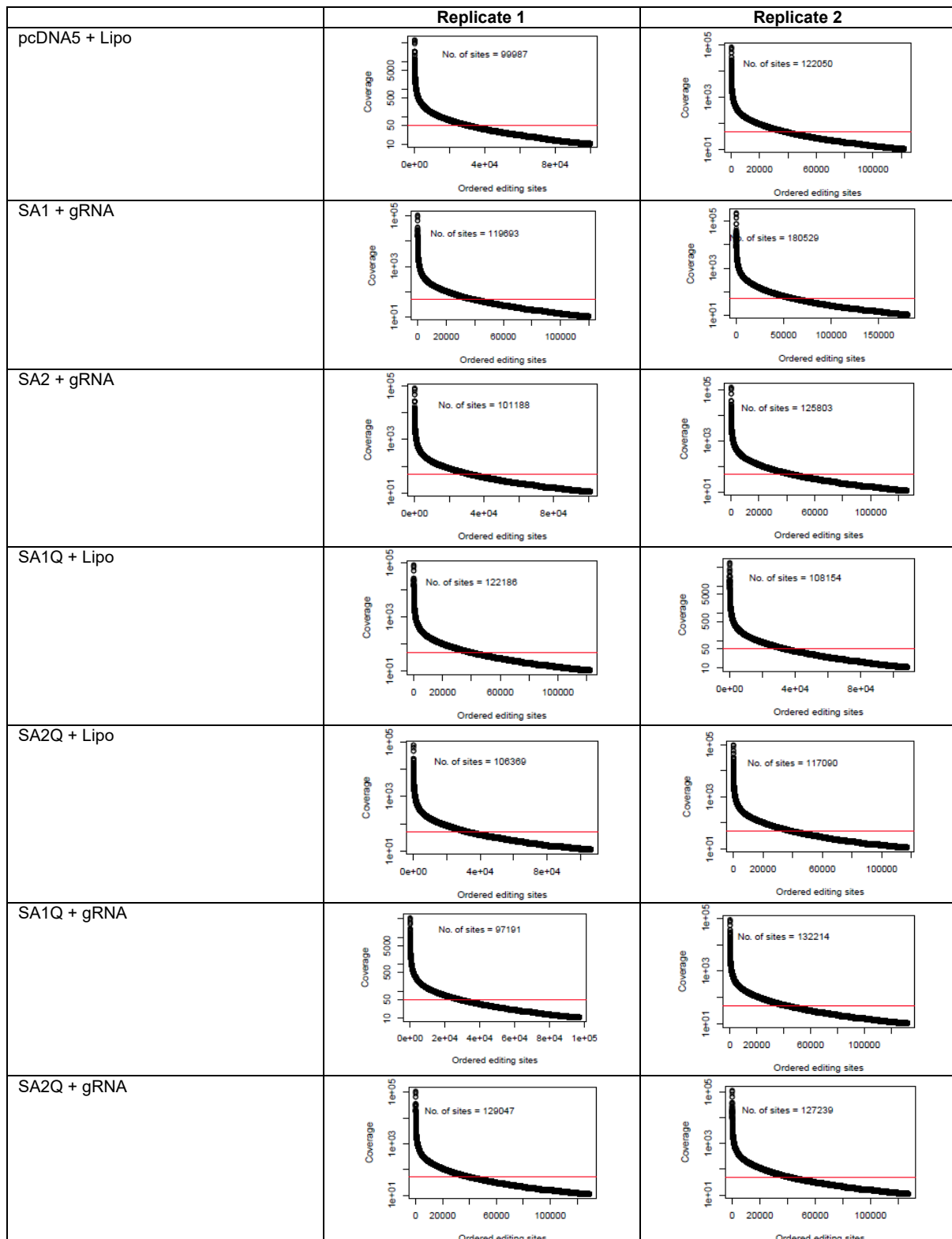




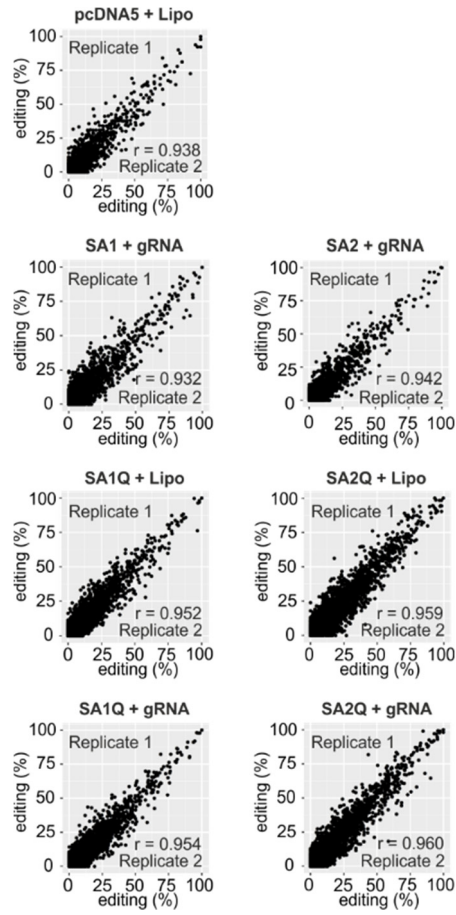
Supplementary Note 3. Protein expression was induced by doxycycline (dox) for 24 h. Cells were incubated with BG-FITC to stain SNAP-ADARs (green) and with Hoechst 33342 to stain nuclei (blue). Microscopy was performed with a Zeiss CellObserverZ1 under 630x total magnification. The scale bar represents 40 μm . FITC-BG/SNAP-tag labeling was done as described before (Vogel et al., ACS Synth. Biol. 2017, doi: 10.1021/acssynbio.7b00113). N=3 independent experiments were performed with similar results.

Supplementary Note 4 (NGS quality data, SNAP-ADAR gene sequences, target sites on endogenous transcripts)

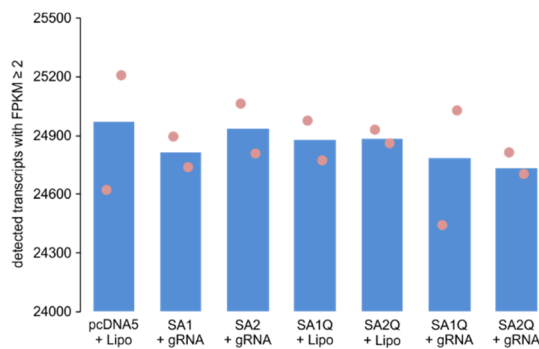
Additional NGS quality data



Detected editing sites ranked by coverage for each experiment. For testing significant editing differences, a coverage cut-off of 50 (red line) for the sum of each experiment with its replicate was applied. This typically yielded around 50.000 sites / experiment to be analyzed.



Scatter plots of editing levels of all called editing sites of replicate 1 against replicate 2 for the indicated editing experiments show good replicability with correlation ranging from 0.932-0.960.



Number of transcript covered in RNA sequencing was performed with two replicates of each sample. Shown are number of detected transcripts with a FPKM value ≥ 2 for both replicates combined (light blue bars) or separated (pink dots).

Sequences of editing enzymes and editing targets

```

10      20      30      40      50      60
1      ATGGGGAAGGTGAAGGTCGGAGTCAACGGATTTGGTCGTATTGGGCGCCTGGTCACCAGG
1      M G K V K V G V N G F G R I G R L V T R
70      80      90      100     110     120
61     GCTGCTTTTAACTCTGGTAAAGTGGATATTTGTTGCCATCAATGACCCCTTCATTGACCTC
21     A A F N S G K V D I V A I N D P F I D L
130     140     150     160     170     180
121    AACTACATGGTTTACATGTTCCAATATGATTCCACCCATGGCAAATCCATGGCACCCTC
41     N Y M V Y M F Q Y D S T H G K F H G T V
190     200     210     220     230     240
181    AAGGCTGAGAACGGGAAGCTTGTTCATCAATGAAATCCCATCACCATCTCCAGGAGCGA
61     K A E N G K L V I N G N P I T I F Q E R
250     260     270     280     290     300
241    GATCCCTCCAAAATCAAGTGGGGCGATGCTGGCGCTGAGTACGTCGTGGAGTCCACTGGC
81     D P S K I K W G D A G A E Y V V E S T G
310     320     330     340     350     360
301    GTCTTCACCACCATGGAGAAGGCTGGGGCTCATTGTCAGGGGGGAGCCAAAAGGGTCATC
101    V F T T M E K A G A H L Q G G A K R V I
370     380     390     400     410     420
361    ATCTCTGCCCCCTCTGCTGATGCCCCATGTTTCGTCATGGGTGTGAACCATGAGAAGTAT
121    I S A P S A D A P M F V M G V N H E K Y
430     440     450     460     470     480
421    GACAAAGCAGCCTCAAGATCATCAGCAATGCTCCTGCACCACCAACTGCTTAGCACCCCTG
141    D N S L K I I S N A S C T T N C L A P L
490     500     510     520     530     540
481    GCCAAGGTTCATCCATGACAACCTTTGGTATCGTGGGAAGGACATCATGACCAAGTCCATGCC
161    A K V I H D N F G I V E G L M T T V H A
550     560     570     580     590     600
541    ATCACTGCCACCCAGAAGACTGTGGATGGCCCCCTCCGGGAACCTGTGGCGTGATGGCCCGC
181    I T A T Q K T V D G P S G K L W R D G R
610     620     630     640     650     660
601    GGGGCTCTCCAGAACATCATCCCTGCCTCTACTGGCGCTGCCAAGGCTGTGGGCAAGGTC
201    G A L Q N I I P A S T G A A K A V G K V
670     680     690     700     710     720
661    ATCCCTGAGCTGAACGGGAAGCTCACTGGCATGGCCTTCCCGTGTCCCCACTGCCAACGTG
221    I P E L N G K L T G M A F R V P T A N V
730     740     750     760     770     780
721    TCAGTGGTGGACCTGACCTGCCGTCAGAAAAACCTGCCAAATATGATGACATCAAGAAG
241    S V V D L T C R L E K P A K Y D D I K K
790     800     810     820     830     840
781    GTGGTGAAGCAGGCGTCGGAGGGCCCCCTCAAGGGCATCCTGGGCTACACTGAGCACCAG
261    V V K Q A S E G P L K G I L G Y T E H Q
850     860     870     880     890     900
841    GTGGTCTCCTCTGACTTCAACAGCGACACCCACTCCTCCACCTTTGACGCTGGGGCTGGC
281    V V S S D F N S D T H S S T F D A G A G
910     920     930     940     950     960
901    ATTGCCCTCAACGACCACCTTGTCAAGCTCATTTCCTGGTATGACAACGAATTGGCTAC
301    I A L N D H F V K L I S W Y D N E F G Y
970     980     990     1000
961    AGCAACAGGGTGGTGGACCTCATGGCCACATGGCCTCCAAGGACTAA
321    S N R V V D L M A H M A S K E *

```

Open reading frame of GAPDH transcript isoform 1 (NM_002046.5). All 16 adenosine-containing triplets (yellow and cyan) were tested for editing. Most of the triplets (yellow), sites could be chosen with no resulting amino acid change. Only for 4 triplets (cyan), editing of the corresponding site lead to amino acid change. However, these changes happen in the variable region of the protein and thus, are supposed not to disturb protein activity.

10 20 30 40 50 60
 1 GGAGACGCCATCCACGCTGTTTTGACCTCCATAGAAAGACACCGGGACCGATCCAGCCTCC
 1
 70 80 90 100 110 120
 61 GGACTCTAGCGTTTAACTTAAGCTTGGTACCGAGCTCGGATCCACCATGGACAAAGACT
 20 M D K D
 130 140 150 160 170 180
 121 GCGAAATGAAGCGCACCACCTGGATAGCCCTCTGGGCAAGCTGGAAGTGTCTGGGTGCG
 40 C E M K R T T L D S P L G K L E L S G C
 190 200 210 220 230 240
 181 AACAGGGCTGCACCGTATCATCTTCTGGGCAAGGAACATCTGCCGCCGACGCCGTGG
 60 E Q G L H R I I F L G K G T S A A D A V
 250 260 270 280 290 300
 241 AAGTGCCTGCCCCAGCCGCGTGTGGGCGGACCAGAGCCACTGATGCAGGCCACCGCCT
 80 E V P A P A A V L G G P E P L M Q A T A
 310 320 330 340 350 360
 301 GGCTCAACGCCTACTTTCCACGCTGAGGCCATCGAGGAGTTCCCTGTGCCAGCCCTGC
 100 W L N A Y F H Q P E A I E E F P V P A L
 370 380 390 400 410 420
 361 ACCACCCAGTGTTCACGAGGAGAGCTTTACCCGCCAGGTGCTGTGGAAGTGTGAAAG
 120 H H P V F Q Q E S F T R Q V L W K L L K
 430 440 450 460 470 480
 421 TGGTGAAGTTCGAGAGGTCATCAGCTACAGCCACCTGGCCGCCCTGGCCGGCAATCCCG
 140 V V K F G E V I S Y S H L A A L A G N P
 490 500 510 520 530 540
 481 CCGCCACCGCCGCGTGAACCCGCTGAGCGGAAATCCCGTGCCCATCTGTATCCCT
 160 A A T A A V K T A L S G N P V P I L I P
 550 560 570 580 590 600
 541 GCCACCGGGTGGTGCAGGCGCACCTGGACGTGGGGGGCTACGAGGGCGGGCTCGCCGTGA
 180 C H R V V Q G D L D V G G Y E G G L A V
 610 620 630 640 650 660
 601 AAGAGTGGCTGCTGGCCACGAGGGCCACAGACTGGGCAAGCCTGGGCTGGGTCTGCAG
 200 K E W L L A H E G H R L G K P G L G P A
 670 680 690 700 710 720
 661 GCGGAGGCGCCAGGGTCTGGCGGGCAGTAAGGCAGAACGCATGGGTTTCACAGAGG
 220 G G G A P G S G G S K A E R M G F T E
 730 740 750 760 770 780
 721 TAACCCAGTGACAGGGGCCAGTCTCAGAAAGTATGCTCCTCCTCAAGTCCCCAG
 240 V T P V T G A S L R R T M L L L S R S P
 790 800 810 820 830 840
 781 AAGCACAGCCAAAGACACTCCCTCTCACTGGCAGCACCTTCCATGACCAGATAGCCATGC
 260 E A Q P K T L P L T G S T F H D Q I A M
 850 860 870 880 890 900
 841 TGAGCCACCGGTGCTTCAACTCTGACTAACAGCTTCCAGCCCTCCTTGTCTCGCCGCA
 280 L S H R C F N T L T N S F Q P S L L G R
 910 920 930 940 950 960
 901 AGATTCTGGCCGCATCATTATGAAAAAGACTCTGAGGACATGGGTGTCTGTCTGTCAGCT
 300 K I L A A I I M K K D S E D M G V V V S
 970 980 990 1000 1010 1020
 961 TGGGAACAGGGAATCGCTGTGTAAGGAGATTCTCTCAGCCTAAAAGGAGAAACTGTCA
 320 L G T G N R C V K G D S L S L K G E T V
 1030 1040 1050 1060 1070 1080
 1021 ATGACTGCCATGCAGAAATAATCTCCCGGAGAGGCTTCATCAGGTTTCTCTACAGTGAGT
 340 N D C H A E I I S R R G F I R F L Y S E
 1090 1100 1110 1120 1130 1140
 1081 TAATGAAATACAACCTCCAGACTGCGAAGGATAGTATATTGAACTGCTAAGGAGAGG
 360 L M K Y N S Q T A K D S I F E P A K G G
 1150 1160 1170 1180 1190 1200
 1141 AAAAGCTCCAAATAAAAAAGACTGTGTCTATCCATCTGTATATCAGCACTGCTCCGTGTG
 380 E K L Q I K K T V S F H L Y I S T A P C
 1210 1220 1230 1240 1250 1260
 1201 GAGATGGCGCCCTCTTTGACAAGTCTGCAGCGACCGTGTATGAAAGCACAGAATCCC
 400 G D G A L F D K S C S D R A M E S T E S
 1270 1280 1290 1300 1310 1320
 1261 GCCACTACCCTGTCTTCGAGAATCCCAAACAAGGAAAGCTCCGCACCAAGGTGGAGAACG
 420 R H Y P V F E N P K Q G K L R T K V E N
 1330 1340 1350 1360 1370 1380
 1321 GAGAAGGCACAATCCCTGTGGAATCCAGTGACATGTGCTACGTGGGATGGCATTCGGC
 440 G E G T I P V E S S D I V P T W D G I R
 1390 1400 1410 1420 1430 1440
 1381 TCGGGGAGAGACTCCGTACCATGTCTGTAGTGACAAAATCCTACGCTGGAACGTGTGG
 460 L G E R L R T M S C S D K I L R W N V L
 1450 1460 1470 1480 1490 1500
 1441 GCCTGCAAGGGGCACTGTTGACCCACTTCTGCAGCCATTTATCTCAAATCTGTACAT
 480 G L Q G A L L T H F L Q P I Y L K S V T
 1510 1520 1530 1540 1550 1560
 1501 TGGGTTACCTTTTCAGCCAAGGGCATCTGACCCGTGCTATTTGCTGTCTGTGACAAGAG
 500 L G Y L F S Q G H L T R A I C C R V T R
 1570 1580 1590 1600 1610 1620
 1561 ATGGGAGTGCATTTGAGGATGGACTACGACATCCCTTTATTGTCAACCCCAAGTTG
 520 D G S A F E D G L R H P F I V N H P K V
 1630 1640 1650 1660 1670 1680

```

1621 GCAGAGTCAGCATATATGATTCCAAAAGGCAATCCGGGAAGACTAAGGAGACAAGCGTCA
540 G R V S I Y D S K R Q S G K T K E T S V
      1690 1700 1710 1720 1730 1740
1681 ACTGGTGTCTGGCTGATGGCTATGACCTGGAGATCCTGGACGGTACCAGAGGCACTGTGG
560 N W C L A D G Y D L E I L D G T R G T V
      1750 1760 1770 1780 1790 1800
1741 ATGGGCCACGGAATGAATTGTCCCGGGTCTCCAAAAGAACATTTTCTTCTATTTAAGA
580 D G P R N E L S R V S K K N I F L L F K
      1810 1820 1830 1840 1850 1860
1801 AGCTCTGCTCCTTCCGTTACCGCAGGGATCTACTGAGACTCTCCTATGGTGAGGCCAAGA
600 K L C S F R Y R R D L L R L S Y G E A K
      1870 1880 1890 1900 1910 1920
1861 AAGCTGCCCGTGACTACGAGACGGCCAAGAAGTACTTCAAAAAGGCCTGAAGGATATGG
620 K A A R D Y E T A K N Y F K K G L K D M
      1930 1940 1950 1960 1970 1980
1921 GCTATGGGAACTGGATTAGCAAACCCAGGAGGAAAAGAACTTTTATCTCTGCCAGTAT
640 G Y G N W I S K P Q E E K N F Y L C P V
      1990 2000 2010 2020 2030 2040
1981 CTAGATGACTGCCTGTTCCGTAGCCGACACGGGCCCGTTTAAACCCGCTGATCAGCCTCG
660 S R *
      2050 2060 2070 2080 2090 2100
2041 ACTGTGCCTTCTAGTTGCCAGCCATCTGTTGTTTGCCCTCCCCCGTGCCTTCCTTGACC
680

```

Sequence of SNAP-ADAR1 as expressed from the 293 genome with chosen editing sites (yellow).

10 20 30 40 50 60
1 GGAGACGCCATCCACGCTGTTTTGACCTCCATAGAAAGACACCGGGACCGATCCAGCCTCC
1
70 80 90 100 110 120
61 GGACTCTAGCGTTTAACTTAAGCTTGGTACCGAGCTCGGATCCACCATGGACAAAGACT
20 M D K D
130 140 150 160 170 180
121 GCGAAATGAAGCGCACCACCTGGATAGCCCTCTGGGCAAGCTGGAAGTGTCTGGGTGCG
40 C E M K R T T L D S P L G K L E L S G C
190 200 210 220 230 240
181 AACAGGCCTGCACCGTATCATCTTCTGGGCAAGGAACATCTGCCGCCGACGCCGTGG
60 E Q G L H R I I F L G K G T S A A D A V
250 260 270 280 290 300
241 AAGTGCCTGCCCCAGCCGCCGTGTGGGCGGACCAGAGCCACTGATGCAGGCCACCGCCT
80 E V P A P A A V L G G P E P L M Q A T A
310 320 330 340 350 360
301 GGCTCAACGCCTACTTTCCACGCTGAGGCCATCGAGGAGTTCCCTGTGCCAGCCCTGC
100 W L N A Y F H Q P E A I E E F P V P A L
370 380 390 400 410 420
361 ACCACCCAGTGTTCACGAGGAGAGCTTTACCCGCCAGGTGTGTGGAAACTGTGAAAG
120 H H P V F Q Q E S F T R Q V L W K L L K
430 440 450 460 470 480
421 TGGTGAAGTTCGAGAGGTCATCAGCTACAGCCACCTGGCCGCCCTGGCCGGCAATCCCG
140 V V K F G E V I S Y S H L A A L A G N P
490 500 510 520 530 540
481 CCGCCACCGCCCGCTGAAAACCGCCCTGAGCGGAAATCCCGTGCCCATCTGTATCCCT
160 A A T A A V K T A L S G N P V P I L I P
550 560 570 580 590 600
541 GCCACCGGTGGTGCAGGCGCACCTGGACGTGGGGGCTACGAGGGCGGGCTCGCCGTGA
180 C H R V V Q G D L D V G G Y E G G L A V
610 620 630 640 650 660
601 AAGAGTGGCTGTGGCCACGAGGGCCACAGACTGGGCAAGCCTGGGCTGGGTCTCGCAG
200 K E W L L A H E G H R L G K P G L G P A
670 680 690 700 710 720
661 GCGGAGGCGCCAGGGTCTGGCGGGCAGTAAGAAGCTTGCCAAGGCCCGGGCTGCGC
220 G G G A P G S G G S K K L A K A R A A
730 740 750 760 770 780
721 AGTCTGCCCTGGCCGCAATTTTTAACTTGCACCTGGATCAGACGCCATCTCGCCAGCCTA
240 Q S A L A A I F N L H L D Q T P S R Q P
790 800 810 820 830 840
781 TTCCAGTGAGGGTCTTCAGCTGCATTTACCGCAGGTTTTAGCTGACGCTGTCTCAGCC
260 I P S E G L Q L H L P Q V L A D A V S R
850 860 870 880 890 900
841 TGGTCTGGGTAAGTTTGGTGACCTGACCGACAACCTTCTCCTCCCTCACGCTCGCAGAA
280 L V L G K F G D L T D N F S S P H A R R
910 920 930 940 950 960
901 AAGTGCTGGGTGAGTCTGATGACAACAGGCACAGATGTTAAAGATGCCAAGGTGATAA
300 K V L A G V M T T G T D V K D A K V I
970 980 990 1000 1010 1020
961 GTGTTTCTACAGGAACAAATGTATTAATGGTGAATACATGAGTATCGTGGCCCTGCAT
320 S V S T G T K C I N G E Y M S D R G L A
1030 1040 1050 1060 1070 1080
1021 TAAATGACTGCCATGCAGAAATAATATCTCGGAGATCCTTGCTCAGATTTCTTTATACAC
340 L N D C H A E I I S R R S L L R F L Y T
1090 1100 1110 1120 1130 1140
1081 AACTTGAGCTTTACTTAAATAACAAGATGATCAAAAAAGATCCATTTTCAGAAATCAG
360 Q L E L Y L N N K D D Q K R S I F Q K S
1150 1160 1170 1180 1190 1200
1141 AGCGAGGGGGTTTAGGCTGAAGGAGAATGTCCAGTTTCATCTGTACATCAGCACCTCTC
380 E R G G F R L K E N V Q F H L Y I S T S
1210 1220 1230 1240 1250 1260
1201 CCTGTGGAGATGCCAGAATCTTCTCACCACATGAGCCAATCCTGGAAGAACCAGCAGATA
400 P C G D A R I F S P H E P I L E E P A D
1270 1280 1290 1300 1310 1320
1261 GACACCCAAATCGTAAAGCAAGAGGACAGCTACGGACCAAAATAGAGTCTGGTGAGGGGA
420 R H P N R K A R G Q L R T K I E S G E G
1330 1340 1350 1360 1370 1380
1321 CGATTCCAGTGCCTCCAAATGCGAGCATCCAACGTGGGACGGGTGCTGCAAGGGGAGC
440 T I P V R S N A S I Q T W D G V L Q G E
1390 1400 1410 1420 1430 1440
1381 GGCTGCTCACCATGCTCTGCAGTGACAAGATTGCACGCTGGAACGTGGTGGGCATCCAGG
460 R L L T M S C S D K I A R W N V V G I Q
1450 1460 1470 1480 1490 1500
1441 GATCCCTGCTCAGCATTTTCGTGGAGCCCATTTACTTCTCGAGCATCATCCTGGGCAGCC
480 G S L L S I F V E P I Y F S S I I L G S
1510 1520 1530 1540 1550 1560
1501 TTTACCACGGGGACCACCTTTCCAGGGCCATGTACCAGCGGATCTCCAACATAGAGGACC
500 L Y H G D H L S R A M Y Q R I S N I E D
1570 1580 1590 1600 1610 1620
1561 TGCCACCTCTACACCCTCAACAAGCCTTTGCTCAGTGGCATCAGCAATGCAGAAGCAC
520 L P P L Y T L N K P L L S G I S N A E A
1630 1640 1650 1660 1670 1680


```

1621 GGCAGCCAGGGAAGGCCCACTTCAGTGTCAACTGGACGGTAGGCGACTCCGCTATTG
540 R Q P G K A P N F S V N W T V G D S A I
      1690      1700      1710      1720      1730      1740
1681 AGGTCATCAACGCCACGACTGGGAAGGATGAGCTGGGCCGCGCTCCCGCCTGTGTAAGC
560 E V I N A T T G K D E L G R A S R L C K
      1750      1760      1770      1780      1790      1800
1741 ACGCGTTGTACTGTGCTGGATGCGTGTGCACGGCAAGGTTCCCTCCCCTTACTACGCT
580 H A L Y C R W M R V H G K V P S H L L R
      1810      1820      1830      1840      1850      1860
1801 CCAAGATTACCAAGCCCAACGTGTACCATGAGTCCAAGCTGGCGGCAAAGGAGTACCAGG
600 S K I T K P N V Y H E S K L A A K E Y Q
      1870      1880      1890      1900      1910      1920
1861 CCGCCAAGGCGGTCTGTTCACAGCCTTCATCAAGCGGGGCTGGGGCCTGGGTGGAGA
620 A A K A R L F T A F I K A G L G A W V E
      1930      1940      1950      1960      1970      1980
1921 AGCCACCGAGCAGGACCAGTTCTCACTCACGCCCTCTAGATGACTGCCTGTTCCGTAGC
640 K P T E Q D Q F S L T P S R *
      1990      2000      2010      2020      2030      2040
1981 CGACACGGGCCGTTTAAACCCGCTGATCAGCCTCGACTGTGCCTTCTAGTTGCCAGCCA
660

```

Sequence of SNAP-ADAR2 as expressed from the 293 genome with chosen editing sites (yellow).

1 10 20 30 40 50 60
1 GGAGACGCCATCCACGCTGTTTTGACCTCCATAGAAAGACACCGGGACCGATCCAGCCTCC
1
61 70 80 90 100 110 120
GGACTCTAGCGTTTAACTTAAGCTTGGTACCGAGCTCGGATCCACCATGGACAAAGACT
20 M D K D
121 130 140 150 160 170 180
GCGAAATGAAGCGCACCACCTGGATAGCCCTCTGGGCAAGCTGGAAGTGTCTGGGTGCG
40 C E M K R T T L D S P L G K L E L S G C
181 190 200 210 220 230 240
AACAGGGCTGCACCGTATCATCTTCTGGGCAAGGAACATCTGCCGCCGACGCCGTGG
60 E Q G L H R I I F L G K G T S A A D A V
241 250 260 270 280 290 300
AAGTGCCTGCCCCAGCCGCGTGTGGGCGGACCAGAGCCACTGATGCAGGCCACCGCCT
80 E V P A P A A V L G G P E P L M Q A T A
301 310 320 330 340 350 360
GGTCAACGCCTACTTTCCACGCTGAGGCCATCGAGGAGTTCCCTGTGCCAGCCCTGC
100 W L N A Y F H Q P E A I E E F P V P A L
361 370 380 390 400 410 420
ACCACCCAGTGTTCACGAGGAGAGCTTTACCCGCCAGGTGCTGTGGAAACTGCTGAAAG
120 H H P V F Q Q E S F T R Q V L W K L L K
421 430 440 450 460 470 480
TGGTGAAGTTCGAGAGGTCATCAGCTACAGCCACCTGGCCGCCCTGGCCGGCAATCCCG
140 V V K F G E V I S Y S H L A A L A G N P
481 490 500 510 520 530 540
CCGCCACCGCCGCGTGA AAAACCGCCCTGAGCGGAAATCCCGTGCCCATCTGATCCCCCT
160 A A T A A V K T A L S G N P V P I L I P
541 550 560 570 580 590 600
GCCACCGGGTGGTGCAGGCGCACCTGGACGTGGGGGGCTACGAGGGCGGGCTCGCCGTGA
180 C H R V V Q G D L D V G G Y E G G L A V
601 610 620 630 640 650 660
AAGAGTGGCTGCTGGCCACGAGGGCCACAGACTGGGCAAGCCTGGGCTGGGTCTGCGAG
200 K E W L L A H E G H R L G K P G L G P A
661 670 680 690 700 710 720
GCGGAGGCGCCAGGGTCTGGCGGGCAGTAAGGCAGAACGCATGGGTTTCACAGAGG
220 G G G A P G S G G S K A E R M G F T E
721 730 740 750 760 770 780
TAACCCAGTGACAGGGGCCAGTCTCAGAAACTATGCTCCTCCTCAAGTCCCCAG
240 V T P V T G A S L R R T M L L L S R S P
781 790 800 810 820 830 840
AAGCACAGCCAAAGACACTCCCTCTCACTGGCAGCACCTTCCATGACCAGATAGCCATGC
260 E A Q P K T L P L T G S T F H D Q I A M
841 850 860 870 880 890 900
TGAGCCACCGGTGCTTCAACACTCTGACTAACAGCTTCCAGCCCTCCTTGTCTCGCCGCA
280 L S H R C F N T L T N S F Q P S L L G R
901 910 920 930 940 950 960
AGATTCTGGCCGCATCATTATGAAAAAGACTCTGAGGACATGGGTGTCTGTCGTGAGCT
300 K I L A A I I M K K D S E D M G V V V S
961 970 980 990 1000 1010 1020
TGGGAACAGGGAATCGCTGTGTA AAAAGGAGATTCTCTCAGCCTAAAAGGAGAAACTGTCA
320 L G T G N R C V K G D S L S L K G E T V
1021 1030 1040 1050 1060 1070 1080
ATGACTGCCATGCAGAAATAATCTCCCGGAGAGGCTTCATCAGGTTTCTCTACAGTGAGT
340 N D C H A E I I S R R G F I R F L Y S E
1081 1090 1100 1110 1120 1130 1140
TAATGAAATACAACCTCCAGACTGCGAAGGATAGTATATTTGAACCTGCTAAGGGAGAG
360 L M K Y N S Q T A K D S I F E P A K G G
1141 1150 1160 1170 1180 1190 1200
AAAAGCTCCAAATAAAAAAGACTGTGTCTATCCATCTGTATATCAGCACTGCTCCGTGTG
380 E K L Q I K K T V S F H L Y I S T A P C
1201 1210 1220 1230 1240 1250 1260
GAGATGGCGCCCTCTTTGACAAGTCTGCAGCGACCGTGTATGGAAAGCACAGAATCCC
400 G D G A L F D K S C S D R A M E S T E S
1261 1270 1280 1290 1300 1310 1320
GCCACTACCCTGTCTTCGAGAATCCCAAACAAGGAAAGCTCCGCACCAAGGTGGAGAACG
420 R H Y P V F E N P K Q G K L R T K V E N
1321 1330 1340 1350 1360 1370 1380
GACAAGGCACAATCCCTGTGGAATCCAGTGACATGTGCTACGTGGGATGGCATTCGGC
440 G Q G T I P V E S S D I V P T W D G I R
1381 1390 1400 1410 1420 1430 1440
TCGGGGAGAGACTCCGTACCATGTCTGTAGTGACAAAATCCTACGCTGGAACGTGCTGG
460 L G E R L R T M S C S D K I L R W N V L
1441 1450 1460 1470 1480 1490 1500
GCCTGCAAGGGGCACTGTTGACCCACTTCTGCAGCCCATTTATCTCAAATCTGTACAT
480 G L Q G A L L T H F L Q P I Y L K S V T
1501 1510 1520 1530 1540 1550 1560
TGGGTTACCTTTTCAGCCAAGGGCATCTGACCCGTGCTATTTGCTGTCTGTGACAAGAG
500 L G Y L F S Q G H L T R A I C C R V T R
1561 1570 1580 1590 1600 1610 1620
ATGGGAGTGCATTTGAGGATGGACTACGACATCCCTTTATTGTCAACCCCAAGTTG
520 D G S A F E D G L R H P F I V N H P K V
1630 1640 1650 1660 1670 1680

```

1621 GCAGAGTCAGCATATATGATTCCAAAAGGCAATCCGGGAAGACTAAGGAGACAAGCGTCA
540 G R V S I Y D S K R Q S G K T K E T S V
      1690 1700 1710 1720 1730 1740
1681 ACTGGTGTCTGGCTGATGGCTATGACCTGGAGATCCTGGACGGTACCAGAGGCACTGTGG
560 N W C L A D G Y D L E I L D G T R G T V
      1750 1760 1770 1780 1790 1800
1741 ATGGGCCACGGAATGAATTGTCCCGGGTCTC AAAAAGAACATTTTCTTCTATTTAAGA
580 D G P R N E L S R V S K K N I F L L F K
      1810 1820 1830 1840 1850 1860
1801 AGCTCTGCTCCTCCGTTACCGCAGGGATCTACTGAGACTCTCCTATGGTGAGGCCAAGA
600 K L C S F R Y R R D L L R L S Y G E A K
      1870 1880 1890 1900 1910 1920
1861 AAGCTGCCCGTGACTACGAGACGGCCAAGA AACTACTTCAAAAAGGCCTGAAGGATATGG
620 K A A R D Y E T A K N Y F K K G L K D M
      1930 1940 1950 1960 1970 1980
1921 GCTATGGGAACTGGATTAGCAAACCCCA GAGGAAAAGAACTTTTATCTCTGCCCAGTAT
640 G Y G N W I S K P Q E E K N F Y L C P V
      1990 2000 2010 2020 2030 2040
1981 CTAGATGACTGCCTGTTCCGTAG CCGACACGGGCCCGTTTAAACCCGCTGATCAGCCTCG
660 S R *
      2050 2060 2070 2080 2090 2100
2041 ACTGTGCCTTCTAGTTGCCAGCCATCTGTTGTTTGCCCTCCCGTGCCTTCCTTGACC
680

```

Sequence of SNAP-ADAR1Q as expressed from the 293 genome with chosen editing sites (yellow). E/Q site is highlighted in cyan.

10 20 30 40 50 60
1 GGAGACGCCATCCACGCTGTTTTGACCTCCATAGAAAGACACCGGGACCGATCCAGCCTCC
1
70 80 90 100 110 120
61 GGACTCTAGCGTTTAACTTAAGCTTGGTACCGAGCTCGGATCCACCATGGACAAAGACT
20 M D K D
130 140 150 160 170 180
121 GCGAAATGAAGCGCACCACCTGGATAGCCCTCTGGGCAAGCTGGAAGTGTCTGGGTGCG
40 C E M K R T T L D S P L G K L E L S G C
190 200 210 220 230 240
181 AACAGGCCTGCACCGTATCATCTTCTGGGCAAAGGAACATCTGCCGCCGACGCCGTGG
60 E Q G L H R I I F L G K G T S A A D A V
250 260 270 280 290 300
241 AAGTGCCTGCCCCAGCCGCGTGTGGGCGGACCAGAGCCACTGATGCAGGCCACCGCCT
80 E V P A P A A V L G G P E P L M Q A T A
310 320 330 340 350 360
301 GGCTCAACGCCTACTTTCCACGCTGAGGCCATCGAGGAGTTCCCTGTGCCAGCCCTGC
100 W L N A Y F H Q P E A I E E F P V P A L
370 380 390 400 410 420
361 ACCACCCAGTGTTCACGAGGAGAGCTTTACCCGCCAGGTGCTGTGGAAACTGTGAAAG
120 H H P V F Q Q E S F T R Q V L W K L L K
430 440 450 460 470 480
421 TGGTGAAGTTCGAGAGGTCATCAGCTACAGCCACCTGGCCGCTGGCCGCAATCCCG
140 V V K F G E V I S Y S H L A A L A G N P
490 500 510 520 530 540
481 CCGCCACCGCCGCGTGA AAAACCGCCCTGAGCGGAAATCCCGTGCCCATCTGATCCCT
160 A A T A A V K T A L S G N P V P I L I P
550 560 570 580 590 600
541 GCCACCGGTGGTGCAGGCGCACCTGGACGTGGGGGCTACGAGGGCGGGCTCGCCGTGA
180 C H R V V Q G D L D V G G Y E G G L A V
610 620 630 640 650 660
601 AAGAGTGGCTGCTGGCCACGAGGGCCACAGACTGGGCAAGCCTGGGCTGGTCTCGCAG
200 K E W L L A H E G H R L G K P G L G P A
670 680 690 700 710 720
661 GCGGAGGCGCCAGGGTCTGGCGGGCAGTAAGAAGCTTGCCAAGGCCCGGGCTGCGC
220 G G G A P G S G G G S K K L A K A R A A
730 740 750 760 770 780
721 AGTCTGCCCTGGCCGCAATTTTTAACTTGCACTGGATCAGACGCCATCTCGCCAGCCTA
240 Q S A L A A I F N L H L D Q T P S R Q P
790 800 810 820 830 840
781 TTCCAGTGAGGGTCTTCAGCTGCATTTACCGCAGGTTTTAGCTGACGCTGTCTCAGCC
260 I P S E G L Q L H L P Q V L A D A V S R
850 860 870 880 890 900
841 TGGTCTGGGTAAGTTTGGTGACCTGACCGACAACCTTCTCCTCCCTCACGCTCGCAGAA
280 L V L G K F G D L T D N F S S P H A R R
910 920 930 940 950 960
901 AAGTGTGGCTGGAGTCGTCATGACAACAGGCACAGATGTTAAAGATGCCAAGGTGATAA
300 K V L A G V M T T G T D V K D A K V I
970 980 990 1000 1010 1020
961 GTGTTTCTACAGGAACAAATGTATTAATGGTGAATACATGAGTATCGTGGCCTGTCAT
320 S V S T G T K C I N G E Y M S D R G L A
1030 1040 1050 1060 1070 1080
1021 TAAATGACTGCCATGCAGAAATAATATCTCGGAGATCCTTGCTCAGATTTCTTTATACAC
340 L N D C H A E I I S R R S L L R F L Y T
1090 1100 1110 1120 1130 1140
1081 AACTTGAGCTTTACTTAAATAACAAGATGATCAAAAAAGATCCATCTTCAGAAATCAG
360 Q L E L Y L N N K D D Q K R S I F Q K S
1150 1160 1170 1180 1190 1200
1141 AGCGAGGGGGTTTAGGCTGAAGGAGAATGTCCAGTTTCATCTGTACATCAGCACCTCTC
380 E R G G F R L K E N V Q F H L Y I S T S
1210 1220 1230 1240 1250 1260
1201 CCTGTGGAGATGCCAGAATCTTCTCACCACATGAGCCAATCCTGGAAGAACCAGCAGATA
400 P C G D A R I F S P H E P I L E E P A D
1270 1280 1290 1300 1310 1320
1261 GACACCCAAATCGTAAAGCAAGAGGACAGCTACGGACCAAAATAGAGTCTGGT CAGGGGA
420 R H P N R K A R G Q L R T K I E S G Q G
1330 1340 1350 1360 1370 1380
1321 CGATTCCAGTGCCTCCAAATGCGAGCATCCAACGTGGGACGGGTGCTGCAAGGGGAGC
440 T I P V R S N A S I Q T W D G V L Q G E
1390 1400 1410 1420 1430 1440
1381 GGCTGCTCACCATGTCCTGCAGTGACAAGATTGCACGCTGGAACGTGGTGGGCATCCAGG
460 R L L T M S C S D K I A R W N V V G I Q
1450 1460 1470 1480 1490 1500
1441 GATCCCTGCTCAGCATTTTCGTGGAGCCCATTTACTTCTCGAGCATCATCCTGGGCAGCC
480 G S L L S I F V E P I Y F S S I I L G S
1510 1520 1530 1540 1550 1560
1501 TTTACCACGGGGACCACCTTTCCAGGGCCATGTACCAGCGGATCTCCAACATAGAGGACC
500 L Y H G D H L S R A M Y Q R I S N I E D
1570 1580 1590 1600 1610 1620
1561 TGCCACCTCTACACCCTCAACAAGCCTTTGCTCAGTGGCATCAGCAATGCAGAAGCAC
520 L P P L Y T L N K P L L S G I S N A E A
1630 1640 1650 1660 1670 1680

```

1621 GGCAGCCAGGGAAGGCCCACTTCAGTGTCAACTGGACGGTAGGCGACTCCGCTATTG
540 R Q P G K A P N F S V N W T V G D S A I
      1690 1700 1710 1720 1730 1740
1681 AGGTCATCAACGCCACGACTGGGAAGGATGAGCTGGGCCGCGCTCCCGCCTGTGTAAGC
560 E V I N A T T G K D E L G R A S R L C K
      1750 1760 1770 1780 1790 1800
1741 ACGCGTTGTACTGTGCTGGATGCGTGTGCACGGCAAGGTTCCCTCCCCTTACTACGCT
580 H A L Y C R W M R V H G K V P S H L L R
      1810 1820 1830 1840 1850 1860
1801 CCAAGATTACCAAGCCCAACGTGTACCATGAGTCCAAGCTGGCGGCAAAGGAGTACCAGG
600 S K I T K P N V Y H E S K L A A K E Y Q
      1870 1880 1890 1900 1910 1920
1861 CCGCCAAGGCGCGTCTGTTCACAGCCTTCATCAAGCGGGGCTGGGGCCTGGGTGGAGA
620 A A K A R L F T A F I K A G L G A W V E
      1930 1940 1950 1960 1970 1980
1921 AGCCACCGAGCAGGACCAGTTCTCACTCACGCCCTCTAGATGACTGCCTGTTCCGTAGC
640 K P T E Q D Q F S L T P S R *
      1990 2000 2010 2020 2030 2040
1981 CGACACGGGCCCGTTTAAACCCGCTGATCAGCCTCGACTGTGCCTTCTAGTTGCCAGCCA
660

```

Sequence of SNAP-ADAR2Q as expressed from the 293 genome with chosen editing sites (yellow). E/Q site is highlighted in cyan.

10 20 30 40 50 60
1 GGAGACGCCATCCACGCTGTTTTGACCTCCATAGAAAGACACCGGGACCGATCCAGCCTCC
1
70 80 90 100 110 120
61 GGACTCTAGCGTTTTAACTTAAGCTTGGTACCGAGCTCGGATCCACCATGGACAAAGACT
20 M D K D
130 140 150 160 170 180
121 GCGAAATGAAGCGCACCACCCTGGATAGCCCTCTGGGCAAGCTGGAAGTGTCTGGGTGCG
40 C E M K R T T L D S P L G K L E L S G C
190 200 210 220 230 240
181 AACAGGGCTGCACCGTATCATCTTCTGGGCAAGGAACATCTGCCGCCGACGCCGTGG
60 E Q G L H R I I F L G K G T S A A D A V
250 260 270 280 290 300
241 AAGTGCCTGCCCCAGCCGCCGTGTGGGCGGACCAGAGCCACTGATGCAGGCCACCGCCT
80 E V P A P A A V L G G P E P L M Q A T A
310 320 330 340 350 360
301 GGCTCAACGCCTACTTTCCACGCTGAGGCCATCGAGGAGTTCCCTGTGCCAGCCCTGC
100 W L N A Y F H Q P E A I E E F P V P A L
370 380 390 400 410 420
361 ACCACCCAGTGTTCACGAGGAGAGCTTTACCCGCCAGGTGTGTGGAAACTGTGAAAG
120 H H P V F Q Q E S F T R Q V L W K L L K
430 440 450 460 470 480
421 TGGTGAAGTTCGAGAGGTCATCAGCTACAGCCACCTGGCCGCCCTGGCCGGCAATCCCG
140 V V K F G E V I S Y S H L A A L A G N P
490 500 510 520 530 540
481 CCGCCACCGCCCGCTGAAAACCGCCCTGAGCGGAAATCCCGTGCCCATCTGTATCCCT
160 A A T A A V K T A L S G N P V P I L I P
550 560 570 580 590 600
541 GCCACCGGGTGGTGCAGGCGCACCTGGACGTGGGGGGCTACGAGGGCGGGCTCGCCGTGA
180 C H R V V Q G D L D V G G Y E G G L A V
610 620 630 640 650 660
601 AAGAGTGGCTGCTGGCCACGAGGGCCACAGACTGGGCAAGCCTGGGCTGGGTCTCGCAG
200 K E W L L A H E G H R L G K P G L G P A
670 680 690 700 710 720
661 GCGGAGGCGCCAGGGTCTGGCGGGCAGTAAGAAGCTTGCCAAGGCCCGGGCTGCGC
220 G G G A P G S G G G S K K L A K A R A A
730 740 750 760 770 780
721 AGTCTGCCCTGGCCGCAATTTTTAACTTGCACTGGATCAGACGCCATCTCGCCAGCCTA
240 Q S A L A A I F N L H L D Q T P S R Q P
790 800 810 820 830 840
781 TTCCAGTGAGGGTCTTCAGCTGCATTTACCGCAGGTTTTAGCTGACGCTGTCTCACGCC
260 I P S E G L Q L H L P Q V L A D A V S R
850 860 870 880 890 900
841 TGGTCTGGGTAAGTTTGGTGACCTGACCGACAACCTTCTCCTCCCTCACGCTCGCAGAA
280 L V L G K F G D L T D N F S S P H A R R
910 920 930 940 950 960
901 AAGTGCTGGCTGGAGTCGTATGACAACAGGCACAGATGTTAAAGATGCCAAGGTGATAA
300 K V L A G V V M T T G T D V K D A K V I
970 980 990 1000 1010 1020
961 GTGTTTCTACAGGAGGCAAAATGTATTAATGGTGAATACATGAGTGATCGTGGCCCTGCAT
320 S V S T G G K C I N G E Y M S D R G L A
1030 1040 1050 1060 1070 1080
1021 TAAATGACTGCCATGCAGAAATAATATCTCGGAGATCCTTGCTCAGATTTCTTTATACAC
340 L N D C H A E I I S R R S L L R F L Y T
1090 1100 1110 1120 1130 1140
1081 AACTTGAGCTTTACTTAAATAACAAGATGATCAAAAAAGATCCATCTTCAGAAATCAG
360 Q L E L Y L N N K D D Q K R S I F Q K S
1150 1160 1170 1180 1190 1200
1141 AGCGAGGGGGTTTAGGCTGAAGGAGAATGTCCAGTTTCATCTGTACATCAGCACCTCTC
380 E R G G F R L K E N V Q F H L Y I S T S
1210 1220 1230 1240 1250 1260
1201 CCTGTGGAGATGCCAGAATCTTCTCACCACATGAGCCAATCCTGGAAGAACCAGCAGATA
400 P C G D A R I F S P H E P I L E E P A D
1270 1280 1290 1300 1310 1320
1261 GACACCCAAATCGTAAAGCAAGAGGACAGCTACGGACCAAAATAGAGTCTGGTCAGGGGA
420 R H P N R K A R G Q L R T K I E S G Q G
1330 1340 1350 1360 1370 1380
1321 CGATTCCAGTGCCTCCAAATGCGAGCATCCAACGTGGGACGGGGTGCTGCAAGGGGAGC
440 T I P V R S N A S I Q T W D G V L Q G E
1390 1400 1410 1420 1430 1440
1381 GGCTGCTCACCATGTCCTGCAGTGACAAGATTGCACGCTGGAACGTGGTGGGCATCCAGG
460 R L L T M S C S D K I A R W N V V G I Q
1450 1460 1470 1480 1490 1500
1441 GATCCCTGCTCAGCATTTTCGTGGAGCCCATTTACTTCTCGAGCATCATCCTGGGCAGCC
480 G S L L S I F V E P I Y F S S I I L G S
1510 1520 1530 1540 1550 1560
1501 TTTACCACGGGGACCACCTTTCCAGGGCCATGTACCAGCGGATCTCCAACATAGAGGACC
500 L Y H G D H L S R A M Y Q R I S N I E D
1570 1580 1590 1600 1610 1620
1561 TGCCACCTCTACACCCTCAACAAGCCTTTGCTCAGTGGCATCAGCAATGCAGAAGCAC
520 L P P L Y T L N K P L L S G I S N A E A
1630 1640 1650 1660 1670 1680

```

1621   GGCAGCCAGGGAAGGCCCAACTTCAGTGTCAACTGGACGGTAGGCGACTCCGCTATTG
540   R  Q  P  G  K  A  P  N  F  S  V  N  W  T  V  G  D  S  A  I
      1690   1700   1710   1720   1730   1740
1681   AGGTCATCAACGCCACGACTGGGAAGGATGAGCTGGGCCGCGCTCCCGCCTGTGTAAGC
560   E  V  I  N  A  T  T  G  K  D  E  L  G  R  A  S  R  L  C  K
      1750   1760   1770   1780   1790   1800
1741   ACGCGTTGTACTGTGCTGGATGCGTGTGCACGGCAAGGTTCCCTCCCCTTACTACGCT
580   H  A  L  Y  C  R  W  M  R  V  H  G  K  V  P  S  H  L  L  R
      1810   1820   1830   1840   1850   1860
1801   CCAAGATTACCAAGCCCAACGTGTACCATGAGTCCAAGCTGGCGGCAAAGGAGTACCAGG
600   S  K  I  T  K  P  N  V  Y  H  E  S  K  L  A  A  K  E  Y  Q
      1870   1880   1890   1900   1910   1920
1861   CCGCCAAGGCGCGTCTGTTCACAGCCTTCATCAAGCGGGGCTGGGGCCTGGGTGGAGA
620   A  A  K  A  R  L  F  T  A  F  I  K  A  G  L  G  A  W  V  E
      1930   1940   1950   1960   1970   1980
1921   AGCCCACCGAGCAGGACCAGTTCTCACTCACGCCCTCTAGATGACTGCCTGTTCCGTAGC
640   K  P  T  E  Q  D  Q  F  S  L  T  P  S  R  *
      1990   2000   2010   2020   2030   2040
1981   CGACACGGGCCCGTTTAAACCCGCTGATCAGCCTCGACTGTGCCTTCTAGTTGCCAGCCA
660

```

Sequence of SNAP-ADAR2QG as expressed from the 293 genome with chosen editing sites (yellow). E/Q and T/G sites are highlighted in cyan.

```

1      10      20      30      40      50      60
GGCACCGCAGGCCCGGGATGCTAGTGCGCAGCGGGTGCATCCCTGTCCGGATGCTGCGC

61     70      80      90      100     110     120
CTGCGGTAGACGGCCGCCATGTTGCAACCGGGAAGGAAATGAATGGGCAGCCGTTAGGA

121    130     140     150     160     170     180
AAGCCTGCCGGTGACTAACCCCTGCGCTCCTGCCTCGATGGGTGGAGTTCGCGTGTGGCGGG

181    190     200     210     220     230     240
GAAGTCAGGTGGAGCGAGGCTAGCTGGCCCGATTCTCTCCGGGTGATGCTTTTCCTAG

241    250     260     270     280     290     300
ATTATTCTCTGATTGGTTCGTATTGGGCGCCTGGTCACCAGGGTGTCTTTAACTCTGGT

301    310     320     330     340     350     360
AAAGTGGATATTGTTGCCATCAATGACCCCTTCATTGACCTCAACTACATGGTTTACATG
M V Y M
361    370     380     390     400     410     420
TTCCAATATGATTCCACCCATGGCAAATTCATGGCACCGTCAAGGCTGAGAACGGGAAG
121    F Q Y D S T H G K F H G T V K A E N G K
421    430     440     450     460     470     480
CTTGTCAATGAAATCCCATCACCATCTCCAGGAGCGAGATCCCTCCAAAATCAAG
141    L V I N G N P I T I F Q E R D P S K I K
481    490     500     510     520     530     540
TGGGGCGATGCTGGCGCTGAGTACGTCGTTGGAGTCCACTGGCGTCTTCACCACCATGGAG
161    W G D A G A E Y V V E S T G V F T T M E
541    550     560     570     580     590     600
AAGGCTGGGGCTCATTTGCAGGGGGGAGCCAAAAGGGTCATCATCTGCCCCCTCTGCT
181    K A G A H L Q G G A K R V I I S A P S A
601    610     620     630     640     650     660
GATGCCCCATGTTTCGTTCATGGGTGTGAACCATGAGAAGTATGACAACAGCCCTCAAGATC
201    D A P M F V M G V N H E K Y D N S L K I
661    670     680     690 #1 700     710     720
ATCAGCAATGCCTCTGCACCACCACTGCTTAGCACCCCTGGCCAAGGTCATCCATGAC
221    I S N A S C T T N C L A P L A K V I H D
721    730     740     750     760     770     780
AACTTTGGTATCGTGAAGGACTCATGACCACAGTCCATGCCATCACTGCCACCCAGAAG
241    N F G I V E G L M T T V H A I T A T Q K
781    790     800     810     820     830     840
ACTGTGGATGGCCCTCCGGAAACTGTGGCGTATGGCCGGGGGCTCTCCAGAACATC
261    T V D G P S G K L W R D G R G A L Q N I
841    850     860     870     880     890     900
ATCCCTGCCCTCTACTGGCGCTGCCAAGGCTGTGGGCAAGGTCATCCCTGAGCTGAACGGG
281    I P A S T G A A K A V G K V I P E L N G
901    910     920     930     940     950     960
AAGCTCACTGGCATGGCCTTCCGTTGCCACTGCCAACGTGTGAGTGGTGGACCTGACC
301    K L T G M A F R V P T A N V S V V D L T
961    #2 980     990     1000    1010    1020
TGCCGTC TAGAAAAA CCTGCCAAATATGATGACATCAAGAAGGTGGTGAAGCAGGCGTCG
321    C R L E K P A K Y D D I K K V V K Q A S
1021  1030    1040    1050    1060    1070    1080
GAGGGCCCCCTCAAGGGCATCTGGGCTACACTGAGCACCAGGTGGTCTCCTCTGACTTC
341    E G P L K G I L G Y T E H Q V V S S D F
1081  1090    1100    1110    1120    1130    1140
AACAGCGACACCCACTCCTCCACCTTTGACGCTGGGGCTGGCATGCCCCCAACGACCAC
361    N S D T H S S T F D A G A G I A L N D H
1141  1150    1160    1170    1180    1190    1200
TTTGTCAAGCTCATTTCCCTGGTATGACAACGAATTTGGCTACAGCAACAGGGTGGTGGAC
381    F V K L I S W Y D N E F G Y S N R V V D
1201  1210    1220    1230    1240    1250    1260
CTCATGGCCACATGGCCTCCAAGGAGTAAGACCCTGGACCACCAGCCCCAGCAAGAGC
401    L M A H M A S K E *
1261  1270    1280    1290    1300    1310    1320
ACAAGAGGAAGAGAGAGACCCTCACTGCTGGGGAGTCCCTGCCACACTCAGTCCCCCACC
421
1321  1330    1340    1350    1360    1370    1380
ACACTGAATCTCCCTCCTCACAGTTGCCATGTAGACCCCTTGAAGAGGGGAGGGCCTA
441
1381  1390    1400    1410    1420    1430    1440
GGGAGCCGCACCTTGTCATGTACCATCAATAAAGTACCCTGTGCTCAACCAGTTAAAAAA
461
1441  1450
AAAAAAAAAAAAAAAA
481

```

Sequence of GAPDH mRNA isoform 2 (NM_001256799.2) with chosen editing sites (yellow).


```

1      10      20      30      40      50      60
1      GCCTCAAGACCTTGGGCTGGGACTGGCTGAGCCTGGCGGGAGGCGGGGTCCGAGTCACCG
1
61     70      80      90      100     110     120
20     CCTGCCCGCGCGCCCCGGTTTCTATAAAATTGAGCCCGCAGCCTCCCGCTTCGCTCTCTG
121    130     140     150     160     170     180
40     CTCTCTCTGTTTCGACAGTCAGCCGCATCTTCTTTTGGCTGCCAGCCGAGCCACATCGCT
181    190     200     210     220     230     240
60     CAGACACCATGGGGAAGGTGAAGTTCGGAGTCAACGGATTGGTTCGTATTGGGCGCCTGG
      M G K V K V G V N G F G R I G R L
241    250     260     270     280     290     300
80     TCACCAGGGCTGCTTTTAACTCTGGTAAAGTGGATATTGTTGCCATCAATGACCCCTTCA
      V T R A A F N S G K V D I V A I N D P F
301    310     320     330     340     350     360
100    TTGACCTCAACTACATGGTTTACATGTTCCAATATGATTCCACCCATGGCAAATCCATG
      I D L N Y M V Y M F Q Y D S T H G K F H
361    370     380     390     400     410     420
120    GCACCGTCAAGGCTGAGAACGGGAAGCTTGTCAATCAATGGAAATCCCATCACCATCTTCC
      G T V K A E N G K L V I N G N P I T I F
421    430     440     450     460     470     480
140    AGGAGCGAGATCCCTCCAAAATCAAGTGGGCGATGCTGGCGCTGAGTACGTCGTGGAGT
      Q E R D P S K I K W G D A G A E Y V V E
481    490     500     510     520     530     540
160    CCACTGGCGTCTTACCACCCATGGAGAAGGCTGGGGCTCATTTGCAGGGGGAGCCAAAA
      S T G V F T T M E K A G A H L Q G G A K
541    550     560     570     580     590     600
180    GGGTCATCATCTCTGCCCTCTGCTGATGCCCCATGTTTCGTCATGGGTGTAACCATG
      R V I I S A P S A D A P M F V M G V N H
601    610     620     630     640     650     #1
200    AGAAGTATGACAACAGCCTCAAGATCATCAGCAATGCCTCTGCACCACCAACTGCTTAG
      E K Y D N S L K I I S N A S C T T N C L
661    670     680     690     700     710     720
220    CACCCCTGGCCAAGGTCAATCCATGACAACCTTGGTATCGTGAAGGACTCATGACCACAG
      A P L A K V I H D N F G I V E G L M T T
721    730     740     750     760     770     780
240    TCCATGCCATCACTGCCACCCAGAAGACTGTGGATGGCCCTCCGGGAAACTGTGGCGTG
      V H A I T A T Q K T V D G P S G K L W R
781    790     800     810     820     830     840
260    ATGGCCGCGGGCTCTCCAGAACATCATCCCTGCCTCTACTGGCGCTGCCAAGGCTGTGG
      D G R G A L Q N I I P A S T G A A K A V
841    850     860     870     880     890     900
280    GCAAGGTCACTCCCTGAGCTGAACGGGAAGCTCACTGGCATGGCCTTCCGTGCCCCACTG
      G K V I P E L N G K L T G M A F R V P T
901    910     920     930     #2 940     950     960
300    CCAACGTGTCAAGTGGTGGACCTGACCTGCCGCTTAGAAAAACCTGCCAAATATGATGACA
      A N V S V V D L T C R L E K P A K Y D D
961    970     980     990     1000    1010    1020
320    TCAAGAAGGTGGTGAAGCAGGCGTCGGAGGGCCCCCTCAAGGGCATCCTGGGCTACACTG
      I K K V V K Q A S E G P L K G I L G Y T
1021  1030    1040    1050    1060    1070    1080
340    AGCACCAGGTGGTCTCCTCTGACTTCAACAGCGACCCCACTCCTCCACCTTTGACGCTG
      E H Q V V S S D F N S D T H S S T F D A
1081  1090    1100    1110    1120    1130    1140
360    GGGCTGGCATTGCCTCAACGACCACTTTGTCAAGCTCATTTCTGGTATGACAACGAAT
      G A G I A L N D H F V K L I S W Y D N E
1141  1150    1160    1170    1180    1190    1200
380    TTGGCTACAGCAACAGGGTGGTGGACCTCATGGCCACATGGCCTCCAAGGAGTAAGACC
      F G Y S N R V V D L M A H M A S K E *
1201  1210    1220    1230    1240    1250    1260
400    CCTGGACCACCAGCCCCAGCAAGAGCACAAAGAGGAGAGACCCCTCACTGCTGGGGA
1261  1270    1280    1290    1300    1310    1320
420    GTCCCTGCCACACTCAGTCCCCACCACACTGAATCTCCCTCCTCACAGTTGCCATGTAG
1321  1330    1340    1350    1360    1370    1380
440    GACCCCTTGAAGAGGGGAGGGCCTAGGGAGCCGACCTTGTTCATGTACCATCAATAAAG
1381  1390    1400    1410    1420
460    TACCCTGTCTCAACCAGTTAAAAAAAAAAAAAAAAAAAAA

```

Sequence of GAPDH mRNA isoform 1 (NM_002046.5) with chosen editing sites (yellow).

10 20 30 40 50 60
1 ACCGCCGAGACCGCGTCCGCCCGCAGACAGAGCCTCGCCTTTGCCGATCCGCCGCC
1 T A E T A S A P R A Q S L A F A D P P P
70 80 90 100 110 120
61 GTCCACACCCGCCCGCAGCTCACCATGGATGATGATATCGCCGCGCTCGTCGTCGACAAC
21 V H T R R Q L T M D D D I A A L V V D N
130 140 150 160 170 180
121 GGCTCCGGCATGTGCAAGGCCGGCTTCGCGGGCGACGATGCCCCCGGGCCGTCTTCCCC
41 G S G M C K A G F A G D D A P R A V F P
190 200 210 220 230 240
181 TCCATCGTGGGGCGCCCGAGCACCAGGCGGTGATGGTGGGCATGGTCAGAAGGATTCC
61 S I V G R P R H Q G V M V G M G Q K D S
250 260 270 280 290 300
241 TATGTGGGCGACGAGGCCAGAGCAAGAGAGGCATCCTCACCCCTGAAGTACCCCATCGAG
81 Y V G D E A Q S K R G I L T L K Y P I E
310 320 330 340 350 360
301 CACGGCATCGTCACCAACTGGGACGACATGGAGAAAATCTGGCACCACACCTTCTACAAT
101 H G I V T N W D D M E K I W H H T F Y N
370 380 390 400 410 420
361 GAGCTGCGTGTGGCTCCCGAGGAGCACCCCGTGTGCTGACCGAGGCCCCCTGAACCCC
121 E L R V A P E E H P V L L T E A P L N P
430 440 450 460 470 480
421 AAGGCCAACCGCGAGAAGATGACCCAGATCATGTTTGAGACCTTCAACACCCAGCCATG
141 K A N R E K M T Q I M F E T F N T P A M
490 500 510 520 530 540
481 TACGTTGCTATCCAGGCTGTGCTATCCCTGTACGCTTCTGGCCGTACCACTGGCATCGTG
161 Y V A I Q A V L S L Y A S G R T T G I V
550 560 570 580 590 600
541 ATGGACTCCGGTGACGGGGTCAACCACACTGTGCCCATCTACGAGGGTATGCCCTCCCC
181 M D S G D G V T H T V P I Y E G Y A L P
610 620 630 640 650 660
601 CATGCCATCCTGCGTCTGGACCTGGCTGGCCGGGACCTGACTGACTACCTCATGAAGATC
201 H A I L R L D L A G R D L T D Y L M K I
670 680 690 700 710 720
661 CTCACCGAGCGGCTACAGCTTACCACACGGCCGAGCGGGAATCGTGCCTGACATT
221 L T E R G Y S F T T T A E R E I V R D I
730 740 750 760 770 780
721 AAGGAGAAGCTGTGCTACGCTCGCCCTGGACTTCGAGCAAGAGATGGCCACGGCTGCTTCC
241 K E K L C Y V A L D F E Q E M A T A A S
790 800 810 820 830 840
781 AGCTCCTCCCTGGAGAAGAGCTACGAGCTGCCTGACGGCCAGGTCATCACCATTGGCAAT
261 S S S L E K S Y E L P D G Q V I T I G N
850 860 870 880 890 900
841 GAGCGTTCCGCTGCCCTGAGGCACTTCCAGCCTTCCTTCCTGGGCATGGAGTCCCTGT
281 E R F R C P E A L F Q P S F L G M E S C
910 920 930 940 950 960
901 GGCATCCACGAAACTACCTTCAACTCCATCATGAAGTGTGACGTGGACATCCGCAAAGAC
301 G I H E T T F N S I M K C D V D I R K D
970 980 990 1000 1010 1020
961 CTGTACGCCAACACAGTGTGCTGTGCGCGCACCACCATGTACCCTGGCATTGCCGACAGG
321 L Y A N T V L S G G T T M Y P G I A D R
1030 1040 1050 1060 1070 1080
1021 ATGCAGAAGGAGATCACTGCCCTGGCACCAGCACAATGAAGATCAAGATCATTTGCTCCT
341 M Q K E I T A L A P S T M K I K I I A P
1090 1100 1110 1120 1130 1140
1081 CCTGAGCGCAAGTACTCCGTGTGGATCGGCGGCTCCATCCTGGCCTCGCTGTCCACCTTC
361 P E R K Y S V W I G G S I L A S L S T F
1150 1160 1170 1180 1190 1200
1141 CAGCAGATGTGGATCAGCAAGCAGGAGTATGACGAGTCCGGCCCTCCATCGTCCACCGC
381 Q Q M W I S K Q E Y D E S G P S I V H R
1210 1220 1230 1240 1250 1260
1201 AAATGCTTCTAGGCGGACTATGACTTAGTTCGCTTACACCCCTTTCTTGACAAAACCTAAC
401 K C F *
1270 1280 1290 1300 1310 1320
1261 TTGCGCAGAAAACAAGATGAGATTGGCATGGCTTTATTTGTTTTTTTTTTGTTTTTGG
421
1330 1340 1350 1360 1370 1380
1321 TTTTTTTTTTTTTTTTGGCTTACTCAGGATTTAAAAACTGGAACGGTGAAGGTGACAGC
441
1390 1400 1410 1420 1430 1440
1381 AGTCGGTTGGAGCGAGCATCCCCAAAGTTCACAATGTGGCCGAGGACTTTGATTGCACA
461
1450 1460 1470 1480 1490 1500
1441 TTGTTGTTTTTTTAAATAGTCATTCCAAATATGAGATGCGTTGTTACAGGAAGTCCCTTGC
481 C
1510 1520 1530 1540 1550 1560
1501 CATCCTAAAAGCCACCCCACTTCTCTCTAAGGAGAATGGCCAGTCTCTCCCAAGTCCA
501
1570 1580 1590 1600 1610 1620
1561 CACAGGGAGGTGATAGCATTGCTTTCGTGTAATTTATGTAATGCAAAATTTTTTAATC
521
1630 1640 1650 1660 1670 1680

```

1621   TTCGCCTTAATACTTTTTTATTTTGTTTTATTTTGAATGATGAGCCTTCGTGCCCCCCT
541                                     P
          1690      1700      1710      1720      1730      1740
1681   TCCCCCTTTTTTGTCCCCCAACTTGAGATGTATGAAGGCTTTTGGTCTCCCTGGGAGTGG
561
          1750      1760      1770      1780      1790      1800
1741   GTGGAGGCAGCCAGGGCTTACCTGTACACTGACTTGAGACCAGTTGAATAAAAAGTGCACA
581
          1810      1820      1830      1840      1850
1801   CCTTAAAAATGAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA
601

```

Sequence of ACTB mRNA (NM_001101.3) with chosen editing site (yellow).

10 20 30 40 50 60
1 GTCCCTCAACCAAGATGGCGCGGATGGCTTCAGGGCGCATCACGACACCCGGCGCGTACGCGG
1
70 80 90 100 110 120
61 ACCCGCCCTACGGGCACCTCCCGCGCTTTTCTTAGCGCCGAGACGGTGGCCGAGCGGGG
20
130 140 150 160 170 180
121 GACCGGGAAGCATGGCCCCGGGGTTCGGCGGTTGCCTGGGCGGCGCTCGGGCCGTTGTTGT
40 M A R G S A V A W A A L G P L L
190 200 210 220 230 240
181 GGGGCTGCGCGCTGGGGCTGAGGGCGGGATGCTGTACCCCGAGGAGCCCGTTCGCGGG
60 W G C A L G L Q G G M L Y P Q E S P S R
250 260 270 280 290 300
241 AGTGCAAGGAGCTGGACGGCCTCTGGAGCTTCCGCGCCGACTTCTCTGACAACCGACGCC
80 E C K E L D G L W S F R A D F S D N R R
310 320 330 340 350 360
301 GGGGCTTCGAGGAGCAGTGGTACCAGCGCGCGCTGTGGGAGTACAGCCCCACCGTGGACA
100 R G F E E Q W Y R R P L W E S G P T V D
370 380 390 400 410 420
361 TGCCAGTTCCCTCCAGCTTCAATGACATCAGCCAGGACTGGCGTCTGCGGCATTTTGTGCG
120 M P V P S S F N D I S Q D W R L R H F V
430 440 450 460 470 480
421 GCTGGGTGTGGTACGAACGGGAGGTGATCCTGCCGGAGCGATGGACCCAGGACCTGCGCA
140 G W V W Y E R E V I L P E R W T Q D L R
490 500 510 520 530 540
481 CAAGAGTGGTGTGAGGATTGGCAGTGGCCATTCTATGCCATCGTGTGGGTGAATGGGG
160 T R V V L R I G S A H S Y A I V W V N G
550 560 570 580 590 600
541 TCGACACGCTAGAGCATGAGGGGGCTACCTCCCCTTCGAGGCCGACATCAGCAACCTGG
180 V D T L E H E G G Y L P F E A D I S N L
610 620 630 640 650 660
601 TCCAGGTGGGGCCCTGCCCTCCCGGCTCCGAATCACTATCGCCATCAACAACACTCA
200 V Q V G P L P S R L R I T I A I N N T L
670 680 690 700 710 720
661 CCCCCACCCTGCCACGGGACCATCCAATACCTGACTGACACCTCCAAGTATCCCA
220 T P T T L P P G T I Q Y L T D T S K Y P
730 740 750 760 770 780
721 AGGGTTACTTTGTCCAGAACACATATTTTGACTTTTCAACTACGCTGGACTGCAGCGGT
240 K G Y F V Q N T Y F D F F N Y A G L Q R
790 800 810 820 830 840
781 CTGTACTTCTGTACACGACCCACCACCTACATCGATGACATCACCGTACCACCAGCG
260 S V L L Y T T P T T Y I D D I T V T T S
850 860 870 880 890 900
841 TGGAGCAAGACAGTGGGCTGGTGAATTACCAGATCTCTGTCAAGGGCAGTAACCTGTTC
280 V E Q D S G L V N Y Q I S V K G S N L F
910 920 930 940 950 960
901 AGTTGGAAGTGCCTTTTTGGATGCAGAAAACAAAGTTCGTGGCGAATGGGACTGGGACCC
300 K L E V R L L D A E N K V V A N G T G T
970 980 990 1000 1010 1020
961 AGGGCCAACCTTAAGGTGCCAGGTGTCAGCCTCTGGTGGCCGTACCTGATGCACGAACGCC
320 Q G Q L K V P G V S L W W P Y L M H E R
1030 1040 1050 1060 1070 1080
1021 CTGCCTATCTGTATTTCATTTGGAGGTGCAGCTGACTGCACAGACGCTACTGGGGCCTGTGT
340 P A Y L Y S L E V Q L T A Q T S L G P V
1090 1100 1110 1120 1130 1140
1081 CTGACTTCTACACTCCCTGTGGGGATCCGCACTGTGGCTGTCAACAGAGCCAGTTCC
360 S D F Y T L P V G I R T V A V T K S Q F
1150 1160 1170 1180 1190 1200
1141 TCATCAATGGGAAACCTTTCTATTTCCACGGTGTCAACAAGCATGAGGATGGGGACATCC
380 L I N G K P F Y F H G V N K H E D A D I
1210 1220 1230 1240 1250 1260
1201 GAGGGAAGGGCTTCGACTGGCCGCTGCTGGTGAAGGACTTCAACCTGCTTCGCTGGCTTG
400 R G K G F D W P L L V K D F N L L R W L
1270 1280 1290 1300 1310 1320
1261 GTGCCAACGCTTTCCGTACCAGCCACTACCCCTATGCAGAGGAAGTATGCAGATGTGTG
420 G A N A F R T S H Y P Y A E E V M Q M C
1330 1340 1350 1360 1370 1380
1321 ACCGCTATGGGATTTGTGGTCATCGATGAGTGTCCCGCGTGGGCCTGGCGCTGCCGCAGT
440 D R Y G I V V I D E C P G V G L A L P Q
1390 1400 1410 1420 1430 1440
1381 TCTTCAACAACGTTTCTCTGCATCACCACATGCAGGTGATGGAAGAAGTGGTGGCTAGGG
460 F F N N V S L H H H M Q V M E E V V R R
1450 1460 1470 1480 1490 #1
1441 ACAAGAACCACCCCGGCTGCTGATGTGGTCTGTGGCCAACGAGCCTGCGTCCCACC TAG
480 D K N H P A V V M W S V A N E P A S H L
1510 1520 1530 1540 1550 1560
1501 AATCTGCTGGCTACTACTGAAGATGGTGTGCTCACACCAAATCCTTGGACCCCTCCC
500 E S A G Y Y L K M V I A H T K S L D P S
1570 1580 1590 1600 1610 1620
1561 GGCCTGTGACCTTTGTGAGCAACTTAACATATGCAGCAGACAAGGGGGCTCCGTATGTGG
520 R P V T F V S N S N Y A A D K G A P Y V
1630 1640 1650 1660 1670 1680

```

1621 ATGTGATCTGTTTGAACAGCTACTACTCTTGGTATCAGACTACGGGCACCTGGAGTTGA
540 D V I C L N S Y Y S W Y H D Y G H L E L
      1690      1700      1710      1720      1730      1740
1681 TTCAGCTGCAGCTGGCCACCCAGTTTGAGAACTGGTATAAGAAGTATCAGAAGCCCATTA
560 I Q L Q L A T Q F E N W Y K K Y Q K P I
      1750      1760      1770      1780      1790      1800
1741 TTCAGAGCGAGTATGGAGCAGAAACGATTGCAGGGTTTCACCAGGATCCACCTCTGATGT
580 I Q S E Y G A E T I A G F H Q D P P L M
      1810      1820      1830      1840      1850      1860
1801 TCACTGAAGAGTACCAGAAAAAGTCTGCTAGAGCAGTACCATCTGGGTCTGGATCAAAAAC
600 F T E E Y Q K S L L E Q Y H L G L D Q K
      1870      1880      1890      1900      1910      1920
1861 GCAGAAAAATCGTGGTTGGAGAGCTCATTGGAATTTTGCCGATTCATGACTGAACAGT
620 R R K Y V V G E L I W N F A D F M T E Q
      1930      1940      1950      1960      1970      1980
1921 CACCGACGAGAGTCTGGGGAATAAAAAAGGGGATCTTCACTCGGCAGAGACAACCAAAAA
640 S P T R V L G N K K G I F T R Q R Q P K
      1990      2000      2010      2020      2030      2040
1981 GTGCAGCGTTCCTTTTGGCAGAGAGATACTGGAAGATTGCCAATGAAACCAGGTATCCCC
660 S A A F L L R E R Y W K I A N E T R Y P
      #2      2060      2070      2080      2090      2100
2041 ACTCAGTAGCCAAGTCACAATGTTTGGAAAAACAGCCTGTTTACTTGAGCAAGACTGATAC
680 H S V A K S Q C L E N S L F T *
      2110      2120      2130      2140      2150      2160
2101 CACCTGCGTGTCCCTTCCTCCCCGAGTCAGGGCGACTTCCACAGCAGCAGAACAAGTGCC
700
      2170      2180      2190      2200      2210      2220
2161 TCCTGGACTGTTTACGGCAGACCAGAACGTTTCTGGCCTGGGTTTGTGGTCATCTATTC
720
      2230      2240      2250      2260      2270      2280
2221 TAGCAGGGAACACTAAAGGTGGAATAAAAAGATTTTCTATTATGGAATAAAGAGTTGGC
740
      2290      2300      2310      2320
2281 ATGAAAGTGGCTACTGAAAAAAAAAAAAAAAAAAAAAAAAAAAA
760

```

Sequence of GUSB mRNA (NM_000181.3) with chosen editing sites (yellow).

```

10      20      30      40      50      60
1      TCCTAGGCGGGCGGCCGCGCGGGCGGAGGCGAGCAGCGGGCGCGGCAGTGGCGGGCGGCAAG
1
70      80      90      100     110     120
61     GTGGCGGGCGGCTCGGCCAGTACTCCCGGCCCGCCATTTTCGGACTGGGAGCGAGCGCGG
21
130     140     150     160     170     180
121    CGCAGGCACTGAAGGCGGGCGGGGCCAGAGGCTCAGCGGCTCCAGGTGCGGGAGAGA
41
190     200     210     target A/1     target 2
181    GGCCTGCTGAAAATGACTGAATATAAACTTGTGGTAGTTGGAGCTGGTGGCGTAGGCAAG
61      M T E Y K L V V V G A G G V G K
250     260     270     280     290     300
241    AGTGCCTTGACGATACAGCTAATTCAGAATCATTGTGGACGAATATGATCCAACAATA
81      S A L T I Q L I Q N H F V D E Y D P T I
310     320     330     340     350     360
301    GAGGATTCCTACAGGAAGCAAGTAGTAATTGATGGAGAAACCTGTCTCTTGATATTCTC
101     E D S Y R K Q V V I D G E T C L L D I L
370     380     390     400     410     420
361    GACACAGCAGGTCAAGAGGAGTACAGTGC AATGAGGGACCAAGTACATGAGGACTGGGGAG
121     D T A G Q E E Y S A M R D Q Y M R T G E
430     440     450     460     470     480
421    GGCTTTCTTTGTGTTTGGCCATAATAACTAAATCATTGGAAGATATTCACCATTAT
141     G F L C V F A I N N T K S F E D I H H Y
490     500     510     520     530     540
481    AGAGAACAATTAAGAGGTTAAGGACTCTGAAGATGTACCTATGGTCCCTAGTAGGAAAT
161     R E Q I K R V K D S E D V P M V L V G N
550     560     570     580     590     600
541    AAATGTGATTTGCCTTCTAGAACAGTAGACACAAAACAGGCTCAGGACTTAGCAAGAAGT
181     K C D L P S R T V D T K Q A Q D L A R S
610     620     630     640     650     660
601    TATGGAATTCCTTTTATTGAAACATCAGCAAAGACAAGACAGGGTGTGATGATGCCTTC
201     Y G I P F I E T S A K T R Q G V D D A F
670     680     690     700     710     720
661    TATACATTAGTTCGAGAAATTCGAAAACATAAAGAAAAGATGAGCAAAGATGGTAAAAAG
221     Y T L V R E I R K H K E K M S K D G K K
730     740     750     760     770     780
721    AAGAAAAAGAAGTCAAAGACAAAGTGTGTAATTATGTAATAACAATTTGTACTTTTTTCT
241     K K K K S K T K C V I M *
790     800     810     820     830     840
781    TAAGGCATACTAGTACAAGTGGTAATTTTGTACATTACACTAAATTATTAGCATTGTT
261

```

Sequence of KRAS mRNA (NM_004985.4) with chosen editing sites (yellow).

10 20 30 40 50 60
1 GCTGAGCGCGGAGCCGCCCGGTGATTGGTGGGGGCGGAAGGGGGCCGGCGCCAGCGCTG
1
70 80 90 100 110 120
61 CCTTTTCTCCTGCCGGTAGTTTCGCTTTCTGCGCAGAGTCTGCGGAGGGGCTCGGCTG
21
130 140 150 160 170 180
121 CACCGGGGGGATCGCGCCTGGCAGACCCAGACCGAGCAGAGGCGACCCAGCGCTCGG
41
190 200 210 220 230 240
181 GAGAGGCTGCACCGCCGCGCCCCGCCTAGCCCTTCCGGATCCTGCGCGCAGAAAAGTTT
61
250 260 270 280 290 300
241 CATTTGCTGTATGCCATCCTCGAGAGCTGTCTAGGTTAACGTTTCGCACTCTGTGTATATA
81
310 320 330 340 350 360
301 ACCTCGACAGTCTTGGCACCTAACGTGCTGTGCGTAGCTGCTCCTTTGGTTGAATCCCCA
101
370 380 390 400 410 420
361 GGCCCTTGTGGGGCACAAGGTGGCAGGATGTCTCAGTGGTACGAACTTCCAGCAGCTTGA
121 M S Q W Y E L Q Q L D
421 CTCAAAATTCCTGGAGCAGGTTCCACCAGCTTTATGATGACAGTTTTCCCATGGAATCAG
141 S K F L E Q V H Q L Y D D S F P M E I R
481 ACAGTACCTGGCAGAGTGGTTAGAAAAGCAAGACTGGGAGCAGCTGCCAATGATGTTTC
161 Q Y L A Q W L E K Q D W E H A A N D V S
541 ATTTGCCACCATCCGTTTTTCATGACCTCCTGTACAGCTGGATGATCAATATAGTCGCTT
181 F A T I R F H D L L S Q L D D Q Y S R F
601 TTCTTTGGAGAATAAATTCTTGCTACAGCATAACATAAGGAAAAGCAAGCGTAATCTTCA
201 S L E N N F L L Q H N I R K S K R N L Q
661 GGATAATTTTCAGGAAGACCCCAATCCAGATGTCTATGATCATTACAGCTGTCTGAAGGA
221 D N F Q E D P I Q M S M I I Y S C L K E
721 AGAAAGGAAAATTCGAAAACGCCAGAGATTTAATCAGGCTCAGTCGGGGAATATTCA
241 E R K I L E N A Q R F N Q A Q S G N I Q
781 GAGCACAGTGATGTTAGACAAAACAGAAAGAGCTTGACAGTAAAGTCAGAAATGTGAAGGA
261 S T V M L D K Q K E L D S K V R N V K D
841 CAAGGTTATGTGTATAGAGCATGAAATCAAGAGCCTGGAAGATTTACAAGATGAATATGA
281 K V M C I E H E I K S L E D L Q D E Y D
901 CTTCAAATGCAAAACCTTGAGAACAGAGAACACGAGACCAATGGTGTGGCAAAGAGTGA
301 F K C K T L Q N R E H E T N G V A K S D
961 TCAGAAAACAAGAAGCAGCTGTACTCAAGAAGATGATTTAATGCTTGACAATAAGAGAAA
321 Q K Q E Q L L L K K M Y L M L D N K R K
1021 GGAAGTAGTTCACAAAATAATAGAGTTGCTGAATGTCACTGAACTTACCCAGAATGCCCT
341 E V V H K I I E L L N V T E L T Q N A L
1081 GATTAATGATGAAGTAGTGGAGTGAAGCGGAGACAGCAGAGCGCTGTATTGGGGGGCC
361 I N D E L V E W K R R Q Q S A C I G G P
1141 GCCCAATGCTTGCTTGGATCAGCTGCAGAACTGGTTCACTATAGTTGCGGAGAGTCTGCA
381 P N A C L D Q L Q N W F T I V A E S L Q
1201 GCAAGTTCGGCAGCAGCTTAAAAAGTTGGAGGAATTGGAACAGAAATACACCTACGAACA
401 Q V R Q Q L K K L E E L E Q K Y T Y E H
1261 TGACCCTATCACAAAAACAACAAGTGTATGGGACCGCACCTTCAGTCTTTCCAGCA
421 D P I T K N K Q V L W D R T F S L F Q Q
1321 GCTCATTAGAGCTCGTTTGGTGGAAAGACAGCCCTGCATGCCAACGCACCCCTCAGAG
441 L I Q S S F V V E R Q P C M P T H P Q R
1381 GCCGCTGGTCTTGAAGACAGGGTCCAGTTCAGTGTGAAGTTGAGACTGTTGGTGAATTT
461 P L V L K T G V Q F T V K L R L L V K L
1441 GCAAGAGCTGAATATAATTTGAAAGTCAAAGTCTTATTGATAAAGATGTGAATGAGAG
481 Q E L N Y N L K V K V L F D K D V N E R
1501 AAATACAGTAAAAGGATTTAGGAAGTTCACATTTTGGGCACGCACACAAAAGTGATGAA
501 N T V K G F R K F N I L G T H T K V M N
1561 CATGGAGGAGTCCACCAATGGCAGTCTGGCGGCTGAATTCGGCACCTGCAATTGAAAGA
521 M E E S T N G S L A A E F R H L Q L K E
1630 1640 1650 1660 1670 1680

1621 ACAGAAAAATGCTGGCACCAGAACGAATGAGGGTCTCTCATCGTTACTGAAGAGCTTCA
 541 Q K N A G T R T N E G P L I V T E E L H
 1690 1700 1710 1720 1730 1740
 1681 CTCCTTAGTTTTGAAACCAATTGTGCCAGCCTGGTTTGGTAATTGACCTCGAGACGAC
 561 S L S F E T Q L C Q P G L V I D L E T T
 1750 1760 1770 1780 1790 1800
 1741 CTCTCTGCCCGTTGTGGTGATCTCCAACGTCAGCCAGCTCCCGAGCGGTTGGCCTCCAT
 581 S L P V V V I S N V S Q L P S G W A S I
 1810 1820 1830 1840 1850 1860
 1801 CCTTTGGTACAACATGCTGGTGGCGGAACCCAGGAATCTGTCCTTCTTCCCTGACTCCACC
 601 L W Y N M L V A E P R N L S F F L T P P
 1870 1880 1890 1900 1910 1920
 1861 ATGTGCACGATGGGCTCAGCTTTTCCAGAAGTGTGAGTTGGCAGTTTTCTTCTGTCCACAA
 621 C A R W A Q L S E V L S W Q F S S V T K
 1930 1940 1950 1960 1970 1980
 1921 AAGAGGTCTCAATGTGGACCAGCTGAACATGTTGGGAGAGAAGCTTCTTGGTCCCTAACGC
 641 R G L N V D Q L N M L G E K L L G P N A
 1990 2000 2010 2020 2030 2040
 1981 CAGCCCCGATGGTCTCATTCCCGTGGACGAGGTTTTGTAAGGAAAATATAAATGATAAAAA
 661 S P D G L I P W T R F C K E N I N D K N
 2050 2060 2070 2080 2090 2100
 2041 TTTTCCCTTCTGGCTTTGGATTGAAAGCATCCTAGAACTCATTAAAAAACACCTGCTCCC
 681 F P F W L W I E S I L E L I K K H L L P
 2110 2120 2130 2140 2150 2160
 2101 TCTCTGGAATGATGGGTGCATCATTGGGCTTCATCAGCAAGGAGCGAGAGCGTGCCTGT
 701 L W N D G C I M G F I S K E R E R A L L
 2170 2180 2190 2200 2210 2220
 2161 GAAGGACCAGCAGCCGGGACCTTCCCTGCTGCGGTTCAGTGAGAGCTCCCGGGAAGGGGC
 721 K D Q Q P G T F L L R F S E S S R E G A
 2230 2240 2250 2260 2270 2280
 2221 CATCACATTACATGGGTGGAGCGGTCCCAGAACGGAGGCGAACCTGACTTCCATGCGGT
 741 I T F T W V E R S Q N G G E P D F H A V
 2290 2300 2310 2320 2330 2340
 2281 TGAACCTACACGAAGAAGAACTTTCTGCTGTTACTTTCCCTGACATCATTGCAATTA
 761 E P Y T K K E L S A V T F P D I I R N Y
 2350 2360 2370 2380 2390 2400
 2341 CAAAGTCATGGCTGCTGAGAATATTCCTGAGAATCCCCTGAAGTATCTGTATCCAAATAT
 781 K V M A A E N I P E N P L K Y L Y P N I
 2410 2420 2430 2440 2450 2460
 2401 TGACAAAGACCATGCCTTTGGAAAGTATTACTCCAGGCCAAAGGAAGCACCAGAGCCAAT
 801 D K D H A F G K Y Y S R P K E A P E P M
 2470 2480 2490 2500 2510 2520
 2461 GGAACCTGATGGCCCTAAAGGAACGGAATATCAAGACTGAGTTGATTCTGTGTCTGA
 821 E L D G P K G T G Y I K T E L I S V S E
 2530 2540 2550 2560 2570 2580
 2521 AGTTCACCCTTCTAGACTTCAGACCACAGACAACCTGCTCCCCATGTCTCCTGAGGAGTT
 841 V H P S R L Q T T D N L L P M S P E E F
 2590 2600 2610 2620 2630 2640
 2581 TGACGAGGTGCTCGGATAGTGGGCTCTGTAGAATTCGACAGTATGATGAACACAGTATA
 861 D E V S R I V G S V E F D S M M N T V *
 2650 2660 2670 2680 2690 2700
 2641 GAGCATGAATTTTTTTCATCTTCTCTGGCGACAGTTTTTCTTCTCATCTGTGATTCCCTC
 881

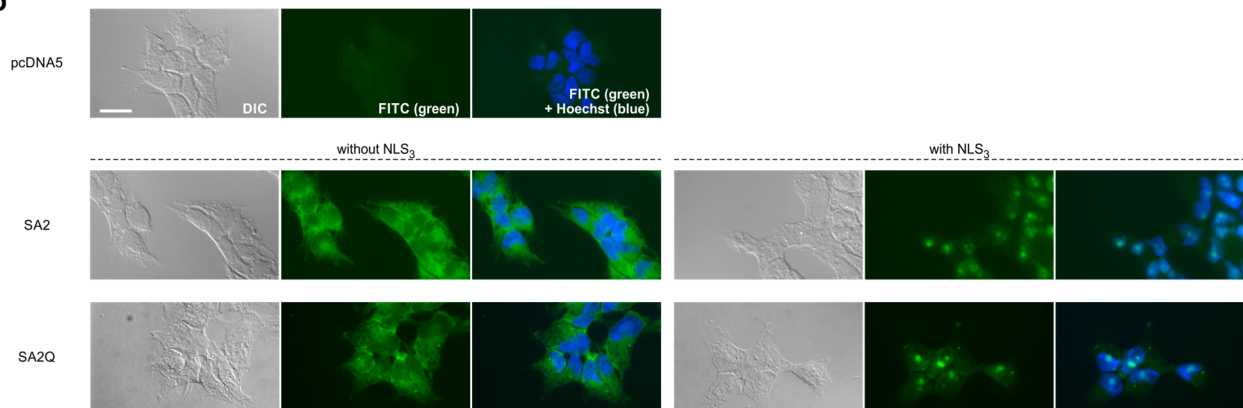
Sequence of STAT1 mRNA (NM_007315.3) with chosen editing site Y701 (yellow).

Non-included data (Man. 4)

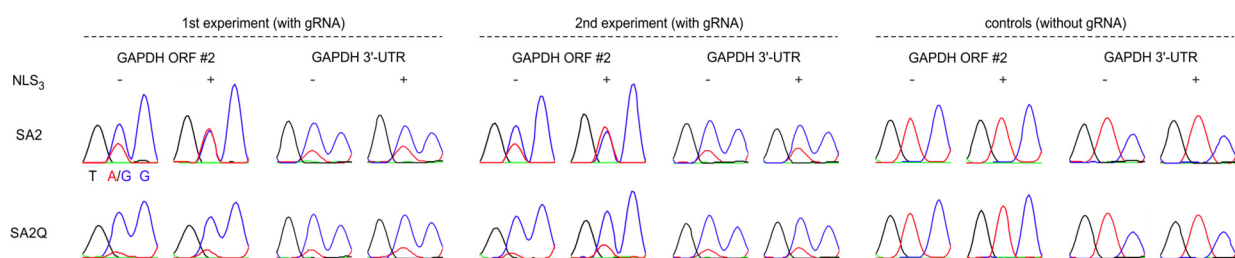
a

Samples	On-target editing levels
4 λ N-ADAR2	CFTR, nt 365: 71% and nt 366: 37%
NLS ₃ -4 λ N-ADAR2	CFTR, nt 365: 45% and nt 366: 21%
4 λ N-ADAR2Q	CFTR, nt 365: 83% and nt 366: 75%
NLS ₃ -4 λ N-ADAR2Q	CFTR, nt 365: 67% and nt 366: 57%

b



c



Effect of nuclear localization on the editing yield. **(a)** NGS data produced by Vallecillo-Viejo et al. (RNA Biol. 2018) was reanalyzed using our pipeline (see online methods). In the study of Vallecillo-Viejo et al., three copies of the *SV40* large T antigen NLS (NLS₃) were added to the N-terminus of 4 λ N-ADAR2 and 4 λ N-ADAR2Q to direct the editing enzymes to the nucleus. Site-directed RNA Editing was performed within an overexpressed reporter mRNA (*CFTR*). Vallecillo-Viejo et al. claimed that sequestering the enzyme in the nucleus decreases off-target editing without affecting on-target editing. However, the reanalysis revealed that nuclear translocation of the 4 λ N-ADAR enzymes results in a reduction of the on-target editing by ~ 20%. **(b)** In our study, three copies of the *SV40* large T antigen NLS (NLS₃) were fused to the N-terminus of SA2 and SA2Q. FITC-staining revealed that both SA2 and SA2Q were predominately localized in the nucleus after NLS₃-fusion. **(c)** Editing was performed in the ORF (site #2) and the 3'-UTR of GAPDH in cells expressing the SA enzymes without and with NLS₃. The editing yields were always decreased when applying nuclear SA enzymes.

Man. 5: accepted

Vogel, P. & Stafforst, T. Site-directed RNA editing with antagomir deaminases - A tool to study protein and RNA function. *ChemMedChem* **9**, 2021-2025 (2014)

DOI: 10.1002/cmdc.201402139

Site-Directed RNA Editing with Antagomir Deaminases — A Tool to Study Protein and RNA Function

Paul Vogel and Thorsten Stafforst*^[a]

RNA-guided machineries perfectly satisfy the demand for rationally programmable tools that manipulate gene function inside the cell. Over the last ten years, various natural machineries have been harnessed, with RNA interference being among the most prominent examples. It is now time to tackle the engineering of novel RNA-guided tools not provided by nature. In this respect, we highlight RNA-guided site-directed RNA editing as a new concept for the manipulation of RNA and protein function. In contrast to currently available tech-

niques, RNA editing allows for the introduction of selected point mutations into the transcriptome without the need for genomic manipulation. In particular, the approach described using chemically stabilized, antagomir-like guideRNAs may offer advantages over others, such as specificity and circumvention of immunogenicity. These new tools have significant potential for the advancement of both basic science and medicinal application, especially in the treatment of genetic diseases.

Introduction

The immediate revelation of the copying mechanism was the most striking when Watson and Crick, 60 years ago, correctly guessed the DNA helix structure.^[1] Applying molecular recognition between two simple base pairs for encoding the entire genetic information organized in a few gigantic macromolecules was unbelievable and still continues to fascinate young life scientists and chemists. Upon the family of bio-macromolecules, base pairing is a unique feature of nucleic acids and is particularly useful as it allows for the rational design of probes that target RNA and DNA with predictable site specificity, binding strength, and binding dynamics in order to precisely manipulate biological function.

Chemists quickly started to chemically alter nucleic acids for various reasons. On one hand, they aimed to unravel the etiology of nucleic acid structure and function.^[2] On the other hand, they were interested in further improving nucleic acids for medicinal applications. In the beginnings, nucleic acid analogues lagged far behind the exaggerated expectations in antisense and antigene technologies. However, we now find numerous designed nucleic acid analogues that are successfully applied as research tools and in clinical trials.

Small chemical modifications on the nucleic acid backbone, including 2'-O-methyl, 2'-fluorine, 2'-O-methoxyethyl (2'-MOE), and phosphothioate for instance typically improve several pharmacological properties of antisense probes, such as the in vivo stability, potency, cellular uptake, and importantly, the immunogenicity.^[3] Modifications have also been shown to modulate the binding strength and specificity of such probes. Examples include the locked nucleic acids (LNAs) that fix the

sugar puckering of the ribose in a conformation optimal for hybridization (Figure 1),^[4] and C5-methyl-cytosine that improves the binding thermodynamics of triplex-forming oligonucleotides.^[5] However, fully artificial analogues, such as peptide nucleic acids (PNAs)^[6] and morpholinos,^[7] have also been successfully applied (Figure 1).

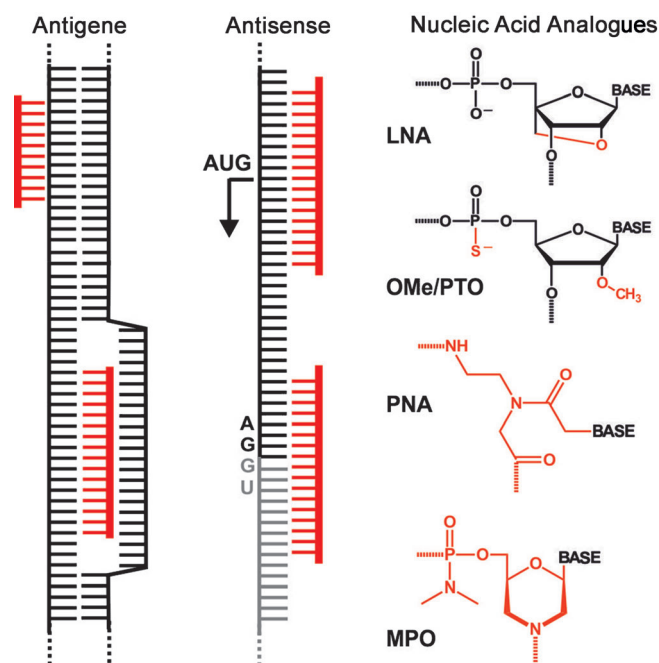


Figure 1. Various nucleic acid analogues have been used to interfere with gene expression by site-specific binding either to a gene (antigene) or to a transcript (antisense) of interest, thereby inhibiting transcription, translation, splicing, and other processes. LNA = locked nucleic acid, OMe/PTO = 2'-O-methyl/phosphothioate oligonucleotide, PNA = peptide nucleic acid, MPO = morpholino oligomer.

[a] P. Vogel, Dr. T. Stafforst
Interfaculty Institute of Biochemistry, University of Tübingen
Auf der Morgenstelle 15, 72076 Tübingen (Germany)
E-mail: thorsten.stafforst@uni-tuebingen.de
Homepage: <http://www.ifib.uni-tuebingen.de/research/stafforst.html>

By applying simple base-pairing rules, antisense and antigene probes can interfere with gene function in a highly rational and predictable way.^[3-8] They can manipulate various processes, including transcription, splicing, translation (initiation), and microRNA function, for instance. The mechanism of action, however, is basically always the same. Owing to their strong binding to specific sites on a gene or transcript, such probes inhibit biological processes by blocking the access of endogenous factors. The general mechanism also accounts for the general limitations. The bare probes can only block active processes in a concentration-dependent manner and for a limited amount of time.

One way to overcome these limitations is the recruitment of catalytically active machineries. Recognizing this, chemists tried to develop gapmers and external guideRNAs that recruit RNaseH or RNaseP activity to support the antisense effects by nuclease digestion of the targeted transcript.^[8] However, the real breakthrough in the antisense field was the discovery of the RNA interference mechanism and the ease by which this RNA-guided machinery can be re-addressed with external guideRNAs, the so-called siRNAs.^[9] In recent years, many cellular RNA-guided machineries turned out to be readily re-addressable with external guideRNAs. Ribosomal RNAs, for instance, are 2'-O-methylated and pseudo-uridylated site-specifically in a snoRNA-dependent manner (Figure 2). Both processes have been shown to be re-addressable towards (pre-)mRNAs by the expression of user-defined snoRNAs and allow for alteration of splice patterns^[10] and for suppressing premature stop codons.^[11] Instead of relying to the natural machineries only, people have recently started to engineer novel RNA-guided machineries. One successful example is the engineering of the CRISPR nucleases that promise to revolutionize genome editing.^[12]

The currently available, natural or engineered RNA-guided reactions either harness nucleases or focus on reactions like transglycosylases or 2'-O-methyltransferases that allow only for a very limited manipulation of the gene content.^[9-12] We were wondering whether it is possible to engineer and assemble an RNA-guided machinery that enables us to reprogram single bases at the RNA-level and thus to manipulate RNA processing or to introduce point mutations into proteins. If practical, such a tool could be very valuable to study the role of single point mutations (SNP) in disease or in protein function. With the advent of high-throughput sequencing, personalized medicine comes into reach that takes the genetic predisposition of the individual into account.

Unfortunately, efficient methods to study the effect of point mutations are lacking. It is still cumbersome and expensive to

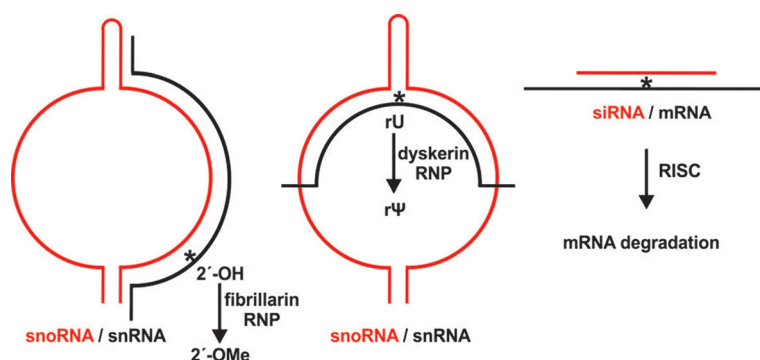


Figure 2. Natural RNA-guided machineries interfere with nucleic acid function to target 2'-O-methylation, uridine to pseudouridine transformation, or mRNA degradation. Riboproteins containing a respective enzyme (fibrillarin, dyskerin or Ago2 inside RISC) are directed site-specifically to their targets (small nuclear RNA (snRNA) or mRNA) by guideRNAs (small nucleolar RNAs (snoRNA) or mi/siRNAs). All three processes are readily redirected to user-defined transcripts with external guideRNAs.

generate animal or even cell models that can conditionally switch between two protein isoforms differing only in a single point mutation. However, humans and other higher animals contain an RNA editing system that allows exactly that. The human adenosine deaminases acting on RNA (hADAR) are a class of enzymes that catalyze the hydrolysis of adenosine to inosine.^[13] Since inosine is biochemically read as guanosine, editing in the open reading frame (ORF) of a transcript can result in the substitution of a single amino acid. Around 30 sites in the human transcriptome, in particular neuroreceptor genes, are targets of very efficient and highly selective editing reactions. Besides that, RNA editing silences intronic Alu repeats, and interferes with RNA processing and microRNA action by editing of splice sites and of (pri/pre)-microRNAs.^[13] Since editing occurs massively in the human brain and seems to be regulated in a complex way, people recently speculated about a yet underestimated role of RNA editing in the development and maintenance of the human cognitive functions.^[14] Apparently, malfunctioning of the editing system leads to behavioral phenotypes and psychiatric disorders.

Engineering RNA-Guided Deaminases

From an engineer's point of view, the broad scope of the editing reaction is intriguing. Beside numerous RNA processing signals, 12 out of the 20 canonical amino acid codons are potential targets for A-to-I RNA editing, including nearly all polar residues that are essential for phosphorylation, enzyme catalysis, metal binding, and protein glycosylation (Ser, Thr, Tyr, His, Lys, Arg, Asp, Glu, Asn, Gln, Met/Start, Val, Stop).^[13,15] Thus, the site-directed editing of a single residue in a protein of interest is expected to have a strong impact if a functionally essential residue is chosen.

Unfortunately, targeting >20000 sites in the human transcriptome, the natural editing system is not RNA-guided but rather applies promiscuous N-terminal dsRNA binding domains (dsRBD) for substrate recognition. The few highly specific editing reactions in the coding frames of neuroreceptor genes, however, are activated by intronically located dsRNA signals.^[13]

Thus, editing activity cannot be redirected by simple administration of external guideRNAs. However, to take the advantage of RNA guidance with its high specificity and ease of rational design, we decided to re-engineer the protein-guided ADAR into an RNA-guided deaminase (Figure 3). For this, we removed all natural substrate binding domains (dsRBD) and substituted them with a single SNAP-tag domain,^[16] an engineered O⁶-alkylguanine-DNA-alkyltransferase.^[15a] The SNAP-tag fusion with the catalytic deaminase domains allows for the simple and efficient formation of highly defined, covalent one-to-one conjugates with any user-defined guideRNA that carries a (5'-terminal) O⁶-benzylguanine modification. The guideRNA in the assembled conjugates fulfills two tasks: first, it steers the conjugate towards the user-defined RNA transcript, and second, it forms the secondary RNA structure required for the efficient activation of one specific adenosine base under concurrent suppression of over-reaction at neighboring off-target bases.^[15a]

Initially, we optimized the system for the transcript repair of tryptophan (UGG) to Stop (UAG) nonsense point mutations in reporter genes as cyan fluorescent protein (CFP), green fluorescent protein (GFP) and luciferase in vitro. When purified mRNA, guideRNA and SNAP-ADAR is incubated, the conjugate assembles and edits the mRNA in up to quantitative yield with high selectivity and over a broad range of reaction conditions.^[15a] We comprehensively studied the prerequisites of the guideRNA for a robust editing reaction (Figure 4).^[15b] We found that a minimal length of 13 nt (SNAP-ADAR2) and 11 nt (SNAP-ADAR1) is required. The targeted adenosine is best situated in the middle of the RNA duplex. Also the counter base opposite the targeted adenosine has an influence on the editing reaction. Most codons accept pyrimidines only and prefer mismatching with cytosine. Mismatching of the target adenosine with purines, in particular with guanosine, leads to complete inhibition of editing. We exploited guanosine mismatching to successfully suppress overediting at off-target bases. This was even feasible when the two adenosine bases were separated by only one intervening nucleotide, as we recently demonstrated with the repair of a glycine to serine missense mutation in GFP.^[15b] Thus, the guideRNA not only steers the deaminase toward the user-defined transcript, but the architecture of the guideRNA also allows reasonable control over the editing reaction. Notably, this control is not limited to inactivate off-site targets but rather enables to activate inherently inactive codons as 5'-GAG.^[15a]

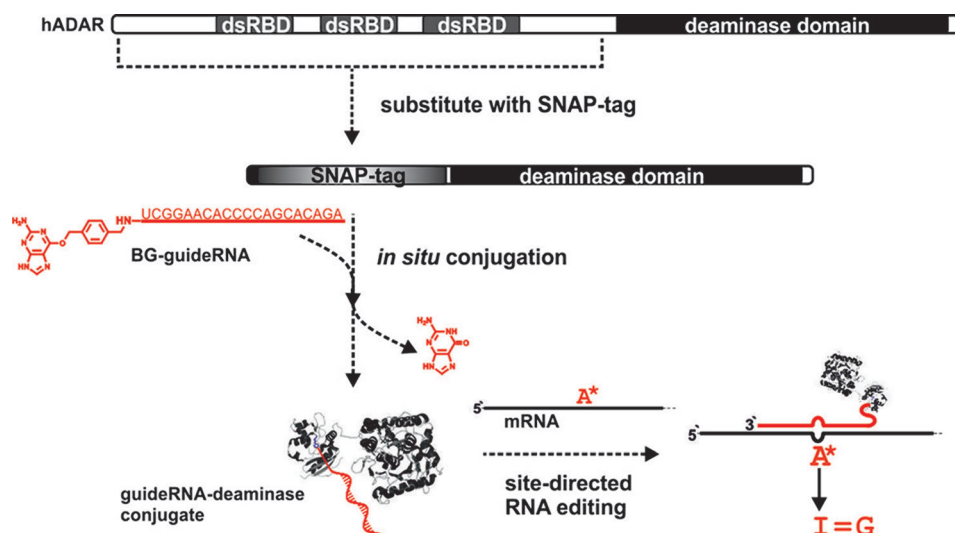


Figure 3. Engineering an RNA-guided RNA editing machine. The natural substrate binding domains (dsRBDs) have been substituted with a SNAP-tag domain that allows for the assembly of defined one-to-one conjugates with O⁶-benzylguanine (BG)-modified guideRNAs. The guideRNA steers the deaminase domain to user-defined mRNAs and activates a single, specific adenosine for editing.

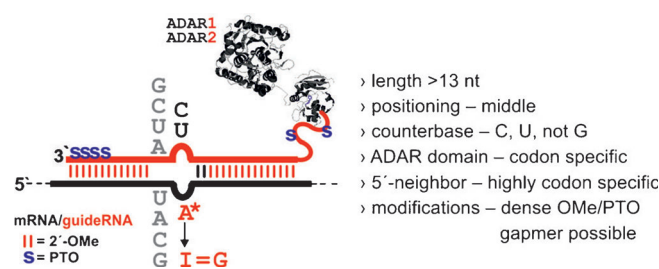


Figure 4. The secondary structure of the guideRNA/mRNA duplex determines the outcome of an editing reaction and needs to be optimized for any given codon and sequence context.

In very challenging adenosine-rich sequence contexts, chemical modification of the guideRNA comes into play. We could demonstrate that careful incorporation of two 2'-O-methyl groups fully suppresses the overediting of the direct neighboring adenosine in a 5'-CAA codon that we targeted to repair the factor 5 Leiden missense mutation.^[15c] Chemical modification of the guideRNA was also helpful when we applied site-directed RNA editing inside the living cell. For this, SNAP-ADAR and a fluorogenic reporter gene, carrying a Trp→Stop nonsense mutation, were transiently overexpressed in 293T cells (Figure 5). Transcript repair was successfully stimulated by lipofection of a 20 nt-long guideRNA that was extensively modified in an antagomir-like fashion with global 2'-O-methyl and terminal phosphothioate groups, leaving only a small gap of three natural ribonucleotides opposite the targeted adenosine. Compared with the unmodified guideRNA, chemical modification clearly improved the robustness of the editing reaction.^[15c] We were surprised by the ability of ADAR deaminases to accept such densely modified guideRNAs as substrates. Other guideRNAs that get incorporated into riboproteins, siRNAs for instance, typically tolerate much less chemical modification.^[17]

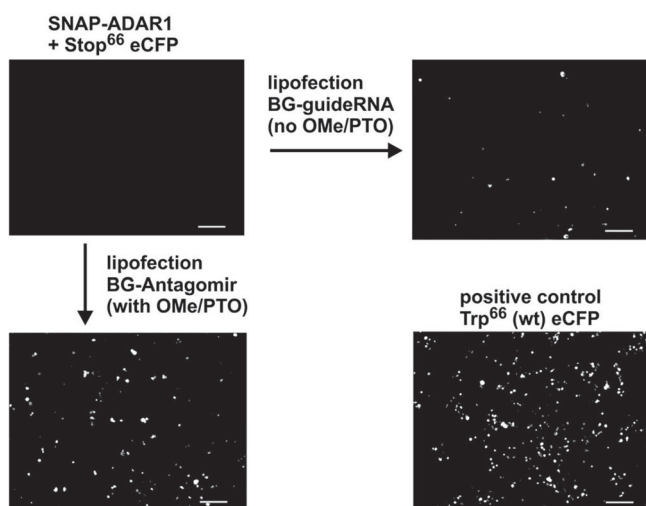
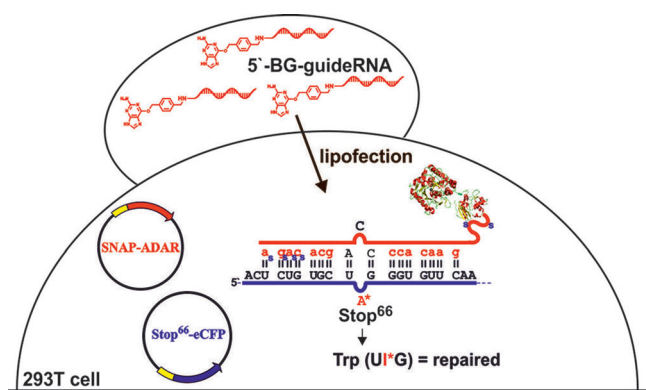


Figure 5. Repair of a nonsense point mutation in the *eCFP* gene by site-directed RNA editing. The 2'-*O*-methyl/phosphothioate oligonucleotide (OMe/PTO)-modified guideRNA gives a substantially improved editing yield compared with the unmodified guideRNA of the same sequence. Modifications: small red letters indicate 2'-OMe; "s" indicates PTO.

However, the susceptibility of SNAP-ADAR for antagomir-like guideRNAs is very promising for future application, as antagomirs have been used in various settings, including cultured cells, living animals, and even living brains^[18] to inhibit specific microRNA functions.^[19] They are characterized by a low toxicity, low immunogenicity, and due to their high stability, they give long-lasting effects even after single administration.^[20] Since 2'-*O*-methylation directly opposite the targeted adenosine fully inhibits editing,^[15c] antagomir-like guideRNAs can be expected to have an improved editing selectivity as random binding of the guideRNA to a partly complementary RNA will less often lead to editing at the off-site. This could prove particularly beneficial when targeting a low-abundant mRNA, which may require stronger binding and thus longer guideRNAs.

The extent to which the guideRNA can be elongated without eliciting antisense effects remains to be determined. However, from the early failings with antisense probes directed to the ORF, we know that the helicase activity of the translating ribosome is very robust. Chemical modification of the probes will not only improve their potency and the endurance of the editing, but may also improve specificity and decrease possible

off-site effects. A guideRNA that cannot conjugate to SNAP-ADAR is unable to repair the reporter transcript.^[15c] This demonstrates two things: first, that SNAP-ADAR alone, lacking the dsRBDs, has difficulties to process random dsRNA, and second, this shows that SNAP-tag technology can indeed provide the assembly of catalytically competent, covalent antagomir-deaminase conjugates inside the living cell. However, off-target effects that may arise from overexpression of SNAP-ADAR have not yet been investigated. Future experiments will first have to clarify the level to which SNAP-ADAR needs to be expressed to achieve optimal editing and what guideRNA length and architecture is required to edit an RNA species with low to normal copy numbers.

Future Directions

Site-directed RNA editing has high potential for application in basic biology. Most intriguing is the possibility of generating animal models that conditionally express SNAP-ADAR in specific tissues. Various point mutations in different genes could be studied from that same animal by administration of different guideRNAs. Transcript repair may also find application in medicine as a means to attenuate otherwise untreatable genetic disease phenotypes. Many applications could benefit from more advanced and flexible delivery methods. SNAP-deaminases could be readily transduced even into non-dividing cells with virus particles and may result in a long-term expression of the machinery. The guideRNAs on the other hand could be 3'-terminally conjugated with cholesterol or folate to harness receptor-mediated uptake. Furthermore, new features are conceivable, including editing of RNA processing elements, photo-control, evolved deaminases, and genetically fully encodable variants. With respect to the latter, others have recently complemented our approach and achieved to steer RNA editing by applying the genetically encodable protein/RNA interaction between the λ N peptide and the BoxB RNA element.^[21] The very first attempt to redirect RNA editing activity, almost 20 years ago,^[22] failed due to difficulties to control overediting. Whether the Box-dependent strategy enables the editing of challenging codons, for instance those that are adenosine-rich, remains to be shown. The comparably long guideRNAs (≈ 60 nt) may potentially offer more problems with off-target editing and immunogenic responses. That said, the enormous progress that various groups are currently making on engineering RNA-guided machineries is fascinating and opens new avenues for biotechnology and engineering. In particular, our approach may also appeal to the medicinal chemist as it opens new perspectives for using strongly chemically altered guideRNAs in redirecting engineered riboproteins.

Acknowledgements

The authors thank the Deutsche Forschungsgemeinschaft (DFG) (grant numbers STA 1053/3-1 and STA 1053/3-2) and the University of Tübingen (Germany) for generous financial support.

Keywords: bioengineering · editing · genetic diseases · nucleic acid analogues · RNA

- [1] J. D. Watson, F. H. C. Crick, *Nature* **1953**, *171*, 737–738.
- [2] A. Eschenmoser, *Science* **1999**, *284*, 2118–2124.
- [3] J. C. Burnett, J. J. Rossi, *Chem. Biol.* **2012**, *19*, 60–71.
- [4] D. A. Braasch, D. R. Corey, *Chem. Biol.* **2001**, *8*, 1–7.
- [5] a) L. E. Xodo, G. Manzini, F. Quadrioglio, G. A. v. d. Marel, J. H. v. Boom, *Nucleic Acid Res.* **1991**, *19*, 5625–5631; b) N. T. Thuong, C. Hélène, *Angew. Chem.* **1993**, *105*, 697–723; *Angew. Chem. Int. Ed. Engl.* **1993**, *32*, 666–690.
- [6] M. Egholm, O. Buchardt, L. Christensen, C. Behrens, S. M. Freier, D. A. Driver, R. H. Berg, S. K. Kim, B. Norden, P. E. Nielsen, *Nature* **1993**, *365*, 566–568.
- [7] J. E. Summerton in *Peptide Nucleic Acids, Morpholinos and Related Antisense Biomolecules*, (Eds.: C. G. Janson, M. J. Daring), Springer, New York, **2006**, pp. 89–113.
- [8] R. Kole, A. R. Krainer, S. Altman, *Nat. Rev. Drug Discovery* **2012**, *11*, 125–140.
- [9] Y. Dorsett, T. Tuschl, *Nat. Rev. Drug Discovery* **2004**, *3*, 318–329.
- [10] X. Zhao, Y.-T. Yu, *Nat. Methods* **2008**, *5*, 95–100.
- [11] J. Karijolich, Y.-T. Yu, *Nature* **2011**, *474*, 395–398.
- [12] a) M. Jinek, K. Chylinski, I. Fonfara, M. Hauer, J. A. Doudna, E. Charpentier, *Science* **2012**, *337*, 816–821; b) H. Wang, H. Yang, C. S. Shivalila, M. M. Dawlaty, A. W. Cheng, F. Zhang, R. Jaenisch, *Cell* **2013**, *153*, 910–918.
- [13] K. Nishikura, *Annu. Rev. Biochem.* **2010**, *79*, 321–349.
- [14] J. B. Li, G. M. Church, *Nat. Neurosci.* **2013**, *16*, 1518–1522.
- [15] a) T. Stafforst, M. F. Schneider, *Angew. Chem.* **2012**, *124*, 11329–11332; *Angew. Chem. Int. Ed.* **2012**, *51*, 11166–11169; b) M. F. Schneider, J. Wet-
tengel, P. C. Hoffmann, T. Stafforst, *Nucleic Acids Res.* **2014**, *42*, e87; c) P. Vogel, M. F. Schneider, J. Wettengel, T. Stafforst, *Angew. Chem.* **2014**, *126*, 6382–6386; *Angew. Chem. Int. Ed.* **2014**, *53*, 6267–6271.
- [16] A. Keppler, S. Gendreizig, T. Gronemeyer, H. Pick, H. Vogel, K. Johnsson, *Nat. Biotechnol.* **2002**, *21*, 86–89.
- [17] J. B. Bramsen, M. B. Laursen, A. F. Nielsen, T. B. Hansen, C. Bus, N. Langkjær, B. R. Babu, T. Højland, M. Abramov, A. Van Aerschot, D. Odadzic, R. Smicius, J. Haas, C. Andree, J. Barman, M. Wenska, P. Srivastava, C. Zhou, D. Honcharenko, S. Hess, E. Müller, G. V. Bobkov, S. N. Mikhailov, E. Fava, T. F. Meyer, J. Chattopadhyaya, M. Zerial, J. W. Engels, P. Herdewijn, J. Wengel, J. Kjems, *Nucleic Acids Res.* **2009**, *37*, 2867–2881.
- [18] F. Ruberti, C. Barbato, C. Cogoni, *Exp. Neurol.* **2012**, *235*, 419–426.
- [19] a) A. Bonauer, G. Carmona, M. Iwasaki, M. Mione, M. Koyanagi, A. Fischer, J. Burchfield, H. Fox, C. Doebele, K. Ohtani, E. Chavakis, M. Potente, M. Tjwa, C. Urbich, A. M. Zeiher, S. Dimmeler, *Science* **2009**, *324*, 1710–1713; b) M. Trajkovski, J. Hausser, J. Soutschek, B. Bhat, A. Akin, M. Zavalan, M. H. Heim, M. Stoffel, *Nature* **2011**, *474*, 649–653; c) R. A. Boon, K. Lekushi, S. Lechner, T. Seeger, A. Fischer, S. Heydt, D. Kaluza, K. Treguer, G. Carmona, A. Bonauer, A. J. G. Horrevoets, N. Didier, Z. Girmatsion, P. Biliczki, J. R. Ehrlich, H. A. Katus, O. J. Muller, M. Potente, A. M. Zeiher, H. Hermeking, S. Dimmeler, *Nature* **2013**, *495*, 107–110.
- [20] J. Krützfeldt, N. Rajewsky, R. Braich, K. G. Rajeev, T. Tuschl, M. Manoharan, M. Stoffel, *Nature* **2005**, *438*, 685–689.
- [21] M. F. Montiel-Gonzalez, I. Vallecillo-Viejo, G. A. Yudowski, J. J. C. Rosenthal, *Proc. Natl. Acad. Sci. USA* **2013**, *110*, 18285–18290.
- [22] T. M. Woolf, J. M. Chase, D. T. Stinchcomb, *Proc. Natl. Acad. Sci. USA* **1995**, *92*, 8298–8302.

Received: April 22, 2014

Published online on June 20, 2014

Man. 6: accepted

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

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

Philipp Reautschnig*, Paul Vogel*, and Thorsten Stafforst

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

ABSTRACT

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

Abbreviations: ASO, Antisense oligonucleotide; CD, Cluster of differentiation; CFTR, Cystic fibrosis transmembrane conductance regulator; CRISPR/Cas9, Clustered regularly interspaced short palindromic repeats/CRISPR-associated 9; FDA, US Food and Drug Administration; GalNAc, N-acetyl galactosamine; IVT-mRNA, In-vitro transcribed mRNA; MHC, Major histocompatibility complex; miRNA, microRNA; MOE, 2'-O-methoxyethyl; mRNA, messenger RNA; ψ , pseudouridine; PS, Phosphothioate; RNAi, RNA interference; siRNA, Short interfering RNA; SSO, Splice-switching oligonucleotide; SMN2, Survival of motor neuron 2; TALEN, Transcription activator-like effector nuclease; TLR, Toll-like receptor; $T_H1/2$ cell, Type 1/2 T helper cell; T_R1 , Type 1 regulatory T cell; VEGF, Vascular endothelial growth factor; VEGFR-1, Vascular endothelial growth factor receptor 1; ZFN, Zinc finger nuclease

ARTICLE HISTORY

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

KEYWORDS

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

Introduction

During the last 15 y the diverse roles of RNA in regular but also pathological cellular processes became increasingly clear. RNA is not only a short-lived messenger and part of the translational machinery but RNA contributes significantly to the regulation and diversification of the genetic information. There is now increasing insight into the mechanistic role of defective RNA processing, including (alternative) splicing, modification, translation, and decay for the etiology of various diseases.^{1–4} However, not only mis-regulation and defective processing cause disease, but even RNA species themselves can initiate disease processes independent of their protein-coding function. Nucleotide repeat diseases are typical examples.⁵ To employ this new mechanistic knowledge and to translate it into therapy requires drugs that reliably target nucleic acids in a sequence-specific manner. However, there are only few small molecule drugs that target nucleic acids and those are limited in their capacity of sequence addressing. In contrast, oligonucleotide analogs provide a basis for the rational design of highly sequence-specific drugs to target virtually any cellular nucleic acid in a specific manner.⁶ Classical drugs like small molecules target enzymes and receptors to block or alter their specific functions. In contrast, the interference at the nucleic acid level would allow to manipulate the transcriptome and the proteome itself. This is

not limited to the simple up- or down-regulation of target gene expression. Most appealing is the possibility of actively creating new transcript and protein isoforms with altered properties and functions, for instance by re-programming a protein-coding stretch, or by altering splice sites, modification patterns, polyadenylation states, miRNA binding sites, etc.⁷ Affecting the cell by targeting its nucleic acids clearly enlarges the scope of currently available therapeutic interventions including the causal treatment of some genetic diseases.

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

CONTACT Thorsten Stafforst  thorsten.stafforst@uni-tuebingen.de

*These authors equally contributed to this work.

Published with license by Taylor & Francis Group, LLC © Philipp Reautschnig, Paul Vogel, and Thorsten Stafforst

This is an Open Access article distributed under the terms of the Creative Commons Attribution-Non-Commercial License (<http://creativecommons.org/licenses/by-nc/3.0/>), which permits unrestricted non-commercial use, distribution, and reproduction in any medium, provided the original work is properly cited. The moral rights of the named author(s) have been asserted.

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

In principle, interference with the genetic information could be achieved permanently at the DNA- or transiently at the RNA-level. In this review we will focus on the RNA-level. Even though novel approaches for genome engineering are currently keenly explored,⁹ we believe that it would be foolish to carelessly discard the RNA alternative. With respect to ethical issues and safety aspects, the transient and thus reversible nature of RNA manipulation could turn out as a blessing in disguise. Both, the therapeutic effects and the potential adverse effects, are likely to be tunable. Furthermore, manipulations are conceivable that are inaccessible or difficult to realize on the genome level per se. This includes amino acid changes or transcript level changes that would kill a cell if they are permanently enforced. Potentially lethal interventions on kinases, apoptosis

factors, transcription or translation factors could be realized on the RNA-level suddenly, transiently or partially to obtain a therapeutic effect, for instance. Manipulation at the RNA-level might also be much more efficient compared to HDR-dependent genomic knock-in, which remained persistently inefficient in vivo, in particular in postmitotic tissues like the brain.⁹ For many genetic diseases, which are caused by loss-of-function mutations, a patient would benefit more from a drug that can restore a small fraction (like 5%) of functional gene product in a large fraction of a the tissue than from a drug that can restore full gene function (100%) but only in a small fraction of the tissue. A typical example is cystic fibrosis.¹⁰

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

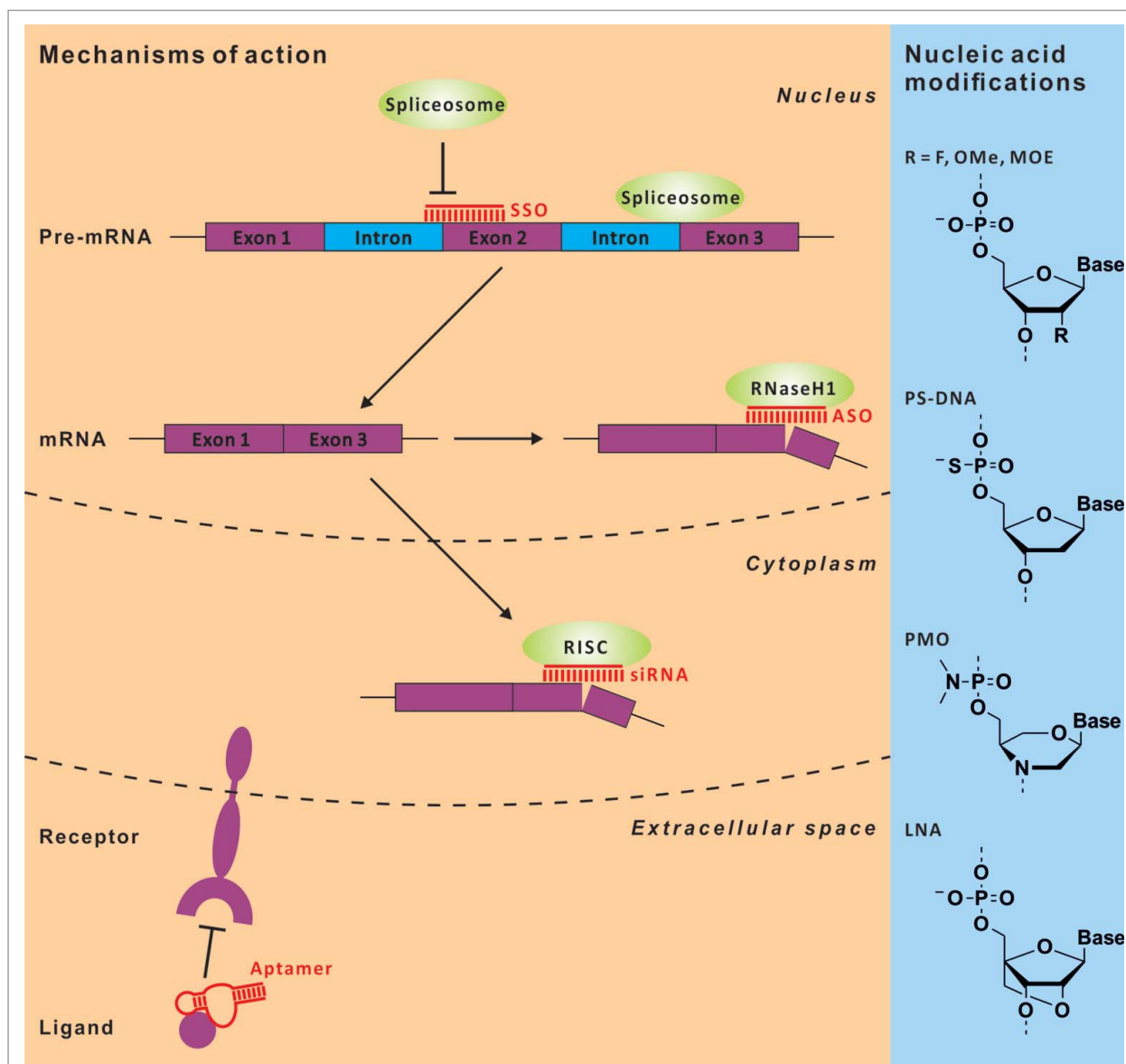


Figure 1. Chemically stabilized, short oligonucleotides can employ various mechanisms for their therapeutic effects ranging from blocking ligand – receptor binding, RNA degradation via RISC or RNaseH(1) recruitment, and alteration of splicing. The classical modes of action are shown on the left panel,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

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

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

Update on established approaches

RNaseH-dependent antisense oligonucleotides

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

Medicinal chemists have undertaken great efforts to improve ASO design regarding nuclease resistance, circulation half-life, target affinity (potency), and tissue specificity. The first ASOs tested in clinical trials, also referred to as 1st generation ASOs, have been modified by oxygen-to-sulfur substitutions in the phosphate backbone. ASOs with such a phosphothioate (PS) backbone show enhanced nuclease resistance and prolonged plasma half-life due to non-specific binding to plasma proteins preventing them from rapid renal filtration. However, numerous toxicities were also associated with that type of modification.⁶ In 1998, fomivirsen was the first FDA-approved ASO and was applied for the treatment of human cytomegalovirus-induced retinitis in HIV patients.¹⁴⁻¹⁶ Marketed as Vitravene, the 21 nt PS-oligonucleotide was administered by intravitreal injection to target the immediate early region 2 of the viral mRNA. Since the approval of fomivirsen, several ASOs belonging to the 1st generation are under clinical review. For instance, targeting the mRNA of intercellular adhesion molecule 1 and the insulin receptor substrate 1 are advanced in the treatment of pouchitis^{17,18} and vascular disorders in the eye,¹⁹⁻²² respectively. The RNaseH-mediated degradation of Akt-1 mRNA to impede tumor proliferation²³ is currently tested for clinical application.²⁴⁻²⁶

Due to the early success with 1st generation ASO, further medicinal chemistry was explored to improve half-life and potency of the drugs in order to reduce the administered dose, the application frequency, the costs, and to minimize adverse effects.²⁷ This resulted in the 2nd generation ASOs, also referred to as gapmers. A typical gapmer is a 20 nt oligonucleotide comprising a PS backbone and 5 flanking 2'-O-methoxyethyl (MOE) groups at both termini. Due to the unmodified internal DNA gap, such ASOs remain good substrates for RNaseH, whereas the terminal MOE modifications increase nuclease resistance and enhances the binding of the ASO to the target mRNA.²⁸ 2nd generation ASOs entered clinical trials for various therapeutic applications. The most prominent representative of the 2nd generation is the MOE gapmer mipomersen as the second FDA-approved RNaseH-dependent ASO. The compound targets apolipoprotein B-100 mRNA and is subcutaneously administered to treat familiar hypercholesterolemia. The genetic disorder is caused by the loss of low-density lipoprotein (LDL) receptor function leading to high LDL cholesterol plasma concentration and early cardiovascular disease. Phase III trials had demonstrated an efficient decrease of LDL cholesterol by lowering ApoB-100 amount in patients obtaining mipomersen.²⁹⁻³¹ The treatment obviously profited from the general pharmacokinetics of systemically administered ASOs which preferably accumulate in the liver where ApoB-100 synthesis takes place.⁸ Recently, an RNase-dependent ASO³² has reached clinical phase III to reduce transthyretin expression in patients suffering from familial amyloid polyneuropathy.³³⁻³⁵ Chemotherapy combined with RNaseH-mediated degradation of clusterin mRNA is a potential therapeutic option in the treatment of prostate³⁶⁻³⁸ and lung cancer.^{39,40}

Generation 2.5 ASO are derived from the traditional gapmer design. For this, the MOE modifications are replaced by 2',4'-constrained ethyl (cEt) bridges in the flanking nucleotides. It was found that cEt-modified oligonucleotides provide the same superior target affinity, but increased nuclease resistance as compared to locked nucleic acid (LNA)-containing oligonucleotides.⁴¹ One of the generation 2.5 ASOs targets the mRNA of signal transducer and activator of transcription 3⁴² and is currently tested for the treatment of various cancer types.⁴³⁻⁴⁶

Most recently, a new chemistry has been developed that strongly increases the liver-specific uptake of oligonucleotide drugs, including ASO and siRNA therapeutics. For this, ASOs⁴⁷ and siRNAs⁴⁸ are conjugated with triantennary N-acetyl galactosamine (GalNAc₃). GalNAc₃ mediates liver-specific uptake through the asialoglycoprotein receptor (ASGPR) that is exclusively expressed on hepatocytes. Marketed as ligand-conjugated antisense (LICA) technology (Ionis Pharmaceuticals), it could be shown that the conjugation increases the potency of MOE gapmers up to 10-fold for inhibiting the expression of hepatic genes in mice.⁴⁹ When using a GalNAc₃-conjugated cEt gapmer, the RNaseH-mediated mRNA degradation was enhanced around 60-fold as compared to the corresponding 2nd generation MOE ASO. Additionally, Ionis Pharmaceuticals announced that its LICA drug targeting apolipoprotein(a) was 30-fold more potent in a phase I study than the unconjugated MOE gapmer.^{50,51}

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

Table 1: Clinical Trials using RNA-Therapeutics*						
RNaseH-dependent Antisense Oligonucleotides						
Compound	Application Route	Target	Indication	Company / Initiator	Phase / Status	Trial ID / Reference
Fomivirsen (Vitravene® , ISIS 2922)	IVT	IE2	CMV retinitis	Ionis Pharmaceuticals /Ciba Vision	FDA Approved	14-16
Mipomersen (KYNAMRO® , ISIS 301012)	SC	ApoB-100	HoFH	Ionis Pharmaceuticals / Genzyme	FDA Approved	29-31
Custirsen (OGX-011)	IV	Clusterin	Prostate cancer	Ionis Pharmaceuticals / OncoGenex Technologies	III - Completed	NCT01188187 ³⁶
Custirsen (OGX-011)	IV	Clusterin	Prostate cancer	Ionis Pharmaceuticals / OncoGenex Technologies	III - Active	NCT01578655 ³⁷
Custirsen (OGX-011)	IV	Clusterin	Lung cancer	Ionis Pharmaceuticals / OncoGenex Technologies	III - Recruiting	NCT01630733 ⁴⁰
Aganirsen (GS-101)	Eye drops	IRS-1	Corneal neovascularization in graft patients	Gene Signal	III - Completed	EudraCT 2008-005388-33 ²¹
Aganirsen (GS-101)	Eye drops	IRS-1	Ischemic central retinal vein occlusion	Gene Signal	II/III - Recruiting	EudraCT 2014-000239-18 ²²
Alicaforsen (ISIS 2302)	Enema	ICAM-1	Pouchitis	Ionis Pharmaceuticals / Atlantic Healthcare	III - Recruiting	NCT02525523 ¹⁸
IONIS-TTRRx (ISIS 420915)	SC	TTR	FAP, TRR Amyloidosis	Ionis Pharmaceuticals / GlaxoSmithKline	II/III - Active	NCT01737398 ³³
IONIS-TTRRx (ISIS 420915)	SC	TTR	FAP, TRR Amyloidosis	Ionis Pharmaceuticals / GlaxoSmithKline	III - Enrolling by invitation	NCT02175004 ³⁴
RX-0210 (Arhexin®)	IV	Akt-1	Pancreatic cancer	Rexahn Pharmaceuticals	II - Completed	NCT01028495 ²⁴
RX-0210 (Arhexin®)	IV	Akt-1	Renal cell cancer	Rexahn Pharmaceuticals	IB/II - Recruiting	NCT02089334 ²⁵
IONIS-STAT3-2.5Rx (AZD9150, ISIS 481464)	IV	STAT3	Gastrointestinal or ovarian cancer	Ionis Pharmaceuticals / AstraZeneca	II - Recruiting	NCT02417753 ⁴³
IONIS-STAT3-2.5Rx (AZD9150, ISIS 481464)	IV	STAT3	Metastatic Squamous Cell Carcinoma of the Head and Neck	Ionis Pharmaceuticals / AstraZeneca	IB/II - Recruiting	NCT02499328 ⁴⁴
IONIS-STAT3-2.5Rx (AZD9150, ISIS 481464)	IV	STAT3	Lymphoma	Ionis Pharmaceuticals / AstraZeneca	I/II - Active	NCT01563302 ⁴⁵
IONIS-STAT3-2.5Rx (AZD9150, ISIS 481464)	IV	STAT3	Lymphoma	Ionis Pharmaceuticals / AstraZeneca	IB - Not yet recruiting	NCT02549651 ⁴⁶
IONIS-APO(a)-LRx (ISIS 681257)	SC	ApoA	Elevated Lipoprotein(a)	Ionis Pharmaceutical / Akcea Therapeutics	I - Completed	NCT02414594 ⁵¹
Splice-switching Oligonucleotides						
Compound	Application Route	Target	Indication	Company / Initiator	Phase / Status	Trial ID / Reference
PRO051/GSK2402968 (drisapersen/kyndrisa®)	SC	Dystrophin exon 51 (skipping)	DMD	BioMarin Pharmaceuticals (developed by Prosenza Therapeutics)/GlaxoSmithKline)	III - Recruiting	NCT01803412 ⁵⁶
PRO051/GSK2402968 (drisapersen/kyndrisa®)	SC	Dystrophin exon 51 (skipping)	DMD	BioMarin Pharmaceuticals (developed by Prosenza Therapeutics)/GlaxoSmithKline)	IIIB - Enrolling by invitation	NCT02636686 ⁵⁷
PRO044/BMN 044	SC/IV	Dystrophin exon 44 (skipping)	DMD	BioMarin Pharmaceuticals (developed by Prosenza Therapeutics)/GlaxoSmithKline)	II - Enrolling by invitation	NCT02329769 ²⁴⁵
PRO045/BMN 045	SC	Dystrophin exon 45 (skipping)	DMD	BioMarin Pharmaceuticals (developed by Prosenza Therapeutics)	IIIB - Active	NCT01826474 ²⁴⁶
PRO053/BMN 053	SC/IV	Dystrophin exon 53 (skipping)	DMD	BioMarin Pharmaceuticals (developed by Prosenza Therapeutics)	I/II - Active	NCT01957059 ²⁴⁷

Compound	Application Route	Target	Indication	Aptamers Indication	Company / Initiator	Phase / Status	Trial ID / Reference
AVI-4685 (eteplirsen)	IV	Dystrophin exon 51 (skipping)	DMD	DMD	Sarepta Therapeutics	III - Recruiting	NCT02255552 ⁵⁹
SRP-4045	IV	Dystrophin exon 45 (skipping)	DMD	DMD	Sarepta Therapeutics	III - Not yet recruiting	NCT02500381 ²⁴⁸
SRP-4053	IV	Dystrophin exon 53 (skipping)	DMD	DMD	Sarepta Therapeutics	III - Not yet recruiting	NCT02500381 ²⁴⁸
DS-5141b	SC	Dystrophin exon 45 (skipping)	DMD	DMD	Daiichi Sankyo / Orphan Disease Treatment Institute	I/II - Recruiting	NCT02667483 ⁶⁵
NS-065/NCNP-01	IV	Dystrophin exon 53 (skipping)	DMD	DMD	Nippon Shinyaku / National Center of Neurology and Psychiatry	II	not yet registered ⁶²
IONIS-SMNRx / ISIS 396443 / ASO-10-27 (nusinersen)	IT	SNM2 exon 7 (retention)	SMA	SMA	Ionis Pharmaceuticals / Biogen	III - Recruiting	NCT02193074 ⁷⁰
IONIS-SMNRx / ISIS396443 / ASO-10-27 (nusinersen)	IT	SNM2 exon 7 (retention)	SMA	SMA	Ionis Pharmaceuticals / Biogen	III - Active	NCT02292537 ⁷¹
Aptamers							
Indication							
Peqaptanib sodium (Macugen [®])	IVT	VEGF	AMD	AMD	Evetech Pharmaceuticals, Pfizer	FDA - Approved	^{75,76}
RB006 (pegnivacogin) as component of REG1	IV	Coagulation factor IXa	CAD	CAD	Regado Bioscience	III - Terminated	NCT01848106 ⁷⁸
E10030 (Fovista [®])	IVT	PDGF	AMD	AMD	Ophthotech	III - Recruiting	NCT01944839 ⁸⁰ NCT01940887 ⁸¹
ARC1905 (Zimura [®])	IVT	complement component C5	AMD	AMD	Ophthotech	I/III - Recruiting	NCT02686658 ⁸³
NOX-A12 (olaptesed pegol)	IV	CXCL12/SDF-1	MM	MM	Noxzon Pharma	IIA - Completed	NCT01521533 ⁸⁸
NOX-A12 (olaptesed pegol)	IV	CXCL12/SDF-1	CLL	CLL	Noxzon Pharma	IIA - Active	NCT01486797 ⁸⁹
NOX-E36 (emapticap pegol)	IV	CCL2/MCP-1	DM2, Albuminuria	DM2, Albuminuria	Noxzon Pharma	IIA - Completed	NCT01547897 ⁹¹
NOX-H94 (lexaptapid pegol)	IV	Hepcidin	Anemia	Anemia	Noxzon Pharma	IIA - Completed	NCT01691040 ⁹²
RNA Interference							
Indication							
Bevasiranib	IVT	VEGF	Macular Degeneration	Macular Degeneration	OPKO Health	III - Terminated	NCT00499590 ¹⁰¹
AGN 211745	IVT	VEGFR-1	CNV, AMD	CNV, AMD	Allergan	II - Terminated	NCT00395057 ¹⁰⁴
ISNP	IVT	p53	AKI	AKI	Quark Pharmaceuticals	I - Terminated	NCT00683553 ¹⁰⁰
PF-04523655	IVT	DDIT4	Diabetic Retinopathy, Diabetes Complications	Diabetic Retinopathy, Diabetes Complications	Quark Pharmaceuticals	II - Terminated	NCT00701181 ¹⁰²
PRO-040201	IV	ApoB	Hypercholesterolemia	Hypercholesterolemia	Arbutus Biopharma	I - Terminated	NCT00927459 ¹⁰³
ND-L02-s0201	IV	HSP47	Hepatic Fibrosis	Hepatic Fibrosis	Nitto Denko	I - Ongoing	NCT02227459 ¹¹⁴
ALN-AS1	SC	ALAS1	AIP	AIP	Alnylam Pharmaceuticals	I - Recruiting	NCT02452372 ¹¹⁶
pbi-shRNA STMN1 LP	IT	STMN1	Advanced / Metastatic Cancer, Solid Tumors	Advanced / Metastatic Cancer, Solid Tumors	Graddals	I - Ongoing	NCT01505153 ¹¹⁹
Bamosiran	Eye drops	ADRB2	OAG, Ocular Hypertension	OAG, Ocular Hypertension	Sylentis	II - Completed	NCT02250612 ¹²¹
ARB-001467	IV	Three sites of the HBV genome	Chronic Hepatitis B	Chronic Hepatitis B	Arbutus Biopharma	II - Recruiting	NCT02631096 ¹¹⁵
ALN-TTRsc/ reusiran	SC	TTR	FAC	FAC	Alnylam Pharmaceuticals	III - Recruiting	NCT02319005 ¹¹⁷
ALN-A13	SC	AT	Hemophilia	Hemophilia	Alnylam Pharmaceuticals	I - Recruiting	NCT02035605 ¹¹⁸
siRNA-transfected PBMCs (APN401)	IV	Cbl-b	Several types of cancer	Several types of cancer	Comprehensive Cancer Center of Wake Forest University	I - Recruiting	NCT02166255 ¹²⁰
QPI-1007	IVT	CASP2	NAION	NAION	Quark Pharmaceuticals	II/III - Recruiting	NCT02341560 ¹²²
MRX34	IV	miRNA34a	Several types of cancer	Several types of cancer	Mirna Therapeutics	I - Recruiting	NCT01829971 ¹²⁸
Targoirs	IV	miRNA16	Malignant Pleural Mesothelioma	Malignant Pleural Mesothelioma	University of Sydney	I - Recruiting	NCT02369198 ¹²⁷
AZD4076	SC	miRNA103/107	NASH	NASH	AstraZeneca	I - Recruiting	NCT02612662 ²⁴⁹
MiraVirsen	SC	miRNA122	Hepatitis C	Hepatitis C	Santaris Pharma A/S	II - Ongoing	NCT02031133 ¹²⁵

(continued)

Table 1. (Continued)

Compound	Application Route	Target	mRNA Therapy		Company / Initiator	Phase / Status	Trial ID / Reference
			Indication	Indication			
CV9104	ID	—	Prostate Cancer	Prostate Cancer	CureVac	II - Ongoing	NCT01817738 ²⁰⁸
CV9202	ID	—	Stage IV NSCLC	Stage IV NSCLC	CureVac	I - Ongoing	NCT01915524 ²⁰⁶
CV7201	IM	—	Rabies	Rabies	CureVac	I - Ongoing	NCT02241135 ²⁰⁴
CV9104	ID	—	Prostate Carcinoma	Prostate Carcinoma	CureVac	II - Ongoing	NCT02140138 ²⁰⁹
IVAC_W_bre1_uid	Not available	—	TNBC	TNBC	BioNtech	I - Not yet Recruiting	NCT02316457 ²⁵⁰
IVAC_W_bre1_uid/IVAC_M_uid	—	—	—	—	—	—	—
IVAC MUTANOME RBL001/RBL002	IN	—	Melanoma	Melanoma	BioNtech	I - Ongoing	NCT02035956 ²¹⁰
Lipo-MERIT	IV	—	Melanoma	Melanoma	BioNtech	I - Recruiting	NCT02410733 ¹⁷²
AGS-003	ID	—	mRCC	mRCC	Argos Therapeutics	III - Ongoing	NCT01582672 ¹⁹⁶
Dendritic cell vaccine (plus temozolomide chemotherapy)	ID	—	GBM	GBM	University Hospital, Antwerp	II - Recruiting	NCT02649582 ²⁰⁰
Dendritic cell vaccine (plus chemotherapy)	IV	—	MPM	MPM	University Hospital, Antwerp	II - Recruiting	NCT02649829 ²⁰¹
pp65 Dendritic cell vaccine	SC	—	GBM	GBM	University of Florida	II - Not yet Recruiting	NCT02465268 ²⁰²
ZFN Modified CD4+ T Cells	Infusion	CCR5	HIV	HIV	University of Pennsylvania	I - Recruiting	NCT02388594 ²⁵¹

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

Splice-switching oligonucleotides

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

Dystrophin, the protein encoded by the DMD gene, is crucial for the integrity of muscle tissue.⁵² In rare cases, newborn males harbor a defect dystrophin gene on their X chromosome. The patients suffer from successive muscle wasting resulting in a premature death due to respiratory or cardiac failure. In most cases, the loss-of-protein-function results from exonic out-of-frame deletions. In many cases the reading frame can be restored by skipping the aberrant exon by addressing a SSO to an internal exonic splicing enhancer.⁵³ The resulting truncated dystrophin protein retains partial function and gives the less severe Becker muscular dystrophy phenotype.⁵⁴ Several SSOs have been developed that are clinically evaluated for the skipping of exons 44, 45, 51, and 53, including drisapersen and eteplirsen (Table 1). Recently, both companies submitted new drug applications for their lead compounds drisapersen⁵⁵⁻⁵⁷ and eteplirsen,^{58,59} both amenable to exon 51 skipping. In case of drisapersen, the FDA rejected the application due to major concerns about the efficacy and safety of the drug.⁶⁰ The high dosage required led to severe adverse effects including renal and vascular injury. To improve efficacy and safety other SSO chemistries might be more successful. Whereas drisapersen is a 20 nt 2'-O-methoxy phosphorothioate RNA analog, eteplirsen is a 30 nt phosphorodiamidate oligomer, a so-called morpholino. The final decision on the efficacy and safety evaluation by the FDA is still pending for eteplirsen. Additionally, a new, morpholino-based SSO for exon 53 skipping is currently under clinical evaluation (NS-065/NCNP-01).^{61,62} For the future, we can hope in new chemistries. A SSO that relies on 2'-O,4'-C-ethylene-bridged nucleosides (ENA oligonucleotides)⁶³ which mediate nuclease resistance and improved binding affinity to RNA has now entered a clinical phase I/II trial for the treatment of DMD (DS-4151b).^{64,65}

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

and safety of the drug candidate,⁶⁹ two phase III trials were recently initiated for evaluating IONIS-SMN_{Rx}.^{70,71}

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

Aptamers

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

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

Currently, several aptamers for the local treatment of eye diseases are in late clinic trials (II and III), for instance the aptamers Fovista⁷⁹⁻⁸¹ and Zimura,^{82,83} which target PDFG (it is a growth factor) and C5, respectively. In combination with VEGF inhibitors they might find application in the treatment of AMD in the near future. To overcome the prevalent problems with toxicity and immunogenicity, NOXXON Pharma develops so-called Spiegelmer therapeutics.⁸⁴ These drugs apply stereochemically inverted nucleotides based on L-ribose instead of the natural D-ribose, can be evolved via SELEX, and are suggested to be resistant against nucleases⁸⁵ and invisible for the immune system.⁸⁶ Currently, 3 Spiegelmer aptamers⁸⁶⁻⁹² are in clinical phase II studies (Table 1).

Therapeutic RNAi

RNA interference (RNAi) is a mechanism of posttranscriptional gene regulation that was discovered in 1998.¹² RNAi can

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

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

However, after recent successes in clinical trials, showing the efficacy of RNAi therapeutics to reduce transthyretin¹⁰⁷ and PCSK9¹⁰⁸ in patients, the interest in RNAi is currently growing and even big pharma including Sanofi and Roche started to invest again.⁹⁸ The initial drawbacks in clinical trials were mostly related to the low efficacy of the drugs, off-target issues and immune-related toxicity.¹⁰⁹ Off-target effects include immune-reactions induced by the siRNA/miRNA precursors, and up- and downregulation of non-target mRNAs due to saturation of the RNAi machinery and off-target binding of the siRNA.¹¹⁰ There is now increasing success in tackling all those issues. Current innovations include chemical modification / sequence optimization of siRNAs and its precursors, and new solutions to the delivery problem. The latter include various forms of (lipid) nanoparticles and bioconjugates. The details of this progress are comprehensively reviewed elsewhere.¹¹⁰⁻¹¹³ Briefly, clinical trials seem more successful when they are confined to readily accessible organs like the liver, cancer, and immune-privileged areas like the eye.¹¹⁴⁻¹²² Whereas the eye is a good target for naked siRNAs, treatment of the liver benefited from lipid-based nanoparticles and the above-mentioned GalNac₃ conjugates.¹¹⁶ In particular the GalNac₃ approach has significantly improved the efficacy of siRNA-conjugates, allowing now the weekly administration of liver-targeting siRNA via subcutaneous injection in non-human primates to knock-down antithrombin to clinically relevant levels.¹¹ Notable in this approach is that it allows to knockdown an essential protein (like antithrombin) in a tunable and reversible manner,

whereas the permanent knock-out of antithrombin (for instance at the DNA-level) is lethal.¹¹ Overall, more than 20 siRNA drugs in various formulations are in clinical trials now (up to phase III, Table 1).¹²³ RNAi-therapy clearly has the potential to tackle currently undruggable diseases and to appear in the clinics soon.

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

Emerging concepts for therapy

Therapeutic mRNA

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

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

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

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

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

- b) Rigorous purification of the mRNA product from unincorporated nucleoside triphosphates, small abortive transcripts, remaining DNA templates, and in particular dsRNA via HPLC (High performance liquid chromatography) was shown to dramatically reduce immunogenicity of the transcripts and can increase the translation 10- to 1000-fold.^{140,141}
- c) Synthetic cap analog structures like ARCA (anti-reverse-cap-analog) can further decrease immune response and improve translation. In contrast to older cap analogs, ARCA is always incorporated in correct orientation.^{142,143} A new ARCA variant contains a phosphothioate that resists enzymatic decapping and can increase the half-life of the mRNA.¹⁴⁴
- d) Computational sequence design allows to reduce the number of particularly immune-stimulatory nucleotides and combinations (like UW, with W = A or U).¹⁴⁵⁻¹⁴⁷ Furthermore, transcript stability can be optimized by the introduction of 3'-UTRs (or some elements) taken from other mammalian or viral genes as well as addition of Poly(A)-tails.¹⁴⁸⁻¹⁵³

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

IVT-mRNA could turn out as a valuable tool for genome editing. Genome editing holds great promise for the treatment of various diseases by a permanent repair of a gene via a site-directed knock-in or knockout.¹⁶² However, the respective nucleases that induce the required double-strand DNA breaks including ZFNs, Talens, and CRISPR/Cas, should not be persistently expressed as this would dramatically increase the chance of off-target genome editing.⁹ Consequently, its delivery as an mRNA is beneficial compared to a DNA vector and also circumvents the typical safety risks of viral and non-viral DNA-based methods like genomic insertion and antivector immunogenicity. Encoding of genome editing tools via IVT-mRNAs has already been widely used to generate transgenic animals.¹⁶³⁻¹⁶⁹ In a proof-of-concept study, gene function was restored via homology-directed promoter exchange in a surfactant-B-deficient mouse model by in-vivo-delivery of the ZFN in form of an IVT-mRNA. However, this required the additional delivery of the repair template (with the promoter) in

form of an AAV6 (Adeno-associated-virus serotype 6).¹⁷⁰ Successful promoter exchange was demonstrated and resulted in a prolonged life of the treated mice. IVT-mRNA encoded Talen have been used successfully to disrupt the CCR5 (CC chemokine receptor type 5) gene via non-homologous-end-joining in the T-cell line PM1. As the loss of CCR5 function confers resistance toward R5-tropic HIV-1 infection, side-directed nucleases are promising to target this infectious disease.¹⁷¹ An initial clinical phase I study is currently starting.¹⁷² As IVT-mRNA is a young field, this study represents the first clinical study that uses IVT-mRNAs, but more are likely to follow soon.

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

Oligonucleotides for vaccination and desensitization

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

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

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

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

Even though cumbersome, the ex-vivo strategy allows to optimize and control mRNA transfection and immune stimulation more carefully. The in-vivo approach, however, is potentially more simple and elegant, but encounters additional problems. Whereas all IVT-mRNA strategies require stable and highly translatable transcripts, the in-vivo strategy requires additionally the immune-stimulatory effect that counteracts translation. It was found that complexation of IVT-mRNA with protamine enhances immunogenicity via TLR-7 activation and simultaneously improves stability, however, with the downside of low antigen expression.²⁰³ Anyway, a combination of protamine-complexed IVT-mRNA together with naked IVT-mRNA of the same sequence turned out to satisfy both needs at the same time: high translation efficiency and immune stimulation. Those self-adjuvanting mRNAs are currently in phase I and II clinical trials against prostate cancer, late stage

lung cancer, and rabies; pre-clinical trials against influenza have been performed.^{183,186,204-209} We wish to mention that also other approaches that apply naked or formulated IVT-mRNAs are in clinical trials, for instance for targeting other cancer entities.^{172,210} Furthermore, non-coding RNA can also be used as a vaccination adjuvant replacing the classical alum salts as adjuvant of protein- or peptide-based vaccines.²¹¹

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

Finally, mRNA vaccines could also be used in allergy treatment to desensitize the immune system against a specific antigen. Desensitization against type-I allergies is typically accomplished through repeated intra-dermal, intra-nodal, or sub-lingual application of allergens. Whereas a strong Immunglobuline E and CD8⁺ T-cell responses is intended during vaccination, desensitization aims to change the T_{H1}/T_{R1} to T_{H2} cell ratio toward T_{H1}/T_{R1} to fine-tune the immune response and to induce tolerance.¹⁷⁸ Application of low-dose IVT-mRNA could be used for that purpose, and there is pre-clinical data that prove efficacy and suggest a long-term protective effect.²¹⁸ One can expect first clinical trials to start within the next few years. Applying mRNA as an anti-allergic vaccine has several advantages compared to the classical allergen extract (like standardized cat extract) or DNA-based vaccines.^{219,220} IVT-mRNA is

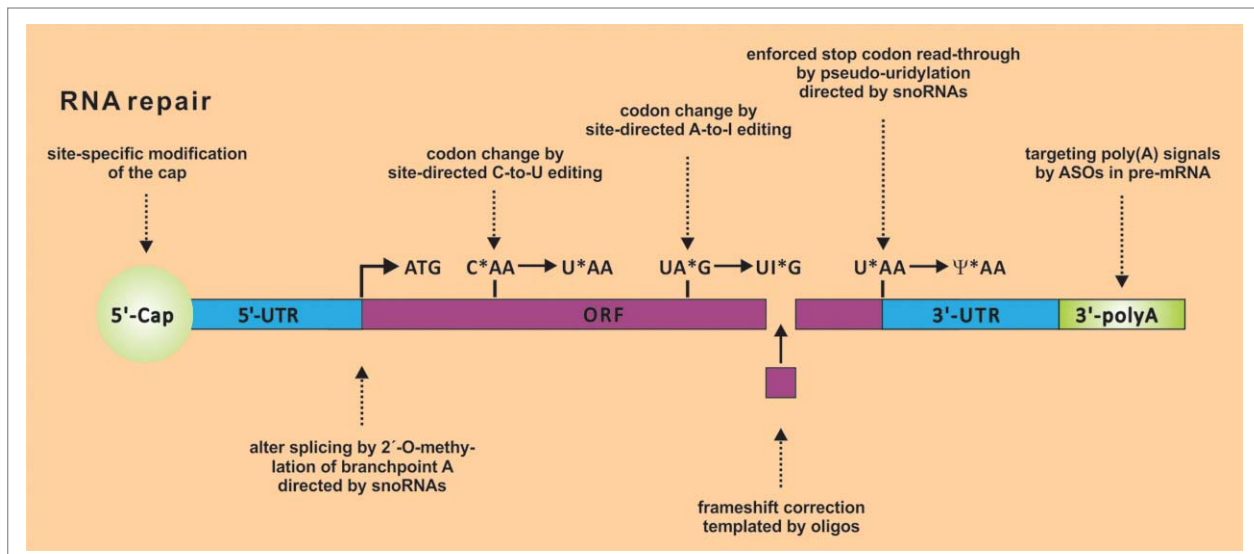


Figure 2. Overview on selected enzymatic processes that could be harnessed to restore gene function by repairing or re-programming mRNA site-specifically. Site-directed A-to-I editing, 2'-O-methylation, pseudouridylation, and frameshift correction via expression or administration of short guideRNAs has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

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

RNA repair

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

Very recently, we and others have engineered artificial RNA-guided editing machineries that allow to re-program genetic information at the RNA level.²²³⁻²²⁵ For this, adenosine-to-inosine (A-to-I) RNA editing enzymes^{226,227} are directed toward specific sites on selected transcripts and allow for the precise posttranscriptional manipulation of the genetic information. The manipulation results from the fact that inosine is biochemically interpreted as guanosine. Thus, formal A-to-G conversions become accessible, in a highly site-specific manner. The specificity comes from the guide-RNA that addresses the editing enzymes and can be readily programmed in rational way, simply by applying Watson-Crick pairing rules.²²⁸ Even though only A-to-G mutations are accessible the scope of manipulations is large. Twelve out of the 20 canonical amino acids can be manipulated, comprising almost all of the polar ones which are essential for protein function.²²³ Furthermore, START and STOP codon, splice elements, polyadenylation signals, and viral RNA are potential targets.^{226,227} We and others have shown that such strategies work inside mammalian cell culture²²⁹ and even in a simple organism²³⁰ and allow the repair of disease-relevant genes, like the CFTR mRNA.²²⁵

Other people have recently shown the possibility of redirecting snoRNA-guided RNA modification machineries, like the 2'-O-methylation²³¹ and the pseudouridylation machinery.¹³⁹ The first modification allows interference with splicing, the second allows the read-through of premature STOP codons. Mammalian cells harbor a plethora of RNA modifying and processing enzymes. There is no need to restrict ourselves to the usage of nucleases, like RISC, RNaseH, and RNaseP.²³² Just to give a few examples, there are RNA editing and modifying enzymes inside the cell that can change nucleotides (A-to-I, C-

to-U²³³, U-to-ψ, A-to-m6A,²³⁴ and many more for the tRNAs²³⁵), that add the cap²³⁶ and the poly(A)-tail,²³⁷ RNAs can be precisely processed, for instance by the CCA-adding enzymes,^{238,239} TUTases,²⁴⁰ etc.²⁴¹ Thus, even complex repair processes are conceivable, including the repair of insertion and deletion mutations at the RNA-level. In this respect, we want to recall a largely overseen work from 2004, done by Paul Zamecnik, the pioneer of antisense therapy, in his early nineties shortly before he passed away. He demonstrated the possibility of repairing the terrible Δ508 deletion mutation in the CFTR gene, the main cause of cystic fibrosis, simply by administration of 2 chemically stabilized RNA oligomers.²⁴² In cell culture, the efficiency of mRNA repair was sufficient to restore the chloride channel function. Unfortunately, he was unable to elucidate the mechanism, but he could clearly demonstrate the repair to take place at the mRNA. Such a complex repair requires a concerted nuclease, ligase (and polymerase) activity at a specific site on an mRNA molecule. In summary, it seems that numerous endogenous enzymes stand ready inside the cell for RNA repair processes. We just have to learn how to make use of them.^{243,244} If successful, one can establish novel platforms for therapeutic intervention.

Conclusions

While splice-switching oligomers and aptamers are still struggling on their ways to the clinic, major progress has been made for RNaseH-dependent ASOs and for therapeutic RNAi with chemically stabilized siRNAs. This is due to the development of new chemistries that improve efficacy and delivery of the drugs to some specific organs. An impressive example is the development of the GalNAC₃ conjugation that clearly improves liver targeting and might allow for the administration of siRNA and ASO by subcutaneous administration in the future. However, overcoming problems with delivery and efficacy remains elusive for many organs and will require massive basic research in the future.

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

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

Disclosure of potential conflicts of interest

No potential conflicts of interest were disclosed.

Funding

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

References

- Klungland A, Dahl JA. Dynamic RNA modifications in disease. *Curr Opin Genet Dev* 2014; 26:47-52; PMID:25005745; <http://dx.doi.org/10.1016/j.gde.2014.05.006>
- Schoenberg DR, Maquat LE. Regulation of cytoplasmic mRNA decay. *Nat Rev Genet* 2012; 13:246-59; PMID:22392217; <http://dx.doi.org/10.1038/nrg3254>
- Chabot B, Shkreta L. Defective control of pre-messenger RNA splicing in human disease. *J Cell Biol* 2016; 212:13-27; PMID:26728853; <http://dx.doi.org/10.1083/jcb.201510032>
- Scheper GC, van der Knaap MS, Proud CG. Translation matters: protein synthesis defects in inherited disease. *Nat Rev Genet* 2007; 8:711-23; PMID:17680008; <http://dx.doi.org/10.1038/nrg2142>
- Nalavade R, Griesche N, Ryan DP, Hildebrand S, Krauss S. Mechanisms of RNA-induced toxicity in CAG repeat disorders. *Cell Death Dis* 2013; 4:e752; PMID:23907466; <http://dx.doi.org/10.1038/cddis.2013.276>
- Bennett CF, Swayze EE. RNA targeting therapeutics: molecular mechanisms of antisense oligonucleotides as a therapeutic platform. *Annu Rev Pharmacol Toxicol* 2010; 50:259-93; PMID:20055705; <http://dx.doi.org/10.1146/annurev.pharmtox.010909.105654>
- Kole R, Krainer AR, Altman S. RNA therapeutics: beyond RNA interference and antisense oligonucleotides. *Nat Rev Drug Discov* 2012; 11:125-40; PMID:22262036; <http://dx.doi.org/10.1038/nrd3625>
- Geary RS. Antisense oligonucleotide pharmacokinetics and metabolism. *Expert Opin Drug Metab Toxicol* 2009; 5:381-91; PMID:19379126; <http://dx.doi.org/10.1517/17425250902877680>
- Cox DB, Platt RJ, Zhang F. Therapeutic genome editing: prospects and challenges. *Nat Med* 2015; 21:121-31; PMID:25654603; <http://dx.doi.org/10.1038/nm.3793>
- Myra Stern and Eric FWF Alton. Use of Liposomes in the Treatment of Cystic Fibrosis. *Gene Therapy in Lung Disease* CRC Press, 2002; 383–396; <http://dx.doi.org/10.3109/9780203908822-18>
- Sehgal A, Barros S, Ivanciu L, Cooley B, Qin J, Racie T, Hettinger J, Carioto M, Jiang Y, Brodsky J, et al. An RNAi therapeutic targeting antithrombin to rebalance the coagulation system and promote hemostasis in hemophilia. *Nat Med* 2015; 21:492-7; PMID:25849132; <http://dx.doi.org/10.1038/nm.3847>
- Fire A, Xu S, Montgomery MK, Kostas SA, Driver SE, Mello CC. Potent and specific genetic interference by double-stranded RNA in *Caenorhabditis elegans*. *Nature* 1998; 391:806-11; PMID:9486653; <http://dx.doi.org/10.1038/35888>
- Wu H, Lima WF, Zhang H, Fan A, Sun H, Crooke ST. Determination of the role of the human RNase H1 in the pharmacology of DNA-like antisense drugs. *J Biol Chem* 2004; 279:17181-9; PMID:14960586; <http://dx.doi.org/10.1074/jbc.M311683200>
- Vitravene Study Group. A randomized controlled clinical trial of intravitreal fomivirsen for treatment of newly diagnosed peripheral cytomegalovirus retinitis in patients with AIDS. *Am J Ophthalmol* 2002; 133:467-74; PMID:11931780; [http://dx.doi.org/10.1016/S0002-9394\(02\)01327-2](http://dx.doi.org/10.1016/S0002-9394(02)01327-2)
- Vitravene Study Group. Randomized dose-comparison studies of intravitreal fomivirsen for treatment of cytomegalovirus retinitis that has reactivated or is persistently active despite other therapies in patients with AIDS. *Am J Ophthalmol* 2002; 133:475-83; PMID:11931781; [http://dx.doi.org/10.1016/S0002-9394\(02\)01326-0](http://dx.doi.org/10.1016/S0002-9394(02)01326-0)
- Vitravene Study Group. Safety of intravitreal fomivirsen for treatment of cytomegalovirus retinitis in patients with AIDS. *Am J Ophthalmol* 2002; 133:484-98; PMID:11931782; [http://dx.doi.org/10.1016/S0002-9394\(02\)01332-6](http://dx.doi.org/10.1016/S0002-9394(02)01332-6)
- Greuter T, Biedermann L, Rogler G, Sauter B, Seibold F. Alicaforsen, an antisense inhibitor of ICAM-1, as treatment for chronic refractory pouchitis after proctocolectomy: A case series. *United European Gastroenterol J* 2016; 4:97-104; PMID:26966529; <http://dx.doi.org/10.1177/2050640615593681>
- Randomized Study of Topical Alicaforsen Enema in Antibiotic Refractory Pouchitis. (2015). Clinical Phase Trial III
- Cloutier F, Lawrence M, Goody R, Lamoureux S, Al-Mahmood S, Colin S, Ferry A, Conduzorgues JP, Hadri A, Cursiefen C, et al. Anti-angiogenic activity of aganirsin in nonhuman primate and rodent models of retinal neovascular disease after topical administration. *Invest Ophthalmol Vis Sci* 2012; 53:1195-203; PMID:22323484; <http://dx.doi.org/10.1167/iovs.11-9064>
- Cursiefen C, Viaud E, Bock F, Geudelin B, Ferry A, Kadlecová P, Lévy M, Al Mahmood S, Colin S, Thorin E, et al. Aganirsin antisense oligonucleotide eye drops inhibit keratitis-induced corneal neovascularization and reduce need for transplantation: the I-CAN study. *Ophthalmology* 2014; 121:1683-92; PMID:24811963; <http://dx.doi.org/10.1016/j.ophtha.2014.03.038>
- A multicentre double-blind randomized study to investigate the efficacy and tolerability of GS-101 eye drops, an antisense oligonucleotide, versus placebo on inhibition of corneal neovascularization, a major risk factor of corneal graft rejection: The I-GRAFT study. (2009). Clinical Trial Phase III
- The STRONG Study. (<http://strong-nvg.com/the-study/>). Clinical Trial Phase II/III
- Yoon H, Kim DJ, Ahn EH, Gellert GC, Shay JW, Ahn CH, Lee YB. Antitumor activity of a novel antisense oligonucleotide against Akt1. *J Cell Biochem* 2009; 108:832-8; PMID:19693774; <http://dx.doi.org/10.1002/jcb.22311>
- A Safety and Efficacy Study of RX-0201 Plus Gemcitabine in Metastatic Pancreatic Cancer. (2009). Clinical Trial Phase II
- Dose-Finding, Safety and Efficacy Study of RX-0201 Plus Everolimus in Metastatic Renal Cell Cancer. (2014). Clinical Trial Phase IB/II
- Tagawa ST, Chatta GS, Mazhari R, Benaim E. Archexin, a novel AKT-1-specific inhibitor for the treatment of metastatic renal cancer: Preliminary phase I data. *ASCO Annual Meeting Proceedings* 2016; (suppl 2S; abstr 550);
- MacLeod AR. Antisense therapies for cancer: bridging the pharmacogenomic divide. *Drug Discovery Today: Therapeutic Strategies* 2013; 10:e157-e63
- Teplova M, Minasov G, Tereshko V, Inamati GB, Cook PD, Manoharan M, Egli M. Crystal structure and improved antisense properties of 2'-O-(2-methoxyethyl)-RNA. *Nat Struct Mol Biol* 1999; 6:535-9; PMID:10360355; <http://dx.doi.org/10.1038/9304>
- Raal FJ, Santos RD, Blom DJ, Marais AD, Charnig M-J, Cromwell WC, Lachmann RH, Gaudet D, Tan JL, Chasan-Taber S, et al. Mipomersen, an apolipoprotein B synthesis inhibitor, for lowering of LDL cholesterol concentrations in patients with homozygous familial hypercholesterolaemia: a randomised, double-blind, placebo-controlled trial. *The Lancet* 2010; 375:998-1006; [http://dx.doi.org/10.1016/S0140-6736\(10\)60284-X](http://dx.doi.org/10.1016/S0140-6736(10)60284-X)
- Stein EA, Dufour R, Gagne C, Gaudet D, East C, Donovan JM, Chin W, Tribble DL, McGowan M. Apolipoprotein B synthesis inhibition with mipomersen in heterozygous familial hypercholesterolemia: results of a randomized, double-blind, placebo controlled trial to assess efficacy and safety as add-on therapy in patients with coronary artery disease. *Circulation* 2012; 126:2283-92; PMID:23060426; <http://dx.doi.org/10.1161/CIRCULATIONAHA.112.104125>
- Santos RD, Duell PB, East C, Guyton JR, Moriarty PM, Chin W, Mittleman RS. Long-term efficacy and safety of mipomersen in patients with familial hypercholesterolaemia: 2-year interim results of an open-label extension. *Eur Heart J* 2015; 36:566-75; PMID:24366918; <http://dx.doi.org/10.1093/eurheartj/ehf549>
- Ackermann EJ, Guo S, Booten S, Alvarado L, Benson M, Hughes S, Monia BP. Clinical development of an antisense therapy for the treatment of transthyretin-associated polyneuropathy. *Amyloid* 2012; 19:43-4; PMID:22494066; <http://dx.doi.org/10.3109/13506129.2012.673140>
- Efficacy and Safety of IONIS-TTR Rx in Familial Amyloid Polyneuropathy. (2012). Clinical Phase trial II/III

34. Open-Label Extension Assessing Long Term Safety and Efficacy of IONIS-TTR Rx in Familial Amyloid Polyneuropathy (FAP). (2012). Clinical Phase Trial III
35. Ionis Pharmaceuticals Press Release. (<http://ir.isispharm.com/phoenix.zhtml?c=222170&p=irol-newsArticle&ID=2105651>). 2015
36. Comparison of Docetaxel/Prednisone to Docetaxel/Prednisone in Combination With OGX-011 in Men With Prostate Cancer (SYNERGY). (2010). Clinical Trial Phase III
37. Comparison of Cabazitaxel/Prednisone Alone or in Combination With Custirsen for 2nd Line Chemotherapy in Prostate Cancer (AFFINITY). (2012). Clinical Trial Phase III
38. OncoGenex Press Release. (<http://ir.oncogenex.com/releasedetail.cfm?ReleaseID=933276>). 2015
39. Laskin JJ, Nicholas G, Lee C, Gitlitz B, Vincent M, Cormier Y, Stephenson J, Ung Y, Sanborn R, Pressnail B, et al. Phase I/II trial of custirsen (OGX-011), an inhibitor of clusterin, in combination with a gemcitabine and platinum regimen in patients with previously untreated advanced non-small cell lung cancer. *J Thorac Oncol* 2012; 7:579-86; PMID:22198426; <http://dx.doi.org/10.1097/JTO.0b013e31823f459c>
40. A Multinational, Randomized, Open-Label Study of Custirsen In Patients With Advanced or Metastatic (Stage IV) Non-Small Cell Lung Cancer. (2012). Clinical Trial Phase III
41. Seth PP, Vasquez G, Allerson CA, Berdeja A, Gaus H, Kinberger GA, Prakash TP, Migawa MT, Bhat B, Swayze EE. Synthesis and biophysical evaluation of 2', 4'-constrained 2' O-methoxyethyl and 2', 4'-constrained 2' O-ethyl nucleic acid analogues. *J Org Chem* 2010; 75:1569-81; PMID:20136157; <http://dx.doi.org/10.1021/jo902560f>
42. Hong D, Kurzrock R, Kim Y, Woessner R, Younes A, Nemunaitis J, Fowler N, Zhou T, Schmidt J, Jo M, et al. AZD9150, a next-generation antisense oligonucleotide inhibitor of STAT3 with early evidence of clinical activity in lymphoma and lung cancer. *Sci Translat Med* 2015; 7:314ra185-314ra185; <http://dx.doi.org/10.1126/scitranslmed.aac5272>
43. AZD9150, a STAT3 Antisense Oligonucleotide, in People With Malignant Ascites. (2015). Clinical Trial Phase II
44. Study to Assess MEDI4736 With Either AZD9150 or AZD5069 in Relapsed Metastatic Squamous Cell Carcinoma of Head & Neck. (2015). Clinical Trial Phase IB/II
45. Phase 1/2, Open-label, Dose-escalation Study of IONIS-STAT3Rx, Administered to Patients With Advanced Cancers. (2012). Clinical Trial Phase I/II
46. MEDI4736 Alone and in Combination With Tremelimumab or AZD9150 in Adult Subjects With Diffuse Large B-cell Lymphoma (D4190C00023). (2015). Clinical Trial Phase IB
47. Prakash PT, Seth PP, Swayze EE, Graham MJ. Compositions and methods for modulating apolipoprotein (a) expression (US. Patent No. Nine,181,550). 2015
48. Nair JK, Willoughby JL, Chan A, Charisse K, Alam MR, Wang Q, Hoekstra M, Kandasamy P, Kel'in AV, Milstein S, et al. Multivalent N-acetylgalactosamine-conjugated siRNA localizes in hepatocytes and elicits robust RNAi-mediated gene silencing. *J Am Chem Soc* 2014; 136:16958-61; PMID:25434769; <http://dx.doi.org/10.1021/ja505986a>
49. Prakash TP, Graham MJ, Yu J, Carty R, Low A, Chappell A, Schmidt K, Zhao C, Aghajan M, Murray HF, et al. Targeted delivery of antisense oligonucleotides to hepatocytes using triantennary N-acetyl galactosamine improves potency 10-fold in mice. *Nucleic Acids Res* 2014; 42:8796-807; PMID:24992960; <http://dx.doi.org/10.1093/nar/gku531>
50. Ionis Pharmaceuticals Press Release. (<http://ir.isispharm.com/phoenix.zhtml?c=222170&p=irol-newsArticle&ID=2110180>) 2015
51. Safety, Tolerability, Pharmacokinetics, and Pharmacodynamics of IONIS APO(a)-LRx in Healthy Volunteers With Elevated Lipoprotein(a). (2015). Clinical Trial Phase I
52. Hoffman EP, Brown RH, Kunkel LM. Dystrophin: the protein product of the Duchenne muscular dystrophy locus. *Cell* 1987; 51:919-28; PMID:3319190; [http://dx.doi.org/10.1016/0092-8674\(87\)90579-4](http://dx.doi.org/10.1016/0092-8674(87)90579-4)
53. Aartsma-Rus A, Fokkema I, Verschuuren J, Ginjaar I, van Deutekom J, van Ommen GJ, den Dunnen JT. Theoretic applicability of antisense-mediated exon skipping for Duchenne muscular dystrophy mutations. *Hum Mutat* 2009; 30:293-9; PMID:19156838; <http://dx.doi.org/10.1002/humu.20918>
54. Koenig M, Beggs A, Moyer M, Scherpf S, Heindrich K, Bettecken T, Meng G, Müller CR, Lindlöf M, Kaariainen H, et al. The molecular basis for Duchenne versus Becker muscular dystrophy: correlation of severity with type of deletion. *Am J Hum Genet* 1989; 45:498; PMID:2491009
55. Voit T, Topaloglu H, Straub V, Muntoni F, Deconinck N, Campion G, De Kimpse SJ, Eagle M, Guglieri M, Hood S, et al. Safety and efficacy of drisapersen for the treatment of Duchenne muscular dystrophy (DEMAND II): an exploratory, randomised, placebo-controlled phase 2 study. *Lancet Neurol* 2014; 13:987-96; PMID:25209738; [http://dx.doi.org/10.1016/S1474-4422\(14\)70195-4](http://dx.doi.org/10.1016/S1474-4422(14)70195-4)
56. A Study of the Safety, Tolerability & Efficacy of Long-term Administration of Drisapersen in US & Canadian Subjects. (2013). Clinical Trial Phase III
57. Extension Study of Drisapersen in DMD Subjects. (2015). Clinical Trial Phase IIIB
58. Mendell JR, Rodino-Klapac LR, Sahenk Z, Roush K, Bird L, Lowes LP, Alfano L, Gomez AM, Lewis S, Kota J, et al. Eteplirsen for the treatment of Duchenne muscular dystrophy. *Ann Neurol* 2013; 74:637-47; PMID:23907995; <http://dx.doi.org/10.1002/ana.23982>
59. Confirmatory Study of Eteplirsen in DMD Patients (PROMOVI). (2014). Clinical Trial Phase III
60. Drisapersen FDA Report. (<http://www.fda.gov/downloads/AdvisoryCommittees/CommitteesMeetingMaterials/Drugs/PeripheralandCentralNervousSystemDrugsAdvisoryCommittee/UCM473737.pdf>). 2015
61. Komaki H, Nagata T, Saito T, Masuda S, Takeshita E, Tachimori H, et al. GP 251-Exon 53 skipping of the dystrophin gene in patients with Duchenne muscular dystrophy by systemic administration of NS-065/NCNP-01: A phase 1, dose escalation, first-in-human study. *Neuromuscular Disorders* 2015; 25:S261-S2; <http://dx.doi.org/10.1016/j.nmd.2015.06.276>
62. Nippon Shinyaku Press Release. (<https://www.nippon-shinyaku.co.jp/english/news/?id=2920>). 2016
63. Koizumi M. 2'-O, 4'-C-Ethylene-bridged nucleic acids (ENA) as next-generation antisense and antigene agents. *Biol Pharm Bull* 2004; 27:453-6; PMID:15056846; <http://dx.doi.org/10.1248/bpb.27.453>
64. Daiichi Sankyo Press Release. (http://www.daiichisankyo.com/media_investors/media_relations/press_releases/detail/006412/160225_635_E.pdf). 2016
65. Study of DS-5141b in Patients With Duchenne Muscular Dystrophy. (2016). Clinical Trial Phase I/II
66. Lefebvre S, Bürglen L, Reboullet S, Clermont O, Burlet P, Viollet L, Benichou B, Cruaud C, Millasseau P, Zeviani M, et al. Identification and characterization of a spinal muscular atrophy-determining gene. *Cell* 1995; 80:155-65; PMID:7813012; [http://dx.doi.org/10.1016/0092-8674\(95\)90460-3](http://dx.doi.org/10.1016/0092-8674(95)90460-3)
67. Lorson CL, Hahnen E, Androphy EJ, Wirth B. A single nucleotide in the SMN gene regulates splicing and is responsible for spinal muscular atrophy. *Proc Natl Acad Sci U S A* 1999; 96:6307-11; PMID:103395583; <http://dx.doi.org/10.1073/pnas.96.11.6307>
68. Hua Y, Vickers TA, Okunola HL, Bennett CF, Krainer AR. Antisense masking of an hnRNP A1/A2 intronic splicing silencer corrects SMN2 splicing in transgenic mice. *Am J Hum Genet* 2008; 82:834-48; PMID:18371932; <http://dx.doi.org/10.1016/j.ajhg.2008.01.014>
69. Ionis Pharmaceuticals. (<http://ir.ionispharma.com/phoenix.zhtml?c=222170&p=irol-newsArticle&ID=2097778>). 2015
70. A Study to Assess the Efficacy and Safety of IONIS-SMN Rx in Infants With Spinal Muscular Atrophy. (2014). Clinical Trial Phase III
71. A Study to Assess the Efficacy and Safety of IONIS-SMN Rx in Patients With Later-onset Spinal Muscular Atrophy. (2014). Clinical Trial Phase III
72. Disterer P, Kryczka A, Liu Y, Badi YE, Wong JJ, Owen JS, Khoo B. Development of therapeutic splice-switching oligonucleotides. *Hum Gene Ther* 2014; 25:587-98; PMID:24826963; <http://dx.doi.org/10.1089/hum.2013.234>
73. Keefe AD, Pai S, Ellington A. Aptamers as therapeutics. *Nat Rev Drug Discov* 2010; 9:537-50; PMID:20592747; <http://dx.doi.org/10.1038/nrd3141>

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

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

149. Peixeiro I, Silva AL, Romao L. Control of human β -globin mRNA stability and its impact on β -thalassemia phenotype. *Haematologica* 2011; 96:905-13; PMID:21357703; <http://dx.doi.org/10.3324/haematol.2010.039206>
150. Waggoner SA, Liebhaber SA. Regulation of α -globin mRNA stability. *Exp Biol Med* (Maywood) 2003; 228:387-95; PMID:12671183
151. Withers JB, Beemon KL. The structure and function of the rous sarcoma virus RNA stability element. *J Cell Biochem* 2011; 112:3085-92; PMID:21769913; <http://dx.doi.org/10.1002/jcb.23272>
152. Gallie DR. The cap and poly(A) tail function synergistically to regulate mRNA translational efficiency. *Genes Dev* 1991; 5:2108-16; PMID:1682219; <http://dx.doi.org/10.1101/gad.5.11.2108>
153. Peng J, Murray EL, Schoenberg DR. In vivo and in vitro analysis of poly(A) length effects on mRNA translation. *Methods Mol Biol* 2008; 419:215-30; PMID:18369986; http://dx.doi.org/10.1007/978-1-59745-033-1_15
154. Balmayer ER, Geiger JP, Aneja MK, Berezanskyy T, Utzinger M, Mykhaylyk O, Rudolph C, Plank C. Chemically modified RNA induces osteogenesis of stem cells and human tissue explants as well as accelerates bone healing in rats. *Biomaterials* 2016; 87:131-46; PMID:26923361; <http://dx.doi.org/10.1016/j.biomaterials.2016.02.018>
155. Elangovan S, Khorsand B, Do AV, Hong L, Dewerth A, Kormann M, Ross RD, Sumner DR, Allamargot C, Salem AK. Chemically modified RNA activated matrices enhance bone regeneration. *J Control Release* 2015; 218:22-8; PMID:26415855; <http://dx.doi.org/10.1016/j.jconrel.2015.09.050>
156. Plews JR, Li J, Jones M, Moore HD, Mason C, Andrews PW, Na J. Activation of pluripotency genes in human fibroblast cells by a novel mRNA based approach. *PloS one* 2010; 5:e14397; PMID:21209933; <http://dx.doi.org/10.1371/journal.pone.0014397>
157. Warren L, Manos PD, Ahfeldt T, Loh YH, Li H, Lau F, Ebina W, Mandal PK, Smith ZD, Meissner A, et al. Highly efficient reprogramming to pluripotency and directed differentiation of human cells with synthetic modified mRNA. *Cell Stem Cell* 2010; 7:618-30; PMID:20888316; <http://dx.doi.org/10.1016/j.stem.2010.08.012>
158. Mandal PK, Rossi DJ. Reprogramming human fibroblasts to pluripotency using modified mRNA. *Nat Protoc* 2013; 8:568-82; PMID:23429718; <http://dx.doi.org/10.1038/nprot.2013.019>
159. Preskey D, Allison TF, Jones M, Mamchaoui K, Unger C. Synthetically modified mRNA for efficient and fast human iPS cell generation and direct transdifferentiation to myoblasts. *Biochem Biophys Res Commun* 2015; PMID:26449459; <http://dx.doi.org/10.1016/j.bbrc.2015.09.102>
160. Simeonov KP, Uppal H. Direct reprogramming of human fibroblasts to hepatocyte-like cells by synthetic modified mRNAs. *PloS One* 2014; 9:e100134; PMID:24963715; <http://dx.doi.org/10.1371/journal.pone.0100134>
161. Kariko K, Muramatsu H, Keller JM, Weissman D. Increased erythropoiesis in mice injected with submicrogram quantities of pseudouridine-containing mRNA encoding erythropoietin. *Mol Ther* 2012; 20:948-53; PMID:22334017; <http://dx.doi.org/10.1038/mt.2012.7>
162. Yin H, Xue W, Chen S, Bogorad RL, Benedetti E, Grompe M, Koteliansky V, Sharp PA, Jacks T, Anderson DG. Genome editing with Cas9 in adult mice corrects a disease mutation and phenotype. *Nat Biotechnol* 2014; 32:551-3; PMID:24681508; <http://dx.doi.org/10.1038/nbt.2884>
163. Wang X, Yu H, Lei A, Zhou J, Zeng W, Zhu H, Niu Y, Shi B, Cai B, Liu J, et al. Generation of gene-modified goats targeting MSTN and FGF5 via zygote injection of CRISPR/Cas9 system. *Sci Rep* 2015; 5:13878; PMID:26354037; <http://dx.doi.org/10.1038/srep13878>
164. Kaneko T, Mashimo T. Simple Genome Editing of Rodent Intact Embryos by Electroporation. *PloS one* 2015; 10:e0142755; PMID:26556280; <http://dx.doi.org/10.1371/journal.pone.0142755>
165. Wang T, Hong Y. Direct gene disruption by TALENs in medaka embryos. *Gene* 2014; 543:28-33; PMID:24713411; <http://dx.doi.org/10.1016/j.gene.2014.04.013>
166. Yang D, Zhang J, Xu J, Zhu T, Fan Y, Fan J, Chen YE. Production of apolipoprotein C-III knockout rabbits using zinc finger nucleases. *J Vis Exp* 2013; 81; e50957; PMID:24301055; <http://dx.doi.org/10.3791/50957>
167. Flisikowska T, Thorey IS, Offner S, Ros F, Lifke V, Zeitler B, Rottmann O, Vincent A, Zhang L, Jenkins S, et al. Efficient immunoglobulin gene disruption and targeted replacement in rabbit using zinc finger nucleases. *PloS One* 2011; 6:e21045; PMID:21695153; <http://dx.doi.org/10.1371/journal.pone.0021045>
168. Williams DA, Thrasher AJ. Concise review: lessons learned from clinical trials of gene therapy in monogenic immunodeficiency diseases. *Stem Cells Transl Med* 2014; 3:636-42; PMID:24682287; <http://dx.doi.org/10.5966/sctm.2013-0206>
169. Schirmbeck R, Reimann J, Kochanek S, Kreppel F. The immunogenicity of adenovirus vectors limits the multispecificity of CD8 T-cell responses to vector-encoded transgenic antigens. *Mol Ther* 2008; 16:1609-16; PMID:18612271; <http://dx.doi.org/10.1038/mt.2008.141>
170. Mahiny AJ, Dewerth A, Mays LE, Alkhaled M, Mothes B, Malaeksefat E, Loretz B, Rottenberger J, Brosch DM, Reautschnig P, et al. In vivo genome editing using nuclease-encoding mRNA corrects SP-B deficiency. *Nat Biotechnol* 2015; 33:584-6; PMID:25985262; <http://dx.doi.org/10.1038/nbt.3241>
171. Allers K, Hutter G, Hofmann J, Lodenkemper C, Rieger K, Thiel E, Schneider T. Evidence for the cure of HIV infection by CCR5Delta32/Delta32 stem cell transplantation. *Blood* 2011; 117:2791-9; PMID:21148083; <http://dx.doi.org/10.1182/blood-2010-09-309591>
172. Evaluation the Safety and Tolerability of i.v. Administration of a Cancer Vaccine in Patients With Advanced Melanoma (LIPOMERIT). (2015). *Clinical Trial Phase I*
173. Wissink EM, Fogarty EA, Grimson A. High-throughput discovery of post-transcriptional cis-regulatory elements. *BMC Genomics* 2016; 17:177; PMID:26941072; <http://dx.doi.org/10.1186/s12864-016-2479-7>
174. Desmet CJ, Ishii KJ. Nucleic acid sensing at the interface between innate and adaptive immunity in vaccination. *Nat Rev Immunol* 2012; 12:479-91; PMID:22728526; <http://dx.doi.org/10.1038/nri3247>
175. Brubaker SW, Bonham KS, Zanoni I, Kagan JC. Innate immune pattern recognition: a cell biological perspective. *Annu Rev Immunol* 2015; 33:257-90; PMID:25581309; <http://dx.doi.org/10.1146/annurev-immunol-032414-112240>
176. Clark GJ, Angel N, Kato M, Lopez JA, MacDonald K, Vuckovic S, Hart DN. The role of dendritic cells in the innate immune system. *Microbes Infect* 2000; 2:257-72; PMID:10758402; [http://dx.doi.org/10.1016/S1286-4579\(00\)00302-6](http://dx.doi.org/10.1016/S1286-4579(00)00302-6)
177. Jacobson JM, Routy JP, Welles S, DeBenedette M, Tcherepanova I, Angel JB, Asmuth DM, Stein DK, Baril JG, McKellar M, et al. Dendritic Cell Immunotherapy for HIV-1 Infection Using Autologous HIV-1 RNA: A Randomized, Double-Blind, Placebo-Controlled Clinical Trial. *J Acquir Immune Defic Syndr* 2016; 72:31-8; PMID:26751016; <http://dx.doi.org/10.1097/QAI.0000000000000926>
178. Weiss R, Scheiblhofer S, Thalhamer J. Allergens are not pathogens: why immunization against allergy differs from vaccination against infectious diseases. *Hum Vaccin Immunother* 2014; 10:703-7; PMID:24280693; <http://dx.doi.org/10.4161/hv.27183>
179. Weiss R, Scheiblhofer S, Roesler E, Ferreira F, Thalhamer J. Prophylactic mRNA vaccination against allergy. *Curr Opin Allergy Clin Immunol* 2010; 10:567-74; PMID:20856111; <http://dx.doi.org/10.1097/ACI.0b013e32833fd5b6>
180. Fotin-Mlecsek M, Duchardt KM, Lorenz C, Pfeiffer R, Ojkic-Zrna S, Probst J, Kallen KJ. Messenger RNA-based vaccines with dual activity induce balanced TLR-7 dependent adaptive immune responses and provide antitumor activity. *J Immunother* 2011; 34:1-15; PMID:21150709; <http://dx.doi.org/10.1097/CJI.0b013e3181f7d8e8>
181. Fotin-Mlecsek M, Zanzinger K, Heidenreich R, Lorenz C, Kowalczyk A, Kallen KJ, et al. mRNA-based vaccines synergize with radiation therapy to eradicate established tumors. *Radiat Oncol* 2014; 9:180; PMID:25127546; <http://dx.doi.org/10.1186/1748-717X-9-180>
182. Schmidt MA, Goodwin TJ. Personalized medicine in human space flight: using Omics based analyses to develop individualized countermeasures that enhance astronaut safety and performance. *Metabolomics* 2013; 9:1134-56; PMID:24273472; <http://dx.doi.org/10.1007/s11306-013-0556-3>

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

- Immunol Res 2015; 2015:797421; PMID:26557723; <http://dx.doi.org/10.1155/2015/797421>
219. Nelson HS, Oppenheimer J, Vatsia GA, Buchmeier A. A double-blind, placebo-controlled evaluation of sublingual immunotherapy with standardized cat extract. *J Allergy Clin Immunol* 1993; 92:229-36; PMID:8349933; [http://dx.doi.org/10.1016/0091-6749\(93\)90166-D](http://dx.doi.org/10.1016/0091-6749(93)90166-D)
 220. Roesler E, Weiss R, Weinberger EE, Fruehwirth A, Stoecklinger A, Mostböck S, Ferreira F, Thalhamer J, Scheibelhofer S. Immunize and disappear-safety-optimized mRNA vaccination with a panel of 29 allergens. *J Allergy Clin Immunol* 2009; 124:1070-7 e1-11; PMID:19665781; <http://dx.doi.org/10.1016/j.jaci.2009.06.036>
 221. Ball T, Sperr WR, Valent P, Lidholm J, Spitzauer S, Ebner C, Kraft D, Valenta R. Induction of antibody responses to new B cell epitopes indicates vaccination character of allergen immunotherapy. *Eur J Immunol* 1999; 29:2026-36; PMID:10382766; [http://dx.doi.org/10.1002/\(SICI\)1521-4141\(199906\)29:06%3c2026::AID-IMMU2026%3e3.0.CO;2-2](http://dx.doi.org/10.1002/(SICI)1521-4141(199906)29:06%3c2026::AID-IMMU2026%3e3.0.CO;2-2)
 222. Van Ree R, Van Leeuwen WA, Dieges PH, Van Wijk RG, De Jong N, Brewczynski PZ, Kroon AM, Schilte PP, Tan KY, Simon-Licht IF, et al. Measurement of IgE antibodies against purified grass pollen allergens (Lol p 1, 2, 3 and 5) during immunotherapy. *Clin Exp Allergy* 1997; 27:68-74; PMID:9117883; <http://dx.doi.org/10.1046/j.1365-2222.1997.d01-416.x>
 223. Stafforst T, Schneider MF. An RNA-deaminase conjugate selectively repairs point mutations. *Angewandte Chemie* 2012; 51:11166-9; PMID:23038402; <http://dx.doi.org/10.1002/anie.201206489>
 224. Schneider MF, Wettengel J, Hoffmann PC, Stafforst T. Optimal guideRNAs for re-directing deaminase activity of hADAR1 and hADAR2 in trans. *Nucleic Acids Res* 2014; 42:e87; PMID:24744243; <http://dx.doi.org/10.1093/nar/gku272>
 225. Montiel-Gonzalez MF, Vallecillo-Viejo I, Yudowski GA, Rosenthal JJ. Correction of mutations within the cystic fibrosis transmembrane conductance regulator by site-directed RNA editing. *Proc Natl Acad Sci U S A* 2013; 110:18285-90; PMID:24108353; <http://dx.doi.org/10.1073/pnas.1306243110>
 226. Bass BL. RNA editing by adenosine deaminases that act on RNA. *Annu Rev Biochem* 2002; 71:817-46; PMID:12045112; <http://dx.doi.org/10.1146/annurev.biochem.71.110601.135501>
 227. Nishikura K. Functions and regulation of RNA editing by ADAR deaminases. *Annu Rev Biochem* 2010; 79:321-49; PMID:20192758; <http://dx.doi.org/10.1146/annurev-biochem-060208-105251>
 228. Vogel P, Stafforst T. Site-directed RNA editing with antagomir deaminases—a tool to study protein and RNA function. *ChemMedChem* 2014; 9:2021-5; PMID:24954543; <http://dx.doi.org/10.1002/cmdc.201402139>
 229. Vogel P, Schneider MF, Wettengel J, Stafforst T. Improving site-directed RNA editing in vitro and in cell culture by chemical modification of the guideRNA. *Angewandte Chemie* 2014; 53:6267-71; PMID:24890431; <http://dx.doi.org/10.1002/anie.201402634>
 230. Hanswillemenke A, Kuzdere T, Vogel P, Jékely G, Stafforst T. Site-Directed RNA Editing in Vivo Can Be Triggered by the Light-Driven Assembly of an Artificial Riboprotein. *J Am Chem Soc* 2015; 137:15875-81; PMID:26594902; <http://dx.doi.org/10.1021/jacs.5b10216>
 231. Zhao X, Yu Y-T. Targeted pre-mRNA modification for gene silencing and regulation. *Nat Methods* 2008; 5:95-100; PMID:18066073; <http://dx.doi.org/10.1038/nmeth1142>
 232. Kole R, Krainer AR, Altman S. RNA therapeutics: beyond RNA interference and antisense oligonucleotides. *Nat Rev Drug Discov* 2012; 11:125-40; PMID:22262036; <http://dx.doi.org/10.1038/nrd3625>
 233. Blanc V, Davidson NO. C-to-U RNA editing: mechanisms leading to genetic diversity. *J Biol Chem* 2003; 278:1395-8; PMID:12446660; <http://dx.doi.org/10.1074/jbc.R200024200>
 234. Liu N, Pan T. N6-methyladenosine-encoded epitranscriptomics. *Nat Struct Mol Biol* 2016; 23:98-102; PMID:26840897; <http://dx.doi.org/10.1038/nsmb.3162>
 235. Machnicka MA, Milanowska K, Oglou OO, Purta E, Kurkowska M, Olchowik A, Januszewski W, Kalinowski S, Dunin-Horkawicz S, Rother KM, et al. MODOMICS: a database of RNA modification pathways—2013 update. *Nucleic Acids Res* 2013; 41:D262-D7; PMID:23118484; <http://dx.doi.org/10.1093/nar/gks1007>
 236. Muttach F, Rentmeister A. A Biocatalytic Cascade for Versatile One-Pot Modification of mRNA Starting from Methionine Analogues. *Angewandte Chemie International Edition* 2016; 55:1917-20; <http://dx.doi.org/10.1002/anie.201507577>
 237. Vickers TA, Wyatt JR, Burckin T, Bennett CF, Freier SM. Fully modified 2' MOE oligonucleotides redirect polyadenylation. *Nucleic Acids Res* 2001; 29:1293-9; PMID:11238995; <http://dx.doi.org/10.1093/nar/29.6.1293>
 238. Xiong Y, Steitz TA. Mechanism of transfer RNA maturation by CCA-adding enzyme without using an oligonucleotide template. *Nature* 2004; 430:640-5; PMID:15295590; <http://dx.doi.org/10.1038/nature02711>
 239. Cho HD, Verlinde CL, Weiner AM. Reengineering CCA-adding enzymes to function as (U, G)- or dC/dCda-adding enzymes or poly (C, A) and poly (U, G) polymerases. *Proc Natl Acad Sci U S A* 2007; 104:54-9; PMID:17179213; <http://dx.doi.org/10.1073/pnas.0606961104>
 240. Trippe R, Guschina E, Hossbach M, Urlaub H, Lührmann R, Benecke B-J. Identification, cloning, and functional analysis of the human U6 snRNA-specific terminal uridylyl transferase. *RNA* 2006; 12:1494-504; PMID:16790842; <http://dx.doi.org/10.1261/rna.87706>
 241. Martin G, Keller W. RNA-specific ribonucleotidyl transferases. *RNA* 2007; 13:1834-49; PMID:17872511; <http://dx.doi.org/10.1261/rna.652807>
 242. Zamecnik PC, Raychowdhury MK, Tabatadze DR, Cantiello HF. Reversal of cystic fibrosis phenotype in a cultured Δ508 cystic fibrosis transmembrane conductance regulator cell line by oligonucleotide insertion. *Proc Natl Acad Sci U S A* 2004; 101:8150-5; PMID:15148387; <http://dx.doi.org/10.1073/pnas.0401933101>
 243. Exploratory Study to Evaluate QR-010 in Subjects With Cystic Fibrosis ΔF508 CFTR Mutation. (2015). Clinical Trial Phase I
 244. Dose Escalation Study of QR-010 in Homozygous ΔF508 Cystic Fibrosis Patients. (2015). Clinical Trial Phase IB
 245. Open Label, Extension Study of PRO044 in Duchenne Muscular Dystrophy (DMD). (2014). Clinical Trial Phase II
 246. Phase IIB Study of PRO045 in Subjects With Duchenne Muscular Dystrophy. (2013). Clinical Trial Phase 2B
 247. A Phase I/II Study of PRO053 in Subjects With Duchenne Muscular Dystrophy (DMD). (2013). Clinical Phase Trial I/II
 248. Study of SRP-4045 and SRP-4053 in DMD Patients (ESSENCE). (2015). Clinical Trial Phase III
 249. A Study to Assess the Safety and Tolerability of Single Doses of AZD4076 in Healthy Male Subjects. (2015). Clinical Trial Phase I
 250. RNA-Immunotherapy of IVAC_W_bre1_uID and IVAC_M_uID (TNBC-MERIT). (2014). Clinical Trial Phase I
 251. A Phase I Study of T-Cells Genetically Modified at the CCR5 Gene by Zinc Finger Nucleases SB-728mR in HIV-Infected Patients. (2015). Clinical Trial Phase I